Modeling of Daily Precipitation Process in the Context of Climate Change

Hoang Lam Nguyen

Doctor of Philosophy

Department of Civil Engineering

McGill University, Montreal

March 2023

A thesis submitted to McGill University in partial fulfillment

of the requirements for the degree of Doctor of Philosophy

Acknowledgements

I would like to express my profound gratitude to my supervisor, Professor Van-Thanh-Van Nguyen for providing me an excellent opportunity to work under his supervision. I will remember his encouragement and patience with me during my research.

I am sincerely thankful to all professors of the Water Resources Group, Department of Civil Engineering at McGill University for all the knowledge I gained in their graduate courses.

My sincere thanks also extend to my colleagues and friends in my research team for their encouragement, discussion, and friendship.

I gratefully acknowledge the support from the Department of Water Resources Engineering of University of Danang - University of Science and Technology, especially Dr. Ngoc-Duong Vo for providing me the data for this research.

Last but not least, sincere thanks to my wife, my little son, and my parents whose understand and encouragement helped me to overcome difficulties during my study and research and have always inspired me to complete my study.

Abstract

Understanding the spatial and temporal variations of the precipitation process is essential for the planning, design, and management of various water resources systems (e.g., urban drainage systems, flood protection dams, etc.). Furthermore, in recent years, climate change impacts on precipitation have been considered as one of the most critical issues for water resources management worldwide. Hence, it is essential to establish the linkage between the large-scale climate variables in the atmosphere with the precipitation characteristics at a local site of interest for impact and adaptation studies. The present study is therefore carried out in order to develop appropriate methods for improving the accuracy of precipitation estimation at a gauged or ungauged local site in the context of a changing climate. This study can be divided into five main parts.

The first part of this research aims to develop a new statistical downscaling (SD) model for describing the linkage between large-scale climate predictors and observed daily precipitation characteristics at a local site. The proposed SD model, referred hereafter as SDGAM, is based on the Generalized Additive Modeling (GAM) method. The feasibility and accuracy of the SDGAM are assessed using the National Center for Environmental Prediction (NCEP) re-analysis data and the observed daily precipitation data available for the 1961–2000 period at ten gauged sites located in Southern Quebec and Ontario, Canada. Results of this illustrative application have indicated that the proposed SDGAM model could provide more accurate results than those given by the currently popular SDSM method in practice.

The second part of this research is to propose a new statistical downscaling approach based on the combination of the *spatial downscaling* method to link large-scale climatic variables provided by Global Climate Models (GCMs) to daily extreme precipitations at a local site using the SDGAM and the *temporal downscaling* procedure to describe the relationships between daily extreme precipitations with sub-daily extreme precipitations using the scaling General Extreme Value distribution and the scaling behavior of the empirical Probability Weighted Moments (GEV/PWM). The proposed approach was assessed using precipitation data from 5 minutes to 24 hours at 10 representative stations across Canada. It was found that the annual maximum precipitation series in Canada displayed different scaling behaviors depending on the locations considered. Intensity-Duration-Frequency (IDF) relations were then constructed for historical period of 1961-2000 and future periods of 2030s, 2060s and 2090s for different Representative Concentration Pathways (RCP 2.6, RCP 4.5 and RCP 8.5).

The third part of this research aims to estimate daily precipitation series for ungauged sites in Vietnam. Initially, daily rainfall series data of 155 stations across Vietnam were employed to identify different homogeneous rainfall regions using the Principal Component Analysis (PCA) method. Daily precipitation series at ungauged sites were then estimated using a proposed two-stage interpolation method to describe the persistence in rainfall occurrences and amounts for the identified rainfall homogenous regions. The jackknife technique was used to represent the ungauged site condition. Results of this study have shown that Vietnam can be identified into 7 homogeneous rainfall regions. In addition, the proposed estimation procedure can provide the estimated daily precipitation series that are statistically similar to the observed data.

The fourth part of this research is to investigate the presence of trends in daily annual maximum precipitation series using the historical rainfall records available from a network of 175 high-quality stations across Canada and the downscaled regional gridded data from the National Aeronautics Space Administration (NASA) Earth Exchange Global Daily Downscaled Projections (NEX-GDDP). The Mann-Kendall non-parametric test was adopted for trend detection of historical observed data and the trends were estimated using Sen's method. The trends were computed for two different periods: historical period from 1950 to 2005 (for all datasets) and future period from 2006 to 2100 (for NEX-GDDP dataset). Results of this study have indicated an increasing trend for most stations across Canada, approximately 55% for the observed historical records, and around 80% for the downscaled regional gridded data. In addition, it was found that the CanESM2 model provided the best results in terms of the mean and standard deviation of daily annual maximum precipitation time series for Canada. In particular, the gridded data for British Colombia (BC) showed a widespread variation among the 21 GCMs considered in NEX-GDDP. Furthermore, a positive trend was found for more than 90% stations for the future period.

The final part of this research is to perform a detailed analysis of the variability in time and in space of the daily annual maximum rainfalls and extreme temperatures over the Montreal region for the present and future climates using the data from two different sources: the Pacific Climate Impacts Consortium (PCIC) and the National Aeronautics Space Administration (NASA) Earth Exchange Global Daily Downscaled Projections (NEX-GDDP). More specifically, the evaluation was based on the climate simulation outputs from ten different Global Climate Models downscaled (i) by PCIC to a regional 1/12-degree grid using the BCCAQ and BCSD methods; and (ii) by NASA to a regional 1/4-degree grid. For the present climate, historical data for the 1961-1990 period from observed weather stations in the Montreal region were also used for this evaluation. For the future climates, climate projections corresponding to the RCP 4.5 scenario for the 2006 – 2100 period were considered. Results of this study have indicated that the downscaled regional gridded data from PCIC are generally more robust and more accurate than those given by NEX-GDDP. However, the downscaled data are different from the observed data at a given station. A bias correction is hence required before these data could be used in planning and design of urban infrastructures.

Résumé

La connaissance sur les variations spatiales et temporelles du processus de précipitation est essentielle pour la planification, la conception et la gestion de divers systèmes de ressources en eau (par exemple, les systèmes de drainage urbain, les barrages de protection contre les inondations, etc.). En outre, ces dernières années, les impacts du changement climatique sur les précipitations ont été considérés comme l'un des problèmes les plus critiques pour la gestion des ressources en eau dans le monde. Par conséquent, il est essentiel d'établir le lien entre les variables climatiques à grande échelle dans l'atmosphère et les caractéristiques des précipitations sur les sites locaux pour les études d'impact et d'adaptation. La présente étude est donc réalisée dans le but de développer des méthodes appropriées pour améliorer la précision de l'estimation des précipitations à un site local jaugé ou non jaugé dans le contexte du changement climatique. Cette étude peut être divisée en cinq grandes parties.

La première partie de cette recherche vise à proposer un nouveau modèle statistique de réduction d'échelle pour décrire le lien entre les prédicteurs climatiques à grande échelle et les caractéristiques des précipitations quotidiennes observées sur un site local. Le modèle proposé, appelé ci-après SDGAM, est basé sur les méthodes d'ajustement du modèle additif généralisé (GAM). La faisabilité et la précision de l'approche suggérée sont évaluées à l'aide des données de réanalyse du Centre national de prévision environnementale (NCEP) et des données de précipitations quotidiennes observées disponibles pour la période 1961-2000 à dix sites jaugés situés dans le sud du Québec et en Ontario, Canada. Les résultats de cette application illustrative ont indiqué que le modèle SDGAM proposé pourrait fournir des résultats plus précis que ceux fournis par la méthode SDSM actuellement la populaire en pratique.

La deuxième partie de cette recherche présente une approche de réduction d'échelle statistique basée sur la combinaison de la méthode de réduction d'échelle spatiale pour relier les variables climatiques à grande échelle fournies par les modèles du climat global (MCG) aux précipitations extrêmes quotidiennes sur un site local en utilisant SDGAM et la procédure de réduction d'échelle temporelle pour décrire les relations entre les précipitations extrêmes quotidiennes avec des précipitations extrêmes sous-journalières en utilisant la distribution des valeurs extrêmes générales (VEG) et le comportement de mise à l'échelle des moments pondérés par la probabilité (VEG/MPP). Le modèle proposé a été évalué à l'aide de données sur les précipitations de 5 minutes à 24 heures à 10 stations représentatives à travers le Canada. On avait constaté que les séries de précipitations maximales annuelles au Canada présentait de multiples comportements d'échelle selon l'emplacement des stations considérées. Les relations intensité-durée-fréquence (IDF) ont ensuite été construites pour la période historique de 1961-2000 et les périodes futures des années 2030, 2060 et 2090 pour différentes voies de concentration représentatives (RCP 2.6, RCP 4.5 et RCP 8.5).

La troisième partie de cette recherche vise à générer des séries de précipitations quotidiennes pour des sites non jaugés au Vietnam. Initialement, les données des séries de précipitations quotidiennes de 155 stations à travers le Vietnam ont été utilisées pour identifier des régions de précipitations homogènes à l'aide de la méthode d'analyse en composantes principales (ACP). Des séries de précipitations quotidiennes en des sites non jaugés sont ensuite générées à l'aide d'une nouvelle méthode d'interpolation en deux étapes pour décrire la dépendance dans l'occurrence des pluies et la quantité de précipitations pour des régions homogènes de précipitations identifiées. La technique du jackknife a été utilisée pour représenter l'état du site non jaugé. Les résultats de cette étude ont montré que le Vietnam peut être identifié en 7 régions de précipitations homogènes. De plus, la méthode d'estimation proposée peut fournir des séries de précipitations quotidiennes qui sont statistiquement semblable aux séries les données observées.

La quatrième partie de cette recherche examine la présence des tendances dans les séries de précipitations annuelles maximales quotidiennes en utilisant les données historiques disponibles aux 175 stations d'observation de haute qualité à travers le Canada et des données régionales maillées à échelle réduite fournies par l'Administration nationale de l'espace aéronautique (NASA) dans le cadre du projet de Projections mondiales à échelle réduite d'Earth Exchange (NEX-GDDP). Le test non paramétrique de Mann-Kendall a été adopté pour la détection des tendances des données observées historiques et les tendances ont été estimées à l'aide de la méthode de Sen. Les tendances ont été calculées pour deux périodes différentes : la période historique de 1950 à 2005 (pour tous les jeux de données) et la période future de 2006 à 2100 (pour le jeu de données NEX-GDDP). Les résultats ont montré une tendance à la hausse à travers le Canada pour la plupart des stations, environ 55% pour les données observées historiques et environ 80% pour les données maillées régionales à échelle réduite. L'étude a également révélé que CanESM2 offre les meilleurs résultats en termes de moyenne et d'écart type des séries chronologiques quotidiennes de précipitations maximales annuelles pour le Canada. En particulier, les données maillées en Colombie-Britannique (C.-B.) ont montré une grande variabilité parmi les 21 MCG de NEX-GDDP. En plus, on avait identifié une tendance positive pour plus de 90% des stations pour la période future.

La dernière partie de cette recherche effectue une analyse détaillée de la variabilité dans le temps et dans l'espace des précipitations maximales annuelles quotidiennes et des températures extrêmes sur la région de Montréal pour les climats présents et futurs en utilisant les données de deux

sources différentes: le Pacific Climate Impacts Consortium (PCIC) et la National Aeronautics Space Administration (NASA) Earth Exchange Global Daily Downscaled Projections (NEX-GDDP). Plus précisément, l'évaluation était basée sur les sorties de simulation climatique de dix modèles climatiques mondiaux différents réduits (i) par le PCIC à une grille régionale de 1/12 degré en utilisant les méthodes BCCAQ et BCSD; et (ii) par la NASA à une grille régionale de 1/4 de degré. Pour le climat actuel, les données historiques pour la période 1961-1990 provenant des stations météorologiques observées dans la région de Montréal ont également été utilisées pour cette évaluation. Pour les climats futurs, des projections climatiques correspondant au scénario RCP 4.5 pour la période 2006–2100 ont été considérées. Les résultats de cette étude ont indiqué que les données maillées régionales à échelle réduite de PCIC sont généralement plus robustes et plus précises que celles fournies par NEX-GDDP. Les données réduites sont cependant différentes des données observées à une station donnée. Une correction de biais est donc nécessaire avant que ces données puissent être utilisées dans la planification et la conception des infrastructures urbaines.

Dedication

To my family.

Table of Content

List of Figures

List of Tables

List of Symbols

F(x)	Cumulative distribution function
P	Daily precipitation
N	Sample size
X	Large-scale atmospheric predictors given by GCM simulations
R	Daily precipitation intensity
X_T	Quantile corresponding to a return period T
X_i^{obs}	Observed quantiles
X_i^{pro}	Projected quantiles
α	Scale parameter of GEV distribution
β	Scaling exponent
λ	Scale factor
ε	Modeling error term
π	Daily probability of wet-day
$\Gamma(\cdot)$	Gamma function
f_{Oi}	smooth function
C_{Ok}	The correction coefficients for Occurrence process
C_{Ak}	The correction coefficients for amount process
r_i	Uniform distributed random number
S_e^i	The standard error of month i^{th}
f	Bias correction coefficient
Z	Normally distributed random number
ξ	Location parameter of GEV distribution

κ	Shape parameter of GEV distribution
η	Modeling error in the precipitation model
θ	Parameter values in likelihood function

List of Abbreviations

AMP	Annual Maximum Precipitation
AMS	Annual Maximum Series
ANUSPLIN	Australian National University Spline
BCCAQ	Bias Correction/Constructed Analogues with Quantile mapping
BCSD	Bias Correction Spatial Disaggregation
CanESM2	The second-generation Canadian Earth System Model
CGCM	Canadian Global Coupled Model
CMIP5	Coupled Model Inter-Comparison Project Phase 5
CORDEX	Coordinated Regional Downscaling Experiment
CSA	Canadian Standard Association
DD	Dynamical Downscaling
GAM	Generalized Additive Models
GCM	General Circulation Models
GEV	Generalized Extreme Values
IDF	Intensity-Duration-Frequency
IDW	Inverse Distance Weighted
IPCC	Intergovernmental Panel on Climate Change
NA	North America
NASA	National Aeronautics and Space Administration
NCEP	National Center for Environmental Prediction
NEX-GDDP	NASA Earth Exchange Global Daily Downscaled Projections
NCM	Non-Central Moments

MOM	Method Of Moments
ON	Ontario
PCA	Principal Component Analysis
PCIC	Pacific Climate Impacts Consortium
PWM	Probability Weighted Moments
QC	Quebec
RCM	Regional Climate Model
RCP	Representative Concentration Pathways
REML	restricted maximum likelihood
RMSE	Root Mean Square Error
RRMSE	Relative Root Mean Square Error
SD	Statistical Downscaling
SDGAM	Statistical Downscaling model using GAM
SDII	Mean precipitation amount at wet days
SDSM	Statistical DownScaling Model
SI	Statistical Indices
STD	Standard Deviation
US	United States
WMO	World Meteorological Organization

Chapter 1: General introduction

1.1 Problem statement

Understanding the spatial-temporal variations of precipitation process is essential for the planning, design, and management of various water resources systems. For instance, daily precipitation time series are commonly used to assess the availability of water resources in a region, and in particular the extreme rainfall amount for a given return period is required for flood design of various hydraulic structures, (e.g., urban drainage systems, flood protection dams, etc.) (Hershfield, 1961; WMO, 2009). Recently, climate change impacts on precipitation have been considered as one of the most critical issues for water resources management around the world (IPCC, 2007; IPCC, 2014). Hence, it is essential to establish the linkage between the large-scale climate variables in the atmosphere with the precipitation characteristics at local sites for impact and adaptation studies.

General Circulation Models (GCMs) have been commonly used for evaluating the effects of climate change under different scenarios of greenhouse gas emissions on the hydrological regime. Although these GCMs have been recognized to be able to represent the main features of the global distribution of basic climate parameters (Randall et al., 2007), they still cannot reproduce well details of regional climate conditions at temporal and spatial scales of relevance to hydrological impacts and adaptation studies (Nguyen et al., 2006). This is because outputs from

GCMs are usually at resolution that is too coarse (as illustrated in Figure 1-1) for many climate change impact studies, generally greater than 2.5° for both latitude and longitude (approximately 250km). To refine the GCM coarse grid resolution climate projection data to much finer spatial resolutions (regional or local scales) for the reliable assessment of climate change impacts, different downscaling methods have been proposed to resolve this scale discrepancy (Wilby et al., 2002; Fowler et al., 2007; Nguyen and Nguyen, 2008; Maraun et al., 2010; Khalili and Nguyen, 2016; Gooré Bi et al., 2017).

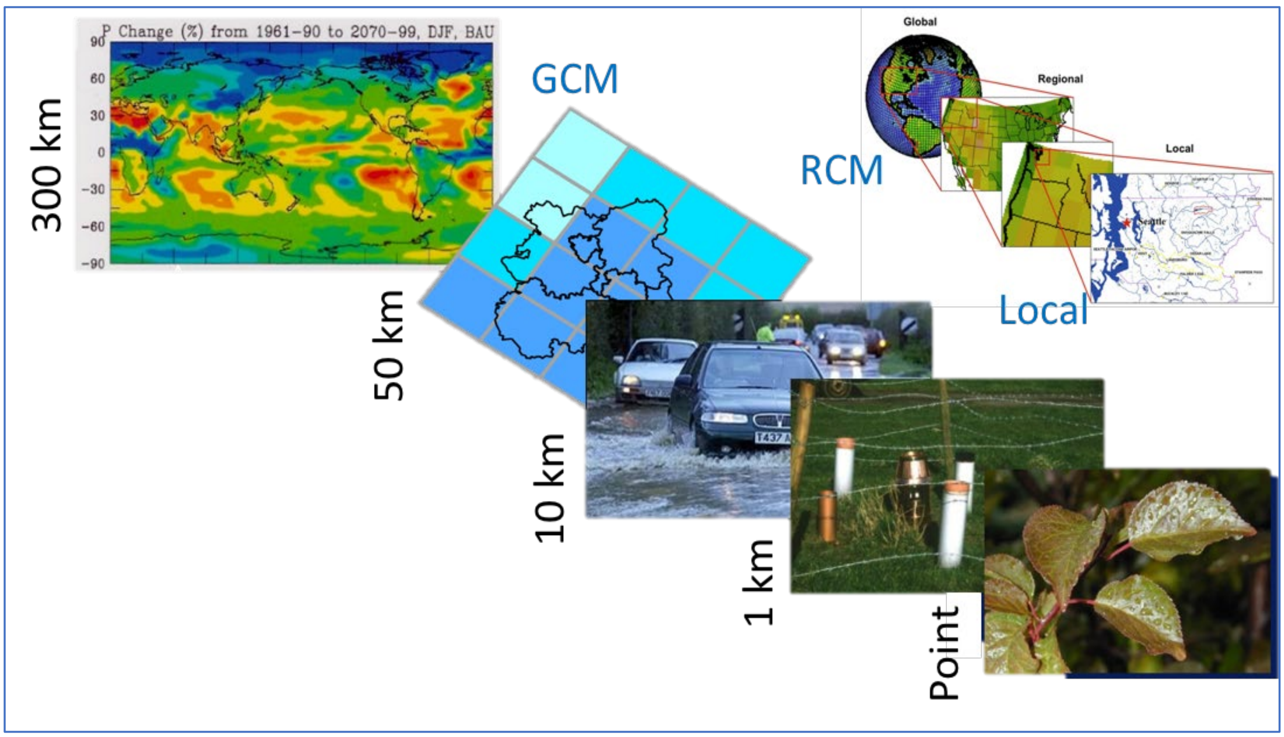


Figure 1-1. Spatial downscaling (Source: P. Gachon & Earthsystemcog.org)

In general, two broad categories of these downscaling procedures currently exist: dynamical downscaling (DD) techniques, involving the extraction of regional scale information from large-scale GCM data based on the modeling of regional climate dynamical processes (Denis et al., 2002; Lenderink et al., 2007), and statistical downscaling (SD) procedures that relied on the empirical relationships between large-scale atmospheric variables and surface environment parameters (Wilby et al., 2004; Diaz-Nieto and Wilby, 2005; Nguyen and Nguyen, 2008; Wilby and Dawson, 2013; Gaur and Simonovic, 2017). It has been widely recognized that the SD methods offer several practical advantages over the DD procedures, especially in terms of flexible adaptation to specific study purposes, and inexpensive computing resource requirement (Xu, 1999; Prudhomme et al., 2002; Wilby et al., 2004; Nguyen et al., 2006). In addition, SD methods are able to account for the observed climate and weather data available at studied sites.

The SD methods can be classified into three sub-categories based on the statistical techniques used: weather typing approaches (Hay et al., 1991; Bárdossy, 1997; Goodess, 1998; Schnur and Lettenmaier, 1998), stochastic weather generators (Richardson, 1981; Semenov and Barrow, 1997); and regression methods (Wilby et al., 2002; Wilby and Dawson, 2013). The major disadvantage of the stochastic weather generators is related to the arbitrary manner of determining the model parameters for future climate conditions, while its of weather classification schemes in the weather typing approaches are somewhat subjective. Of these three approaches, the regression-based SD procedures are more popular because they are relied on the directly derived statistical relationships between large-scale climate predictors and local-scale parameters. The most popular one in this sub-category is the Statistical Downscaling Model SDSM (Wilby et al., 2002) which describes the daily precipitation process including two separate components: the modeling of the occurrence of rainy days using a linear regression technique, and the modeling of the precipitation amount on a rainy day. However, linear regression model in occurrence process fails to describe the probability of a wet day as the value is outside of the range [0,1]. Furthermore, another limitation of SDSM is less accurate in estimating the variance of daily precipitation amounts.

Hence, it is necessary to develop an improved SD model for describing more accurately the daily precipitation processes at a given site.

In addition, in most practical applications, precipitation data at the locations of interest are often limited or unavailable, consequently the existing statistical downscaling approaches proposed for gaged sites cannot be employed. The estimation and prediction of hydrological variables such as precipitation and flow with climate change conditions for these ungauged sites remains a crucial challenge for managing and planning water resources (Sivapalan, 2003). Although several studies have been proposed to assess the impacts of climate change on water resources for ungauged locations (Creutin and Obled, 1982; Besaw et al., 2010; Candela et al., 2012; Yeo and Nguyen, 2014; Bae and Oh, 2017; Nguyen et al., 2018), there is still no general agreement on what the best approach is. Consequently, it is essential to develop a new SD approach to describing more accurately the linkages between the large-scale climate variables given by GCM simulation outputs and the expected daily precipitation characteristics at locations with limited or without historical rainfall data.

1.2 Objectives of the study

In view of the aforementioned issues, the overall objective of the proposed research is to develop innovative modeling approaches to describe accurately statistical and physical properties of the daily precipitation series at a single site or at many sites concurrently in the context of climate change for cases with sufficient rainfall records (gauged sites) and for cases where data are limited or unavailable (ungauged sites) in order to provide suitable tools for high-quality climate change impact assessment studies. More specifically, the proposed study aims at the following objectives:

i) To develop a new SD model for describing accurately the linkage between large-scale climate variables and the local characteristics of the daily precipitation process at a given gauged location;

ii) To develop a new statistical approach to modeling sub-daily extreme rainfall processes in order to improve the accuracy in the estimation of the Intensity-Duration-Frequency (IDF) relations at a given gauged site in the context of a changing climate;

iii) To develop a new SD approach for downscaling the daily precipitation process at an ungauged location based on the rainfall records available at other sites within a given homogeneous region;

iv) To evaluate the accuracy and reliability of regional climate simulations for present and future periods for Canada; and

v) To evaluate the spatial and temporal variability of temperature and precipitation extremes over Montreal region for present and future climates.

1.3 Organization of the thesis and chapter overview

The thesis consists of eight chapters. Chapter 1 provides the general introduction to the current issues related to rainfall modeling in the context of climate change and describes the main objectives of this research. Chapter 2 presents an overview of existing SD models and proposes a new downscaling model for generating daily precipitation series at a single gauged site. Chapter 3 describes a new SD approach (the so-called spatial-temporal downscaling approach) for modeling the sub-daily annual maximum precipitation (AMP) processes in the context of climate change. Chapter 4 proposes a new approach for generating daily precipitation series at an ungauged site based on rainfall information available within the same homogeneous region. Chapter 5 evaluates the accuracy and reliability of regional climate simulations over Canada using the NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP) data for present and future climates. Chapter 6 presents the assessment of the variability of precipitation and temperature extremes over Montreal region for present and future climates. The major conclusions and recommendations for further studies are summarized in Chapter 7. Finally, the statement of originality is detailed in Chapter 8.

Chapter 2: A statistical downscaling model for daily precipitation process at a local site

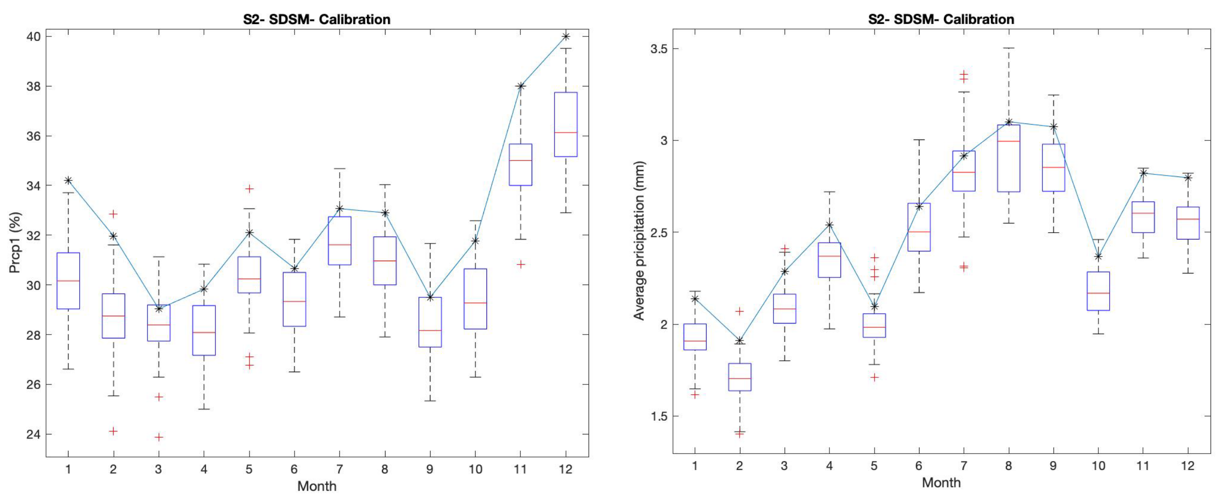
2.1 Introduction

As mentioned in previous section, climate change has been recognized as having a profound impact on the hydrologic cycle at different temporal and spatial scales (Zhang et al., 2011; Arnbjerg-Nielsen et al., 2013; Zhang et al., 2019). Global Climate Models (GCMs) have been commonly used in various studies for assessing the potential impacts of climate change. However, resolutions of outputs from these models are considered too coarse (generally greater than 200km) and hence are not suitable for climate change impact studies at a regional or local scale (Nguyen and Nguyen, 2008). Therefore, it is necessary to develop the linkage between daily climate variables at global scale and the daily precipitation at a local site of interest. If this linkage could be established, then the projected change of climate conditions given by a GCM could be used to predict the resulting change of the local precipitation and the resulting runoff characteristics. Different downscaling techniques have been proposed to downscale these global GCM information to the precipitation series at a local site in several previous studies (Yarnal et al., 2001; Nguyen and Yeo, 2011).

Generally, downscaling techniques can be classified into two broad categories: statistical downscaling (SD) and dynamical downscaling (DD). The DD techniques involve the extraction of regional scale information from large-scale GCM data based on the modeling of regional climate dynamical processes (Denis et al., 2002; Lenderink et al., 2007). Being comprehensive physical models, they are able to provide a more detailed physical understanding of the relationship between the large-scale atmospheric variables and the regional weather conditions. The main disadvantage of DD is the fact that it is computationally intensive and too coarse for local site studies (Xu, 1999). On the other hand, statistical downscaling (SD) procedures rely on the empirical relationships between large-scale atmospheric variables and surface environment parameters (Nguyen and Nguyen, 2008; Wilby and Dawson, 2013). Furthermore, SD methods are flexible to adapt to specific study purposes, and inexpensive computing resource requirement (Wilby et al., 2004; Nguyen et al., 2006). Because of these practical advantages, SD methods have been commonly used in many climate change impact studies in practice.

Depending on the selected statistical techniques, SD methods can be further categorized into three main groups: weather typing, stochastic weather generation, and regression based (Kilsby et al., 1998; Wilks and Wilby, 1999; Yarnal et al., 2001; Fowler et al., 2007; Hessami et al., 2008). Firstly, the weather typing approach classifies days into number of discrete weather conditions; however, this classification is somewhat subjective, and this approach is also computationally intensive for large amount of input observed data (Von Storch et al., 1993). Secondly, stochastic weather generation generates synthetic data series that have similar statistical properties as observed data (Richardson, 1981). The challenge of stochastic approach is to establish the linkage between parameters of these models and large-scale climate variables. Finally, regression-based method establishes empirical relationships between global climate predictors and local predictands (e.g., temperature and precipitation). This approach is simple and straightforward; however, the limitation is related to stationary assumption of regression model parameters.

In general, there is still no general agreement about which downscaling method is the most appropriate approach for describing the observed precipitation characteristics for a given site in the context of climate change, depending mainly on the specific study objectives and the specific climatology of a particular study area (Nguyen and Nguyen, 2008). However, the Statistical Downscaling Model SDSM (Wilby et al., 2002) has been considered as the most popular since it is recommended by the Intergovernmental Panel on Climate Change (IPCC). Significant limitations of the model have been recognized in some previous studies such as: i) the linear multiple regression model used for modeling the precipitation occurrence process could produce some unrealistic results since the probability of rainfall occurrence could be outside of the range [0,1]; and ii) the observed variance of rainfall occurrences and amounts for every month cannot be accurately reproduced (see for instance the results for Dorval station shown in Figure 2-1).



(i) (ii)

Figure 2-1. SDSM model at Dorval station for the period 1961-1980 (i) Occurrence; (ii) Amount; Black markers: Observed data

The present study proposes therefore a new statistical model, hereafter referred to as SDGAM, using the Generalized Additive Modeling (GAM) methods in order to address the shortcomings of the current popular SDSM model. The feasibility and accuracy of the suggested approach are evaluated using the National Center for Environmental Prediction (NCEP) re-analysis data and the observed daily precipitation data available for the 1961–2000 period at ten gauged sites located in Southern Quebec and Ontario, Canada.

2.2 Methodology

2.2.1 Theoretical background

Traditionally, regression analysis is used to describe the linear relationship between the random variable Y (dependent variables) and the random variables X (independent variables) as follow:

$$Y = \beta_0 + \sum_{i=1}^{n} \beta_i X_i + \varepsilon \qquad (2\text{-}1)$$

in which β_i denotes the regression parameters, and ε is the error term. However, the relationship between dependent and independent variables cannot always be represented by a linear behavior, for example the relation between the precipitation and the atmospheric predictors. Hence, one could consider the Generalized Additive Model (GAM), which was first introduced by Hastie and Tibshirani (1986), as an extension of the linear regression methods by replacing the linear relation by the smooth function f_j, as follow:

$$Y = \alpha + \sum_{i=1}^{n} f_i(X_i) + \varepsilon \qquad (2\text{-}2)$$

where α denotes the intercept, X_i denotes independent variables, and ε is error term.

It has been shown that GAM has several advantages over linear regression models because of its flexibility because of the smooth function. Furthermore, data transformation is not required due to the smooth functions f_j. For instance, f_j could be represented by the smooth splines which are curves composed of polynomial functions connected at points named knots. Smooth parameters can be automatically estimated using restricted maximum likelihood (REML) (Wood, 2006). The GAMs have been successfully adopted in some fields of water resources. Villarini and Serinaldi (2012) used GAMs to forecast seasonal rainfall in Romania. Jones et al. (2013) evaluated changes in the frequency and magnitude of extreme daily rainfall in Northern Ireland region. Chebana et al. (2014) applied for regional frequency analysis to estimate flood quantiles at ungauged sites in Canada. Laanaya et al. (2017) proved that GAM outperforms logistic and linear regressions in modeling water temperatures.

In this study, a new statistical downscaling approach, called SDGAM, will be proposed using the GAM for the modeling of the daily rainfall process. Details of the proposed method are provided the following section. The performance of the SDGAM will be assessed using the GAMs package developed by RStudio (RStudio Team, 2020).

2.2.2 Proposed statistical downscaling model for daily precipitation process - SDGAM

- Precipitation Occurrence Process:

$$\widehat{\pi_i} = C_{Ok}(a_0 + \sum_{a=1}^{n} f_{Oi}(X_i)) \tag{2-3}$$

in which f_{Oi}: smooth function

X_i: the large-scale atmospheric predictors given by GCM simulations

C_{Ok}: the correction coefficients for the rainfall occurrence process

r_i is a uniform distributed random number, if $r_i \leq \widehat{\pi_i}$, precipitation occurs at day i

- Precipitation Amount Process (R_i)

$$Y = f.C_{Ak}(\alpha + \sum_{i=1}^{n} f_i(X_i) + \eta_i) \quad (2\text{-}4)$$

in which α: intercept

f_i: smooth function

X_i: the large-scale atmospheric predictors given by GCM simulations

C_{Ak}: the correction coefficients for amount process

$\eta_i = Z * S_e^i$

S_e^i: the standard error of month i^{th}

f: the bias correction coefficient, coming from the deviation of the simulated mean given by GCMs and the estimated mean given by the NCEP re-analysis data. The value of f is set to 1 in the calibration step of the SDGAM model.

$$f = \frac{total\ mean\ by\ NCEP\ for\ calibration\ period}{total\ mean\ by\ GCMs\ for\ calibration\ period}$$

In both precipitation occurrence and amount processes, the correction coefficients C_{Ok} and C_{Ak} represent the difference between the mean of the observed data and the mean of the simulated results based on the regression of GAM for the percentage of wet-day and precipitation amounts, respectively. These coefficients are automatically computed during the calibration of the SDGAM model such that an adequate agreement between the simulated results and the historical data was found. Initially, the values of these coefficients are set to 1 in the calibration step. Figure 2-2 illustrated the steps in the SDGAM model.

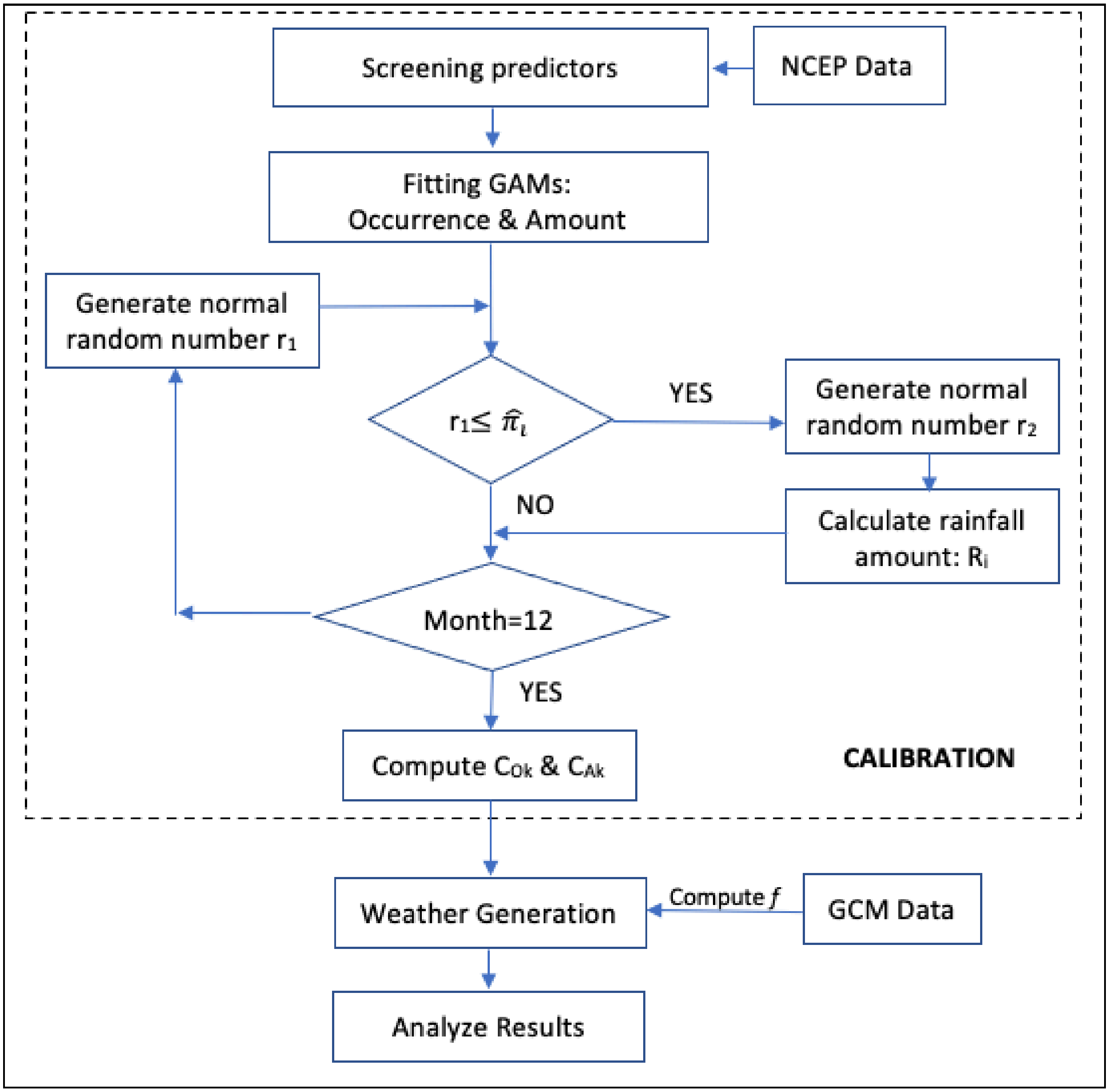


Figure 2-2. Scheme of SDGAM model

2.3 Illustrative application

2.3.1 Data

To assess the accuracy and feasibility of the proposed SDGAM model, a case study was conducted using the NCEP re-analysis data (Kalnay et al., 1996) and the observed daily precipitation data available at 10 stations located in Southern Quebec and Ontario regions, Canada (see Figure 2-3). For comparison purposes, both SDSM and SDGAM models are considered for this study. More specifically, the observed daily precipitation data for the period from 1961 to 2000 were used as detailed in Table 2-1. The 40-year record length are divided into 2 periods: calibration period from 1961 to 1980 and validation period from 1981 to 2000. The NCEP re-analysis data are composed of 26 daily atmospheric variables for the same periods that are selected at grid box covering each of the stations considered (Table 2-2).

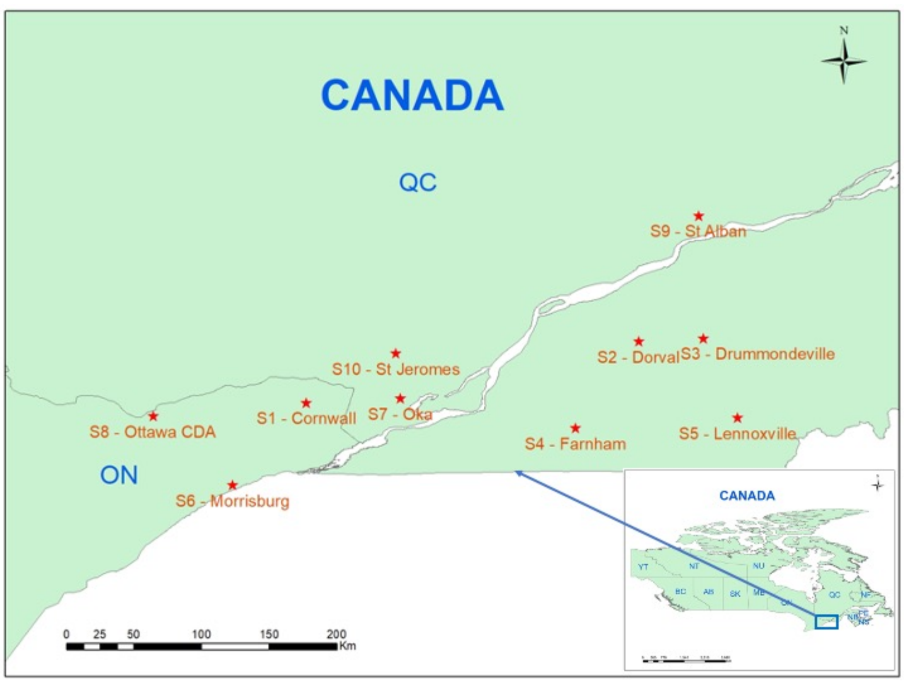


Figure 2-3. Selected stations in Southern Quebec and Ontario, Canada

Table 2-1. Information of rain-gaged stations in Southern Quebec and Ontario, Canada

Code	Site name	Province	Latitude	Longitude	Elevation	Starting
S1	Cornwall	ON	45.47	-74.70	64.0	1961
S2	Dorval	QC	45.88	-72.48	36.0	1961
S3	Drummondeville	QC	45.90	-72.05	82.3	1961
S4	Farnham	QC	45.30	-72.90	68	1961
S5	Lennoxville	QC	45.37	-71.82	181	1961
S6	Morrisburg	ON	44.92	-75.19	81.7	1961
S7	Oka	QC	45.50	-74.07	91.4	1961
S8	Ottawa CDA	ON	45.38	-75.72	79.2	1961
S9	St Alban	QC	46.72	-72.08	76.2	1961
S10	St Jeromes	QC	45.80	-74.10	169.5	1961

Table 2-2. List of atmospheric variables of NCEP re-analysis data in the grid-box

Variable	**Level of measurement**		
Mean sea level pressure	Surface		
Airflow strength	Surface	500 hPa	850 hPa
Zonal velocity	Surface	500 hPa	850 hPa
Meridional velocity	Surface	500 hPa	850 hPa
Vorticity	Surface	500 hPa	850 hPa

Variable	Level of measurement		
Wind direction	Surface	500 hPa	850 hPa
Divergence	Surface	500 hPa	850 hPa
Specific humidity	Near surface	500 hPa	850 hPa
Geopotential height		500 hPa	850 hPa

2.3.2 Evaluation statistical indices

The evaluation of the performance of SDGAM model was carried out in comparison with the SDSM model using different statistical indices as detailed in Table 2-3. These indices were selected to represent the basic statistical properties of the daily precipitation process: average and variance of precipitation, frequency of precipitation occurrence, intensity of precipitation amount, extreme events. For purposes of illustration, results of this comparative study were presented for Dorval station (S2) in the following section, and results of all other stations were presented in Appendix A.

Table 2-3. Evaluation statistics and indices

Categories	Code	Description	Unit	Time scale
Basic variable	Precip_m	Average of precipitation	mm/day	Month
	Precip_std	Standard deviation of precipitation	mm/day	Month
Frequency	PRCP1	Percentage of wet days	%	Season
Intensity	SDII	Mean precipitation amount at wet days	mm/day	Season

Categories	Code	Description	Unit	Time scale
Extreme	CDD	Maximum number of consecutive dry days	days	Season
	PREC90P	90th percentile of rain day amount	mm/day	Season
	AMS	Daily annual maximum precipitation	mm/day	Year
Annual	TAP	Total annual precipitation	mm	Year

In addition, the root-mean-square error (RMSE) was used to compare the performance of the proposed models as given below:

$$RMSE = \sqrt{\frac{1}{N}\sum(SI_{model} - SI_{Observed})^2}$$

where *SI* indicates the value of the statistical indices and *N* is the number of sample size. The smaller value of the RMSE indicates the better accuracy of the model considered.

2.4 Results and discussions

As mentioned above, the comparison of the performance of the SDGAM and SDSM for Dorval station are presented in this section, and the results for other stations are given the Appendix A. In addition, it should be noted that the same screening procedure for selecting the significant climate predictors in the SDSM was also used for the SDGAM. More specifically, the significant climate predictors identified by this screening procedure for both models for Dorval station were the surface zonal velocity, the 850 hPa meridional velocity, the surface precipitation, and the specific humidity at 5000 hPa height. The smooth functions of these predictors for Dorval station are shown in Figure 2-4. It can be seen that the proposed SDGAM model displays the non-linear relations between the climate predictors and the dependent variables. Therefore, the SDGAM could be more flexible in comparison with the current SDSM based on the linear regression.

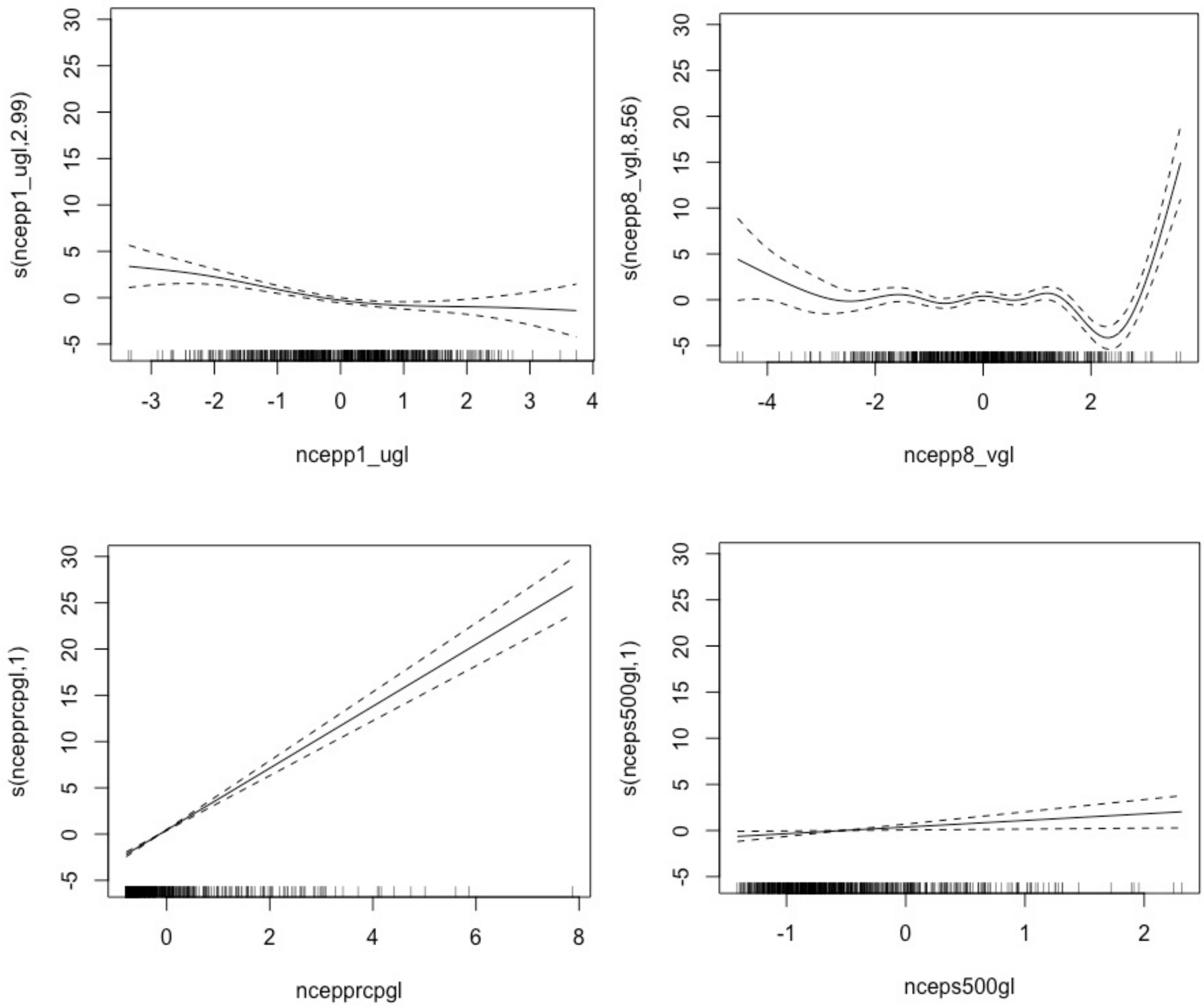


Figure 2-4. Plots of the smooth functions of each variable used in the generalized additive models (GAMs) for Dorval station. Solid lines: fitted smooth curves, Dashed lines: confidence intervals of the predictions

2.4.1 Numerical analysis

Table 2-4 and Table 2-5 demonstrated the computed values of the RMSEs of monthly mean of precipitation (Preceip_m) and monthly standard deviation of precipitation (Precip_std) for the

SDSM and SDGAM models for the calibration (1961-1980) and validation periods (1981-2000) at Dorval station. In these tables, bold numbers denoted the case when the RMSE value of the SDGAM is higher than the value of the SDSM; that is, the SDGAM is less accurate than the SDSM. Regarding the mean and standard deviation of precipitation, it could be seen that proposed SDGAM can provide a significant improvement over the SDSM for all months during the calibration period, and for most months for the validation period (except for October for Precip_m and for January, October and November for Precip_std). In general, it was found that for both calibration and validation steps, the proposed SDGAM model could provide more accurate results than the SDSM in terms of Precip_m and Precip_std.

Table 2-4. RMSEs of Mean of precipitation at Dorval station

Month	Calibration		Validation	
	SDSM	SDGAM	SDSM	SDGAM
Jan	0.839	0.671	0.969	0.952
Feb	0.788	0.715	0.821	0.740
Mar	0.824	0.708	0.954	0.930
Apr	0.984	0.802	1.236	1.123
May	0.830	0.725	0.998	0.930
Jun	1.154	1.083	1.297	1.228
Jul	1.560	1.409	1.652	1.377
Aug	1.299	1.068	1.585	1.477

Month	Calibration		Validation	
	SDSM	SDGAM	SDSM	SDGAM
Sep	1.692	1.551	1.219	1.127
Oct	0.949	0.799	1.021	**1.023**
Nov	0.929	0.867	1.289	1.223
Dec	0.813	0.714	0.883	0.828

Table 2-5. RMSEs of Standard deviation of precipitation at Dorval station

Month	Calibration		Validation	
	SDSM	SDGAM	SDSM	SDGAM
Jan	1.947	1.565	2.006	**2.168**
Feb	1.785	1.628	1.964	1.609
Mar	1.992	1.595	2.318	1.717
Apr	1.939	1.498	2.365	2.086
May	1.743	1.436	2.318	2.258
Jun	2.612	2.327	3.672	3.554
Jul	3.042	2.717	3.297	2.772
Aug	2.873	2.575	3.344	3.070
Sep	4.380	3.655	2.876	2.298
Oct	1.868	1.426	3.163	**3.627**

Month	Calibration		Validation	
	SDSM	SDGAM	SDSM	SDGAM
Nov	1.743	1.422	3.695	**3.901**
Dec	2.046	1.578	1.855	1.382

Regarding the Prcp1, SDII and CDD indices, the proposed SDGAM model could provide significant improvements over the SDSM model. In particular, the SDGAM produced a more accurate result in the simulation of the maximum number of consecutive dry days (CDD), one of the most difficult indices to capture in the modeling process. Regarding the most difficult extreme precipitation index (Prec90p), the SDGAM generally cannot produce an improvement over the SDSM (the same performance for the calibration period but less accuracy for validation period as shown in Table 2-6). Similar results were found for other stations as presented in Appendix A.

Table 2-6. RMSEs of seasonal indices about frequency, intensity, and extreme of precipitation over calibration and validation period Dorval station (S2)

Indices	Season	Calibration		Validation	
		SDSM	SDGAM	SDSM	SDGAM
Prcp1 (%)	Spring	6.380	5.533	6.059	6.027
	Summer	5.435	4.850	6.022	**6.027**
	Fall	6.041	5.998	5.277	5.219
	Winter	5.920	5.076	6.176	5.692
SDII	Spring	1.595	1.543	1.488	1.321

Indices	Season	Calibration		Validation	
		SDSM	SDGAM	SDSM	SDGAM
(mm/wet-day)	Summer	1.550	1.281	2.151	2.070
	Fall	2.393	2.037	2.276	1.858
	Winter	1.396	1.066	1.643	**1.732**
CDD (days)	Spring	7.339	7.277	4.000	3.417
	Summer	5.289	5.287	4.777	**4.832**
	Fall	4.257	3.780	4.729	**4.749**
	Winter	4.414	4.333	4.024	3.746
Prec90p (mm/day)	Spring	7.078	**7.550**	6.271	5.758
	Summer	5.426	5.078	8.632	**8.914**
	Fall	8.058	**8.097**	7.379	**8.158**
	Winter	4.408	4.048	5.962	**7.096**

Table 2-7 shows the RMSEs for annual maximum series (AMS) and total yearly precipitation for all stations. Bold values indicated the better results for SDSM. It could be seen that SDGAM model performs better for all station for calibration period and two-third stations for validation period. In brief, the proposed SDGAM model was able to describe well seasonal features of the extreme precipitation, as well as its frequency and intensity for both calibration and validation periods for rain-gaged stations located in Southern Quebec and Ontario, Canada.

Table 2-7. RMSEs of Daily Annual Maximum Precipitation (AMS) and Total Annual Precipitation (TAP) for calibration and validation period Dorval station (S2)

Station	AMS				TAP			
	Calibration		Validation		Calibration		Validation	
	SDSM	SDGAM	SDSM	SDGAM	SDSM	SDGAM	SDSM	SDGAM
S1	20.28	13.90	14.83	13.23	130.67	118.27	131.44	121.51
S2	15.55	14.15	17.83	**19.72**	112.38	104.01	102.17	92.12
S3	13.59	12.11	19.48	**20.58**	211.90	200.44	221.42	198.77
S4	20.53	17.69	21.99	18.49	116.20	95.76	216.40	**216.88**
S5	11.42	10.52	17.36	15.66	108.21	90.94	162.70	**200.72**
S6	19.14	13.85	19.15	16.30	137.31	123.57	118.75	86.64
S7	13.35	11.72	12.74	**18.36**	105.58	90.05	104.89	99.23
S8	12.00	10.77	17.76	16.98	114.80	105.53	87.94	**99.29**
S9	19.58	16.79	15.35	**16.02**	148.46	146.49	87.96	86.83
S10	14.21	13.43	14.29	13.90	110.40	99.10	131.07	**134.49**

2.4.2 Graphical analysis

A graphical comparison of the accuracy of the SDGAM and SDSM models using the box plots (the closeness between the estimated median value of the model and the observation) and the robustness of the model (the size of the Inter-Quartile Range box) were carried out in this study**.** For purposes of illustration, below figures show the results for the monthly indices of monthly percentage of wet-day (Prcp1) and monthly mean of precipitation (Precip_m) for Dorval Station for the calibration (Figure 2-5 and Figure 2-6) and validation (Figure 2-7 and Figure 2-8) periods. It can be seen that the proposed SDGAM model could reproduce more accurate results than those given by the SDSM for Dorval Station. Figure 2-5 and Figure 2-6 demonstrated that the monthly average of observed daily rainfalls is within the Inter-Quartile Range box of monthly average of generated data for both Precip_m and Prcp1 indices for every single month. The accuracy of the results for the percentage of wet days index (Prcp1) and average precipitation (Precip_m) by the SDGAM could indicate that the use of the GAM modeling approach was more appropriate than the ordinary linear regression used in the SDSM for modeling the precipitation occurrence process.

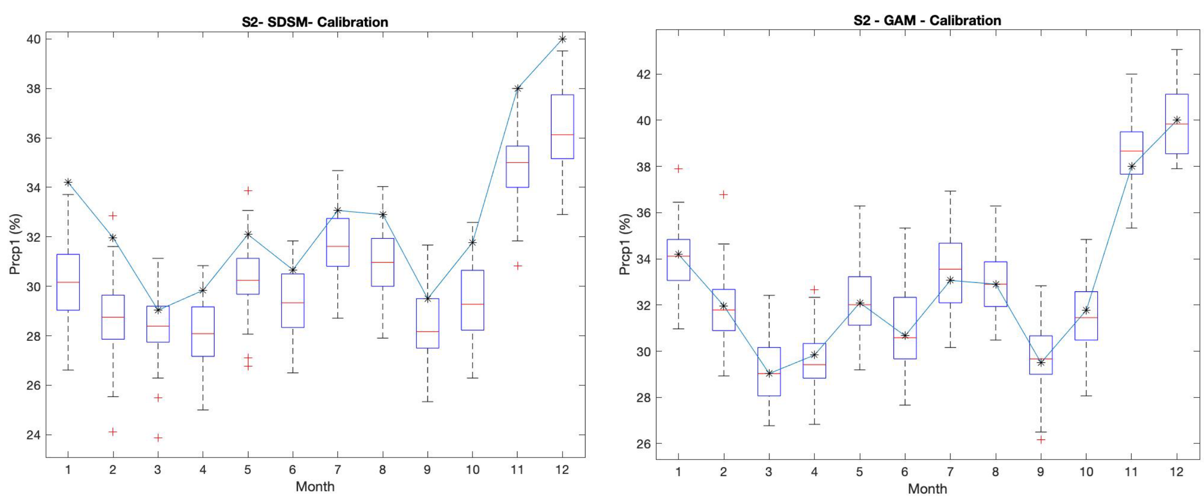


Figure 2-5. Boxplots of monthly percentage of wet-days for SDSM (left) and SDGAM (right) for Dorval station (Black star markers indicate monthly average values of precipitation data, and boxplots indicate model results)

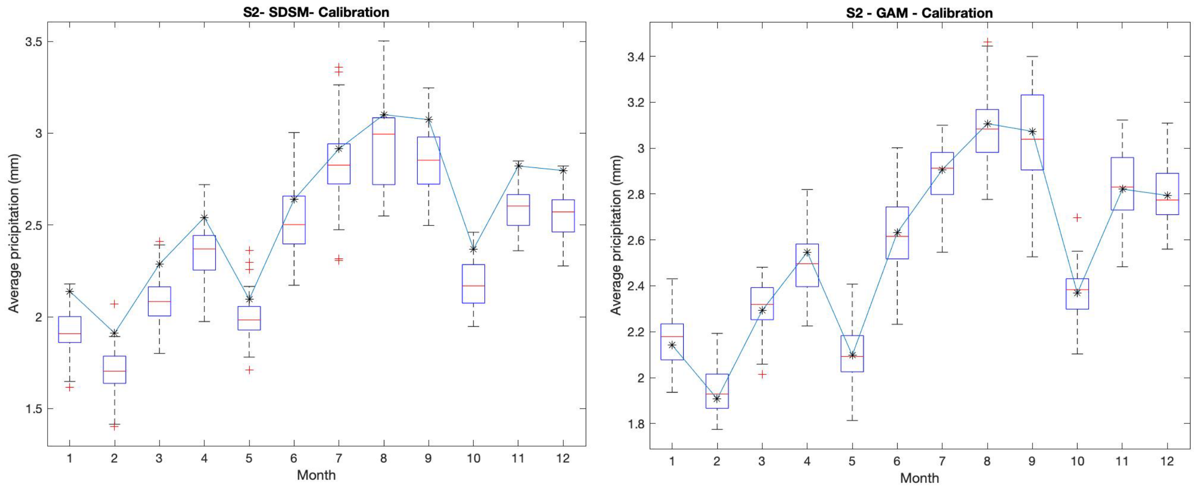


Figure 2-6. Boxplot of monthly means of precipitation for SDSM (left) and SDGAM (right) for Dorval station (Black star markers indicate monthly average values of precipitation data, and boxplots indicate model results)

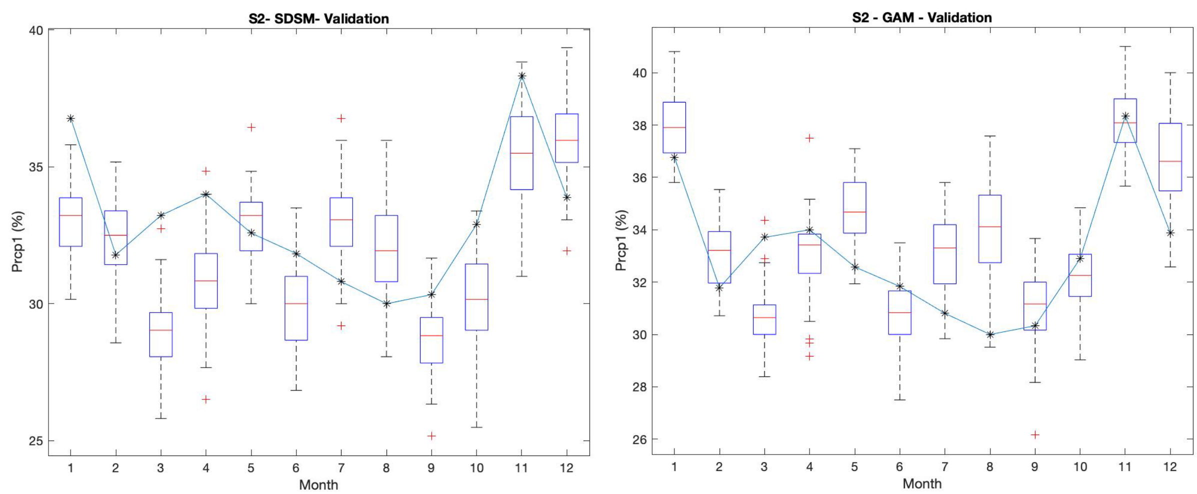


Figure 2-7. Boxplot of percentage of wet days for SDSM (left) and SDGAM (right) for Dorval station for validation period. Black markers: Observed data

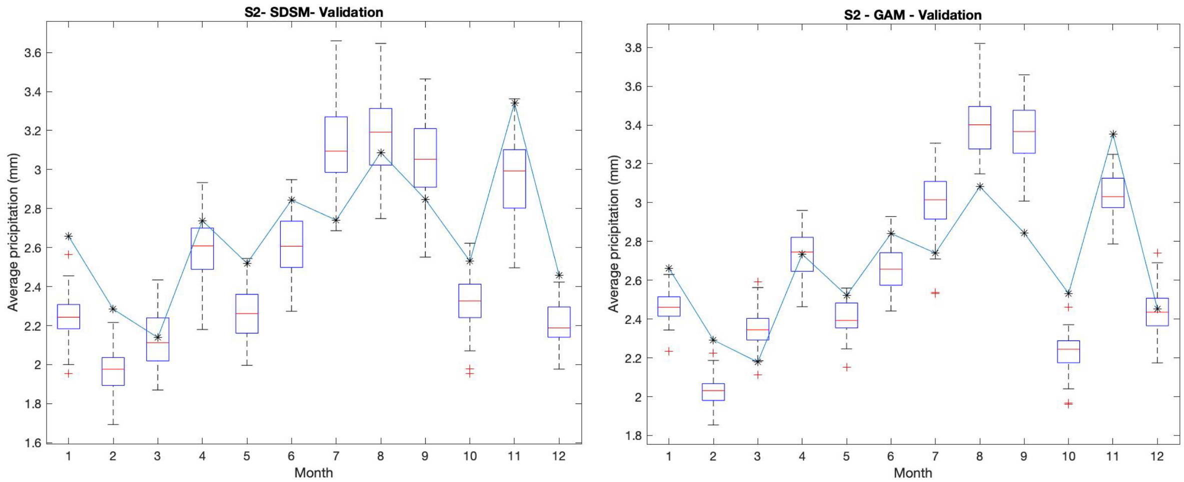
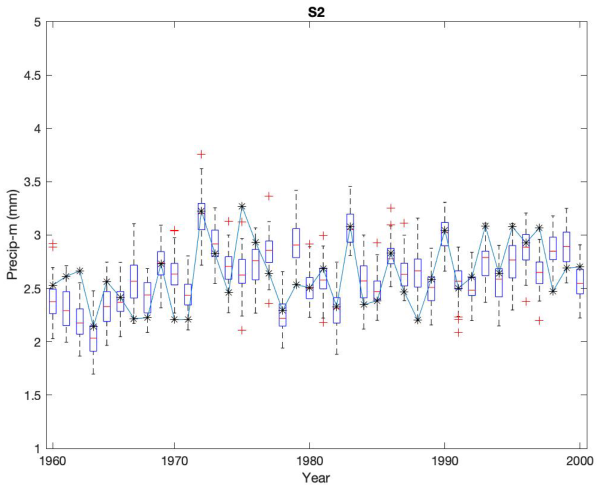
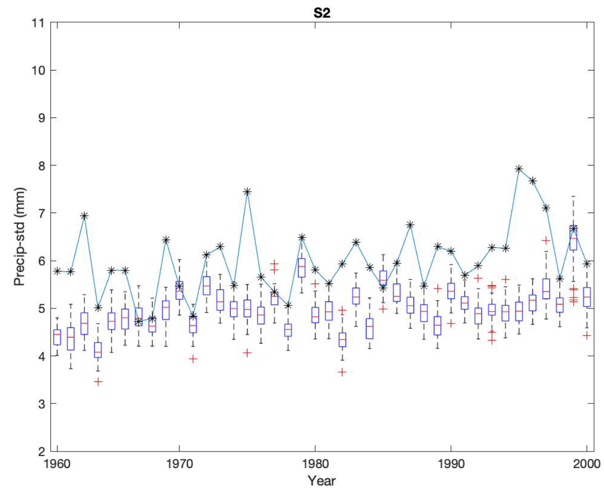


Figure 2-8. Boxplot of monthly means of precipitation for SDSM (left) and SDGAM (right) for Dorval station for validation period. Black markers: Observed data

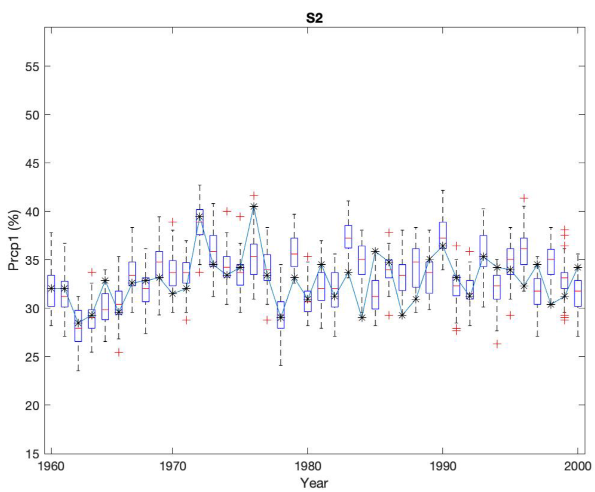
Figure 2-9 presents the boxplots of six common annual indices: mean of precipitation (Precip_m), standard deviation of precipitation (Precip_std), percentage of wet-day (Prcp1), mean of precipitation on wet-days (SDII), consecutive of dry days (CDD), 90th quantiles of wet-days (Prec90p), Annual Maximum Series (AMS), and Total Annual Precipitation (TAP) for the entire record length from 1961 to 2000 at Dorval station (S2). It can be seen that the SDGAM model can capture well the Precip_m, Prcp1, SDII and CDD indices but Precip_std and Prec90p indices are still underestimated. Similar results for all other stations can be found in Appendix A (*Figure A-3* to *Figure A-11*).



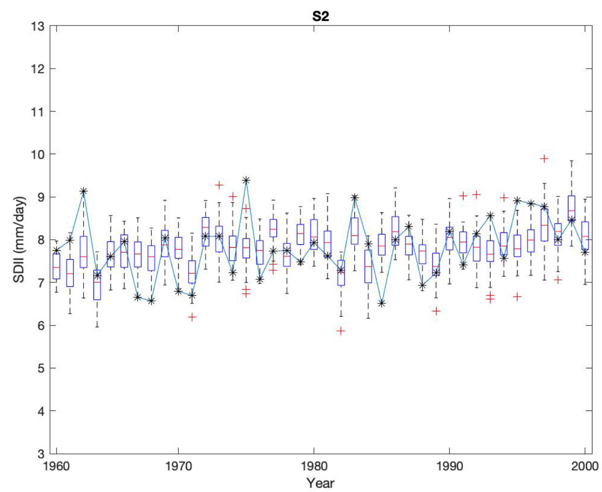
Precip-m



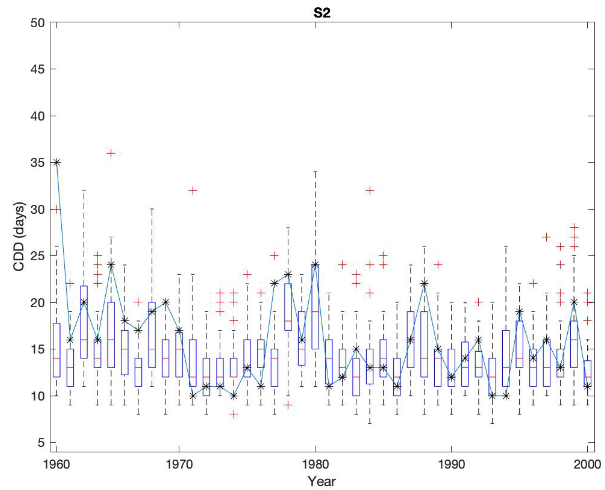
Precip-std



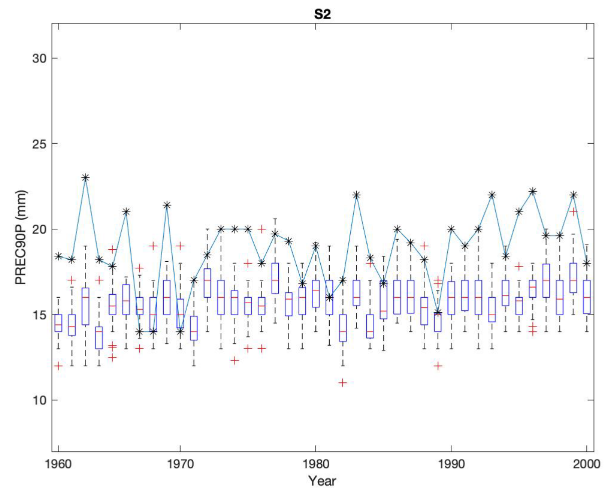
Prcp1



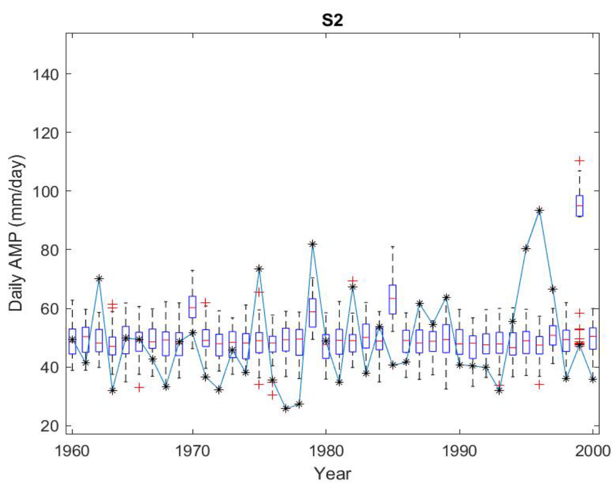
CDII



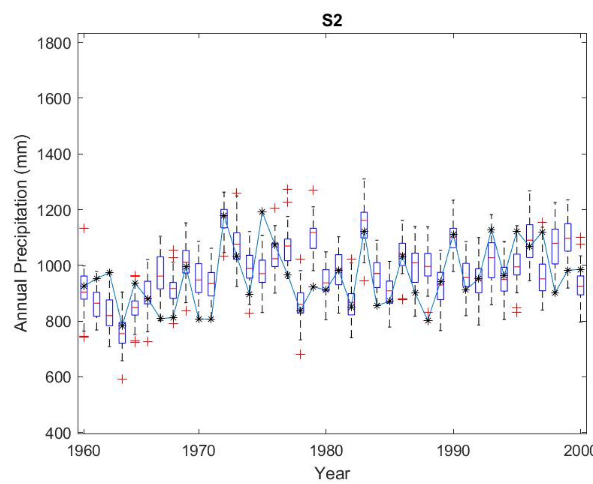
CDD



Prec90p



AMS



TAP

Figure 2-9. Boxplots of annual statistics and indices of SDGAM model: Precip_m, Precip_std, Prcp1, SDII, CDD, Prec90p, AMS and TAP for 1961-2000 period at Dorval station (S2)

2.5 Conclusions

A new downscaling model (SDGAM) has been developed to accurately simulate the daily precipitation processes at a single site in the context of climate change. The proposed SDGAM model was based on the combination of the precipitation occurrence and the precipitation amount using the Generalized Additive Model. In brief, the proposed model was able to describe well many features of the daily precipitation process, including its occurrence frequency, intensity, and extremes for both calibration and validation periods for data from 10 rain-gauged stations located in Southern Quebec and Ontario, Canada. In addition, it has been demonstrated that the suggested SDGAM model could provide more accurate results than those given the existing SDSM model in the modeling of daily precipitation process based on both numerical and graphical performance criteria.

Chapter 3: Modeling of short-duration extreme precipitations in the context of climate change

3.1 Introduction

Many water management applications (i.e., design of urban storm drainage systems, flood management and infrastructure operations) require information of rainfall intensity-duration-frequency (IDF). In order to construct the IDF curves, annual maximum series (AMS) of different rainfall durations from a few minutes to days are obtained, commonly from 5 minutes to 1 day. However, in most practical applications, short-duration extreme rainfall data are very limited or even unavailable for a given location of interest while the daily extreme rainfall records are often available. For instance, less than 600 stations in Canada record short-duration extreme rainfall from 5 minutes to 24 hours (Environment Canada, 2020), while the number of stations observe the daily rainfall is more than 1700 stations (Mekis et al., 2018). Hence, it is necessary to develop new methods for modeling extreme rainfall processes over a wide range of time scales such that extreme rainfalls needed at sub-daily time scales for constructing IDF relations for a given location can be estimated from the available daily extreme rainfall data.

General Circulation Models (GCMs) have been commonly used for evaluating the effects of climate change under different scenarios of greenhouse gas emissions on the hydrological regime. Although these GCMs have been recognized to be able to represent the main features of the global distribution of basic climate parameters (Randall et al., 2007), they still cannot reproduce well details of regional climate conditions at temporal and spatial scales of relevance to hydrological impacts and adaptation studies (Nguyen et al., 2006). This is because outputs from GCMs are usually at resolution that is too coarse (as illustrated in Figure 1-1) for many climate change impact studies, generally greater than 2.5^o for both latitude and longitude (approximately 250km). To refine the GCM coarse grid resolution climate projection data to much finer spatial resolutions (regional or local scales) for the reliable assessment of climate change impacts, different downscaling methods have been proposed to resolve this scale discrepancy (Wilby et al., 2002; Fowler et al., 2007; Nguyen and Nguyen, 2008; Maraun et al., 2010; Khalili and Nguyen, 2016; Gooré Bi et al., 2017). The SDGAM based on a combination of a Generalized Additive Models (GAMs) for representing the daily rainfall occurrences and the daily precipitation amounts has been proposed to describe the linkage of the large-scale climate variability to the historical observations of the precipitation process at a local site in previous section.

Several probability models have been conducted to describe the distribution of extreme precipitation at a gauged site (Wilks, 1993; Zalina et al., 2002). Unfortunately, these models are not accurate with all time frames, it is therefore requiring need for formulating models that could statistically and simultaneously matches various properties of the precipitation process at different levels of aggregations. Recently, the scale-invariance (or scaling) concept has increasingly become a popular methodology for modeling of several hydrological processes across a wide range of time scales (Hubert, 2001; Schertzer et al., 2010; Lovejoy and Schertzer, 2012). For instances, Nguyen et al. (2002) proposed a scaling General Extreme Value (GEV) based estimation method that can be used to estimate extreme rainfalls for a given return period at a local site for sub-daily time scales (hourly, 30 minutes, etc.) from the statistical properties of extreme rainfalls at a daily scale. Nguyen (2020) proposed a new mathematical framework for modeling extreme rainfall processes over a wide range of temporal scales (i.e., from several minutes to one day) based on the three-parameter Generalized Extreme Value (GEV) distribution and the scaling behavior of the PWMs (known as the GEV/PWM model). The proposed model has been tested with data set of long rainfall records from a network of 74 stations located in diverse climatic conditions across Canada.

Climate variability and change have been recognized to have important impacts on the hydrologic cycle at different temporal and spatial scales. The temporal scales could vary from a very short time interval of 5 minutes (for urban water cycle) to a yearly time scale (for annual water balance computation). The spatial resolutions could be from a few square kilometers (for urban watersheds) to several thousand square kilometers (for large river basins). In this study, a suggested approach is based on the combination of the *spatial downscaling* method to link large-scale climatic variables provided by GCMs to daily extreme precipitations at a local site using the SDGAM and the *temporal downscaling* procedure to describe the relationships between daily extreme precipitations with sub-daily extreme precipitations using the scaling General Extreme Value (GEV) distribution and the scaling behavior of the PWMs. The feasibility and accuracy of this spatial-temporal downscaling approach have been assessed using the AM precipitation data available at 10 stations across Canada and based on different climate change scenario simulation results available for the study region provided by the Canadian GCMs for the current 1961-2000 period as well as for future 2030s, 2060s, and 2090s periods. Results of this numerical application have indicated that, after a bias-correction adjustment, it is feasible to develop an accurate linkage between the daily AMPs spatially downscaled from GCM simulations with the observed daily AMPs at local stations. These results suggest that it is possible to use the climate predictors given by GCM simulations under different climate scenarios for projecting the variability of AM daily precipitations for future periods. On the basis of these results for daily AMPs, the IDF curves for the current 1961-2000 period and for future periods (2030s, 2060s, and 2090s) were constructed using the proposed temporal GEV/PWM method for sub-daily AMPs

3.2 A statistical approach to downscaling of extreme precipitation processes

3.2.1 A spatial downscaling approach using SDGAM

The SDGAM model developed in this study can be used to downscale daily precipitation process at a local site. The modeling of this process in the context of climate change involves two components: the modeling of the daily precipitation occurrences and the modeling of the precipitation amounts

$$\hat{\pi}_i = C_{Ok}(a_0 + \sum_{a=1}^{n} f_{Oi}(X_i))$$

in which f_{Oi}: smooth function

X_i: the large-scale atmospheric predictors given by GCM simulations

C_{Ok}: the correction coefficients for Occurrence process

r_i is a uniform distributed random number, if $r_i \leq \hat{\pi}_i$, precipitation occurs at day i

- Precipitation Amount Process (R_i)

$$Y = f.C_{Ak}(\alpha + \sum_{i=1}^{n} f_i(X_i) + \eta_i)$$

in which α: intercept

f_i: smooth function

X_i: the large-scale atmospheric predictors given by GCM simulations

C_{Ak}: the correction coefficients for amount process

$\eta_i = Z * S_e^i$

S_e^i: the standard error of month i^{th}

f: bias correction coefficient, coming from the deviation of the simulated mean given GCMs and the estimated mean given by the NCEP re-analysis data. The value of f is set to 1 in calibration step of SDGAM model.

$$f = \frac{total\ mean\ by\ NCEP\ for\ calibration\ period}{total\ mean\ by\ GCMs\ for\ calibration\ period}$$

In both precipitation occurrence and amount processes, the correction coefficients C_{Ok} and C_{Ak} represent the difference between the mean of observed data and the mean of simulated results based on the regression of GAM for the percentage of wet-day and precipitation amounts, respectively. These coefficients are automatically computed during the calibration of the SDGAM model such that an adequate agreement between the simulated results and the historical data could be obtained. Initially, the values of these coefficients are set to 1 in the calibration step.

It has been demonstrated in previous chapter that the SDGAM model was able to describe accurately the linkage between the daily predictands (precipitation occurrence and amount) at a given local site and the large-scale climate predictors provided by GCMs. Hence, it can be used to generate "synthetic predictands" that represents the generated local weather

3.2.2 A temporal downscaling method using the scaling-GEV distribution

The GEV distribution has been commonly used to describe the distribution of extreme rainfalls for different durations and to construct the IDF curves. The cumulative distribution function, F(x), for the GEV distribution is given as:

$$F(x) = exp\left[-\left(1-\frac{\kappa(x-\xi)}{\alpha}\right)^{\frac{1}{\kappa}}\right] \quad ; \quad (\kappa \neq 0)$$

where ξ, α, and κ are the location, scale, and shape parameter, respectively.

The probability weighted moment (PWM) estimators (or method of L-moment, L-MOM) can be used for estimation of the GEV parameters in consideration of the scaling property of these PWMs over different rainfall durations. For a distribution of a random variable X that has a quantile function, $x(u)$, the PWM of r^{th}-order can be expressed as (Hosking and Wallis, 1997):

$$\beta_r = \mathrm{E}(\mathrm{X}\{F(X)\}^r) = \int_0^1 x(u)u^r du \tag{3-1}$$

The PWMs of r^{th}-order, β_r, of the GEV distribution are given as follow:

$$\beta_r = M_{1,r,0} = E[X\,\{F(X)\}^r] = (r+1)^{-1}\left(\xi + \frac{\alpha}{\kappa}\{1-(r+1)^{-\kappa}\,\Gamma(1+\kappa)\}\right) \tag{3-2}$$

in which $\xi, \alpha,$ and κ are the location, scale, and shape parameters respectively; and F is the cumulative probability of interest. $\Gamma(.)$ is the gamma function and r must be non-negative.

For a simple scaling process, it can be shown that the relation between the r^{th}-order PWMs of rainfalls for two different rainfall durations t and λt can be expressed as:

$$\beta_r(\lambda t) = \lambda^{\eta_r}\beta_r(t) = \lambda^{\eta}\beta_r(t) \tag{3-3}$$

where $\eta_r = \eta_0$ is the scaling exponent and can be estimated based on the means of different rainfall durations.

This infers that the scaling exponents η_r are constant across all PWM orders r for the same rainfall scaling regime. In other words, the plot of the scaling exponents η_r (y-axis) with the PWM order r (x-axis) should display a horizontal line rather than a linear sloping line as for the case of the ordinary statistical moments (Nguyen et al., 2002).

Furthermore, let $\tau_3(t)$ and $\tau_3(\lambda t)$ denote the L-skewness of the data samples for two different time scales t and λt respectively (Hosking, 1990). L-skewness is the dimensionless version of the third order L-moment. It is obtained by dividing the third-order L-moment by the second-order L-moment. Hence, for a simple scaling process it can be shown that:

$$\tau_3(\lambda t) = \frac{6\beta_2(\lambda t) - 6\beta_1(\lambda t) + \beta_0(\lambda t)}{2\beta_1(\lambda t) - \beta_0(\lambda t)} = \frac{\lambda^{\eta}}{\lambda^{\eta}}.\frac{[6\beta_2(t) - 6\beta_1(t) + \beta_0(t)]}{[2\beta_1(t) - \beta_0(t)]} = \tau_3(t) \tag{3-4}$$

Equation (3-4) indicates that the L-skewness is constant over different time scales.

Consequently, for the simple scaling process, the shape parameter of the GEV distribution κ, which is a function of the L-skewness, is also constant over the time scale, that is,

$$\kappa(\lambda t) = \kappa(t) \tag{3-5}$$

From Eqn. (3-2) and after some mathematical manipulations, the first- and second-order PWMs can be written as follows:

$$\beta_0 = \xi + \frac{\alpha}{\kappa}\{1 - \Gamma(1+\kappa)\} \tag{3-6}$$

$$\beta_1 = \frac{1}{2}\left[\beta_0 + \frac{\alpha}{\kappa}(1 - 2^{-k})\Gamma(1+k)\right] \tag{3-7}$$

On the basis of Eqns. (3-3), (3-5)-(3-7) the location and scale parameters of the GEV distribution for different time scales can be related as follows:

$$\alpha(\lambda t) = \lambda^{\eta}\alpha(t) \tag{3-8}$$

$$\xi(\lambda t) = \lambda^{\eta}\xi(t) \tag{3-9}$$

and the quantiles for different time scales can also be expressed as:

$$X_T(\lambda t) = \lambda^{\eta}X_T(t) \tag{3-10}$$

In summary, based on these equations, for a simple scaling regime, it is possible to derive the distributions and statistical properties of short-duration extreme rainfalls from those of longer durations at a given study site as presented right below.

There are two different manners to downscale extreme rainfall quantiles from daily to sub-daily and/or sub-hourly intervals: the direct and indirect methods. The direct method scales the quantiles of rainfall duration (λt) from those of duration t directly using Eqn. (3-10). Note that the daily extreme rainfall quantiles computed based on different PWM or NCM estimators could be varied. Consequently, the scaled sub-daily and/or sub-hourly extreme rainfall quantiles obtained using the two systems are therefore different. Similarly, though the parameter scaling relationships are identical for the two moment systems, the scaling parameters obtained using two different estimation methods are also different. For the indirect method, the first three PWMs of sub-daily and/or sub-hourly AMSs are first computes using the scaling relationships of PWMs over different rainfall durations. These scaled PWMs are then utilized to solve for the three parameters in order to calculate the rainfall quantiles

3.3 Numerical application

To access the accuracy and feasibility of the proposed spatial-temporal downscaling approach, a case study was conducted using both global GCM climate simulation output CanESM2 and the observed daily precipitation data at 10 stations located in Southern Quebec and Ontario regions, Canada (see Figure 3-1). For comparison purposes, both SDSM and SDGAM models were considered for this study. More specifically, the daily precipitation data for the period from 1961 to 2000 were used as detailed in Table 3-1.

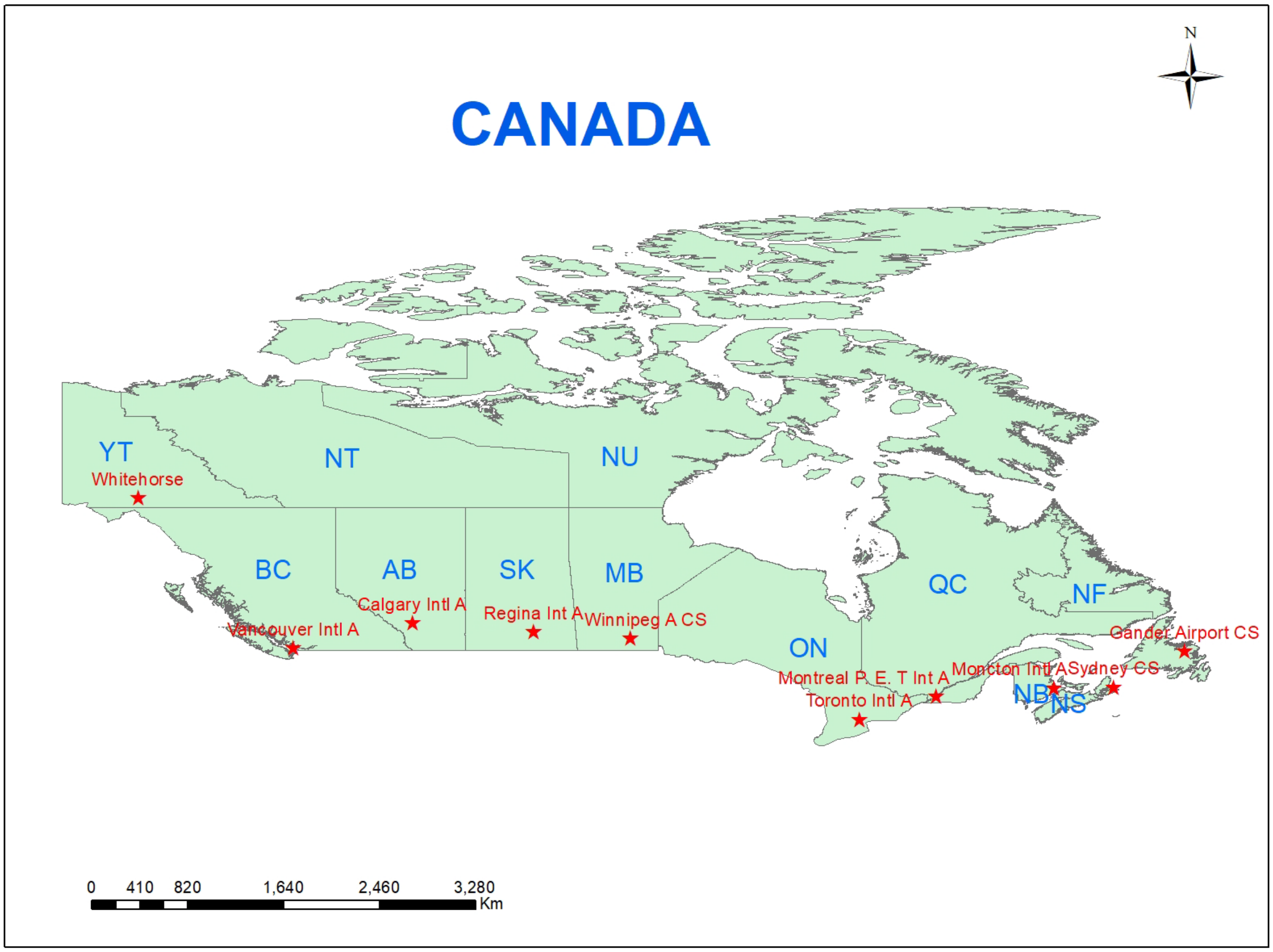


Figure 3-1. Selected stations across Canada

Table 3-1. Information of rain-gaged stations in across Canada

No	Province	Station name	Lat	Lon	Ele (m)	RL (year)
1	AB	CALGARY_INT_L_CS	51.12	-114.00	1081	61
2	BC	VANCOUVER_INTL_A	49.18	-123.18	4	63
3	MB	WINNIPEG_A_CS	49.92	-97.25	238	57
4	NB	MONCTON_INTL_A	46.12	-64.68	70	67
5	NL	GANDER_AIRPORT_CS	48.95	-54.57	151	70

No	Province	Station name	Lat	Lon	Ele (m)	RL (year)
6	NS	SYDNEY_CS	46.17	-60.03	62	53
7	ON	TORONTO_INTL_A	43.68	-79.63	173	64
8	QC	MONTREAL_P.E.T_INTL (Dorval)	45.47	-73.73	32	61
9	SK	REGINA_INT_L_A	50.43	-104.67	577	52
10	YT	WHITEHORSE	60.73	-135.10	707	44

The computational procedure for the suggested spatial-temporal downscaling method in this study can be summarized as follows:

i) Calibrate and validate the SDGAM model using the at-site daily precipitation as predictand and global GCM atmospheric variables as predictors (spatial downscaling);

ii) Generate 50 samples of 40-year daily precipitation series at a given site using the calibrated SDGAM and the corresponding GCM predictors, and extract daily AM precipitation series from these generated samples;

iii) Perform necessary bias correction of the GCM-downscaled daily AM precipitation series;

iv) Establish the scaling relations between the PWMs of the observed at-site AM precipitations for various durations;

v) Construct IDF curves using the adjusted GCM-downscaled AM daily precipitations and the estimated sub-daily AM rainfall amounts given by the calibrated scaling GEV model.

Repeat steps (ii) to (v) to construct IDF curves for future periods (2030s, 2060s, and 2090s).

3.4 Results and discussions

3.4.1 Spatial downscaling and bias-correction

The SDGAM model was calibrated and used to generate daily AMPs for all stations using the climate simulation outputs from CanESM2 under different RCPs (RCP 26, RCP 45, and RCP 85). The probability plots of AMPs downscaled in comparison of observed AMPs for the historical period 1961-2000 for stations S5 and S7 were presented in Figure 3-2 for purposes of illustration. It can be seen that downscaled AMPs are commonly lower than the observed at-site data. The error adjustment functions were established based on data for the 1961-1985 calibration period, then applied for the 1986-2000 validation period to assess their accuracy. In this study, the 4th order adjusted functions were employed to correct the differences between downscaled and observed data, results for S5 and S7 stations as shown in Figure 3-4. Results for all other stations can be found in Appendix B.

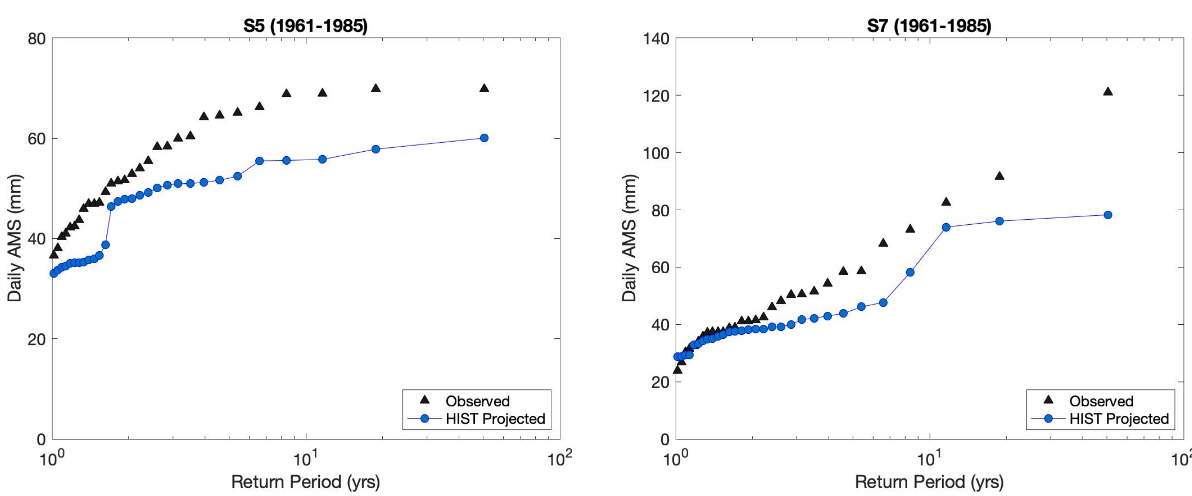


Figure 3-2. Probability plots of observed daily AMPs and Historical Period (HIST) at S5 & S7

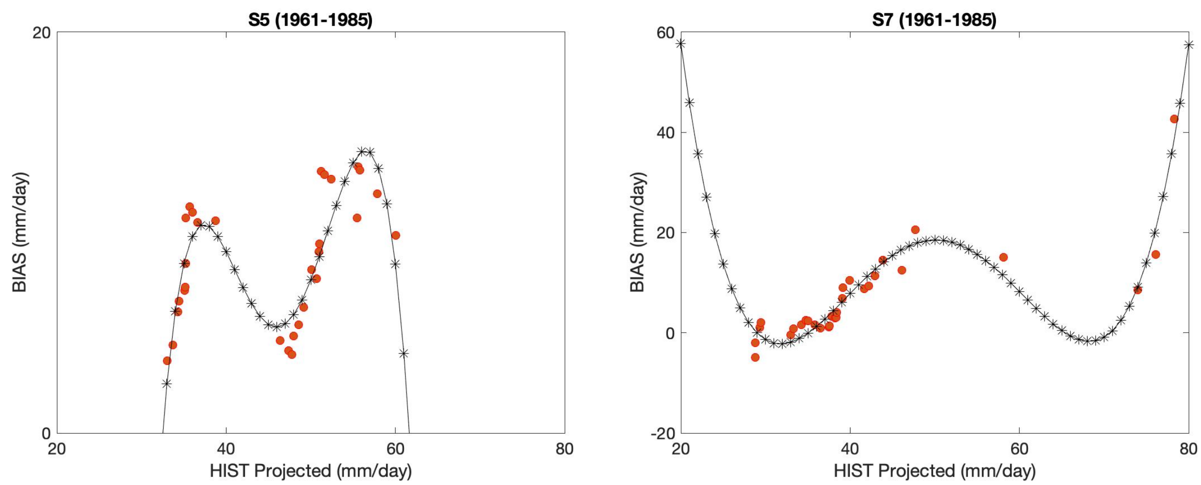


Figure 3-3. Error-Adjustment functions for S5 &S7

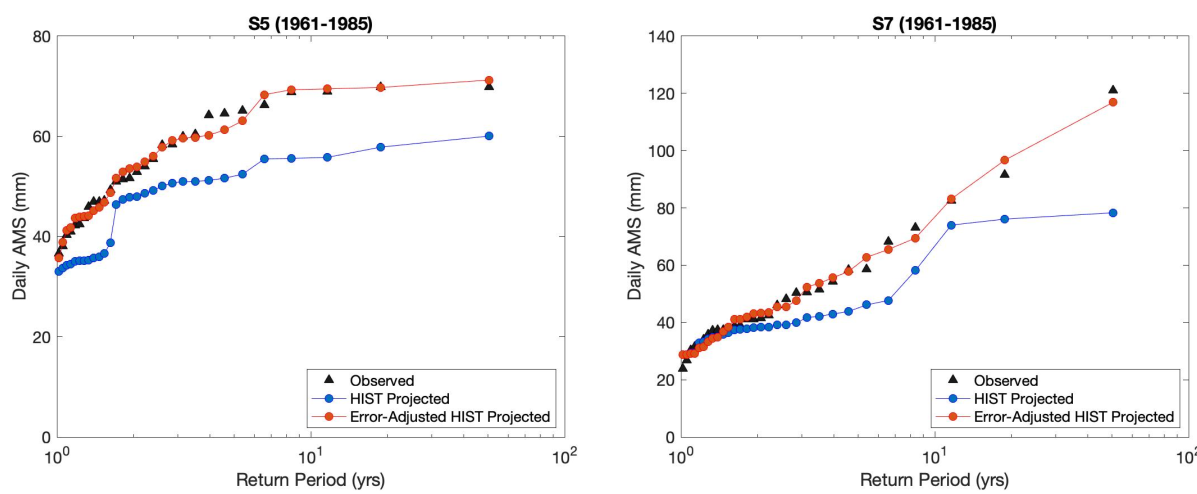


Figure 3-4. Probability plots of observed daily AMPs and Historical Period (HIST) after error-adjustment at S5 & S7

To assess the accuracy of the error adjustment method, the relative-root-mean-square-error (RRMSE) was used and can be as follows:

$$RRMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{X_i^{ob} - X_i^{sc}}{X_i^{ob}}\right)^2}$$

where N is sample size, X_i^{ob} and X_i^{ob} are observed and downscaled quantiles, respectively. Results can be found in Table 3-2. The smaller RRMSE values mean the results of the adjusted downscaled AMPs were improved in compared to the unadjusted downscaled AM amounts for both calibration and validation periods for all stations.

Table 3-2. RRMSE for daily AMPs with and without bias correction for calibration period of 1961-1985 and validation period of 1986-2000

Stations	Calibration		Validation	
	Before	Adjusted	Before	Adjusted
S1	0.637	0.171	0.220	0.202
S2	0.362	0.057	0.335	0.207
S3	0.503	0.077	0.502	0.214
S4	0.246	0.107	0.195	0.340
S5	0.130	0.057	0.204	0.126
S6	0.263	0.064	0.206	0.101
S7	0.152	0.084	0.210	0.117
S8	0.969	0.067	0.640	0.180
S9	0.442	0.113	0.436	0.240
S10	0.308	0.066	0.181	0.178

3.4.2 Temporal downscaling

To assess the scaling behavior of the observed AMP series, the log-log plots of the five rainfall PWMs against duration are prepared for all 10 stations. The log- linearity exhibited in the plot indicates the power law dependency of the rainfall statistical moments. However, Table 3-3 showed that the AM precipitation series in Canada displayed multiple scaling behaviors depending on the location of stations from East to West, for instance, the breaking points at Dorval and Sydney CS are 30 and 360 minutes, respectively. Hence, for a given location, it is possible to determine the PWMs and the distribution of rainfall extremes for short durations (e.g., 30 minutes) using available rainfall data for longer time scales (e.g., 1 day) within the same scaling regime.

Table 3-3. Breaking points (BP) of PWM/GEV for all stations

No	Province	Name	Breaking Point (min)
1	AB	CALGARY_INT_L_CS	15
2	BC	VANCOUVER_INTL_A	120
3	MB	WINNIPEG_A_CS	30
4	NB	MONCTON_INTL_A	120
5	NL	GANDER_AIRPORT_CS	30
6	NS	SYDNEY_CS	360
7	ON	TORONTO_INTL_A	30
8	QC	MONTREAL P.E. INTL (Dorval)	30
9	SK	REGINA_INT_L_A	360
10	YT	WHITEHORSE	15

For purposes of illustration, Figure 3-5 shows the log-log plot of the PWMs versus durations for Sydney_CS (S6) and Dorval (S8) stations.

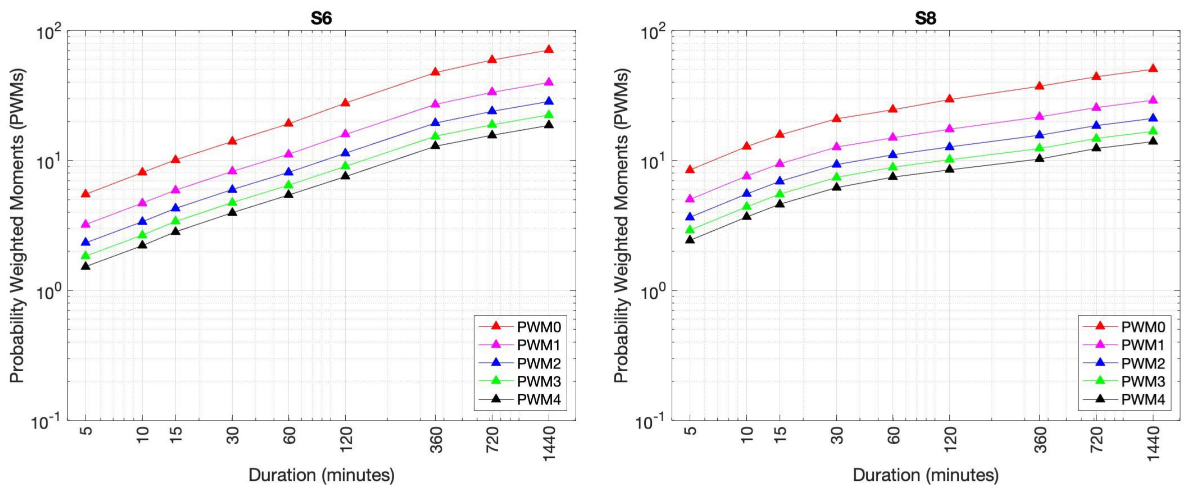


Figure 3-5. Log-log plots of the PWMs versus durations at S6 & S8

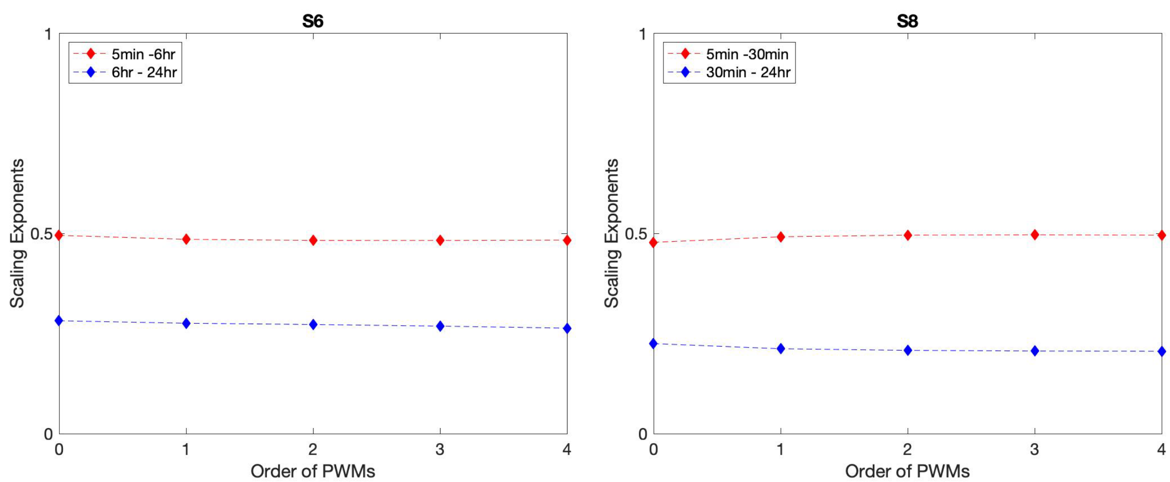


Figure 3-6. Scaling exponents plotted against the order of PWMs at S6 &S8

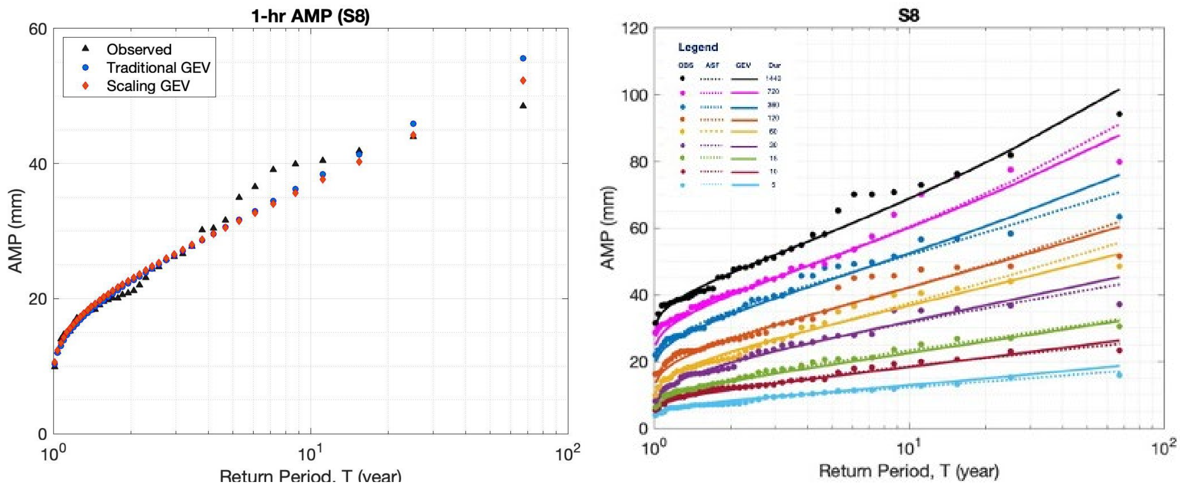


Figure 3-7. Probability plots of Observed AMPs and estimated using traditional and scaling GEV distributions at 1-hr (left) and all duration (right) for Dorval station (S8). Dotted line: Traditional GEV, Solid line: GEV/PWM, circle markers: Observed data

Figure 3-7 illustrates the comparison between observed and estimated precipitation by traditional GEV and scaling GEV for 1-hour duration (left) and all 9 - durations from 5 minutes to 24hr (right). It can be seen from the Figure 3-7 that scaling GEV approach is in very good agreement with the observed data. In addition, Table 3-4 presents numerical IDF relations for Dorval station given by traditional fitted GEV and scaling GEV approaches for the historical period 1961-2000. There is no significant difference between the two methods. Therefore, the scaling GEV approach can be used to estimate sub-daily AM rainfalls from historical or adjusted downscaled daily AM precipitations.

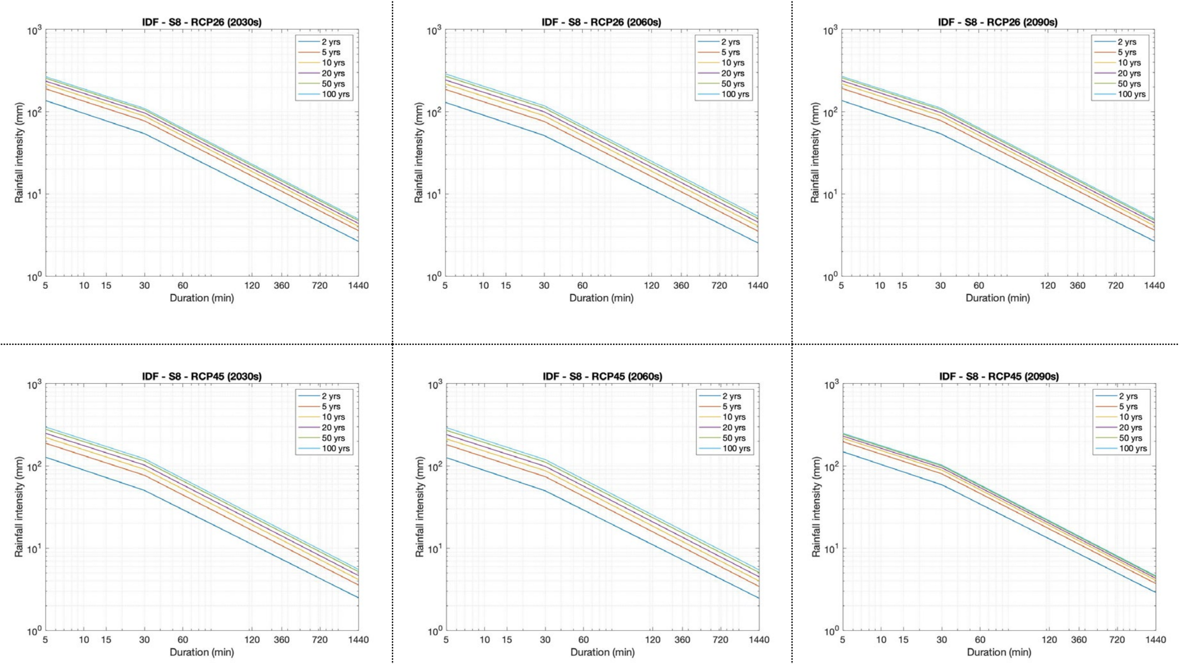
Table 3-4. Numerical IDF curves of AMP estimated by traditional GEV and scaling GEV for Dorval station (1961-2000). Unit of precipitation intensity is mm/hr, Return Period, T in year

Traditional GEV method									
Return Period (T)	**Duration (min): mm/hr**								
	5	**10**	**15**	**30**	**60**	**120**	**360**	**720**	**1440**
2	95.01	72.39	58.66	39.48	22.49	13.59	5.86	3.37	1.94
5	125.80	95.81	79.22	53.95	31.18	17.93	7.50	4.29	2.46
10	146.99	111.36	93.20	63.23	37.42	21.14	8.65	5.03	2.87
20	167.94	126.31	106.90	71.92	43.79	24.50	9.78	5.86	3.32
50	196.01	145.70	125.06	82.86	52.66	29.30	11.30	7.13	3.99
100	217.75	160.26	138.99	90.85	59.78	33.26	12.48	8.27	4.58
Scaling GEV method									
Return Period (T)	**Duration (min): mm/hr**								
	5	**10**	**15**	**30**	**60**	**120**	**360**	**720**	**1440**
2	99.24	69.27	56.13	39.18	22.87	13.35	5.68	3.32	1.94
5	131.50	93.23	76.22	53.99	31.08	17.89	7.44	4.28	2.46
10	154.93	110.10	90.13	63.98	36.74	21.09	8.74	5.01	2.87
20	179.09	127.08	103.94	73.71	42.34	24.31	10.09	5.79	3.32
50	213.01	150.26	122.54	86.50	49.84	28.72	12.00	6.92	3.99
100	240.54	168.58	137.02	96.24	55.65	32.21	13.57	7.88	4.58

3.4.3 IDF curves for the periods of 2030s, 2060s, and 2090s

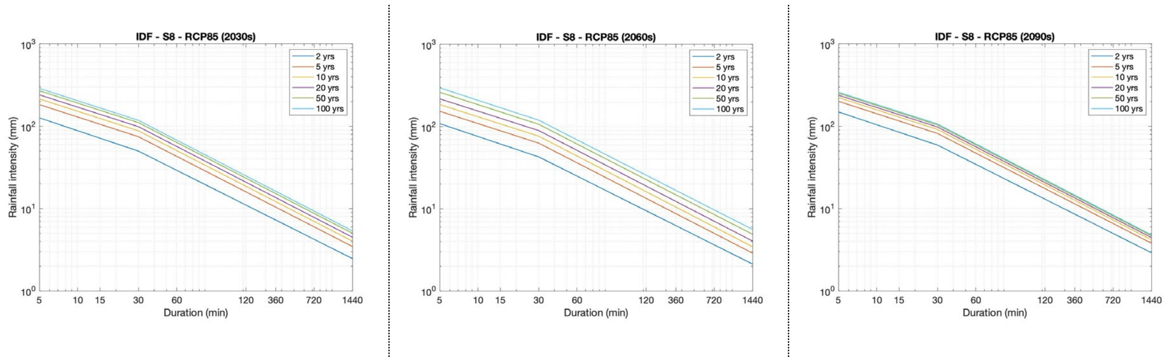
The proposed spatial-temporal SD was used to construct IDF curves for stations located in Canada under different climate change scenarios (RCP 26, RCP 45, and RCP 85) for the current and future periods (2030s, 2060s, and 2090s). For purposes of illustration, Figure 3-8 shows the plots of daily AM precipitations at Dorval station (S8) for the future periods (2030s, 2060s, and 2090s) with different RCP (RCP 26, RCP 45 and RCP 85). It can be seen that there is an increasing trend in the projected extreme rainfalls for future periods.

2030s 2060s 2090s



RCP 26

RCP 45



RCP 85

Figure 3-8. IDF curves for future periods with different RCPs at Dorval Station (S8)

Table 3-5. AMP GCM-projected corresponding to 100-year return period for the current and the future periods. Unit of precipitation intensity is mm/hr

Stations	Current	RCP26			RCP45			RCP85		
		2030s	2060s	2090s	2030s	2060s	2090s	2030s	2060s	2090s
S1	7.05	5.78	4.50	6.95	6.42	5.35	6.32	8.22	5.11	7.22
S2	1.55	2.50	2.09	2.68	3.19	2.54	2.89	2.66	3.08	3.80
S3	3.00	5.03	5.27	6.81	3.35	4.06	3.81	7.10	5.93	4.27
S4	4.06	5.08	4.38	5.40	8.05	5.25	5.83	5.12	5.40	5.95
S5	2.49	3.00	2.98	3.54	3.80	4.17	3.42	3.83	2.87	4.14
S6	6.14	7.24	8.52	8.90	8.95	9.13	7.68	6.75	7.16	5.95
S7	3.67	5.68	5.69	4.29	3.81	5.80	5.45	3.72	4.69	5.12
S8	4.05	4.94	5.36	4.99	5.53	5.45	4.62	5.37	5.61	4.80
S9	1.34	2.23	2.26	2.99	2.32	2.78	3.43	3.01	2.79	3.38
S10	0.90	0.94	0.97	1.06	0.89	0.99	0.90	0.90	0.79	0.96

3.4 Conclusions

A spatial-temporal downscaling approach was proposed in this study to describe the linkage between large-scale climate variables for daily scale to AM precipitations for daily and sub-daily scales at a local site. The feasibility of the proposed downscaling method has been evaluated based on climate simulation outputs from CanESM2 under different RCPs (RCP 26, RCP 45, and RCP 85) using available AM precipitation data for durations ranging from 5 minutes to 24 hours at ten rain-gage stations across Canada. Results have indicated that it is feasible to link daily large-scale climate variables to daily AM precipitations at a given location. In addition, it was found that the AM precipitation series in Canada displayed multiple scaling behaviors depending on the location of stations from east to west regions. Based on this scaling property, the scaling GEV distribution has been shown to be able to provide accurate estimates of sub-daily AM precipitations from GCM-downscaled daily AM amounts. It can be concluded that it is feasible to use the proposed spatial-temporal downscaling method to describe the relationship between large-scale climate predictors for daily scale given by GCM simulation outputs and the daily and sub-daily AM precipitations at a local site. This relationship would be useful for various climate-related impact assessment studies for a given region.

Finally, the proposed downscaling approach was used to construct the IDF relations for a given site for the historical period of 1961-2000 and for future periods (2030s, 2060s, and 2090s) using climate predictors given by CanESM2 simulations. This result has demonstrated the presence of high uncertainty in climate simulations provided by different RCPs. Further studies are planned to assess the feasibility and reliability of the suggested downscaling approach using other GCMs and data from regions with different climatic conditions.

Chapter 4: A statistical approach to estimating missing daily precipitation series at ungauged sites: a case study using data in Vietnam

4.1 Introduction

As described in previous Chapters, a number of studies have been conducted to establish the linkages between the large-scale climate variables given by GCMs and the observed characteristics of the daily precipitation process at a local site using different downscaling methods (Xu, 1999; Yarnal et al., 2001; Nguyen et al., 2006). These downscaling methods, however, are not suitable for dealing with cases where precipitation data at the location of interest are limited or not available. The estimation and prediction of hydrological variables such as precipitation and flow with climate change conditions for these ungauged or partially gauged sites remains a crucial challenge for managing and planning water resources (Sivapalan, 2003). Several studies dealing with the impacts of climate change on water resources at ungauged locations have been conducted in recent years (Besaw et al., 2010; Candela et al., 2012; Gibbs et al., 2012). For instance, Candela et al. (2012) proposed the use of a rainfall-runoff model to assess impacts by climate change on water resource for ungauged location in Northern Spain. Samuel et al. (2012) suggests the bias correction technique with Regional Climate Models (RCMs) and Global Climate Models (GCMs) for simulating precipitation, temperature, and future flows at gauged and ungauged stations. In particular, Wilby et al. (2006) have identified the three sources of uncertainties in climate change impact studies: by GCMs relating to unknown future conditions, by both dynamical and statistical downscaling procedure, and by specific application models.

In the context of regional impact studies, many previous approaches have been proposed for the past decades to supplement limited hydrological data at a local site for the current period or for the assessment for future periods in the context of a changing climate (Wilks, 1998; Mehrotra et al., 2006; Nguyen et al., 2006; Samuel et al., 2012). For instance, a procedure for generating spatially correlated and serially independent random numbers in their stochastic multisite downscaling models in order to preserve the spatial dependency amongst rain-gauge stations in a region was conducted (Wilks, 1998; Mehrotra et al., 2006). Another approach was the nearest neighbor resampling to preserve the spatial correlation of the daily precipitation and temperature data (Buishand and Brandsma, 2001; Beersma and Buishand, 2003). Furthermore, the spatial structure of Fourier Coefficients was applied to describe the spatial variability of rainfall series in a region (Lima and Lall, 2009). These studies, however, did not explicitly consider the similarity or homogeneity of the precipitation series at different sites even though this similarity assessment is an important factor in the understanding of the variability of the precipitation phenomenon in space. It is therefore necessary to assess the similarity of historical rainfall series at different locations to ensure that these observed precipitation measurements are produced from the same storm system (Nguyen et al., 2002; González and Valdés, 2008). Regionalization methods are hence frequently used to transfer rainfall information from one location to the other (Nguyen et al., 2007; Samuel et al., 2012). Regionalization methods have been developed and employed according to two main objectives: considering spatial dependency (homogeneity) and reducing uncertainty. Consequently, for precipitation estimation at an ungauged site, the homogeneity of precipitation processes at different sites is a necessity condition to obtain an accurate rainfall estimate with less uncertainty.

For the determination of precipitation homogeneity, cluster analysis and eigenvector analysis are two common approaches. A popular eigenvector-based method for regional precipitation analysis is Principal Component Analysis (PCA). PCA is a multivariate statistical technique used to simplify the original data by representing in dimensions fewer than original number of variables. The first application of PCA in meteorology and climatology was the end of the 1940s and enormous studies on PCA have published since (Preisendorfer, 1988; Johnson and Hanson, 1995; Baeriswyl and Rebetez, 1997; Nguyen, 2003; Yeo, 2013). This technique allows grouping of stations with similar characteristics and the delimitation of climatic regions, especially while handling a large dataset. PCA can be applied to reduce the dimensionality of the data but still contains most of the information of the original variables. PCA can be performed using either the covariance matrix or the correlation matrix. According to Johnson and Hanson (1995), PCA method can better describe the topographic influence on precipitation phenomenon. Baeriswyl and Rebetez (1997) also found that the PCA was a more accurate procedure for regional precipitation analysis in comparison with cluster analysis for precipitation data in Switzerland.

Vietnam is a developing country in Southeast Asia with limited observed rainfall data, leading to difficulty in construction designing and planning. Most of observed data are are only available at daily scale. This paper applied Principal Components Analysis (PCA) to identify homogenous precipitation regions. Regionalization is applied to define homogenous regions of rainfall for Vietnam. After that, a two-stage interpolation method is proposed to generate daily rainfall series at ungauged sites. Finally, GEV scaling technique is employed to infer the sub-daily and/or sub-hourly extreme rainfalls from daily extreme rainfalls in order to construct IDF curves at the location of interest.

4.2 Data

For this study, a total of 155 observed daily precipitation series with more than 22 years of record across Vietnam were selected (see details of these selected data in Appendix C). These series were selected based on the high quality of the data. In addition, the data with the same concurrent period of record are an important criterion for the main objective of this study that is to estimate the missing data at a location of interest using the available rainfall information from the neighbouring region. Figure 4-1 shows the locations of the selected stations. It can be seen that the density of stations in Northern part is higher than the Southern part, and those stations in the North also have longer record lengths.

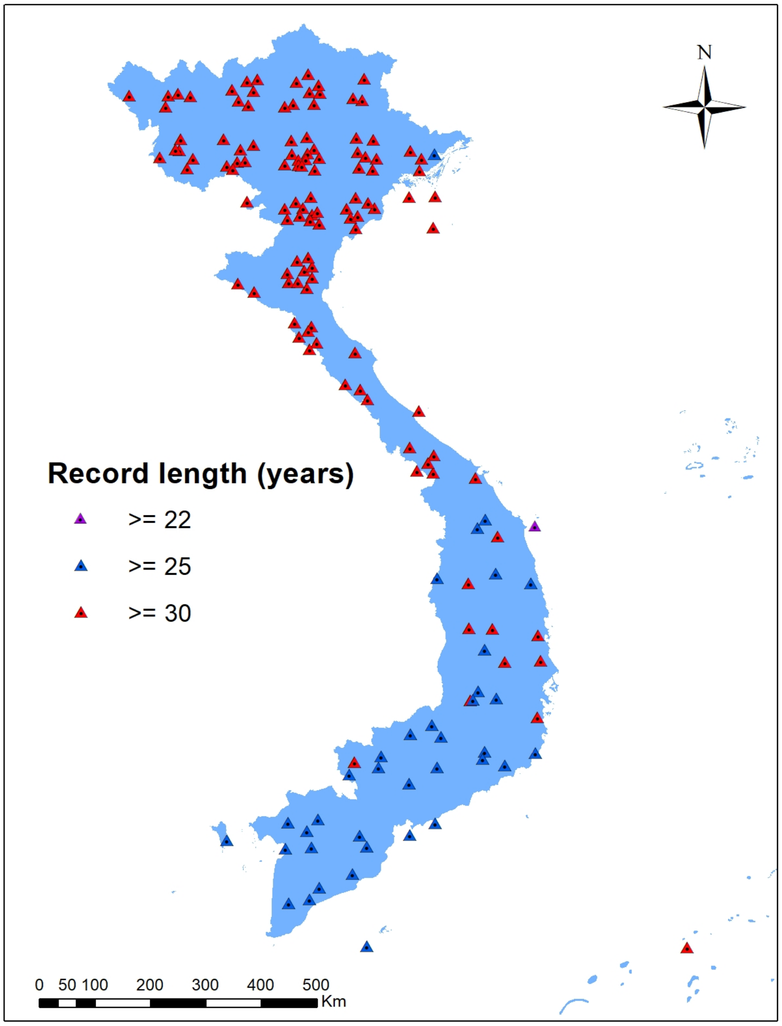
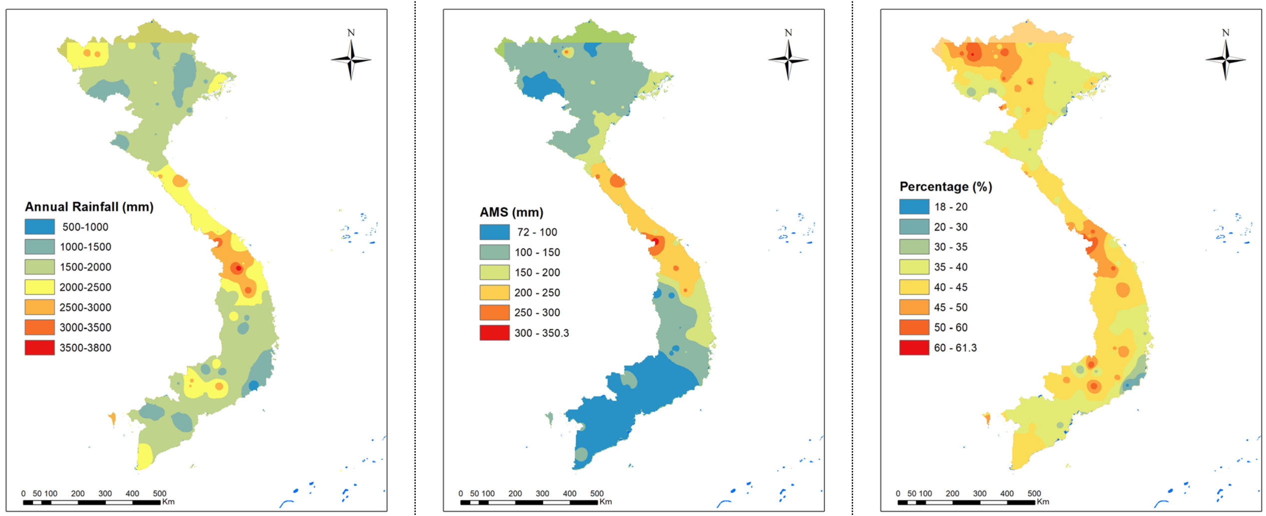


Figure 4-1. Selected rain-gauged stations in Vietnam

Figure 4-2 shows some basic statistics of precipitations across Vietnam. Vietnam is located in the tropical area that receives quite a lot of rainfalls in terms of amounts and extreme values. It can be seen that the highest annual precipitation and the extreme daily precipitation are in the Central part, specifically at Thua Thien Hue province - where the Hai Van pass is located. The annual precipitation here is almost 3800 mm yearly, and the maximum daily precipitation is up to 350 mm. The daily extreme rainfall in the South is smallest with the value of less than 100 mm. The Northwest and area surrounding the Hai Van pass have more than 50% of rain day per year.



(a) (b) (c)

Figure 4-2. Maps of annual precipitation across Vietnam (a): total annual rainfall, (b): daily annual maximum, (c): percentage of rain day. The value of each point is average over record period

4.3 Methodology

4.3.1 Homogeneous regions

Principal Component Analysis (PCA) is a multivariate statistical method that can be employed to reduce the original data by representing in dimensions fewer than the original number of variables. The original dataset of n correlated variables can be transformed into n numbers of uncorrelated principal components (PCs). These PCs are linear transformation of the original variables so that the sums of variances of the original and the new variables are equal. Although the number of PCs and original variables are the same, the first few transformed PCs consist of the majority of the variance in the dataset, reducing the dimensionality of the original dataset. The PCs are sequenced from the highest to the lowest variance as the first PC describes the data's largest proportion of variance. The second highest variance is explained by the second PC and so on. The values of PCs can be obtained from Equations:

$$PC1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = \sum_{j=1}^{n} a_{1j}x_j$$

$$PC2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = \sum_{j=1}^{n} a_{2j}x_j$$

where $x_1, x_2, \ldots x_n$ are the original variables and a_{jj} are the eigenvectors. The eigenvalues are the variances of the PCs. The covariance or correlation matrix of the data set is used to derive the coefficients a_{jj}, which are the eigenvectors. The eigenvalues of the data matrix can be calculated as follow:

$$|C - \lambda I| = 0$$

where C is the correlation or covariance matrix, I is the identity matrix, and λ is the eigenvalue. The PC coefficients are then calculated by Equation:

$$|C - \lambda I| a_{jj} = 0$$

The present study used the Principal Components Analysis (PCA) to identify homogenous precipitation regions. In this study, inter-station correlation coefficient matrices based on the annual maximum monthly mean rainfalls were analyzed using PCA. The principal components (PCs) were rotated, and the rotated component pattern was analyzed. The PCs were chosen based on the Kaiser's rule (Kaiser, 1960). The PCA was carried out in this study using the IBM SPSS software (IBM Corp., 2019). Once the homogenous regions of rainfall were identified, a 2-stage interpolation method is proposed to generate daily rainfall series at ungauged sites based on the daily rainfall data available at stations located within the same homogenous region. The method is detailed in the next section.

4.3.2 Estimation of missing daily precipitation series at an ungauged site

In this study, a two-stage interpolation method is proposed to estimate daily rainfall series at ungauged sites: rainfall occurrence stage and rainfall amount stage. Estimated daily precipitation series is compared with the observed data at the same location to assess the feasibility and accuracy of the proposed method.

+ **Stage 1**: For a calendar day (365 day), let O_i^k be daily precipitation occurrence at station k at day i. O_i^k = 0 if day i is dry, and O_i^k = 1 if day i is wet, threshold of rainfall for wet day is 1 mm. The probability π_i of non-zero precipitation at an ungauged site for day i is defined using IDW as follows:

$$\pi_i = \frac{\sum_{k=1}^{N} w_k\, O_i^k}{\sum_{k=1}^{N} w_k}, \text{ with } w_k = \frac{1}{d_k^2}$$

where N is number of stations within the homogenous region having rainfall data on day i; d_k distance from ungauged site to station k. The value of π_i ranges from 0 to 1. For this study, if value of $\pi_i < 0.5$, there is no rain at ungauged location of interest, rainfall amount at day i $R_i = 0$; if value of $\pi_i \geq 0.5$, rainfall occurs at ungauged location of interest, the rainfall amount R_i at day i is estimated in stage 2.

+ **Stage 2**: if rainfall occurs at location of interest at day i, the rainfall amount R_i is estimated as using IDW technique follows:

$$R_i = \frac{\sum_{k=1}^{N} w_k R_i^k}{\sum_{k=1}^{N} w_k}, \text{ with } w_k = \frac{1}{d_k^2}$$

where R_i^k is rainfall amount at station k at day i; N is number of stations within the region having rainfall data on day i; d_k distance from ungauged site to station k.

By repeating the process for every day of the whole length of record, rainfall series at an ungauged site of interest is calculated based on the rainfall data available at stations located within the same homogenous region. The jackknife technique was used to represent the ungauged site condition. Note that the rainfall is interpolated based on the observed data on the same calendar date.

4.4. Results and discussions

4.4.1 Homogeneous regions

Table 4-1 illustrates the computed total variance explained by each principal component. It can be seen that the first component explains the highest variance (52.26% of the total variance of the system). Based on results of PCA, Vietnam can be divided into seven regions of the homogenous groups of rain gauges as shown in Figure 4-3. Regions 1 to 3 are for the Northern part, Region 4 is for the North-central part, Region 5 is for the South-central part, Region 6 is for the Central Highland area, and Region 7 is for Southern part of Vietnam. It can be seen that this grouping is consistent with climate sub-regions of Vietnam (General Satistics Office of Vietnam, 1999). Results have indicated that the topography of mountainous areas plays a decisive role in the determination of the homogenous rainfall regions in Vietnam.

Table 4-1. Percentage of variances explained by each component computed for monthly amount of rainfall

Principal Component	% of variance	Cumulative variance (%)
1	52.26	52.26
2	18.31	70.58
3	7.57	78.14
4	2.55	80.69
5	1.62	82.31
6	1.19	83.49
7	0.81	84.30

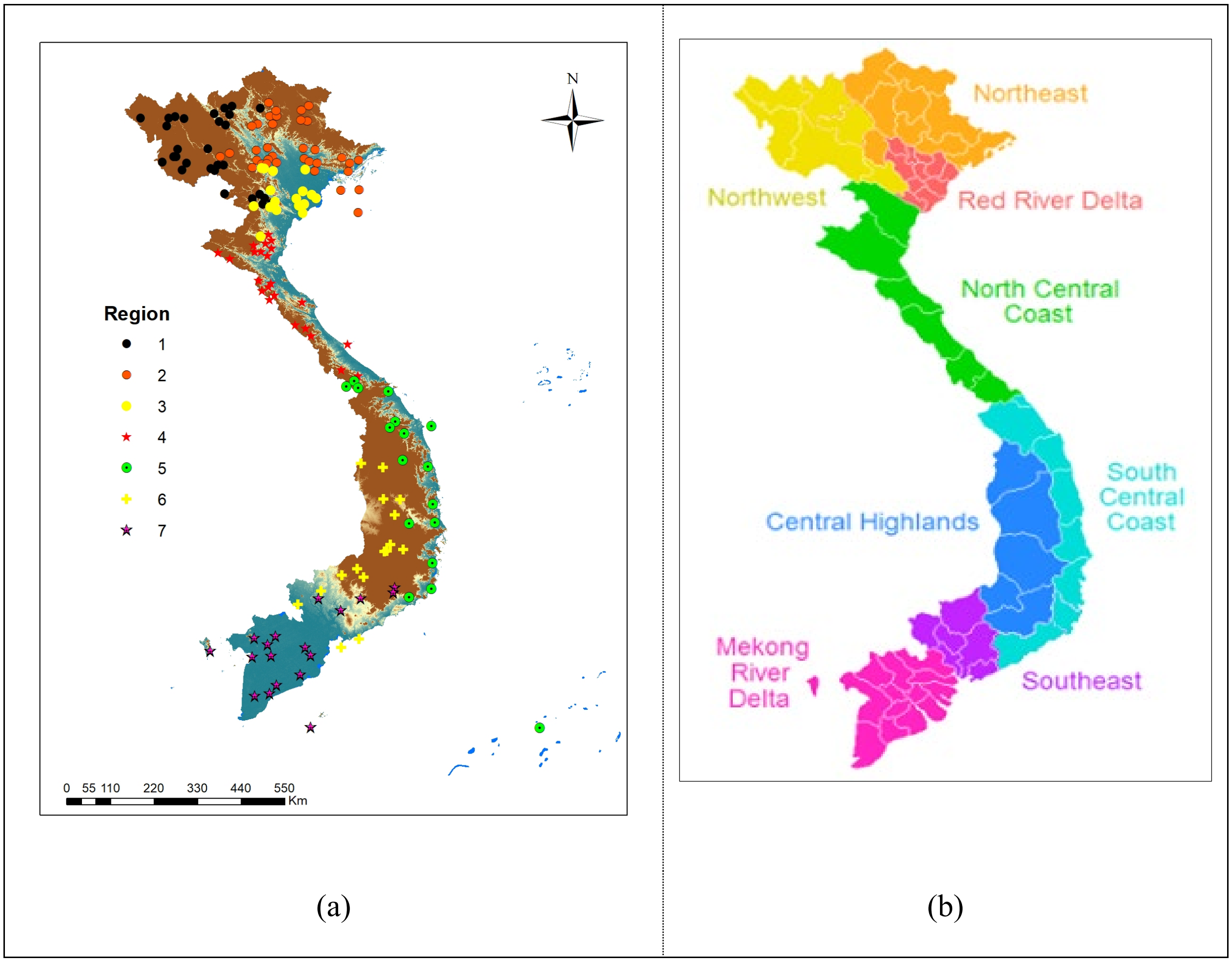


Figure 4-3. Seven homogeneous regions of Vietnam: (a) - Proposed method using PCA vs (b) - according to Vietnam meteorology department

For this study, the PCA works well for the rainfall data of Vietnam in monthly time scale. However, this approach is considered to be sensitive with the time scale. Hence, when apply with different time scales, i.e., daily or weekly, the number of homogeneous groups obtained are not the same as the number obtained for the monthly time scale as shown in Table 4- 2. This is the disadvantage of PCA; therefore, it is recommended to apply PCA with some caution. Further works will be studied to minimize this sensitivity.

Table 4- 2. Number of regions by different time scale

Time scale	No. of regions
Daily	19
Weekly	12
Monthly	7

Once homogenous regions of rainfall obtained, the proposed 2-stage interpolation method is applied to generate daily rainfall series at ungauged sites based on daily rainfall data of stations within the same homogenous region. Results are presented in the following section.

4.4.2 Estimation of precipitation series at an ungauged site

For this study, Region 4 - belongs to North central coast and marked as red stars in Figure 4-3(a) - has been selected for generating precipitation at ungauged sites. This region 4 has been selected due to high quality and uniform of rainfall data including 23 stations with 32 years of record. It is also because of the complexity of rainfall in this region: include stations with the most extreme values in the country. The information and the summary of basic statistics of 23 stations in Region 4 are presented in Table 4-3.

Table 4-3. Statistics of stations of Region 4

No.	Station Code	Lon	Lat	Daily mean (mm)	Daily AMS	Daily Max	No. rain day (%)	Total Annual (mm)
1	146	105.34	19.38	4.14	181.8	376.7	38.0	1544.5

No.	Station Code	Lon	Lat	Daily mean (mm)	Daily AMS	Daily Max	No. rain day (%)	Total Annual (mm)
2	150	105.47	19.45	4.60	174.5	388.5	37.5	1718.2
3	151	105.47	19.27	4.27	159.2	560.7	36.3	1612.1
4	153	105.40	19.59	3.97	120.4	314.7	39.2	1463.0
5	155	104.53	19.03	4.21	136.5	449.5	41.8	1556.4
6	156	105.18	18.54	4.37	175.9	788.4	40.4	1638.6
7	158	105.46	18.48	4.87	208.3	362.0	35.6	1822.2
8	160	105.07	19.34	4.43	129.4	304.1	38.1	1591.3
9	161	105.09	19.19	4.29	123.0	272.4	38.0	1592.6
10	162	105.38	19.10	4.05	161.8	710.1	34.2	1503.8
11	163	105.24	19.19	3.84	141.1	279.5	36.5	1496.5
12	164	104.26	19.17	3.36	101.2	192.0	35.3	1259.5
13	165	105.40	18.40	5.31	215.1	596.7	40.8	1959.0
14	166	105.54	18.21	7.45	266.1	657.2	43.7	2716.3
15	167	105.43	18.11	6.04	245.0	492.6	48.1	2214.4
16	168	105.26	18.31	5.70	164.5	518.8	47.6	2104.0
17	169	106.17	18.05	7.89	280.7	519.1	43.5	2876.9
18	170	106.25	17.45	5.58	229.0	413.7	38.8	1997.2
19	172	106.37	17.29	5.95	248.1	554.6	40.0	2204.4

No.	Station Code	Lon	Lat	Daily mean (mm)	Daily AMS	Daily Max	No. rain day (%)	Total Annual (mm)
20	175	106.01	17.53	6.07	217.2	548.4	45.0	2224.2
21	176	107.20	17.10	5.75	194.0	727.5	39.2	2123.3
22	177	107.05	16.51	6.04	248.7	447.5	43.0	2200.4
23	178	106.44	16.38	5.24	181.6	368.1	50.9	1947.6

Daily rainfall series for all stations in Region 4 were generated using the proposed 2-stage interpolation method. The jackknife technique was used to represent the ungauged site condition for 23 selected sites in Region 4. Basic statistics indices (listed in Table 4-4) have been performed to assess the computed rainfall series, details are presented in Figure 4-4 to Figure 4-8 and Table 4-5 to Table 4-7. Results showed that estimated data are very close to the observed data.

Table 4-4. Statistics indices

No.	Indices	Definition	Unit
1	AP	Annual precipitation	mm
2	AMS	Daily annual maximum precipitation	mm/day
3	WD	Percentage of wet day	%
4	SDII	Mean precipitation amount at wet days	mm
5	CDD	Maximum number of consecutive days	days
6	Prec90pc	90th percentile of rain day amount	mm

It can be seen that there is insignificant difference between estimated and observed data in terms of annual rainfall amount, percentage of wet day and AMS rainfall. In terms of annual rainfall amount, generated data are similar to observed data at almost station except station 12 and 17 as presented in Figure 4-4. Similar results for percentage of wet day at all station except station 5, 16, 17 and 20 are showed in Figure 4-5. Figure 4-6 showed that generated data of daily AMS are lower than observed data at all stations. GEV distribution was used to fit daily annual precipitation of both generated and observed data for all stations, then quantiles different return periods T = 2, 5, 10, 20, 50 and 100 years were calculated and plotted in Figure 4-7. It displayed that the difference of generated and observed data are mainly at extreme values. Note that comparison results were based generated data, no bias correction was applied for this study.

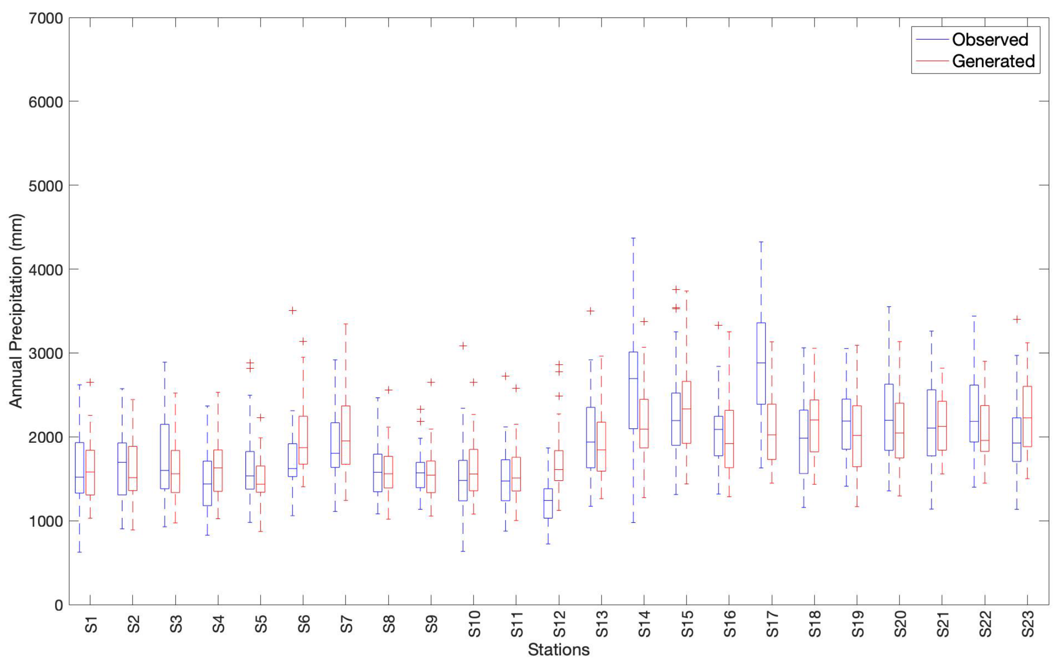


Figure 4-4. Annual precipitation of observed (Blue) and generated (Red) data for all 23 stations. Each boxplot is conducted from data of the station for entire record length

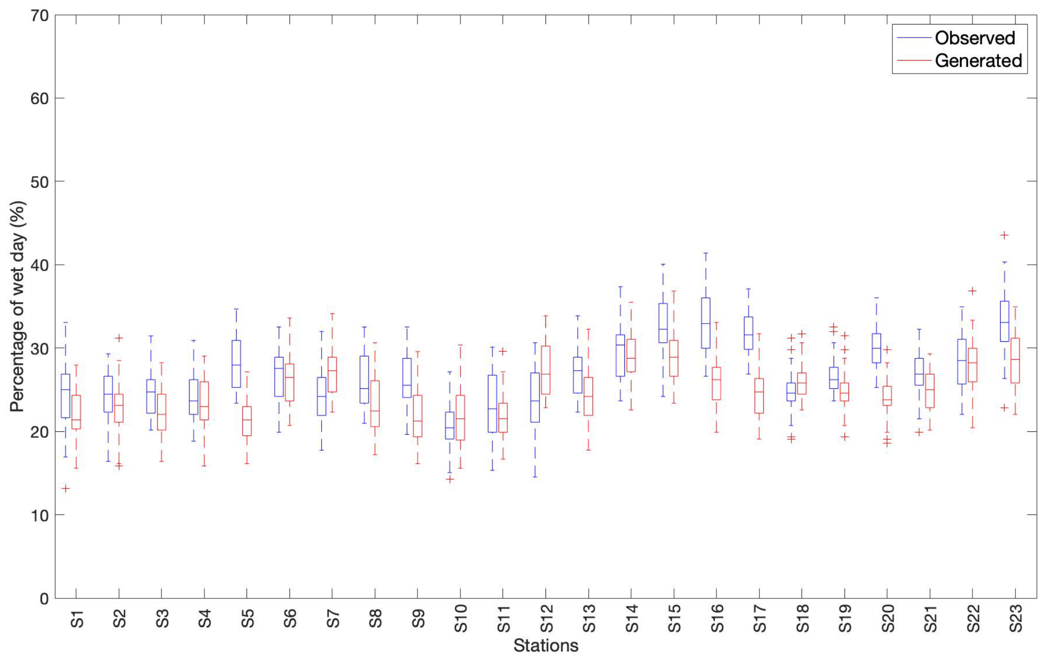


Figure 4-5. Percentage of wet day of observed (Blue) and generated (Red) data for all 23 stations. Each boxplot is conducted from data of the station for entire record length

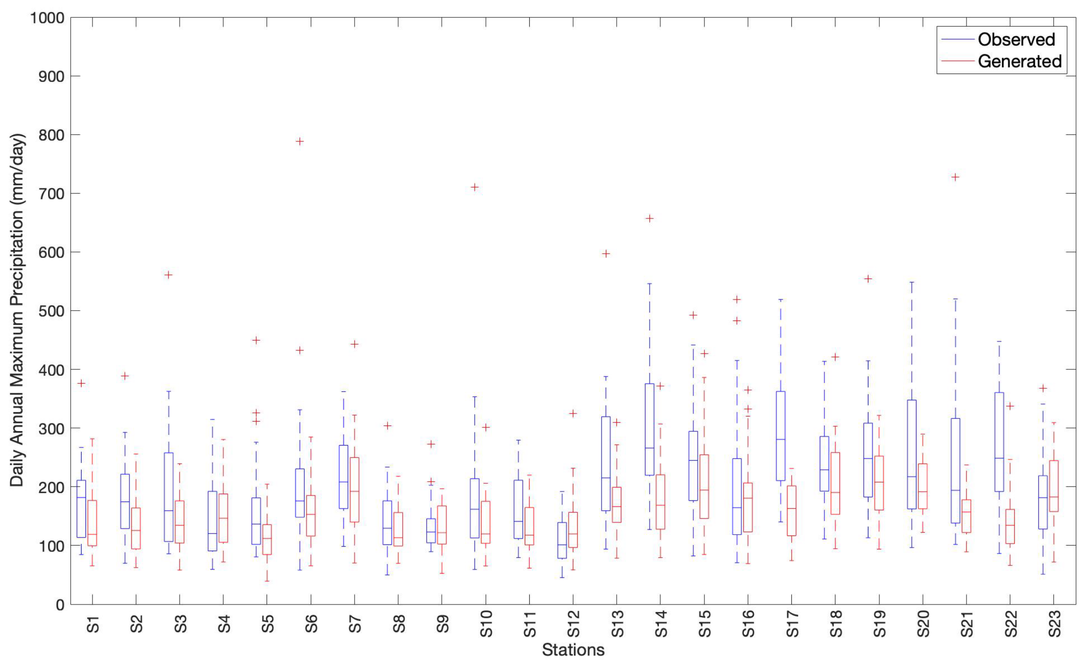


Figure 4-6. Daily AMS precipitation of observed (Blue) and generated (Red) data for all 23 stations. Each boxplot is conducted from data of the station for entire record length

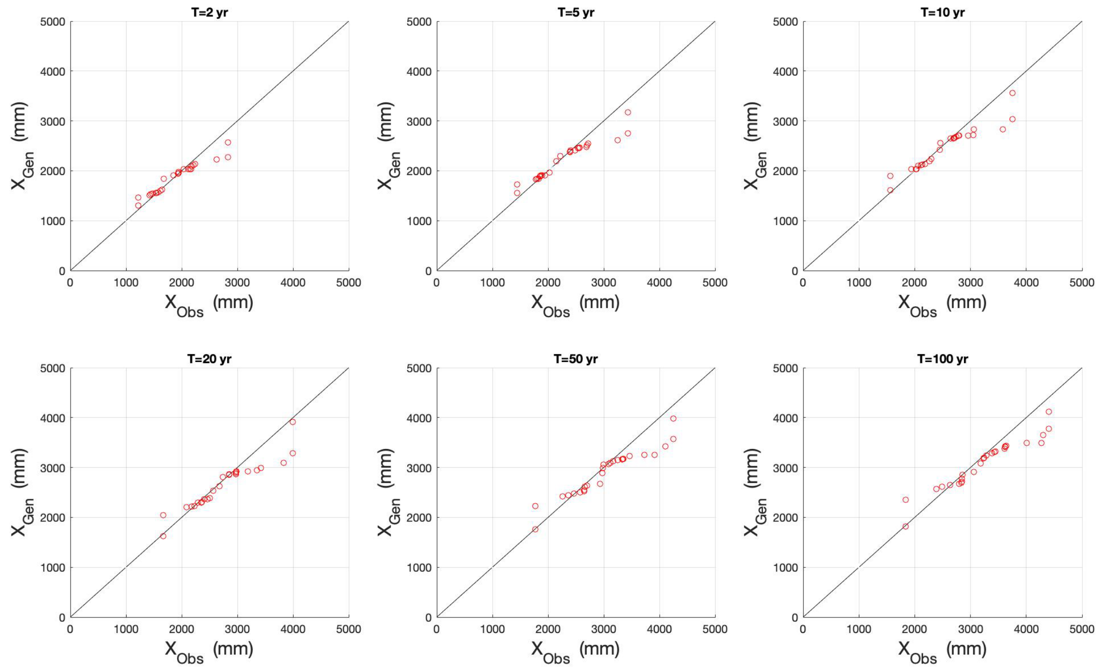
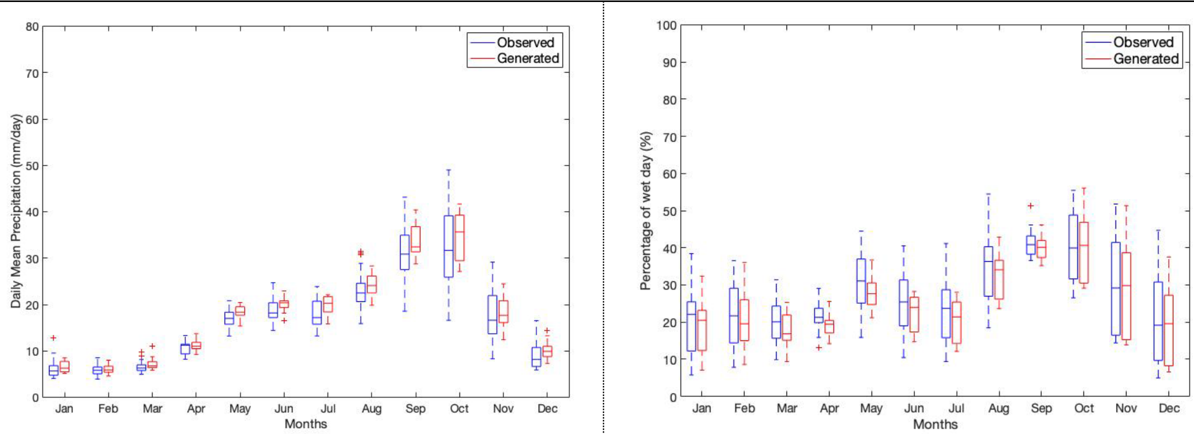


Figure 4-7. Q-Q plot of annual precipitation with return periods T = 2, 5, 10, 20, 50 and 100 years

Figure 4-8 compared results of daily mean precipitation and percentage of wet day of all stations by months. It is found that generated data were very close to observed data. High daily intensity of rainfall as well as numbers of rainy days were from August to November.



(a) (b)

Figure 4-8. Boxplots of daily mean precipitation (a) and percentage of wet day (b) for 12 months (Blue: observed data - Red: generated data). Each boxplot is conducted from data of the month for all stations

Table 4-5 to Table 4-7 compares generated and observed data in terms of mean precipitation amount at wet days (SDII), maximum number of consecutive dry days (CDD), and 90th percentile of rain day amount (Prec90p) by season time scale for all ungauged sites. For this study, Spring is defined from January to March, Summer is from April to June, Fall is from July to September and Winter is from October to December. Highest rainfall intensity at wet days were found in Winter and Spring, which are almost double values in Summer and Fall (Table 4-5). The maximum observed value of 30.1 mm/day and generated values 27.9 mm/day were found in Winter while minimum observed value of 9.4 mm/day and generated value of 10.5 mm/day was in Summer. In contrast, longest period of time with no rainfall was found in Summer with longest observed period of 41.2 days and generated period of 40.3 days while shortest period was in Spring with shortest observed period of 13.3 days and generated period of 14.0 days (Table 4-6). Threshold is defined 1mm to be considered as rainy day. CDD values were decimal because they were averaged all year of record length for a station. Similar to SDII, highest values of Prec90p were in Winter while smallest values were in Summer (Table 4-7). It was found that there is insignificant difference between estimated and observed data for all indices. Hence, it has been demonstrated that the proposed method is feasible and the estimated daily precipitation series are reliable.

Table 4-5. Mean precipitation amount at wet days

SDII (mm): mean precipitation amount at wet days								
Station	Spring		Summer		Fall		Winter	
	Obs	Gen	Obs	Gen	Obs	Gen	Obs	Gen
S1	19.2	20.1	10.1	11.4	12.3	14.4	21.6	22.9
S2	18.2	19.7	10.1	11.3	14.7	13.4	22.2	22.5
S3	21.1	20.3	11.0	11.6	13.0	14.2	23.0	22.9
S4	15.2	20.0	10.6	11.1	12.4	14.2	19.5	22.6
S5	16.0	19.4	9.4	13.1	12.6	13.9	19.2	20.5
S6	19.7	23.0	10.0	12.6	12.5	15.0	21.3	25.5
S7	24.6	23.2	11.6	11.8	13.8	14.1	26.6	25.1
S8	15.1	18.5	10.4	11.1	12.9	13.9	18.8	21.8
S9	15.2	19.4	9.6	11.7	13.1	14.7	19.2	22.0
S10	20.8	21.2	10.6	12.2	13.1	14.5	22.2	23.0
S11	17.9	20.0	10.0	11.6	13.5	14.1	21.0	22.4
S12	14.1	17.7	10.2	10.5	11.7	12.8	14.7	20.4
S13	23.8	24.0	11.3	12.3	13.7	14.6	25.4	25.7
S14	27.4	22.0	16.2	12.7	16.2	14.9	30.1	24.0
S15	20.5	24.7	12.7	13.9	14.2	15.5	22.4	27.9

SDII (mm): mean precipitation amount at wet days								
Station	Spring		Summer		Fall		Winter	
	Obs	Gen	Obs	Gen	Obs	Gen	Obs	Gen
S16	19.2	23.9	10.9	12.7	13.1	14.7	20.9	26.0
S17	27.0	23.7	16.4	14.5	15.8	16.1	28.4	27.4
S18	22.0	22.8	14.0	13.9	12.6	14.9	24.6	26.3
S19	22.4	22.5	13.9	14.1	13.5	13.6	25.4	25.1
S20	20.9	23.7	12.8	15.3	15.1	14.6	23.9	26.5
S21	20.9	22.6	14.8	16.2	17.1	15.5	25.8	26.1
S22	20.5	19.3	16.2	14.5	12.5	14.3	23.5	20.8
S23	15.5	20.6	11.4	15.3	12.4	13.3	16.6	23.5

Table 4-6. Maximum number of consecutive dry days

CDD (days): maximum number of consecutive dry days								
Station	Spring		Summer		Fall		Winter	
	Obs	Gen	Obs	Gen	Obs	Gen	Obs	Gen
S1	22.3	24.3	26.7	27.1	19.6	23.5	18.8	18.9
S2	21.8	25.1	23.9	27.3	20.0	19.6	15.6	18.9
S3	20.8	25.4	20.6	29.1	20.1	23.1	16.9	18.2
S4	24.6	23.7	28.5	25.4	23.1	21.1	14.9	17.3
S5	17.8	27.0	19.3	40.3	16.7	30.9	14.4	18.8

CDD (days): maximum number of consecutive dry days								
Station	Spring		Summer		Fall		Winter	
	Obs	Gen	Obs	Gen	Obs	Gen	Obs	Gen
S6	17.7	17.7	17.9	17.1	17.6	22.5	15.3	18.7
S7	19.2	15.1	18.8	16.3	24.0	19.9	19.6	17.2
S8	26.1	25.1	34.3	28.2	26.8	23.5	15.5	18.0
S9	22.3	26.0	26.7	32.2	20.4	25.8	14.9	21.2
S10	28.7	25.0	29.4	28.9	25.9	22.8	18.0	19.5
S11	24.4	26.3	28.9	29.6	22.6	22.6	20.9	18.4
S12	27.8	18.0	41.2	21.0	28.6	16.9	17.5	14.7
S13	15.1	19.2	16.3	18.8	19.9	24.0	17.2	19.6
S14	13.5	15.3	14.1	15.9	18.9	17.7	18.5	15.5
S15	14.0	14.7	14.3	14.7	16.4	19.7	11.9	20.1
S16	13.3	16.4	15.2	16.7	16.1	21.3	13.6	18.0
S17	14.5	16.7	14.8	19.9	18.2	24.3	19.4	22.6
S18	16.5	14.6	19.6	19.8	21.3	24.1	21.1	21.1
S19	14.6	16.5	19.0	19.6	22.7	21.3	20.6	21.1
S20	13.4	16.9	17.8	20.6	20.1	22.7	15.3	21.1
S21	15.6	15.9	20.3	25.5	29.7	27.6	27.3	22.6
S22	14.0	16.2	21.3	29.6	22.3	24.2	19.3	15.5

CDD (days): maximum number of consecutive dry days								
Station	Spring		Summer		Fall		Winter	
	Obs	Gen	Obs	Gen	Obs	Gen	Obs	Gen
S23	15.8	14.0	32.0	21.3	25.6	22.3	12.7	19.3

Table 4-7. 90th percentile of rain day amount

PREC90P (mm): 90th percentile of rain day amount								
Station	Spring		Summer		Fall		Winter	
	Obs	Gen	Obs	Gen	Obs	Gen	Obs	Gen
S1	60.9	59.2	32.3	35.7	34.9	36.7	64.8	57.7
S2	62.0	58.2	37.4	33.8	41.3	35.0	64.5	57.1
S3	63.5	60.0	36.2	34.6	41.7	35.5	68.5	58.1
S4	52.0	60.0	37.4	34.9	36.7	36.4	56.9	59.0
S5	48.9	52.4	25.8	35.4	33.8	32.6	53.1	47.6
S6	60.6	58.8	29.9	32.3	35.4	37.9	59.5	69.1
S7	69.6	64.9	35.8	31.6	44.7	41.0	81.7	72.9
S8	50.9	55.6	30.5	33.1	35.2	33.6	54.2	55.0
S9	51.0	56.4	31.8	34.5	36.3	33.7	54.2	55.8
S10	65.2	59.8	36.9	36.8	44.2	35.2	66.1	58.5
S11	59.0	57.2	36.3	33.3	39.9	34.5	63.5	57.0
S12	45.3	47.5	32.3	27.7	32.2	31.4	39.8	51.7

PREC90P (mm): 90th percentile of rain day amount								
Station	Spring		Summer		Fall		Winter	
	Obs	Gen	Obs	Gen	Obs	Gen	Obs	Gen
S13	71.2	63.0	31.3	34.3	44.0	41.3	77.4	75.3
S14	82.0	60.1	48.7	32.7	50.7	39.4	96.3	68.5
S15	65.5	69.4	38.0	38.6	38.2	44.9	70.2	83.9
S16	55.7	63.3	31.4	33.7	35.6	40.9	62.2	74.1
S17	80.1	58.6	45.5	35.2	48.2	42.4	81.0	71.3
S18	68.5	63.6	39.5	38.2	39.7	44.7	74.9	75.6
S19	65.8	66.0	41.7	38.0	43.1	40.9	81.7	71.3
S20	57.9	67.8	35.6	40.1	49.4	42.1	69.2	72.4
S21	61.8	56.6	41.3	41.4	62.7	40.8	78.3	68.3
S22	58.5	50.7	47.7	38.5	39.1	38.3	68.8	53.6
S23	48.4	54.6	34.4	42.0	35.7	37.9	42.4	64.2

To assess the important role of the identification of homogenous precipitation regions in the proposed method for estimating missing daily rainfall series, three stations S21-S22-S23 located in Region 4 have been selected for testing. They are considered as in difference groups as presented in Table 4-8. After that, the same estimation approach to generate daily rainfall series for those 3 stations was used considering these stations as located in different regions as shown in Table 4-8. The percentage of wet days and the annual precipitation for all scenarios were compared to the observed data to see the uncertainty given by different region scenarios. Results of this uncertainty are presented in Figure 4-9.

Table 4-8. Scenarios of regions

No.	Scenarios	No.	Scenarios
1	Observed data	4	Region 4+5
2	Region 4	5	Region 2
3	Region 5	6	Region 7

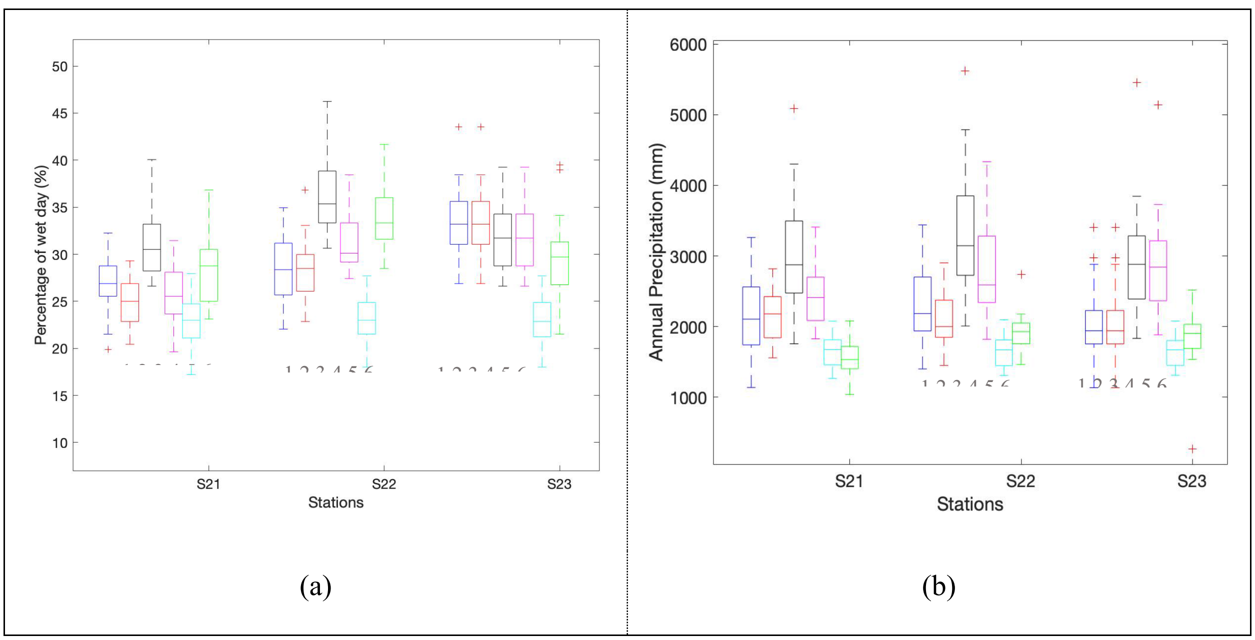


Figure 4-9. Compare percentage of Wet day (a) and Annual precipitation (b) of 3 stations with scenarios

It can be seen that the best estimations obtained when these stations are belongs to Region 4 although it can be seen visually that stations 21 to 23 are very close to Region 5. Homogenous regions are divided based on the same rainfall characteristics, precipitation of stations in the same region are therefore similar. It can be concluded that the determination of homogenous regions is crucial for the proposed method in the estimation of missing daily rainfall data at an ungauged site.

4.5 Conclusions

The estimation of missing daily precipitation series for ungauged sites based on the daily rainfall data of neighboring stations is essential for Vietnam region. The estimated data are useful for various applications in practice such as the construction of IDF relations for design and planning of urban infrastructures for regions where the sub-daily rainfall data are limited or unavailable.

It was found that Vietnam can be divided into 7 sub-regions in terms of meteorology based on PCA of daily series rainfall data from 155 stations across the country. The PCA works well for the estimation of missing daily rainfall data in Vietnam based on the identification of homogeneous regions using data for monthly time scale. However, this approach is considered to be sensitive with the time scale. Hence, it is recommended to apply this approach with some caution.

The two-stage daily rainfall interpolation can be used to estimate daily rainfall data for ungauged sites using rainfall information available at the neighboring stations located within the same homogeneous region. The proposed estimation method can provide good estimates of annual rainfall amounts and the number of rain days; however, there is still some limitation in the estimation of extreme values.

Chapter 5: Evaluation of the reliability of present and future NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP) regional climate simulations over Canada

5.1 Introduction

Global warming is currently a critical issue that every nation has to deal with. It has been recognized that the global climate has significantly changed over past 100 years (IPCC, 2014). These changes might have serious impacts on various hydrologic processes (Miller et al., 2003; Whitfield et al., 2003; Ryu et al., 2011; Assani et al., 2012). To understand and predict the climate change, past trends as well as the projections of future climates for different scenarios have been conducted in many studies (Besaw et al., 2010; Candela et al., 2012; Yeo and Nguyen, 2014; Nguyen et al., 2018). General Circulation Models (GCMs) have been commonly used for evaluating the effects of climate change on the hydrological regime under different scenarios of greenhouse gas emissions. While these GCMs could represent well the main features of the global distribution of basic climate parameters (Randall et al., 2007), they still cannot reproduce accurately the details of regional climate conditions at temporal and spatial scales of relevance to hydrological impacts and adaptation studies (Nguyen et al., 2006; Maraun, 2016). This is because outputs from these GCMs are usually at resolutions that are too coarse for many climate change impact studies, generally greater than 2.5° for both latitude and longitude (approximately 250 km).

To refine the GCM coarse grid resolution climate projection data to much finer spatial resolutions (regional or local scales) for the reliable assessment of climate change impacts, different downscaling techniques have been approached to resolve this scale discrepancy (Wilby et al., 2002; Fowler et al., 2007; Nguyen and Nguyen, 2008; Maraun et al., 2010; Khalili and Nguyen, 2016; Gooré Bi et al., 2017). It can be divided into two main categories: statistical downscaling (SD) and dynamical downscaling (DD). Some downscaled regional gridded datasets can be showed in below Table 5-1. In terms of SD, two commonly-used datasets are: the Pacific Climate Impacts Consortium (PCIC) and the National Aeronautics Space Administration (NASA) Earth Exchange Global Daily Downscaled Projections (NEX-GDDP) (Thrasher et al., 2012; Werner et al., 2019). In terms of DD, coordinated dynamical downscaling comparisons have been undertaken as part of the North American Regional Climate Change Assessment Program (Mearns et al., 2014) and the North America Coordinated Regional Downscaling Experiment (Mearns et al., 2017). For CORDEX, simulations were run at resolutions of approximately 25 km and 50 km.

Table 5-1. Summaries of available precipitation gridded datasets/reanalysis products

Downscaling method	**Dataset**	**Grid**	**Year available**	**Duration**
Dynamic Downscaling	NARCCAP	50km	2014	1971-2000 2041-2070
	NA-CORDEX	25-50km	2017	1950-2100
Statistical downscaling	PCIC	1/12 degree (~10x10km)	2019	1950-2100

Downscaling method	Dataset	Grid	Year available	Duration
	NEX-GDDP	1/4 degree (~25x25km)	2012	1950-2100

(*NARCCAP: The North American Regional Climate Change Assessment Program; CORDEX: the Coordinated Regional Downscaling Experiment; PCIC: The Pacific Climate Impacts Consortium; NEX-GDDP: NASA Earth Exchange Global Daily Downscaled Projections)

Recent studies have been conducted to analyse daily extremes of precipitation at the global scale for both historical observed and gridded downscaled data. Alexander and his colleagues found an increase trend of daily maximum and annual precipitation from more than 600 stations covering the Northern Hemisphere and parts of Australia during the 20th century (Alexander et al., 2006). At the global scale, approximately two-thirds of a global dataset of more than 8000 historical stations with the period of record from 1900 to 2009 indicated an increase trend of daily precipitation (Westra et al., 2013). In terms of analysis of downscaled datasets, Kharin and Zwiers (2000) found a positive trend of daily annual maximum precipitation at majority on the globe in the 20th century using data of the first generation Canadian Global Coupled Model (CGCM1). Min and his colleagues (2008) showed a similar result of precipitation trend using ECHO-G model and the third generation Canadian Global Coupled Model (CGCM3).

For Canada region, several studies of precipitation trends have been conducted; however, these studies were interested in total annual precipitation. For instance, Mekis and Vincent (2011) proves an annual rainfall increase of around 12.5% for the period of 1950 - 2009. Some other researchers also indicated the similar positive trend across Canada in terms of annual precipitation during the late 20th and early 21st centuries (Zhang et al., 2000; Zhang et al., 2011; Thistle and Caissie, 2013; Mekis et al., 2018; Vincent et al., 2018). While many water management applications (i.e., design of urban storm drainage systems, flood management and infrastructure operations) require extreme rainfall information in forms of rainfall intensity-duration-frequency (IDF). In order to construct the IDF curves, annual maximum rainfall series (AMS) of different rainfall durations from a few minutes to days are obtained. However, short-duration extreme rainfall data are very limited or even unavailable for a given location of interest while the daily extreme rainfall records are often available. For instance, Environment Canada provides short-duration extreme rainfall data of nine rainfall durations (from 5 minutes to 24 hours) and IDF relations for approximately 600 stations across Canada (Environment Canada, 2020) whereas the total rain-gauged stations from Environment Canada and their partners is more than 1700 stations (Mekis et al., 2018).

To deal with locations of interest where sub-daily and/or sub-hourly data are limited or unavailable, a scaling method to infer the sub-daily and/or sub-hourly extreme from daily extreme rainfalls has been proposed by Nguyen and his group (Nguyen et al., 2007; Nguyen et al., 2018; Nguyen and Nguyen, 2020). This proposed technique can also be applied for locations where daily precipitation observations are unavailable (ungauged sites) using downscaled regional gridded data. There is no doubt that it could bring many benefits to engineering practices in terms of design and management. It is likely, however, IDF relations obtained from this method are relied on variability of extreme rainfalls. This study aims therefore at performing a trend analysis of daily annual maximum precipitation using historical observed data and downscaled regional gridded data for the present and future climates over Canada.

5.2 Data

5.2.1. Historical observed data

The observation datasets were initially considered for this study: historical data and ANUSPLIN. Historical data are available from Environment Canada's website with the period of record from 1840 to present (Environment Canada, 2020). ANUSPLIN is a gridded observation dataset based on non-parametric fitting technique (Hutchinson et al., 2009). Hutchinson (2009) found daily precipitation for Canada region of ANUSPLIN produce a large error. It is therefore this study only considers the historical data from Environment Canada.

Among approximately 600 stations across Canada (Mekis et al., 2018), 175 stations with the record length of more than 30 years and passed the trend detection test were selected for this study. The high density of stations is located in the southern Ontario. The observed stations are quite limited in the Northern part of Canada. Listing from the West to the East coast, and from the North to South, there are 03 stations from Yukon (YT), 05 stations from Northwest Territories (NT), 37 stations from British Columbia (BC), 16 stations from Alberta (AB), 07 stations from Saskatchewan (SK), 13 stations from Manitoba (MB), 50 stations from Ontario (ON), 25 stations from Quebec (QC), 05 stations from New Brunswick (NB), 05 stations from Nova Scotia (NS), 02 stations from Prince Edward Island (PE) and 07 stations from Newfoundland and Labrador provinces. The location of selected stations is presented in Figure 5-1, more detailed information of these stations can be found in Appendix.

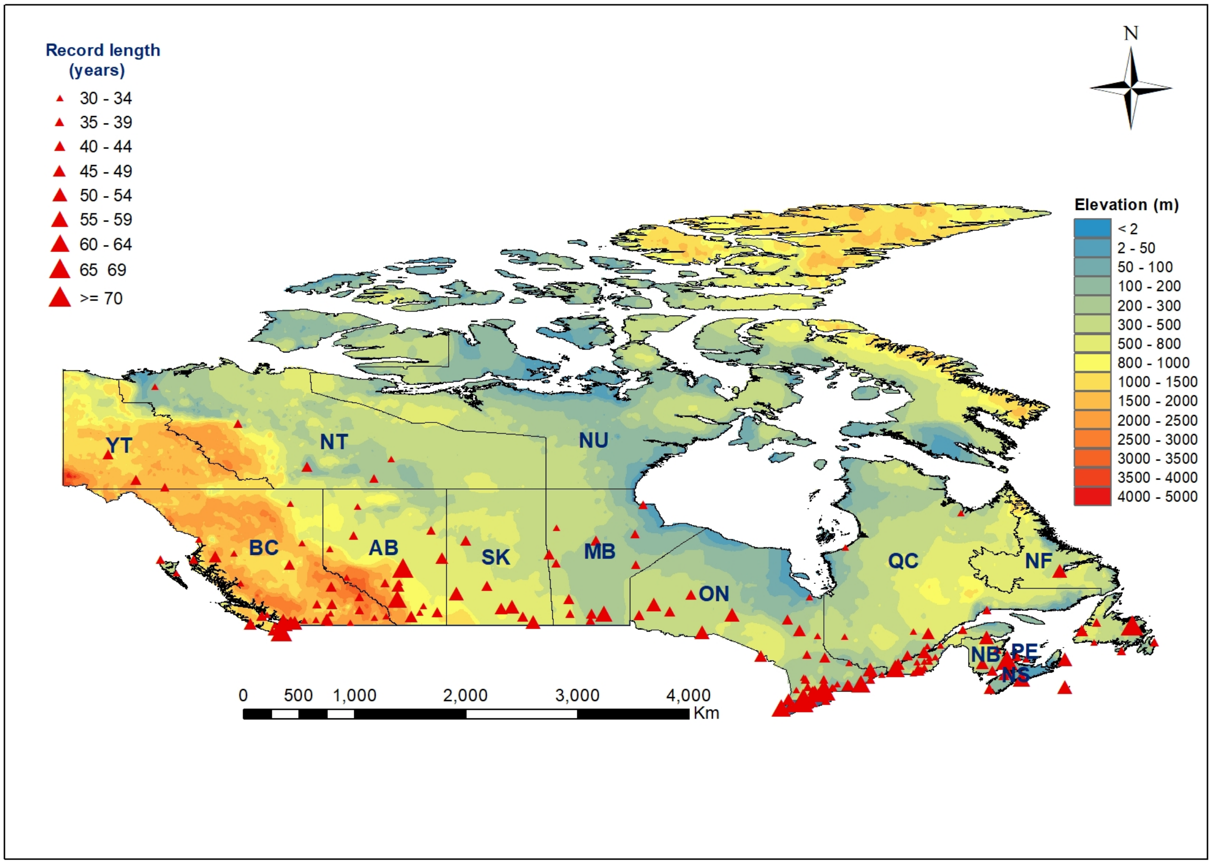


Figure 5-1. Selected stations over Canada

These stations were selected based on the high quality, the adequate length of available historical records, and the representative spatial distribution of the rain-gauges. To ensure the high quality of selected data, only historical observed data provided by the Atmospheric Environmental Service of Environment Canada were employed for this study. Every selected station must be more than 30 years of record and pass the Mann–Kendall test for trend detection. In addition, the raingauges were chosen from different geography locations to partially represent the diverse climatic conditions of Canada.

5.2.2. Downscaled regional gridded data

Due to the limitation of scope and access, only NASA Earth Exchange Global Daily Downscaled Projections (NEX-GDDP) data is analyzed. NEX-GDDP is a daily downscaled dataset (~ 25kmx25km resolution) released in June 2012 by NASA. This dataset was generated from 21 General Circulation Models (GCMs - shown in Table 5-2) runs conducted under the Coupled Model Intercomparison Project Phase 5 (CMIP5) and across two of the four greenhouse gas emissions scenarios known as Representative Concentration Pathways (RCP4.5 and RCP8.5) based on the bias-correction spatial disaggregation downscaling technique (Thrasher et al., 2012). The climate projections include daily maximum temperature, minimum temperature, and precipitation for the historical periods of 1950-2005 and the future period of 2006-2100. Canada is having actions on climate change and projected to archive low emissions level by mid of 21st century (Canada, 2016), it is therefore this study selects projected results of the Representative Concentration Pathway 4.5.

Table 5-2. Information about the 21 Coupled Model Intercomparison Project 5 (CMIP5) general circulation models (GCMs)

Number	**Model**	**Country and institution**
1	ACCESS1-0	Commonwealth Scientific and Industrial Research Organization and Bureau of Meteorology, Australia
2	BCC-CMS1-1	Beijing Climate Center, China
3	BNU-ESM	Institute of global change and Earth System Sciences, Beijing Normal University, China

Number	Model	Country and institution
4	CanESM2	Canadian Centre for Climate Modeling and Analysis, Canada
5	CCSM4	National Center for Atmospheric Research, America
6	CESM1-BGC	National Center for Atmospheric Research, America
7	CNRM-CM5	Centre National de Recherches Meteorologiques, Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique, France
8	CSIRO-Mk3-6-0	Commonwealth Scientific and Industrial Research Organization/Queensland Climate Change Centre of Excellence, Australia
9	GFDL-CM3	Geophysical Fluid Dynamics Laboratory, America
10	GFDL-ESM2G	Geophysical Fluid Dynamics Laboratory, America
11	GFDL-ESM2M	Geophysical Fluid Dynamics Laboratory, America
12	inmcm4	Institute of Numerical Calculation, Russia
13	IPSL-CM5A-LR	Institut Pierre-Simon Laplace, France
14	IPSL-CM5A-MR	Institut Pierre-Simon Laplace, France
15	MIROC5	Atmosphere and Ocean Research Institute, Japan
16	MIROC-ESM	Atmosphere and Ocean Research Institute, Japan

Number	Model	Country and institution
17	MIROC-ESM-CHEM	Atmosphere and Ocean Research Institute, Japan
18	MPI-ESM-LR	Max Planck Institute for Meteorology, Germany
19	MPI-ESM-MR	Max Planck Institute for Meteorology, Germany
20	MRI-CGCM3	Max Planck Institute for Meteorology, Germany
21	NorESM1-M	Norway Consumer Council, Norway

5.3 Methodology

5.3.1. Mann-Kendall test

The results in this paper are based on a popular statistical method for testing whether time series data - the Mann–Kendall nonparametric trend - to evaluate whether there is a monotonic trend in the series. The advantage of this test is that it does not make any assumptions on the distribution of the data, other than that under the null hypothesis, the data are independently distributed in time.

The Mann–Kendall test is a commonly used non-parametric test for evaluating the presence of monotonic trends in time series data (Chandler and Scott, 2011). The test has been applied in analyzing trends of rainfall extremes data (Westra et al., 2013). In this study, the Mann–Kendall analysis was conducted at the significant level of 5% using the MATLAB software (version 2020a).

5.3.2. Sen's method

Sen's method has been commonly used to estimate trends in climate series thanks to its reliability by minimizing sensitivity of outliers in the series in comparison with conventional least-squares methods (Zhang et al., 2000; Fernandes and G. Leblanc, 2005; Zhang et al., 2011).

Slope and intercept were computed according to Sen's method (Sen, 1968) as follow:

$$m_k = \frac{R_j - R_i}{j - i}$$

for ($1 \leq i < j \leq n$) where m is the slope, R denotes the variable, n the number of data, and i, j are indices. The median from all slope s then is calculated: s = Median (m_k).

Trend of precipitation were exanimated for two time period: historical period 1950-2005 and projected period 2006 - 2100 for all stations. To maintain the consistence and accuracy, the trends of NEX-GDDP historical data were computed using the same duration of historical observed data given by Environment Canada.

5.4 Results and discussions

The slopes of 175 historical gauged stations across Canada given by Environment Canada are showed in Figure 5-2. The number of stations for each province by different change intervals of slopes for 175 gauged stations is presented in Table 5-3. According to historical observed data, the number of stations with increase trends is slightly higher than ones with decrease trend (around 55% vs 45%). The highest decrease trend of 40% is Elora_rcs (ON) while the highest increase trend of 50% is at Summerside station (PE). It can be seen that most of Northern and Eastern provinces have increase trends while trends of stations located in ON, MB and BC are quite complicated. SK is the only province that the majority of stations have decrease trend.

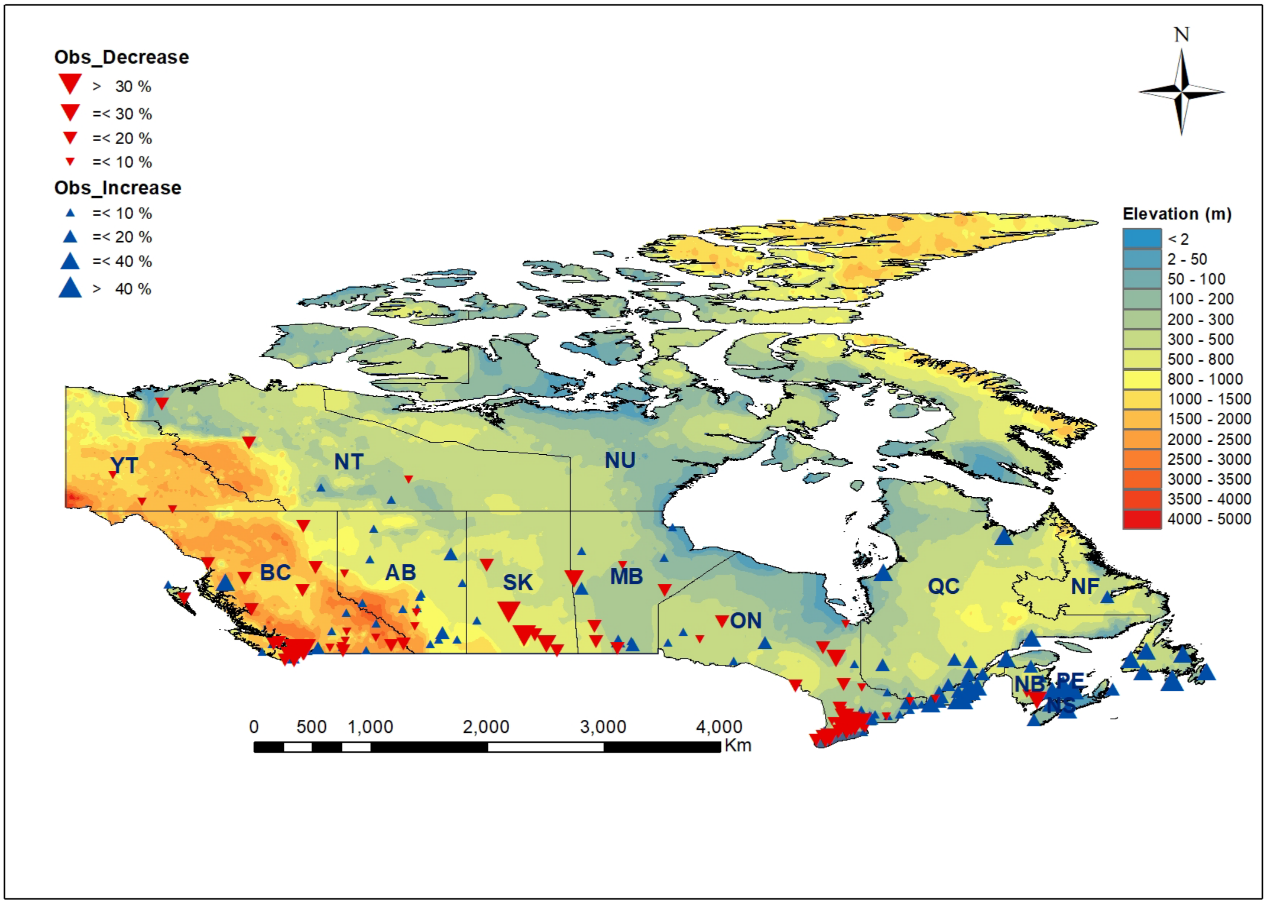


Figure 5-2. Change of slope - Historical observed data. Red triangles: stations with decrease trend, blue triangles: stations with increase trend

Table 5-3. Trend statistics of historical observed data

No.	Province Code	Count of stations by Slope change (%)								
		-[40 30)	-[30 20)	-[20 10)	-[10 0)	[0 10)	[10 20)	[20 30)	[30 40)	[40 50]
1	NL	0	0	0	0	0	1	0	5	1
2	PE	0	0	0	0	0	0	0	0	2
3	NS	0	0	0	0	1	3	0	1	0
4	NB	0	1	0	1	1	1	0	1	0
5	QC	0	0	0	1	0	12	10	2	0
6	ON	1	8	13	9	14	4	1	0	0
7	MB	0	1	4	1	4	3	0	0	0
8	SK	2	1	3	0	1	0	0	0	0
9	AB	0	0	0	3	11	2	0	0	0
10	BC	3	1	15	6	9	2	0	1	0
11	NT	0	0	2	1	2	0	0	0	0
12	YT	0	0	0	3	0	0	0	0	0
	Total	**6**	**12**	**37**	**25**	**43**	**28**	**11**	**10**	**3**

Trend analyses of NEX-GDDP historical data were conducted based on the dataset over 175 grid boxes (called station hereafter) over Canada that historical observed data are available. The slopes of NEX-GDDP dataset for historical period is illustrated in Figure 5-3 and summarized in Table 5-4. It showed that the trends ranging from -10% to + 20%, only one station has the decrease trend of 10.3% at Sparwood station in BC and two stations have increase trends of over 20% (Medicine_hat_rcs station and Nanaimo airport station in BC with 21.8% and 24%, respectively). It can be seen that more than 80% stations have increase trends. AB is the only province that has more station with decrease trends than ones with increase trends. Note that the value of slopes presented in Table 5-4 are the median from 21 GCMs for a single station. Detailed performance of all GCMs is showed in Figure 5-4.

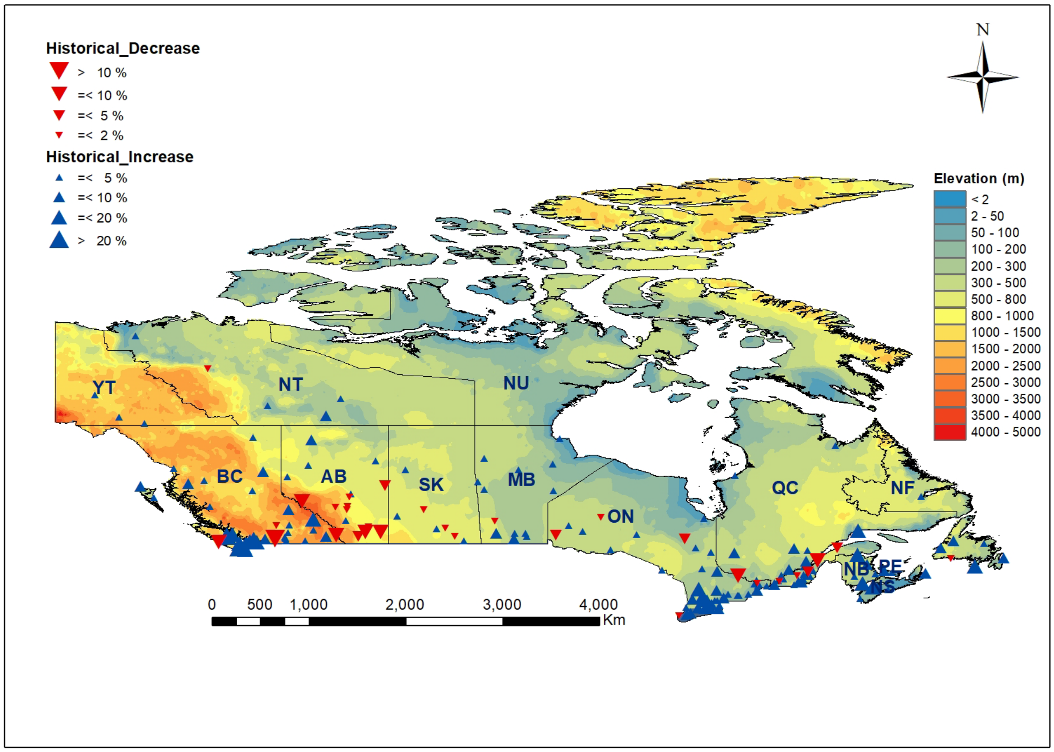


Figure 5-3. Change of slope - Historical NEX-GDDP data. Each value is the median from 21 GCMs. Red triangle: stations with decrease trend, blue triangle: stations with increase trend

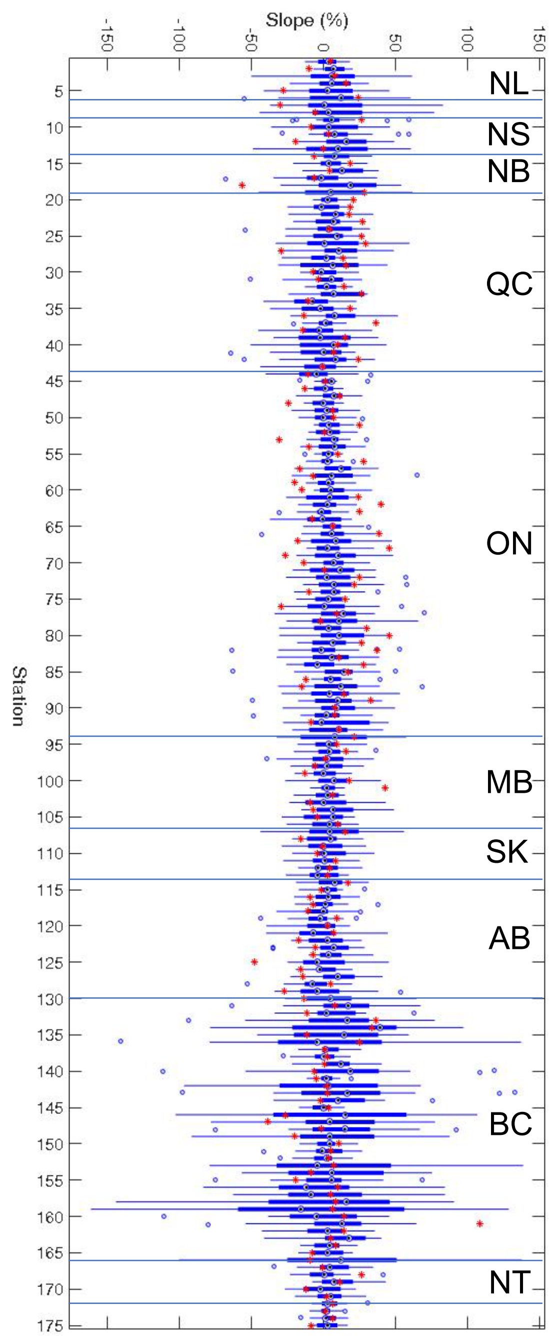
Table 5-4. Trend statistics of NEX-GDDP historical period

No.	Province code	Count of stations by Slope change (%)				
		-[30 - 20)	-[10 - 0)	[0 - 10)	[10 - 20)	[20 - 30]
1	NL	0	1	5	1	0
2	PE	0	0	2	0	0
3	NS	0	0	4	1	0
4	NB	0	0	3	2	0
5	QC	0	5	19	1	0
6	ON	0	6	41	3	0
7	MB	0	1	12	0	0
8	SK	0	3	4	0	0
9	AB	0	10	6	0	0
10	BC	1	3	22	9	2
11	NT	0	1	4	0	0
12	YT	0	0	3	0	0
	Total	**1**	**30**	**125**	**17**	**2**

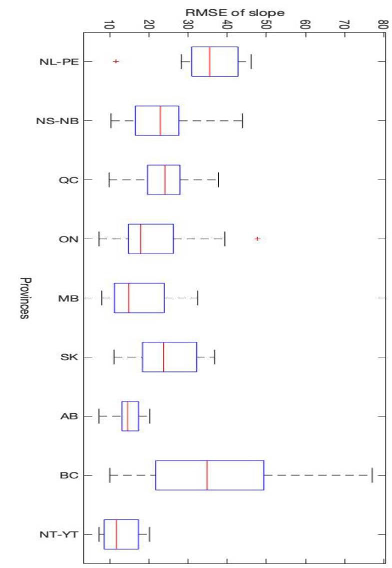
The boxplot (a) of Figure 5-4 presents slopes of historical NEX-GDDP data (blue) in comparison with slopes of historical observed data (red) for all 175 stations. Each station is illustrated by one boxplot that constructed from 21 values of 21 GCMs. Root mean square error (RMSE) of 21 GCMs for each station was calculated and plotted in Figure 5-4 (b), and the average of RMSEs for each province is presented in Figure 5-4 (c). The RMSE is given as below:

$$RMSE = \sqrt{\frac{1}{N}\sum(S_{GCMs} - S_{Observed})^2}$$

where *S* indicates the slopes of stations and *N* is the number of sample size. It could be seen from Figure 5-4 that NEX-GDDP data for BC have widespread slopes from 21 GCMs and average RMSEs of all stations in this province is also highest. It could be explained that downscaling models are limited to capture complex topography of mountainous areas in BC, driving less accurate results for this area.



(a)



(b)

No	Provinces	RMSE
1	NL-PE	34.33
2	NS-NB	23.78
3	QC	23.70
4	ON	20.87
5	MB	17.04
6	SK	24.00
7	AB	14.94
8	BC	35.58
9	NT-YT	12.79

(c)

Figure 5-4. (a) AMS precipitation boxplot for historical period for 175 stations - boxplot (Blue) contains 21 values of 21 GCMs, each value is the average of historical period. Red points are observed data, each point is the average of record period; (b): Boxplot

To assess the performance of 21 GCMs in NEX-GDDP data, a ranking score has been conducted based on the mean (MEAN) and the standard deviation (STD). Steps of calculation are: (*i*) compute the MEAN and STD all stations (historical observed and 21GCMs of historical NEX-GDDP data) over period of time; (*ii*) calculate the absolute bias of 21 GCMs and observe data for both the MEAN and STD; (*iii*) for each station: ranking each GCM according to the bias values calculated in previous step, a rank of 1 (the best model) is given to the GCM having smallest bias, apply for both the MEAN and STD; (*iv*) count the number of stations for each ranking from 1 to 21 for all GCMs; (*v*) ranking score for each GCM: sum of rank of that GCM for 175 stations, apply for both the MEAN and STD; (*vi*) ranking GCMs: lower ranking score, better model. Each cell in Table 5-5 shows the total number of stations each ranking for 21 GCMs by the MEAN and STD (upper number: count by MEAN ranking and lower number: count by STD ranking). For instance, GCM "bcc-csm1-1" is ranked 1 at 6 stations in terms of the MEAN and 63 stations in terms of STD.

According to the Table 5-6, it can be seen that the second-generation Canadian Earth System Model (CanESM2) performs the best among 21 GCMs over Canada with rank 2nd in terms of the MEAN and 1st in terms of the STD. In particular, the GCM "CanESM2" is ranked 1 to 4 at more than 65% of stations and ranked 1 to 3 at more than 75% of stations. Beside "CanESM2", two models from China ("bcc-csm1-1", "BNU-ESM") and two models from Japan ("MIROC-ESM", "MIROC-ESM-CHEM") are also considered well performing over Canada region. GCMs "CESM1-BGC", "MIROC5" and "MRI-CGCM3" are considered as worst models over Canada for this study. However, it can be seen that there are high uncertainties from results of total 21 GCMs that unable to estimated and also require an in-depth evaluation of the performance of these all GCMs, it is recommended to use the median of all GCMs for calculation.

Table 5-5. Count of stations ranking by GCMs in terms of MEAN (upper number) and STD (lower number) of AMS NEX-GDDP historical data

Rank GCM	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
bcc-csm1-1	6	27	25	21	25	18	9	7	5	5	1	1	2	2	2	3	5	4	2	4	1
	63	11	52	2	2	0	0	1	3	2	0	2	1	0	0	2	1	1	9	10	13
BNU-ESM	7	14	17	14	24	33	14	12	3	6	2	6	2	0	5	4	3	2	2	2	3
	32	62	38	1	0	1	0	2	1	0	0	1	0	0	0	3	2	1	4	16	11
CanESM2	30	30	24	29	14	12	6	2	2	0	4	2	0	2	0	2	1	2	4	6	3
	42	60	32	1	2	0	1	0	0	0	2	0	0	0	0	0	0	0	13	9	13
CCSM4	10	10	7	5	4	2	0	2	3	3	1	2	2	2	6	4	10	17	25	32	28
	2	5	10	37	9	4	3	5	3	1	5	5	9	4	8	5	9	20	15	8	8
CESM1-BGC	12	11	4	3	3	3	3	0	1	1	1	5	1	6	6	7	8	13	23	30	34
	30	5	5	8	4	2	1	1	0	1	2	2	3	0	2	3	1	6	7	25	67
CSIRO-Mk3-6-0	0	0	0	0	3	6	8	8	3	9	21	17	24	20	18	11	7	9	6	3	2
	3	29	2	3	7	4	2	3	1	2	2	2	0	6	2	1	2	3	19	62	20
GFDL-CM3	1	0	1	3	2	4	13	15	19	18	30	20	15	12	6	2	1	0	3	4	6
	0	0	19	13	7	6	10	1	4	4	5	2	7	4	4	7	5	23	42	7	5
GFDL-ESM2G	2	0	4	1	7	14	17	23	22	21	14	8	15	7	4	7	3	2	2	0	2
	0	0	2	14	13	9	4	7	10	7	4	5	4	6	6	7	35	34	5	1	2
GFDL-ESM2M	0	2	0	0	4	12	29	26	30	21	14	8	6	6	3	2	3	1	3	4	1
	0	0	0	5	15	17	14	11	9	6	4	6	8	6	14	29	16	6	5	1	3
inmcm4	0	0	0	3	3	1	5	3	4	8	6	14	18	13	19	23	16	17	14	5	3
	0	0	1	3	7	15	17	10	12	5	3	5	9	10	27	24	8	10	5	1	3
IPSL-CM5A-LR	54	21	16	21	16	11	3	2	2	3	1	1	0	1	4	2	0	4	2	3	8
	0	0	0	3	14	12	11	12	7	9	13	8	13	28	12	12	9	3	4	5	0
IPSL-CM5A-MR	0	0	1	1	3	4	9	22	24	27	18	13	9	5	13	10	8	3	2	1	2

Rank / GCM	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
	0	0	1	17	16	7	12	10	7	9	15	21	14	12	5	1	9	11	2	3	3
MIROC5	2	2	5	8	3	3	0	3	3	1	2	1	7	6	9	8	30	24	29	18	11
	1	0	0	4	3	7	6	5	17	16	21	15	25	16	11	8	7	4	6	3	0
NorESM1-M	0	0	0	3	7	12	27	19	20	21	16	12	6	7	3	5	4	3	3	5	2
	0	1	0	4	3	5	9	9	16	25	31	18	14	9	9	7	7	2	2	2	2
ACCESS1-0	0	1	2	0	1	1	3	5	10	15	9	21	17	24	15	9	13	8	6	11	4
	0	0	4	4	10	8	9	17	17	21	15	20	8	16	5	8	5	1	2	3	2
CNRM-CM5	1	2	12	2	2	2	8	5	0	0	3	5	7	11	18	21	22	24	6	14	10
	1	0	2	4	5	3	6	13	19	20	10	12	18	14	8	10	9	7	7	4	3
MIROC-ESM	22	25	22	32	19	14	6	5	3	3	0	0	1	2	0	2	6	3	6	1	3
	0	2	0	1	6	14	13	20	13	12	9	20	10	14	13	8	8	5	4	2	1
MIROC-ESM-CHEM	20	22	30	19	27	15	4	2	5	2	4	0	2	3	4	3	2	3	2	3	3
	0	0	0	16	11	20	15	14	10	13	8	7	6	13	8	14	7	7	4	1	1
MPI-ESM-LR	0	1	0	1	3	4	7	2	6	6	14	15	17	28	10	18	12	10	9	6	6
	0	0	0	7	17	20	18	15	10	7	8	10	5	7	12	11	11	7	6	2	2
MPI-ESM-MR	0	1	0	2	2	4	2	7	7	6	11	19	19	17	27	20	12	8	3	1	7
	0	0	4	18	16	14	11	10	7	4	12	4	11	9	14	9	9	14	4	5	0
MRI-CGCM3	8	6	5	7	3	1	1	5	3	0	2	5	5	1	4	11	9	18	23	22	36
	1	1	2	10	9	7	13	8	12	10	7	7	10	5	11	8	13	10	10	5	16

Table 5-6. Summary ranking scores for 21 GCMs

No.	GCM	MEAN score	STD score	Rank by MEAN	Rank by STD
1	bcc-csm1-1	1117	1062	5	3
2	BNU-ESM	1228	1044	6	2
3	CanESM2	936	1001	2	1
4	CCSM4	2598	1971	18	10
5	CESM1-BGC	2622	2472	19	20
6	CSIRO-Mk3-6-0	2258	2490	12	21
7	GFDL-CM3	1947	2274	10	19
8	GFDL-ESM2G	1730	2231	8	18
9	GFDL-ESM2M	1703	2060	7	13
10	inmcm4	2518	2130	16	16
11	IPSL-CM5A-LR	926	2035	1	12
12	IPSL-CM5A-MR	1957	1880	11	7
13	MIROC5	2703	2087	21	14
14	NorESM1-M	1792	1979	9	11
15	ACCESS1-0	2380	1867	14	5
16	CNRM-CM5	2521	2102	17	15
17	MIROC-ESM	1010	1961	3	9

No.	GCM	MEAN score	STD score	Rank by MEAN	Rank by STD
18	MIROC-ESM-CHEM	1033	1814	4	4
19	MPI-ESM-LR	2410	1875	15	6
20	MPI-ESM-MR	2368	1892	13	8
21	MRI-CGCM3	2665	2180	20	17

The slopes of NEX-GDDP dataset for projection period under RCP 4.5 is presented in Figure 5-5 and Table 5-7. Figure 5-5 illustrated the slopes of all stations by magnitude and spatial spread while Table 5-7 counts the number of stations each province by different intervals of slopes. Similarity to historical period, trends of NEX-GDDP data for projection period are between -3% to 10%. Stations with decrease trend are mainly located in AB, SK and MB. It demonstrated that majority of stations have an increase trend with more than 90% number of stations. Stations with decrease trend are insignificant with less than 3% and mainly located in AB, SK, MB which the highest decrease slope of 2.3% at Flin-flon station (MB), only one station located in ON. 10 stations with high increase trend (more than 10% of slope) are entirely located in BC which the highest value of approximately 16% at Penticton airport station. It can be seen that extreme rainfall events are likely to occur frequently in BC other than other provinces of Canada in the future according to the projection of NEX-GDDP.

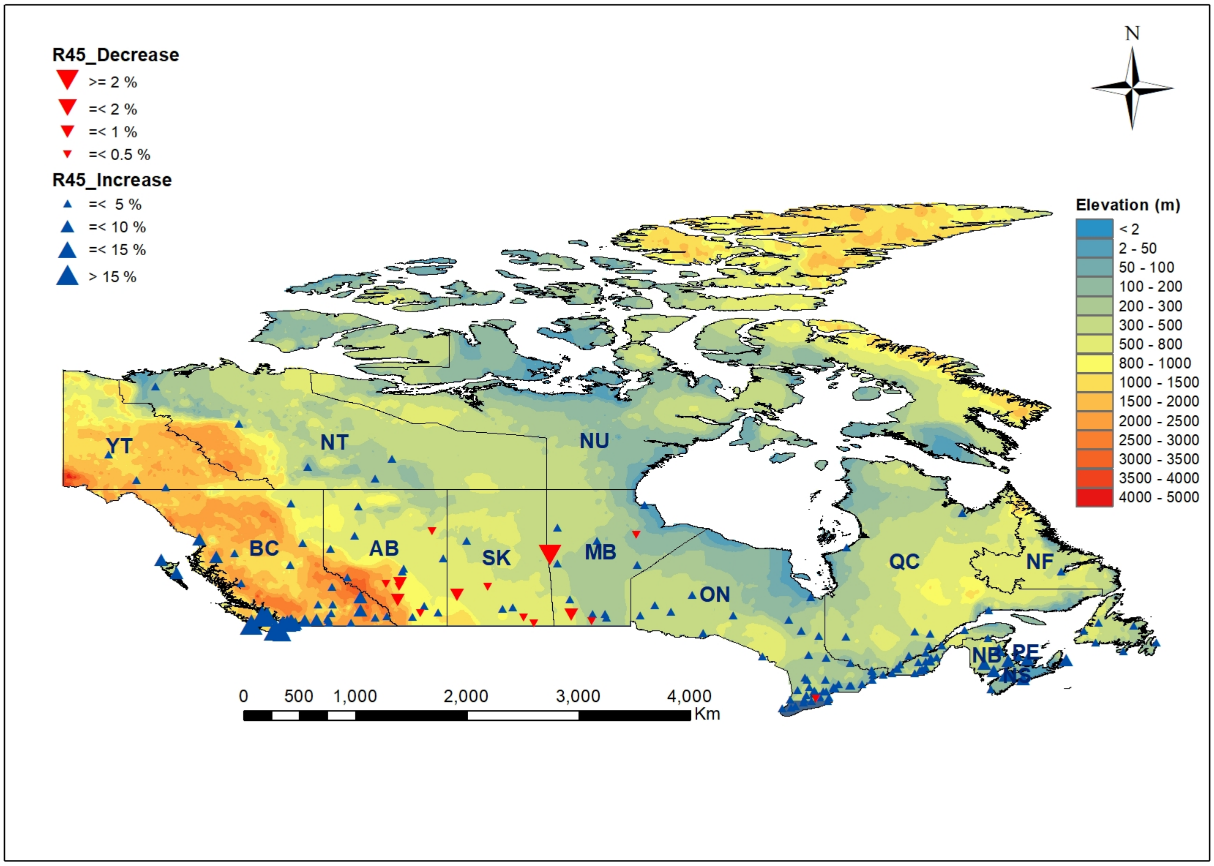


Figure 5-5. Change of slope - Gridded historical data (NEX-R45). Each value is the median from 21 GCMs. Red triangle: stations with decrease trend, blue triangle: stations with increase trend

Table 5-7. NEX-GDDP trend statistics of projection period (R45)

No.	Province code	Count of stations by Slope change Slope (%)		
		-[3 - 0)	[0 - 10)	[10 - 20]
1	NL	0	7	0
2	PE	0	2	0
3	NS	0	5	0

No.	Province code	Count of stations by Slope change Slope (%)		
		-[3 - 0)	[0 - 10)	[10 - 20]
4	NB	0	5	0
5	QC	0	25	0
6	ON	1	49	0
7	MB	4	9	0
8	SK	4	3	0
9	AB	6	10	0
10	BC	0	27	10
11	NT	0	5	0
12	YT	0	3	0
	Total	**15**	**150**	**10**

5.5 Conclusions

This study evaluates the trends of daily annual maximum precipitation using data from 175 high-quality historical observed station across Canada and 25kmx25km resolution downscaled regional gridded data NEX-GDDP for past period from 1950 to 2005 and projection period from 2006 to 2100. The trend is estimated using Sen's method. Overall, it can be concluded that majority of stations have increase trends for all periods of time. According to historical observed data, there is around 40% of stations with decrease trends and most of them are located in the central and western Canada.

Downscaled regional gridded data NEX-GDDP show increase trends for both historical period and projection period with around 80% and 90% of stations, respectively. While trends of NEX-GDDP historical data are mainly from -10% to + 20%, trends of NEX-GDDP projection data are only between -3% to 10%. Most of stations with decrease trend are located in the western Canada and stations with highest trends are mainly in BC. Due to the complex topography of mountainous area, stations in the province of BC have widespread trends given by different GCMs.

Regarding the performance of all 21 GCMs of NEX-GDDP data, CanESM2 model is considered the best model over 175 stations in Canada region. Results of CanESM2 model is ranked 1st for majority of stations in terms of the mean and standard deviation of daily annual maximum precipitation. Models from China and Japan are also considered to produce good results over Canada region.

Chapter 6: Evaluation of variability of precipitation and temperature extremes over Montreal region for present and future climates

6.1 Introduction

In recent years, global climate impacts have been recognized as one of the most the critical issues for many nations and/or regions all over the world. It has been recognized that the global climate has significantly changed over past 100 years (IPCC, 2014). To understand and predict the climate change, past trends as well as the projections of future climates for different scenarios have been conducted in many studies (Creutin and Obled, 1982; Besaw et al., 2010; Candela et al., 2012; Yeo and Nguyen, 2014; Nguyen et al., 2018). In Canada, some studies have indicated an increase trend in both temperature and precipitation with an average increase of around 1.4°C for air temperature and around 12.5% for annual rainfall during the second half of the 20th century (Mekis and Vincent, 2011; Zhang et al., 2011). These changes might have significant impacts on various hydrologic processes (Miller et al., 2003; Whitfield et al., 2003; Ryu et al., 2011; Assani et al., 2012).

General Circulation Models (GCMs) have been commonly used for evaluating the effects of climate change on the hydrological regime under different scenarios of greenhouse gas emissions. While these GCMs could represent well the main features of the global distribution of basic climate parameters (Randall et al., 2007), they still cannot reproduce accurately the details of regional climate conditions at temporal and spatial scales of relevance to hydrological impacts and adaptation studies (Nguyen et al., 2006). This is because outputs from these GCMs are usually at resolutions that are too coarse for many climate change impact studies, generally greater than 2.5^{o} for both latitude and longitude (approximately 250km) as shown in Figure 6-1. To refine the GCM coarse grid resolution climate projection data to much finer spatial resolutions (regional or local scales) for the reliable assessment of climate change impacts, different downscaling methods have been proposed to resolve this scale discrepancy (Wilby et al., 2002; Fowler et al., 2007; Nguyen and Nguyen, 2008; Maraun et al., 2010; Khalili and Nguyen, 2016; Gooré Bi et al., 2017). These downscaling methods can be generally classified into two broad categories: dynamical downscaling (DD) and statistical downscaling (SD). It has been widely recognized that the SD methods offer several practical advantages over the DD procedures, especially in terms of flexible adaptation to specific study purposes, and inexpensive computing resource requirement (Xu, 1999; Prudhomme et al., 2002). In addition, SD methods can be used to spatially disaggregate GCM outputs to regional scales or local/point scales (a single site or multi-sites) (Wilby et al., 2002; Khalili and Nguyen, 2016; Werner and Cannon, 2016). Furthermore, when dealing with a large ensemble of GCMs, the SD methods are often in favor because of their computational efficiency and effectiveness in producing physically plausible hydro-climatology data (Wood, 2004; Werner and Cannon, 2016).

Located on an island in the Saint Lawrence River, Montreal is the biggest city of Quebec province and second-largest city of Canada with the population of approximately 1.9 million (Statistics-Canada, 2016). Every year, the city has experienced frequent extreme weather events such as heavy storm rainfalls and heat waves that cause millions of property losses, and in some cases, the loss of human lives (City-of-Montreal, 2017). These types of extremes events are occurring with increasing frequency. For instance, more than 30 people were killed by a heat wave in Montreal in July 2018 (Cullinane, 2018). Another example is the spring flood in 2017 that affected thousands of people and millions of dollars of damages (Lau, 2017). Furthermore, August 2021 is considered hottest month on record for Montreal consisting of 5 heat wave events with 13 days with the temperature of above the 30-degree - compared to an average of 2 days for the month of August (Graham, 2021). Consequently, information on the spatial and temporal variations of these precipitation and temperature extremes for current and future climates is important for the planning and design of the City's its urban infrastructures to minimize the impacts of these natural disasters. Many studies have been conducted to assess the variability of temperature and precipitation processes in Canada and in other countries (Zhang et al., 2001; Arnbjerg-Nielsen et al., 2013; Thistle and Caissie, 2013; Benmarhnia et al., 2014; City-of-Montreal, 2017) However, very few studies have been carried out specifically on the daily precipitation and temperature extremes for the local City of Montreal region. Therefore, in the present study, a critical evaluation of the spatial and temporal variations of the daily annual maximum rainfalls and daily extreme temperatures over the Montreal region was conducted for the present and future climates using two different datasets that have been statistically downscaled by the Pacific Climate Impacts Consortium (PCIC, 2014) and the National Aeronautics Space Administration Earth Exchange Global Daily Downscaled Projections (NEX-GDDP) (Thrasher et al., 2012). Information of these two datasets will be detailed in section 6.2.

6.2 Numerical application

6.2.1 Data

Figure 6-1 shows a network of seven weather stations in the Montreal region. However, of these seven stations only Montreal-Pierre Elliott Trudeau International Airport (Dorval) and McGill stations have good quality of data with long historical records, other stations have either short historical records or a large number of missing data. Figure 6-1 also indicates the grids of the two downscaled datasets (red: NEX-GDDP; black: PCIC). It can be seen that the NEX-GDDP grid size is approximately nine time larger than the PCIC grid.

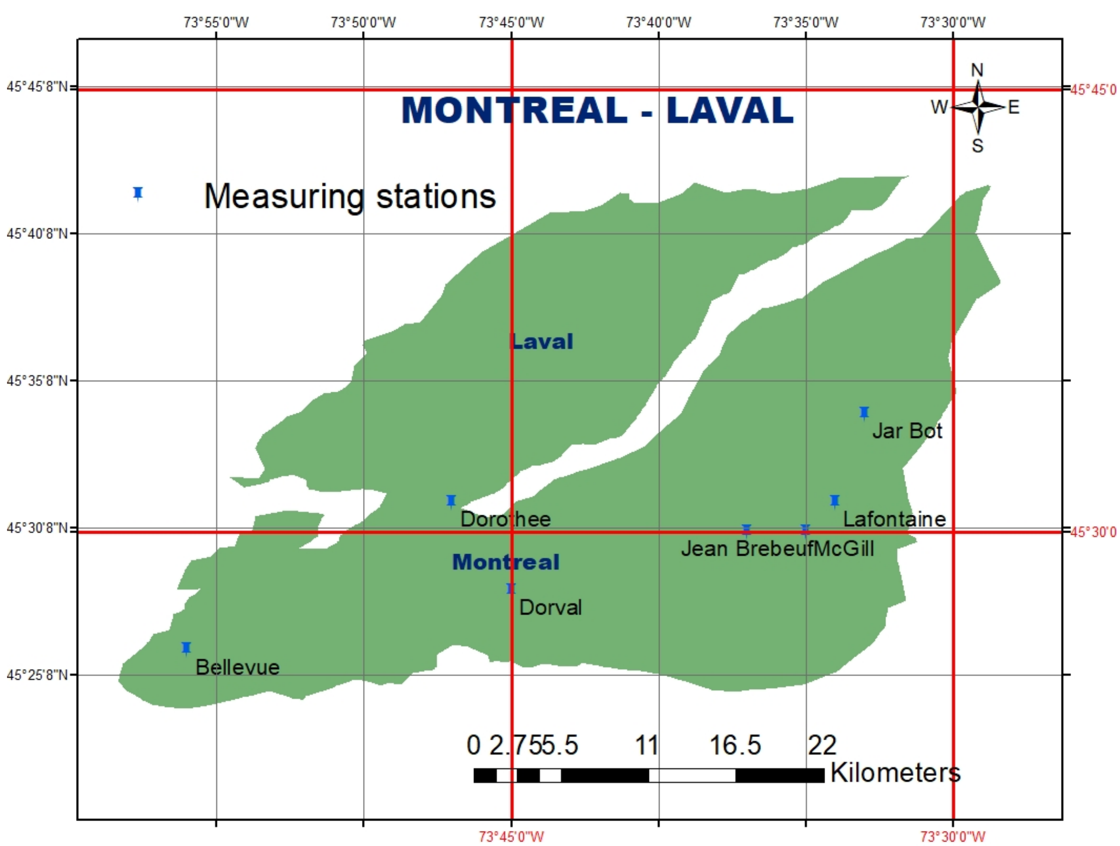


Figure 6-1. Location of measuring stations in Montreal region

Information of PCIC and NEX-GDDP datasets were summarized in Table 6-1 below:

Table 6-1. Summary of PCIC and NEX-GDDP datasets

	PCIC	**NEX-GDDP**	Note
Grid size (degree)	1/12	1/4	
Downscaling method	BCSD*, BCCAQ**	BCSD	
Number of GCMs	24	21	PCIC: 12 for BCCAQ and 12 for BCSD
Variables	T_{max}, T_{min}, Pr	T_{max}, T_{min}, Pr	
Timesteps	Daily	Daily	
Projection duration	1950-2100	1950-2100	
RCP*** scenarios	2.6; 4.5; 8.5	4.5; 8.5	

(BCSD*: bias-correction spatial disaggregation - see Werner and Cannon (2016) for further details; BCCAQ**: Bias Correction/Constructed Analogues with Quantile mapping reordering; RCP***: Representative Concentration Pathway)

In the present study, only gridded daily annual maximum precipitation and daily extreme temperature data were considered. These data were statistically downscaled from 10 GCMs corresponding to the RCP 4.5 scenario (see Table 6-2). For the present climates, the available historical data from Dorval and McGill stations and the PCIC and NEX-GDDP gridded data for the same 1961-1990 period were used. For the future climates, climate projections from the climate models corresponding to the RCP 4.5 scenarios for the 2006 – 2100 period were selected.

Table 6-2. 10 IPCC-CMIP5 climate models used in this study

GCM	Institution
ACCESS1-0	CSIRO (Commonwealth Scientific and Industrial Research Organisation, Australia), and BOM (Bureau of Meteorology, Australia)
CanESM2	Canadian Centre for Climate Modeling and Analysis
CCSM4	National Center for Atmospheric Research
CNRM-CM5	Centre National de Recherches Météorologiques/Centre Européen de Recherche et Formation Avancées en Calcul Scientifique
CSIRO-MK3-6-0	Commonwealth Scientific and Industrial Research Organization in collaboration with the Queensland Climate Change Centre of Excellence
GFDL-ESM2G	NOAA's Geophysical Fluid Dynamics Laboratory
INMCM4	Institute for Numerical Mathematics, Moscow, Russia
MIROC5	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies
MPI-ESM-LR	Max Planck Institute for Meteorology (MPI-M)
MRI-CGCM3	Meteorological Research Institute

6.2.2 Statistical indices

In addition, the root-mean-square error (RMSE) was used to compare the performance of the proposed model as given below:

$$RMSE = \sqrt{\frac{1}{N}\sum(SI_{model} - SI_{Observed})^2}$$

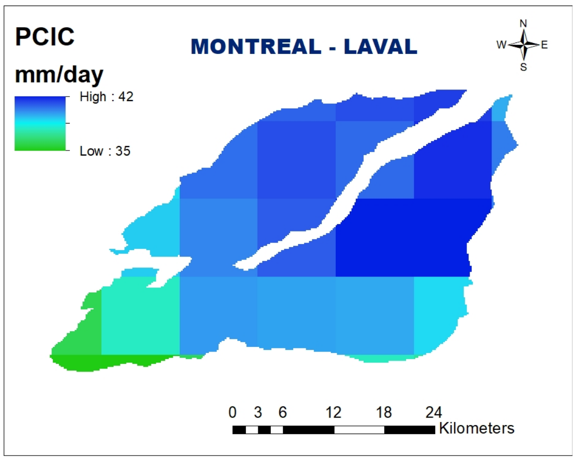
where *SI* indicates the value of the statistical indices and *N* is the number of sample size. The smaller value of the RMSE indicates the better accuracy of the model considered.

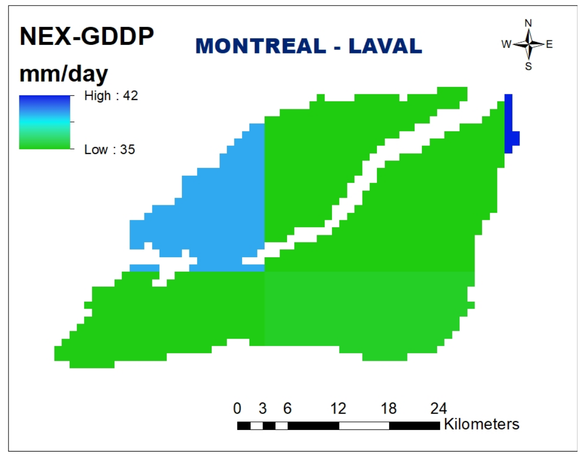
6.3 Results and discussions

6.3.1 Present climate

Figure 6-2 presents the spatial distribution of the downscaled daily annual maximum precipitations (AMPs) over the Montreal region from PCIC and NEX-GDDP datasets based on the average of ten GCMs. It can be seen that the mean precipitation given by NEX-GDDP is smaller than PCIC. More specifically, Table 6-3 shows the means of daily AMPs at Dorval and McGill stations in comparison with PCIC and NEX-GDDP data. Overall, the gridded downscaled data values are smaller than the observed data at a given station. PCIC data are 11.92% and 22.75% lower than observed AMP at Dorval and McGill stations, respectively, while the values from NEX-GDDP data are 29.24% and 35.92%, respectively. It is therefore necessary to perform a bias adjustment before these gridded downscaled data can be used in the planning and design of urban infrastructures.

For purposes of illustration, the results for daily AMP at Dorval Airport are shown in Figure 6-3 using the boxplots, and the results for temperature extremes are presented in Figure 6-4 and Figure 6-5. In addition, Table 6-4 presents the comparison using the root mean square error (RMSE) values for both precipitation and temperature extremes. In general, it can be seen that the PCIC data are more accurate for AMP and somewhat less accurate for temperature extremes as compared to the NEX-GDDP data. However, Figure 6-4 indicates that results given by PCIC are more robust with narrow boxplots in comparison with NEX-GDDP data. Regarding the standard deviation, NEX-GDDP data is more accurate for daily minimum temperature while PCIC data are more robust for daily maximum temperature.





PCIC NEX-GDDP

Figure 6-2. Daily AMPs over the Montreal region downscaled by PCIC and NEX-GDDP

Table 6-3. Mean of daily AMPs at Dorval and McGill stations

No.	Station	Mean of daily AMPs (mm/day)				
		Observed	PCIC	Different (%)	NEX-GDDP	Different (%)
1	Dorval	50.75	44.7	11.92	35.91	29.24
2	McGill	54.76	42.3	22.75	35.09	35.92

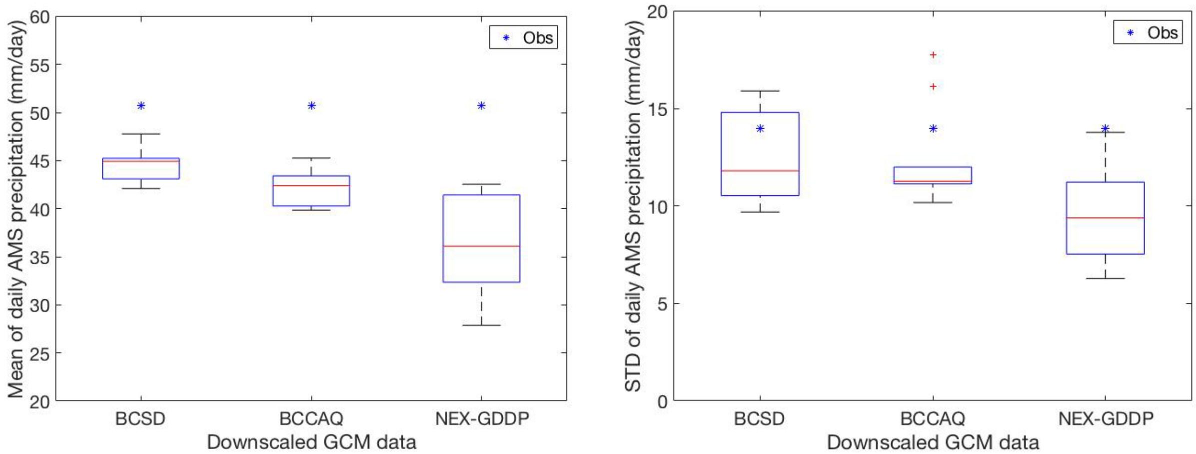


Figure 6-3. Mean (left) and Standard Deviation (right) of daily AMPs at Dorval station based on downscaled gridded data from ten different GCMs

Table 6-4. RMSE of the means of daily AMP and temperature extremes at Dorval station

RCMs	RMSE		
	T_{min}	T_{max}	Precipitation
PCIC-BCSD	6.91	5.16	**2.90**
PCIC-BCCAQ	6.47	4.75	5.00
NEX-GDDP	**5.71**	**3.99**	12.20

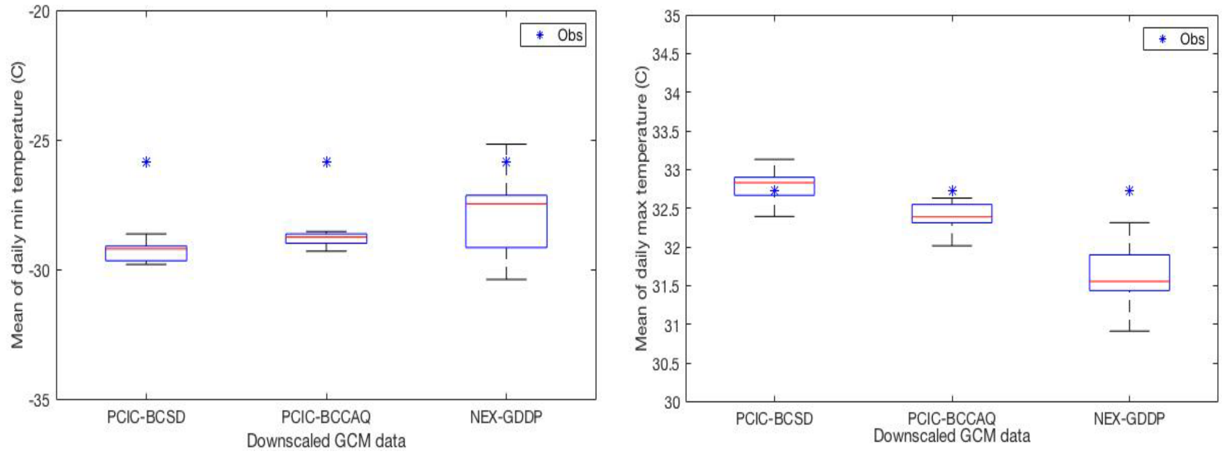


Figure 6-4. Mean of daily minimum (left) and maximum (right) temperatures at Dorval station based on downscaled gridded data from ten different GCMs

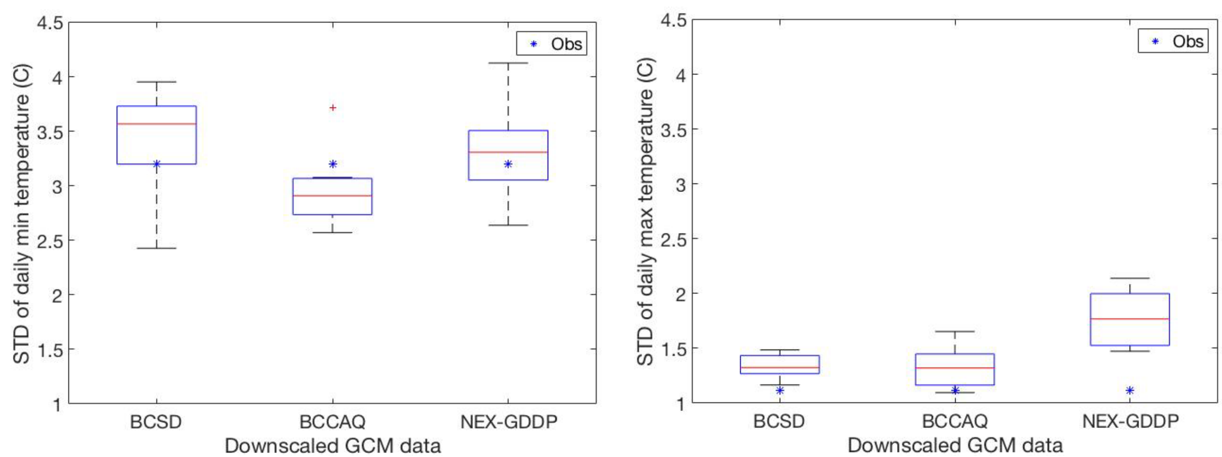


Figure 6-5. Standard deviation of daily minimum (left) and maximum (right) temperatures at Dorval station based on downscaled gridded data from ten different GCMs

6.3.2 Future climate

Daily annual temperature extremes and daily AMPs for the 2006-2100 period downscaled from PCIC and NEX-GDDP were analyzed. It can be seen from Figure 6-6 and Figure 6-7 that there are increasing trends in both temperature extremes and AMP at Dorval station. Table 6-5 shows the values of temperature extremes and AMPs estimated based on the fitted trend regression lines at the year 2006 and 2100. Results are based on an average of all 10 GCMs given by both datasets. It is estimated that precipitation could increase around 10.77% for the 2006-2100 period. In addition, daily maximum temperature is projected to increase around 8.06% in the same period. Daily minimum temperature could have a projected increase of around 16.69%. Hence, the Montreal region could experience more extreme rainfalls and higher maximum and minimum temperatures in the future.

Table 6-5. Increase of temperature and precipitation in 2006-2100 period

Variables	2006	2100	% Increase
Precipitation (mm)	41.47	45.93	10.77
Minimum temperature (°C)	-26.19	-21.82	16.69
Maximum temperature (°C)	33.33	36.02	8.06

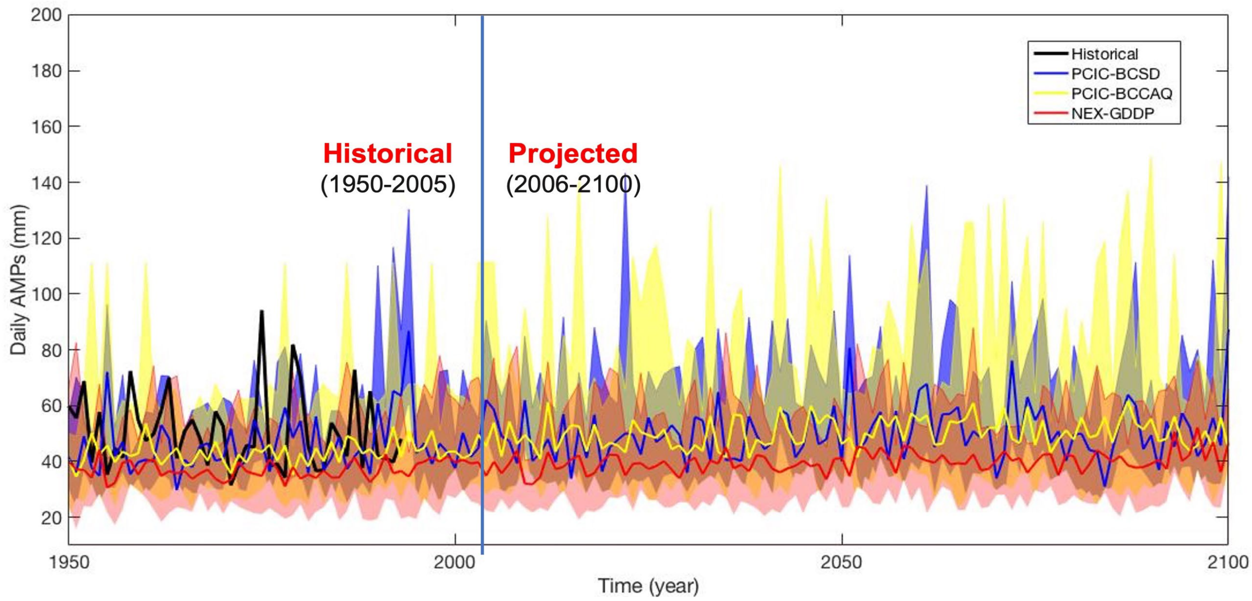


Figure 6-6. Historical and Projected daily AMPs at Dorval station for entire period 1950-2100. The range donates the maximum and minimum values from all GCMs of each dataset, solid lines are median of each dataset.

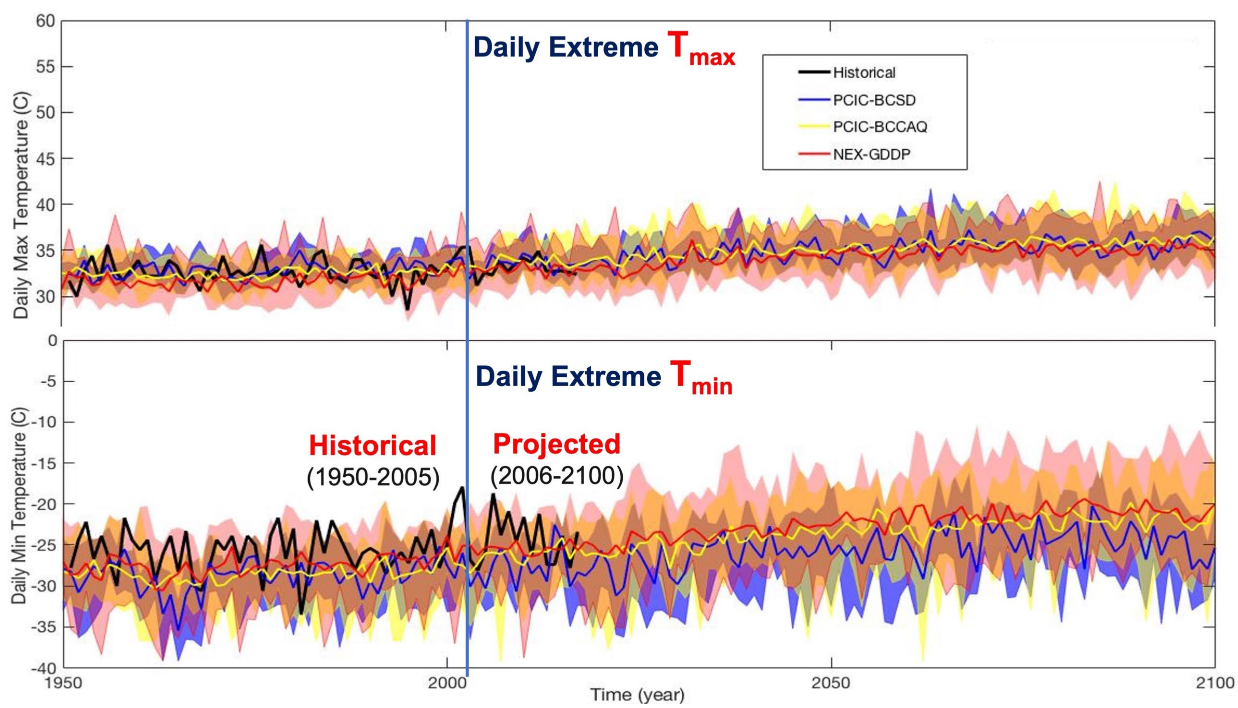


Figure 6-7. Historical and projected daily minimum and maximum temperatures at McGill station for entire period 1950-2100. The range donates the maximum and minimum values from all GCMs of each dataset, solid lines are median of each dataset.

6.4 Conclusions

Major findings of this present study can be summarized as follows:

Many climate projection studies have been commonly conducted at global or large regional scales, the present study has been performed specifically at the City of Montreal scale to provide useful information on the variability in time and in space of annual maximum precipitations and temperature extremes for the design and planning of its urban infrastructures using the regional downscaled climate projection data from ten different GCMs under the RCP 4.5 scenario provided by PCIC and NEX-GDDP. In general, the PCIC data with finer grid size of 1/12 degree (or approximately 10x10 km) could produce more robust results than the NEX-GDDP data with a coarser resolution of ¼ degree (or approximately 25x25 km).

According to the results downscaled by PCIC and NEX-GDDP, there are projected increasing trends in both temperature extremes and AMPs over the Montreal region. The AMP is projected to increase around 10% for the 2006-2100 period. Minimum and maximum temperatures are projected to increase approximately 16% and 8% respectively by the end of this century.

Downscaled gridded data are different from observed data at a given location. It is therefore essential to perform a bias correction of the gridded data before these data could be used in the planning and design of the urban infrastructures.

Chapter 7: Conclusions and recommendations

7.1 Conclusions

The following main conclusions can be drawn from the present study:

1. A statistical downscaling model (called SDGAM) has been proposed for climate change impact assessment studies at a gauged site. The proposed model was based on the combination of the precipitation occurrence and the precipitation amount using the Generalized Additive Modeling method. Results of a numerical application have indicated that the proposed model was able to describe well many features of the daily precipitation process, including its occurrence frequency, intensity, and extremes for both calibration and validation periods for data from 10 rain-gauged stations located in Southern Quebec and Ontario, Canada. In addition, this model could provide a significant improvement over the popular SDSM model in the modeling of daily precipitation process in the context of climate change.

2. A spatial-temporal downscaling approach was proposed in this study to describe the linkage between large-scale climate variables for daily scale to AMP for daily and sub-daily scales at a local site. The proposed method was based on the combination of the spatial downscaling method to link large-scale climatic variables provided by GCMs to daily extreme precipitations at a local site using the SDGAM model developed in this study and the temporal downscaling procedure to describe the relationships between daily

extreme precipitations with sub-daily extreme precipitations using the scaling GEV/PWM model. The feasibility of the proposed downscaling method has been evaluated based on climate simulation outputs from the CanESM2 model under different RCPs (RCP 26, RCP 45, and RCP 85) and using available AMP data for durations ranging from 5 minutes to 24 hours at ten rain-gage stations across Canada. Results have showed that it is feasible to link daily large-scale climate variables to daily AMP at a local site for climate change impact and adaptation studies at a given location of interest.

3. A detailed statistical analysis of AMP series for selected stations representing the diverse climatic conditions across Canada has indicated that these AMP series in Canada displayed different scaling behaviors depending on the location of the station considered. Based on this scaling property, the scaling GEV distribution has been proved to be able to provide accurate estimates of sub-daily AMPs from GCM-downscaled daily AMP amounts.

4. A statistical regionalization method using the Principal Component Analysis (PCA) has been proposed to identify homogeneous regions of precipitation regimes. The feasibility and accuracy of the proposed method has been assessed using the daily precipitation data available from a network of 155 rain-gauge stations across Vietnam. Results of this numerical application have indicated that the suggested regionalization method was able to identify homogeneous precipitation regions which were found physically consistent to the particular climatic features of Vietnam.

5. A statistical estimation approach has been developed in this study to generate daily precipitation series at an ungauged location using rainfall information available within the same homogeneous rainfall region. The proposed approach was based on a two-stage

interpolation method to describe the persistence in rainfall occurrences and rainfall amounts for the rainfall homogenous region. The feasibility and accuracy of the proposed estimation method has been evaluated using daily rainfall data available from a network of 155 raingauges in Vietnam. Results of this assessment have indicated that the proposed procedure could provide an accurate estimate of the daily precipitation series for an ungauged location.

6. A detailed statistical analysis was performed to identify the presence of trends in precipitation series using the historical high-quality rainfall records from a network of 175 stations located across Canada and using the 25kmx25km resolution downscaled regional gridded data from NEX-GDDP for the past period from 1950 to 2005 and for the projection period from 2006 to 2100. It was found that the majority of station data have increase trends, and around 40% of stations located mostly in central and western Canada with decrease trends. For downscaled regional gridded data NEX-GDDP, increase trends was found for both historical and projected periods for more than 80% of stations.

7. Among all 21 GCMs of NEX-GDDP data, the CanESM2 model is considered the best model for Canada, especially in terms of the mean and standard deviation of the annual maximum daily precipitation. Models from China and Japan were also found to be able to produce good results over many locations in Canada.

8. The PCIC data with finer grid size of 1/12 degree (or approximately 10x10 km) could produce more robust results than the NEX-GDDP data with a coarser resolution of ¼ degree (or approximately 25x25 km) over Montreal area. It was also found that these data projected increasing trends in both temperature extremes and AMPs. More specifically, the

AMP was projected to increase around 10% for the 2006-2100 period while minimum and maximum temperatures were projected to increase approximately 16% and 8% respectively by the end of this century for the Montreal region.

7.2 Recommendations for further works

Based on the findings of this study, the following recommendations are suggested for future studies:

1. The regression-based downscaling models for generating daily precipitation process for climate impact studies were found to be sensitive to the selection of the large-scale climate predictors given by the GCMs. However, there is still no general agreement for selecting the best approach that could identify the most significant predictors for these models. Hence, it is essential to develop a new screening method for selecting the most significant predictors that could describe more accurately the linkages between these climate predictors and the observed precipitation characteristics at a local site of interest.

2. The present study has indicated that the performance of the ungauged precipitation model was significantly influenced by the accuracy in the identification of the homogeneous regions of precipitation. For improving our understanding of the spatial and temporal variation of the precipitation process and for improving the accuracy of the precipitation estimation at an ungauged site it is necessary to explore other similarity criteria based on both precipitation regimes and topographic characteristics that could be used to improve the definition of the similarity of the precipitation variability in time and in space.

3. For this study, the PCA works well for rainfall data of Vietnam region in monthly time scale. However, this approach is considered to be sensitive with the time scale of the selected data, it is therefore necessary to develop a more robust method based on, for instance, Ordinary Factor Analysis or Cluster analysis to minimize this sensitivity.

4. Further studies should be conducted to evaluate of the variability in time and in space of the daily annual maximum rainfalls and extreme temperatures over Canada region for different other sources of downscaled regional gridded data beside the National Aeronautics Space Administration (NASA) Earth Exchange Global Daily Downscaled Projections (NEX-GDDP) including the Pacific Climate Impacts Consortium (PCIC), and ANUSPLIN.

5. The methods proposed in this study should be tested with different datasets available worldwide from different climate conditions to assess the feasibility and reliability of these suggested approaches.

Chapter 8: Statement of originality

To the best of the author's knowledge, the followings are the original contributions from the present study:

1. A new statistical downscaling model (SDGAM) has been developed in this study for describing accurately the linkage between large-scale climate predictors and observed daily rainfall characteristics at a local site. This new model was based on the Generalized Additive Modeling (GAM) method. Results of a comparative study using NCEP re-analysis data and observed daily precipitation data in Canada have demonstrated that the SDGAM could provide more accurate results than those given by the currently popular SDSM model. The proposed SDGAM model is therefore could be an essential tool for high-quality climate change impact assessment studies in practice.

2. A novel spatial-temporal downscaling approach was proposed in the present study to provide a more accurate estimation of extreme rainfalls for daily and sub-daily scales at a local site in the context of climate change. The proposed approach was quite useful for improving the accuracy in the construction of the IDF relations at a given site and, consequently, a more accurate estimation of the design storm for urban infrastructures design in a changing climate.

3. An original regionalization method based on the Principal Component Analysis (PCA) technique was proposed for identifying homogeneous regions of precipitation

regimes. Results of an illustrative application using observed daily rainfall data in Vietnam have indicated the feasibility and accuracy of the proposed method.

4. An original statistical approach has been developed in this study to estimate daily precipitation series at a location where the rainfall data are unavailable (an ungauged site) using rainfall information in the same homogeneous region. The proposed approach was based on a two-stage interpolation method to represent the persistence in rainfall occurrences and rainfall amounts within the same homogeneous region. It has been demonstrated that this new approach could provide the estimated daily precipitation series at an ungauged site having similar statistical properties as those of the observed data.

5. A detailed statistical analysis has been performed to identify the presence of trends in the historical records of daily annual maximum precipitation series for different locations and the downscaled regional gridded data from the National Aeronautics Space Administration (NASA) Earth Exchange Global Daily Downscaled Projections (NEX-GDDP) for Canada. Results of this analysis have provided essential information for improving our understanding of the variability of daily precipitation in Canada for the present and future periods.

6. A detailed statistical analysis of the variability in time and in space of the daily annual maximum rainfalls and extreme temperatures over the Montreal region for the present and future climates using the data from two different sources: the Pacific Climate Impacts Consortium (PCIC) and the National Aeronautics Space Administration (NASA) Earth Exchange Global Daily Downscaled Projections

(NEX-GDDP). Results of this analysis have provided valuable information for the planning and design of urban infrastructures for Montreal in the context of a changing climate.

References

Alexander, L.V., Zhang, X., Peterson, T.C., Caesar, J., Gleason, B., Klein Tank, A.M.G., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., Tagipour, A., Rupa Kumar, K., Revadekar, J., Griffiths, G., Vincent, L., Stephenson, D.B., Burn, J., Aguilar, E., Brunet, M., Taylor, M., New, M., Zhai, P., Rusticucci, M. & Vazquez-Aguirre, J.L. 2006. Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research,* 111.

Arnbjerg-Nielsen, K., Willems, P., Olsson, J., Beecham, S., Pathirana, A., Bulow Gregersen, I., Madsen, H. & Nguyen, V.T. 2013. Impacts of climate change on rainfall extremes and urban drainage systems: a review. *Water Sci Technol,* 68**,** 16-28.

Assani, A.A., Landry, R. & Laurencelle, M. 2012. Comparison of Interannual Variability Modes and Trends of Seasonal Precipitation and Streamflow in Southern Quebec (Canada). *River Research and Applications,* 28**,** 1740-1752.

Bae, H.-J. & Oh, J.-H. 2017. Study of method for synthetic precipitation data for ungauged sites using quantitative precipitation model. *Asia-Pacific Journal of Atmospheric Sciences,* 53**,** 403-410.

Baeriswyl, P.-A. & Rebetez, M. 1997. Regionalization of precipitation in Switzerland by means of principal component analysis. *Theoretical and Applied Climatology,* 58**,** 31-41.

Bárdossy, A. 1997. *Introduction to Geostatistics,* Stuttgart, Germany, University of Stuttgart.

Beersma, J.J. & Buishand, T.A. 2003. Multi-site simulation of daily precipitation and temperature conditional on the atmospheric circulation. *Climate Research,* 25**,** 121-133.

Benmarhnia, T., Sottile, M.F., Plante, C., Brand, A., Casati, B., Fournier, M. & Smargiassi, A. 2014. Variability in temperature-related mortality projections under climate change. *Environ Health Perspect,* 122**,** 1293-8.

Besaw, L.E., Rizzo, D.M., Bierman, P.R. & Hackett, W.R. 2010. Advances in ungauged streamflow prediction using artificial neural networks. *Journal of Hydrology,* 386**,** 27-37.

Buishand, T.A. & Brandsma, T. 2001. Multisite simulation of daily precipitation and temperature in the Rhine Basin by nearest-neighbor resampling. *Water Resources Research,* 37**,** 2761-2776.

Canada, E. 2016. Canada's Mid-Century Long-Term Low-Greenhouse Gas Strategy. Gatinear, QC,Canada: Public Inquiries Centre.

Candela, L., Tamoh, K., Olivares, G. & Gomez, M. 2012. Modelling impacts of climate change on water resources in ungauged and data-scarce watersheds. Application to the Siurana catchment (NE Spain). *Sci Total Environ,* 440**,** 253-60.

Chandler, R.E. & Scott, E.M. 2011. *Statistical Methods for Trend Detection and Analysis in the Environmental Sciences*, John Wiley & Sons.

Chebana, F., Charron, C., Ouarda, T.B.M.J. & Martel, B. 2014. Regional Frequency Analysis at Ungauged Sites with the Generalized Additive Model. *Journal of Hydrometeorology,* 15**,** 2418-2428.

City-of-Montreal 2017. Climate Change Adaptation Plan 2015-2020. *In:* supervision, S. d. l. e. u. t. & of Roger Lachance, D. o. t. S. d. l. e. (eds.) *Executive Summary.* Montreal, Quebec, Canada.

Creutin, J.D. & Obled, D. 1982. Objective analyses and mapping techniques for rainfall fields: An objective comparison. *Water Resources Research,* 18**,** 413-431.

Cullinane, S. 2018. *Up to 70 people dead after Quebec heat wave* [Online]. Available: https://www.cnn.com/2018/07/10/americas/quebec-heat-wave-deaths-wxc/index.html [Accessed 20 November 2018].

Denis, B., Laprise, R., Caya, D. & Cote, J. 2002. Downscaling ability of one-way nested regional climate models: the Big-Brother Experiment. *Climate Dynamics,* 18**,** 627-646.

Diaz-Nieto, J. & Wilby, R.L. 2005. A comparison of statistical downscaling and climate change factor methods: impacts on low flows in the River Thames, United Kingdom. *Climate Change,* 69**,** 245-268.

Environment Canada 2020. Rainfall Intensity-Duration-Frequency (IDF) Tables and Graphs. Version 3.10. May 2020 ed.

Fernandes, R. & G. Leblanc, S. 2005. Parametric (modified least squares) and non-parametric (Theil–Sen) linear regressions for predicting biophysical parameters in the presence of measurement errors. *Remote Sensing of Environment,* 95**,** 303-316.

Fowler, H.J., Blenkinsop, S. & Tebaldi, C. 2007. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology,* 27**,** 1547-1578.

Gaur, A. & Simonovic, S.P. 2017. Application of physical scaling towards downscaling climate model precipitation data. *Theoretical and Applied Climatology*.

Gibbs, M.S., Maier, H.R. & Dandy, G.C. 2012. A generic framework for regression regionalization in ungauged catchments. *Environmental Modelling & Software,* 27-28**,** 1-14.

González, J. & Valdés, J.B. 2008. A regional monthly precipitation simulation model based on an L-moment smoothed statistical regionalization approach. *Journal of Hydrology,* 348**,** 27-39.

Goodess, C.M., Palutikof, J.P. 1998. Development of daily rainfall scenarios for southeast Spain using a circulation-type approach to downscaling. *International Journal of Climatology,* 18**,** 1051-1083.

Gooré Bi, E., Gachon, P., Vrac, M. & Monette, F. 2017. Which downscaled rainfall data for climate change impact studies in urban areas? Review of current approaches and trends. *Theoretical and Applied Climatology,* 127**,** 685-699.

Graham, L. 2021. August 2021 was the hottest on record for Montreal. *CTV News*.

Hastie, T. & Tibshirani, R. 1986. Generalized Additive Models. *Statistical Science,* 1**,** 297-318.

Hay, L.E., McCare, G.J., Wolock, D.M. & Ayers, M.A. 1991. Simulation of precipitation by weather type analysis. *Water Resources Research,* 27**,** 493-501.

Hershfield, D.M. 1961. Rainfall frequency atlas of the United States for durations from 30 minutes to 24-hours and return periods from 1 to 100 years, Technical Paper 40 U.S. Weather Bureau, Washington, D.C.

Hessami, M., Gachon, P., Ouarda, T.B.M.J. & St-Hilaire, A. 2008. Automated regression-based statistical downscaling tool. *Environmental Modelling & Software,* 23**,** 813-834.

Hosking, J. 1990. L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. *Royal Statistical Cociety,* 52**,** 105-124.

Hosking, J. & Wallis, J. 1997. Regional Frequency Analysis: An Approach Based on L-Moments. *Cambridge University Press*.

Hubert, P. 2001. Multifractals as a tool to overcome scale problems in hydrology. *Hydrological Sciences Journal,* 46**,** 897-905.

Hutchinson, M.F., McKenney, D.W., Lawrence, K., Pedlar, J.H., Hopkinson, R.F., Milewska, E. & Papadopol, P. 2009. Development and Testing of Canada-Wide Interpolated Spatial Models of Daily Minimum–Maximum Temperature and Precipitation for 1961–2003. *Journal of Applied Meteorology and Climatology,* 48**,** 725-741.

IPCC 2007. Forth Assessment Report.

IPCC 2014. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.) ed. Geneva, Switzerland.

Johnson, G. & Hanson, C. 1995. Topographic and atmospheric influences on precipitation variability over a mountainous watershed. *Journal of Applied Meteorology,* 34**,** 68-87.

Jones, M.R., Blenkinsop, S., Fowler, H.J., Stephenson, D.B. & Kilsby, C.G. 2013. Generalized additive modelling of daily precipitation extremes and their climatic drivers. *National Center for Atmospheric Research.*

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R. & Joseph, D. 1996. The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society,* 77**,** 437-472.

Khalili, M. & Nguyen, V.-T.-V. 2016. An efficient statistical approach to multi-site downscaling of daily precipitation series in the context of climate change. *Climate Dynamics*.

Kharin, V.V. & Zwiers, F.W. 2000. Changes in the Extremes in an Ensemble of Transient Climate Simulations with a Coupled Atmosphere–Ocean GCM. *Journal of Climate,* 13**,** 3760-3788.

Kilsby, C.G., Cowpertwait, P.S.P., O'Connell, P.E. & Jones, P.D. 1998. Predicting rainfall statistics in England and Wales using atmospheric circulation variables. *International Journal of Climatology,* 18**,** 523-539.

Laanaya, F., St-Hilaire, A. & Gloaguen, E. 2017. Water temperature modelling: comparison between the generalized additive model, logistic, residuals regression and linear regression models. *Hydrological Sciences Journal,* 62**,** 1078-1093.

Lau, R. 2017. *Quebec floods: Hundreds of West Island homes may never be rebuilt* [Online]. Available: https://globalnews.ca/news/3548827/quebec-floods-hundreds-of-homes-destroyed-in-montreals-west-island/ [Accessed 20 November 2018].

Lenderink, G., van Ulden, A., van den Hurk, B. & Keller, F. 2007. A study on combining global and regional climate model results for generating climate scenarios of temperature and precipitation for the Netherlands. *Climate Dynamics,* 29**,** 157-176.

Lima, C.H.R. & Lall, U. 2009. Hierarchical Bayesian modeling of multisite daily rainfall occurrence: Rainy season onset, peak, and end. *Water Resources Research,* 45.

Lovejoy, S. & Schertzer, D. 2012. Low-Frequency Weather and the Emergence of the Climate. *American Geophysical Union***,** 231-254.

Maraun, D. 2016. Bias Correcting Climate Change Simulations - a Critical Review. *Current Climate Change Reports,* 2**,** 211-220.

Maraun, D., Wetterhall, F., Ireson, A.M., Chandler, R.E., Kendon, E.J., Widmann, M., Brienen, S., Rust, H.W., Sauter, T., Themeßl, M., Venema, V.K.C., Chun, K.P., Goodess, C.M., Jones, R.G., Onof, C., Vrac, M. & Thiele-Eich, I. 2010. Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics,* 48.

Mearns, L., McGinnis, S., Korytina, D., Arritt, R., Biner, S., Bukovsky, M., Chang, H.-I., Christensen, O., Herzmann, D., Jiao, Y., Kharin, S., Lazare, M., Nikulin, G., Qian, M., Scinocca, J., Winger, K., Castro, C., Frigon, A. & Gutowski, W. 2017. The NA-CORDEX dataset, version 1.0. 2017 ed. Boulder CO: NCAR Climate Data Gateway.

Mearns, L.O., Gutowski, W.J., Jones, R.G., Leung, L.-Y., McGinnis, S., Nunes, A.M.B. & Qian, Y. 2014. The North American Regional Climate Change Assessment Program dataset,. 2014 ed. Boulder, CO: National Center for Atmospheric Research Earth System Grid data portal.

Mehrotra, R., Srikanthan, R. & Sharma, A. 2006. A comparison of three stochastic multi-site precipitation occurrence generators. *Journal of Hydrology,* 331**,** 280-292.

Mekis, E., Donaldson, N., Reid, J., Zucconi, A., Hoover, J., Li, Q., Nitu, R. & Melo, S. 2018. An Overview of Surface-Based Precipitation Observations at Environment and Climate Change Canada. *Atmosphere-Ocean,* 56**,** 71-95.

Mekis, É. & Vincent, L.A. 2011. An Overview of the Second Generation Adjusted Daily Precipitation Dataset for Trend Analysis in Canada. *Atmosphere-Ocean,* 49**,** 163-177.

Miller, N.L., Bashford, K.E. & Strem, E. 2003. Potential impacts of climate change on California hydrology. *Journal of the American Water Resources Association,* 39**,** 771-784.

Min, S.-K., Zhang, X., Zwiers, F.W., Friederichs, P. & Hense, A. 2008. Signal detectability in extreme precipitation changes assessed from twentieth century climate simulations. *Climate Dynamics,* 32**,** 95-111.

Nguyen, T.-H. 2020. *Novel Methods for Estimating Extreme Design Rainfalls at Gauged and Ungauged*

Locations in a Changing Climate. Doctor of Philosophy, McGill University.

Nguyen, T.-H. & Nguyen, V.-T.-V. 2020. Linking climate change to urban storm drainage system design: An innovative approach to modeling of extreme rainfall processes over different spatial and temporal scales. *Journal of Hydro-environment Research,* 29**,** 80-95.

Nguyen, T.-H., Nguyen, V.-T.-V. & Nguyen, H.-L. 2018. A spatio-temporal statistical downscaling approach to deriving extreme rainfall IDF relations at ungauged sites in the context of climate change, 13th International Hydroinformatics conference. *13th International Conference on Hydroinformatics.* Palermo, Italy: Journal of Hydroinformatics.

Nguyen, T.D. 2003. Regional Estimation of Extreme Rainfall Events. *McGill Universtiy*.

Nguyen, V.-T.-V., Nguyen, T.-D. & Gachon, P. 2006. On the linkage of large-scale climate variability with local characteristics of daily precipitation and temperature extremes: an evaluation of statistical. *Advances in Geosciences,* 4**,** 1-9.

Nguyen, V.-T.-V. & Nguyen, T.D. 2008. Statistical Downscaling of Daily Precipitation Process for Climate-Related Impact Studies. *In:* Singh, V. P. (ed.) *Hydrology and Hydraulics.* Water Resources Publications.

Nguyen, V.-T.-V., Nguyen, T.D. & Ashkar, F. 2002. Regional frequency analysis of extreme rainfalls. *Water Sci Technol,* 45**,** 75-81.

Nguyen, V.T.V., Nguyen, T.D. & Cung, A. 2007. A statistical approach to downscaling of sub-daily extreme rainfall processes for climate-related impact studies in urban areas. *Water Science & Technology: Water Supply,* 7**,** 183.

Nguyen, V.T.V. & Yeo, M. 2011. Statistical Downscaling of Daily Rainfall Processes for Climate-Related Impact Assessment Studies. *World Environmental and Water Resources Congress***,** 2-9.

PCIC. 2014. *Statistically Downscaled Climate Scenarios* [Online]. University of Victoria: Pacific Climate Impacts Consortium. Available: https://www.pacificclimate.org/data/statistically-downscaled-climate-scenarios [Accessed 8 August 2018].

Preisendorfer, R.W. 1988. *Principal Component Analysis in Meteorology and Oceanography*, Elsevier.

Prudhomme, C., Reynard, N. & Crooks, S. 2002. Downscaling of global climate models for flood frequency analysis: where are we now? *Hydrological Processes,* 16**,** 1137-1150.

Randall, D.A., Wood, R.A., Bony, S., Colman, R., Fichefet, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., Stouffer, R.J., Sumi, A. & Taylor, K.E. 2007. Climate models and their evaluation. *In:* Solomon, S., et al. (ed.) *Climate change 2007: The Physical science basis. Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change.* Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

Richardson, C.W. 1981. Stochastic simulation of daily precipitation, temperature, and solar radiation. *Water Resources Research,* 17**,** 182-190.

RStudio Team 2020. RStudio: Intergrated Development for R. Studio. Boston, MA: PBC.

Ryu, J.H., Lee, J.H., Jeong, S., Park, S.K. & Han, K. 2011. The impacts of climate change on local hydrology and low flow frequency in the Geum River Basin, Korea. *Hydrological Processes,* 25**,** 3437-3447.

Samuel, J., Coulibaly, P. & Metcalfe, R.A. 2012. Evaluation of future flow variability in ungauged basins: Validation of combined methods. *Advances in Water Resources,* 35**,** 121-140.

Schertzer, D., Tchiguirinskaia, I., Lovejoy, S. & Hubert, P. 2010. No monsters, no miracles: in nonlinear sciences hydrology is not an outlier! *Hydrological Sciences Journal,* 55**,** 965-979.

Schnur, R. & Lettenmaier, D.P. 1998. A case study of statistical downscaling in Australia using weather classification by recursive partitioning. *Journal of Hydrology,* 212-213**,** 362-379.

Semenov, M.A. & Barrow, E.M. 1997. Use of a stochastic weather generator in the development of climate change scenarios. *Climate Change,* 35**,** 397-414.

Sen, P.K. 1968. Estimates of the Regression Coefficient Based on Kendall's Tau. *Journal of the American Statistical Association,* 63**,** 1379-1389.

Sivapalan, M. 2003. Prediction in ungauged basins: a grand challenge for theoretical hydrology. *Hydrological Processes,* 17**,** 3163-3170.

Statistics-Canada. 2016. *Census* [Online]. Available: https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/details/page.cfm?Lang=E&Geo1=CSD&Code1=2466023&Geo2=PR&Code2=24&Data=Count&SearchText=Montreal&SearchType=Begins&SearchPR=01&B1=All&GeoLevel=PR&GeoCode=2466023&TABID=1 [Accessed 8 August 2018].

Thistle, M. & Caissie, D. 2013. Trends in air temperature, total precipitation and streamflow characteristics in eastern Canada. NB, Canada: Fisheries and Oceans Canada.

Thrasher, B., Maurer, E.P., McKellar, C. & Duffy, P.B. 2012. Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping. *Hydrology and Earth System Sciences,* 16**,** 3309-3314.

Villarini, G. & Serinaldi, F. 2012. Development of statistical models for at-site probabilistic seasonal rainfall forecast. *International Journal of Climatology,* 32**,** 2197-2212.

Vincent, L.A., Zhang, X., Mekis, É., Wan, H. & Bush, E.J. 2018. Changes in Canada's Climate: Trends in Indices Based on Daily Temperature and Precipitation Data. *Atmosphere-Ocean,* 56**,** 332-349.

Von Storch, H., Zorita, E. & Cubasch, U. 1993. Downscaling of Global Climate Change Estimates to Regional Scales: An Application to Iberian Rainfall in Wintertime. *American Meteorological Society,* 6**,** 1161-1171.

Werner, A.T. & Cannon, A.J. 2016. Hydrologic extremes – an intercomparison of multiple gridded statistical downscaling methods. *Hydrology and Earth System Sciences,* 20**,** 1483-1508.

Werner, A.T., Schnorbus, M.A., Shrestha, R.R., Cannon, A.J., Zwiers, F.W., Dayon G. & Anslow, F. 2019. A long-term, temporally consistent, gridded daily meteorological dataset for northwestern North America. Scientific Data.

Westra, S., Alexander, L.V. & Zwiers, F.W. 2013. Global Increasing Trends in Annual Maximum Daily Precipitation. *Journal of Climate,* 26**,** 3904-3918.

Whitfield, P.H., Wang, J.Y. & Cannon, A.J. 2003. Modelling Future Streamflow Extremes — Floods and Low Flows in Georgia Basin, British Columbia. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques,* 28**,** 633-656.

Wilby, R.L., Charles, S. & Zorita, E. 2004. Guidelines for use of climate scenarios developed from statistical downscaling methods.

Wilby, R.L. & Dawson, C.W. 2013. The Statistical DownScaling Model: insights from one decade of application. *International Journal of Climatology,* 33**,** 1707-1719.

Wilby, R.L., Dawson, C.W. & Barrow, E.M. 2002. sdsm — a decision support tool for the assessment of regional climate change impacts. *Environmental Modelling & Software,* 17**,** 147-159.

Wilby, R.L., Whitehead, P.G., Wade, A.J., Butterfield, D., Davis, R.J. & Watts, G. 2006. Integrated modelling of climate change impacts on water resources and quality in a lowland catchment: River Kennet, UK. *Journal of Hydrology,* 330**,** 204-220.

Wilks, D.S. 1993. Comparison of Three-Paramete Prrobability Distributions for Representing Annual Extreme and Partial Duration Precipitation Series. *Water Resources Research,* 29**,** 3543-3549.

Wilks, D.S. 1998. Multisite generalization of a daily stochastic precipitation generation model. *Journal of Hydrology,* 210**,** 178-191.

Wilks, D.S. & Wilby, R.L. 1999. The weather generation game: a review of stochastic weather models. *Progress in Physical Geography,* 23**,** 329-357.

WMO 2009. *Guide to hydrological practices, volume II: Management of water resources and application of hydrological practices, 6th edition,* Geneva, Switzerland, World Meteorological Organization,.

Wood, A.W. 2006. *Generalized Additive Models: An Introduction with R*, Chapman and Hall/CRC Press.

Wood, S.N. 2004. Stable and effective multiphe smoothing parmeter estimation for generalized addivtive models. *Journal of the American Statistical Association,* 99**,** 673-686.

Xu, C.-y. 1999. From GCMs to river flow: a review of downscaling methods and hydrologic modelling approaches. *Progress in Physical Geography,* 23**,** 229-249.

Yarnal, B., Comrie, A.C., Frakes, B. & Brown, D.P. 2001. Developments and prospects in synoptic climatology. *International Journal of Climatology,* 21**,** 1923-1950.

Yeo, M. 2013. Statistical modeling of precipitation processes for gaged and ungaged sites in the context of climate change. *McGill Universtiy*.

Yeo, M.-H. & Nguyen, V.-T.-V. 2014. Statistical Approach to Downscaling of Daily Rainfall Process at an Ungauged Site. *Advances in Hydroinformatics.* France: Springer.

Zalina, M., Desa, M., Nguyen, V.-T.-V. & Kassim, A. 2002. Selecting a probability distribution for extreme rainfall series in Malaysia. *Water Science and Technology,* 45**,** 63-68.

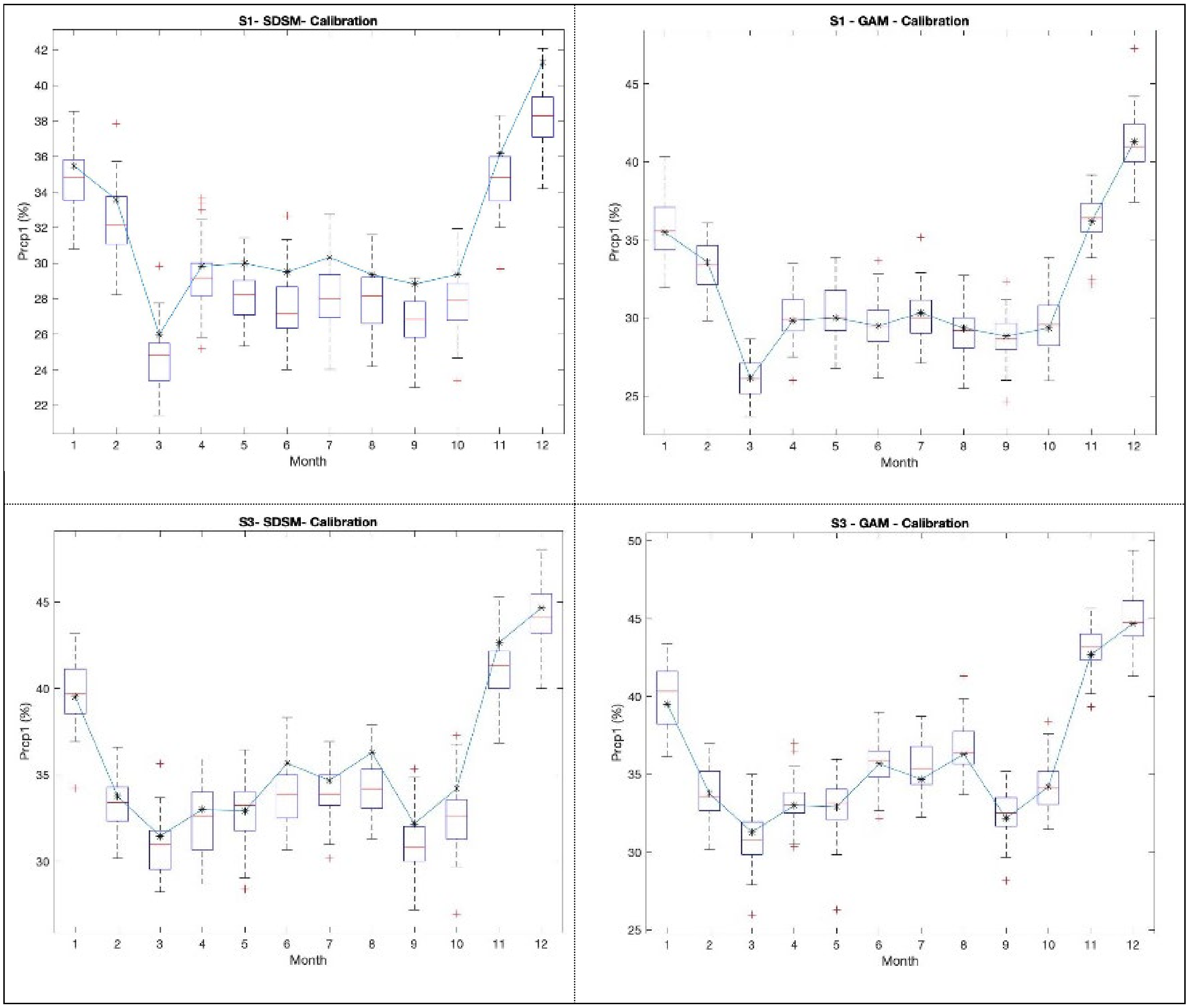
Zhang et al., X. 2019. Changes in Temperature and Precipitation Across Canada. *In:* Bush, E. & Lemmen, D. S. (eds.) *Canada's Changing Climate Report.* Ottawa, Canda: Government of Canada.

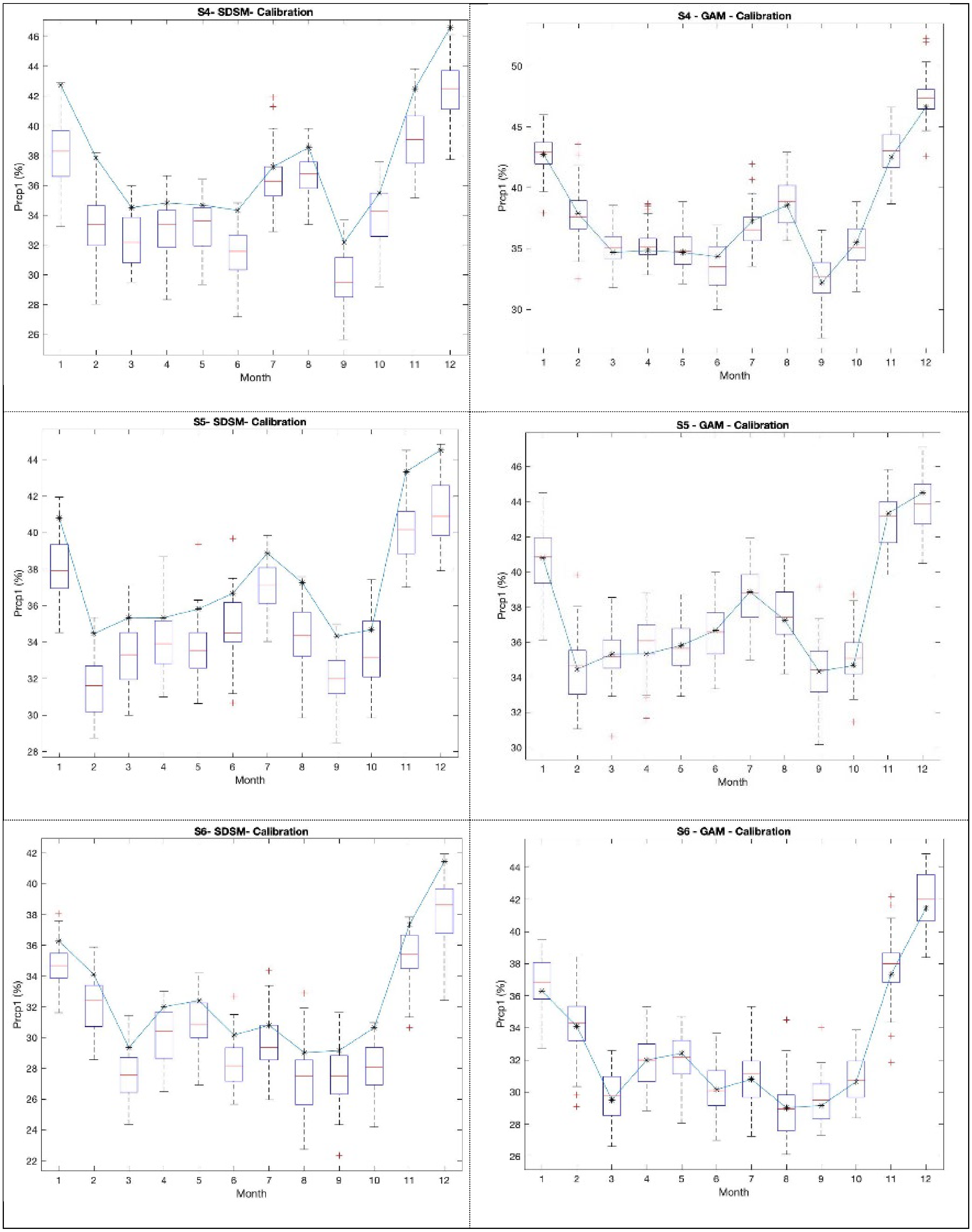
Zhang, X., Brown, R., Vincent, L., Skinner, W., Feng, Y. & Mekis, E. 2011. Canadian climate trends, 1950-2007. *In:* 5, T. T. R. N. (ed.) *Canadian Biodiversity: Ecosystem Status and Trends 2010.* Ottawa, ON: Canadian Councils of Resource Ministers.

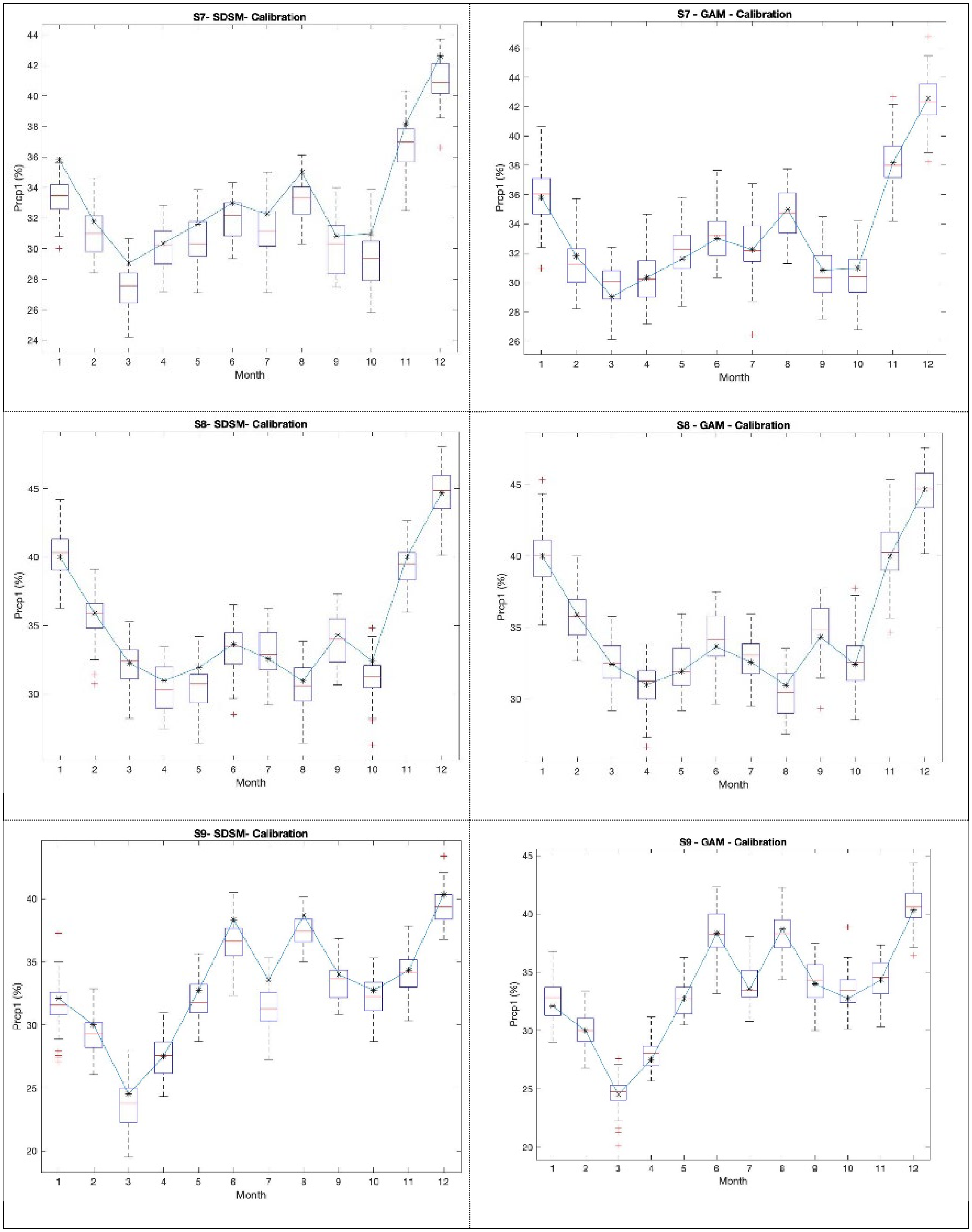
Zhang, X., Hogg, W.D. & Mekis, E.V. 2001. Spatial and Temporal Characteristics of Heavy Precipitation Events over Canada. *American Meteorological Society,* 14**,** 1923-1936.

Zhang, X., Vincent, L.A., Hogg, W.D. & Niitsoo, A. 2000. Temperature and precipitation trends in Canada during the 20th century. *Atmosphere-Ocean,* 38**,** 395-429.

Appendix A: Supplementary materials for chapter 2




S4- SDSM- Calibration
S4 - GAM - Calibration
S5- SDSM- Calibration
S5 - GAM - Calibration
S6- SDSM- Calibration
S6 - GAM - Calibration
Prcp1 (%)
Month


S7- SDSM- Calibration
S7 - GAM - Calibration
S8- SDSM- Calibration
S8 - GAM - Calibration
S9- SDSM- Calibration
S9 - GAM - Calibration
Prcp1 (%)
Month

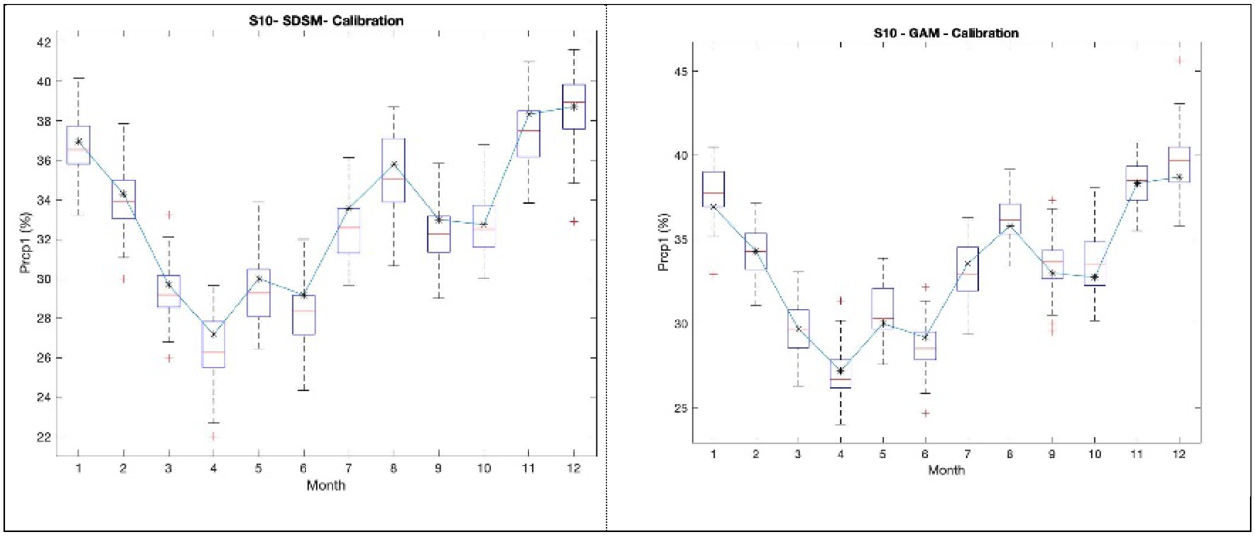
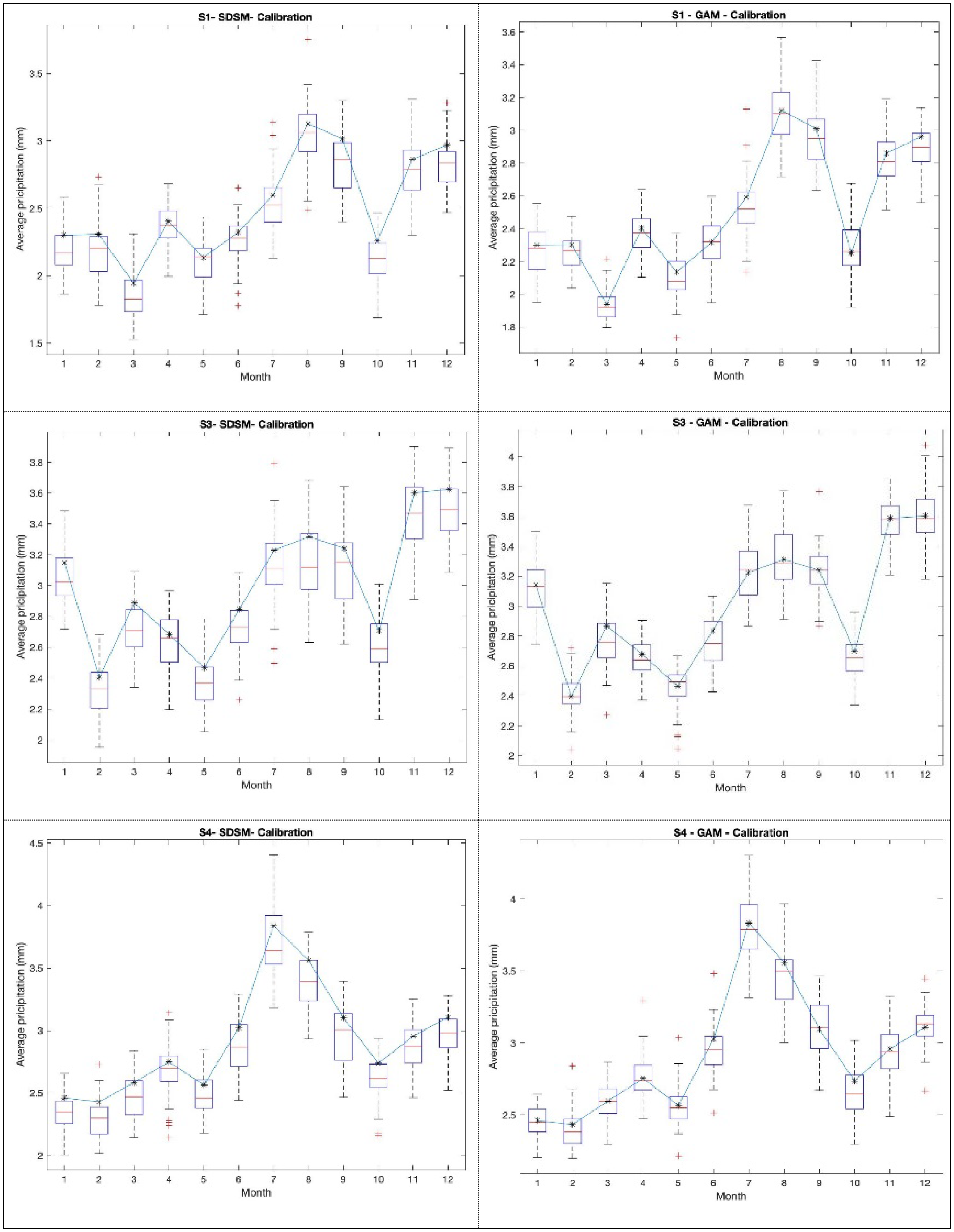
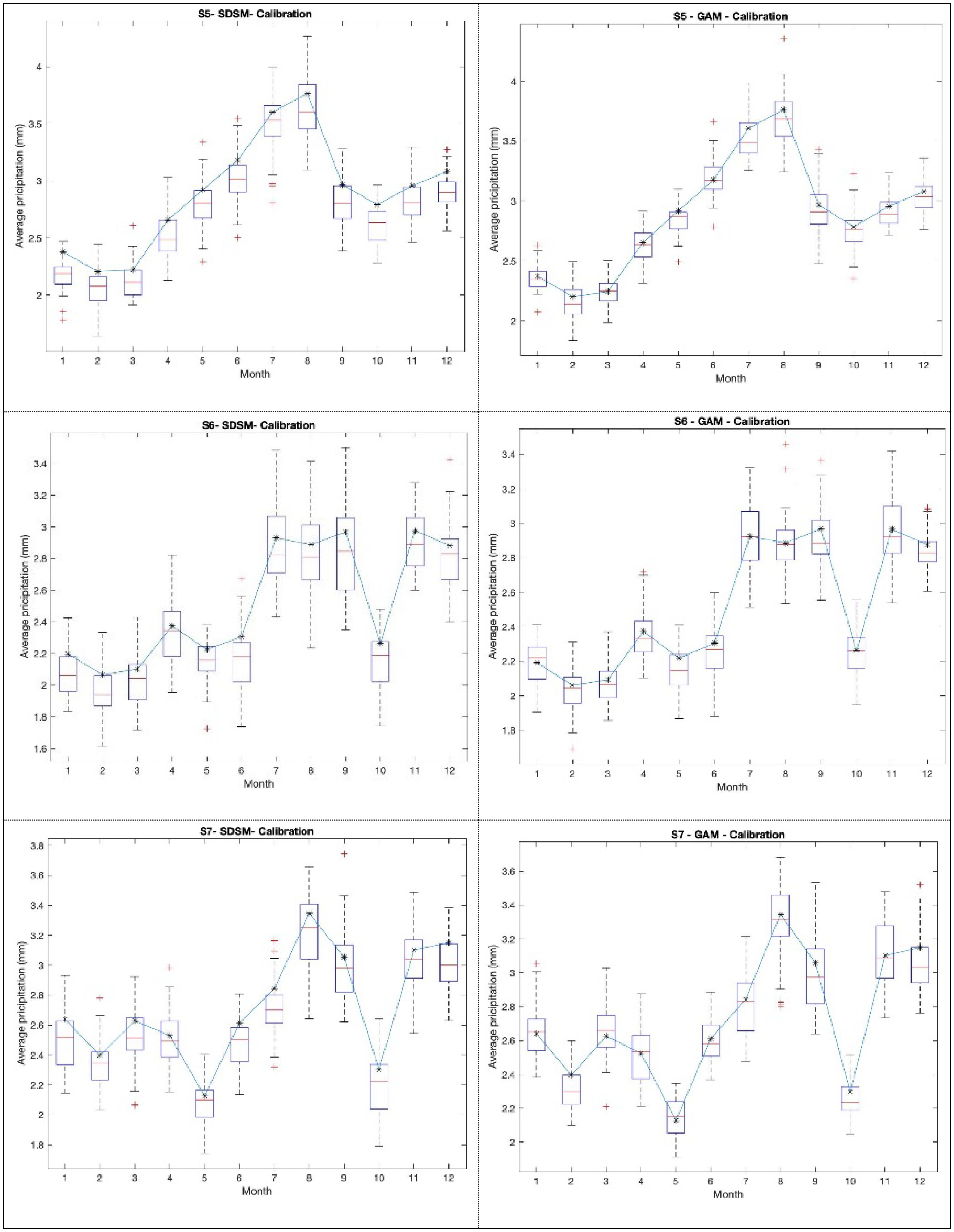


Figure A-1. Boxplots of monthly percentage of wet-day for SDSM (left) and SDGAM (right) for all stations (Black star markers indicate monthly average values of precipitation data)


S1- SDSM- Calibration
S1 - GAM - Calibration
S3- SDSM- Calibration
S3 - GAM - Calibration
S4- SDSM- Calibration
S4 - GAM - Calibration
Average pricipitation (mm)
Month
1
2
3
4
5
6
7
8
9
10
11
12


S5- SDSM- Calibration
Average pricipitation (mm)
Month
S5 - GAM - Calibration
Average pricipitation (mm)
Month
S6- SDSM- Calibration
Average pricipitation (mm)
Month
S6 - GAM - Calibration
Average pricipitation (mm)
Month
S7- SDSM- Calibration
Average pricipitation (mm)
Month
S7 - GAM - Calibration
Average pricipitation (mm)
Month

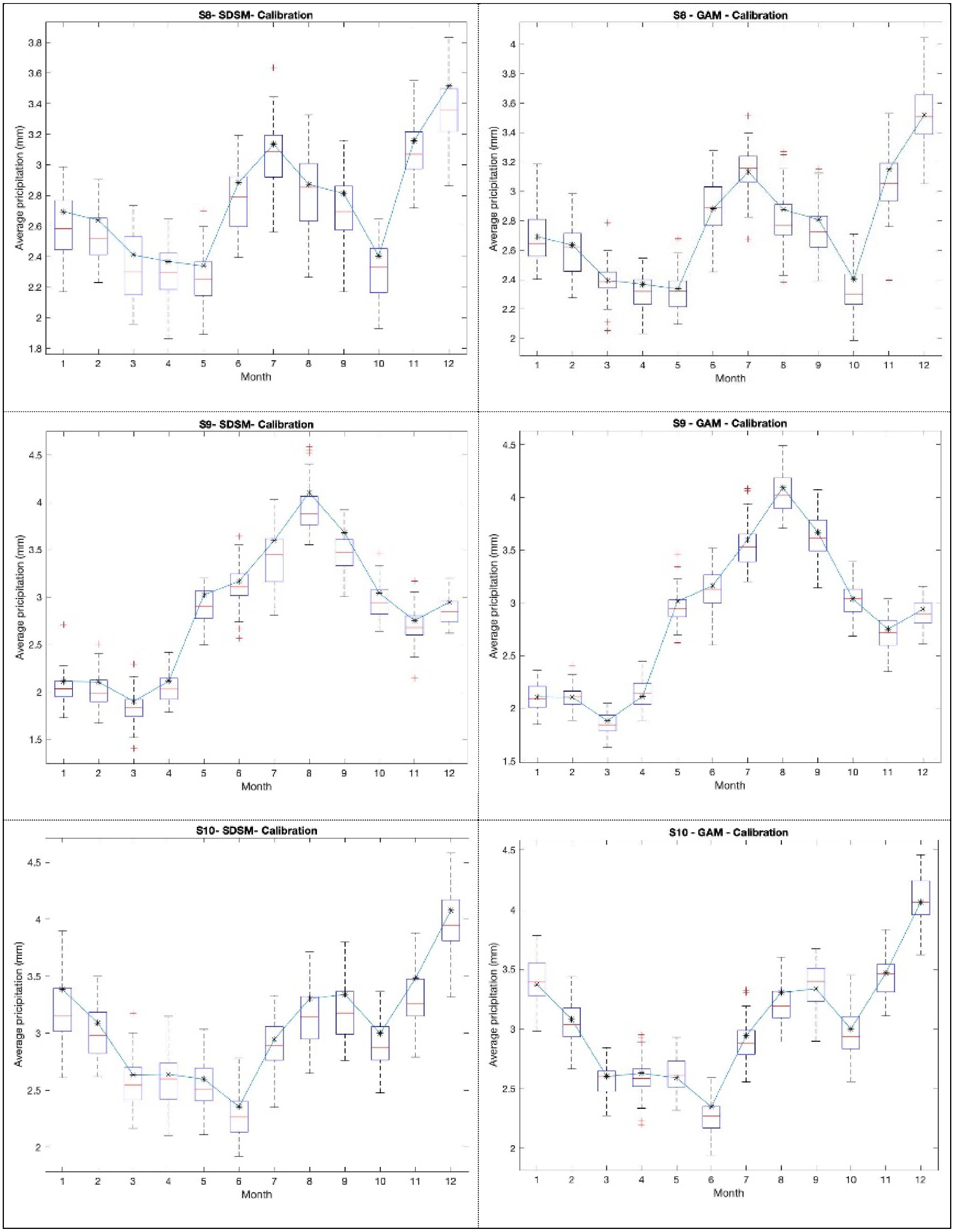


Figure A-2. Boxplot of monthly mean of precipitation for SDSM (left) and SDGAM (right) for all stations (Black star markers indicate monthly average values of precipitation data)

Table A-1. RMSEs of monthly Precip-m and Precip-std for all stations

Indices		Precip-m				Precip-std			
Station	Month	Calibration		Validation		Calibration		Validation	
		SDSM	SDGAM	SDSM	SDGAM	SDSM	SDGAM	SDSM	SDGAM
S1	Jan	1.494	1.357	1.182	1.277	2.778	2.394	2.251	2.534
	Feb	1.047	0.893	1.075	0.912	1.990	1.669	1.971	1.381
	Mar	1.259	1.040	1.007	1.198	2.764	2.058	2.098	2.754
	Apr	1.165	0.939	1.353	1.274	2.145	1.660	2.324	1.999
	May	1.187	0.982	1.152	1.072	2.081	1.722	2.548	2.486
	Jun	1.109	1.006	1.563	1.374	2.684	2.428	2.936	2.759
	Jul	1.348	1.175	1.645	1.496	3.120	2.778	3.223	3.000
	Aug	1.717	1.525	1.480	1.281	3.843	3.280	3.580	2.583
	Sep	1.731	1.454	1.538	1.379	4.690	3.455	3.854	2.600
	Oct	1.137	0.939	1.285	1.149	2.529	2.053	3.127	2.981
	Nov	1.149	0.986	1.189	1.037	2.200	1.792	2.503	1.979
	Dec	1.324	1.149	1.011	0.970	2.317	1.833	1.777	1.565
S3	Jan	1.508	1.359	1.581	1.405	2.317	2.138	2.669	2.454
	Feb	1.374	1.238	1.325	1.311	2.647	2.322	3.032	3.086
	Mar	1.398	1.242	1.304	1.238	3.197	2.739	3.021	2.734
	Apr	1.152	0.907	1.427	1.307	1.773	1.581	2.123	1.721
	May	1.131	1.012	1.516	1.475	2.246	2.013	2.853	2.816
	Jun	1.496	1.318	1.773	1.705	3.028	2.689	3.228	3.330
	Jul	2.128	1.963	1.518	1.352	4.261	3.810	2.992	2.504
	Aug	1.929	1.724	1.933	1.831	3.984	3.485	4.931	4.863

Indices		Precip-m				Precip-std			
Station	Month	Calibration		Validation		Calibration		Validation	
		SDSM	SDGAM	SDSM	SDGAM	SDSM	SDGAM	SDSM	SDGAM
	Sep	1.757	1.590	1.577	1.495	4.128	3.207	3.800	3.297
	Oct	1.237	1.176	1.146	1.035	2.568	2.247	2.393	2.071
	Nov	1.324	1.204	1.734	1.654	2.419	1.925	3.160	2.858
	Dec	1.519	1.407	2.118	2.071	2.246	2.088	3.257	3.216
S4	Jan	0.960	0.844	1.272	1.206	1.538	1.326	2.181	2.070
	Feb	1.009	0.877	1.261	1.103	1.999	1.694	2.353	1.881
	Mar	1.189	1.023	1.253	1.206	2.020	1.701	2.423	2.068
	Apr	1.223	1.074	1.250	1.091	2.100	1.739	2.180	1.773
	May	0.904	0.737	1.134	1.038	1.838	1.496	1.938	1.698
	Jun	1.619	1.476	1.341	1.089	3.243	2.948	2.942	2.524
	Jul	1.976	1.839	1.764	1.612	4.172	3.999	4.568	4.488
	Aug	1.512	1.356	1.948	1.888	3.572	3.527	3.867	3.573
	Sep	1.620	1.427	1.636	1.568	3.293	2.483	3.614	3.286
	Oct	1.057	0.903	1.464	1.357	2.914	2.704	3.342	3.338
	Nov	1.025	0.905	1.534	1.486	1.667	1.451	3.100	3.136
	Dec	0.955	0.878	1.293	1.171	1.603	1.361	2.382	2.401
S5	Jan	1.051	0.884	1.164	1.058	1.705	1.494	2.159	1.578
	Feb	1.025	0.913	1.028	0.899	2.169	1.996	1.883	1.530
	Mar	0.873	0.724	0.896	0.764	1.756	1.527	1.768	1.223
	Apr	1.343	1.151	1.264	1.235	2.221	1.783	2.155	1.572
	May	1.122	0.979	1.214	1.192	1.930	1.650	2.340	1.783
	Jun	1.401	1.324	1.258	1.201	2.647	2.213	2.428	1.924

Indices		Precip-m				Precip-std			
Station	Month	Calibration		Validation		Calibration		Validation	
		SDSM	SDGAM	SDSM	SDGAM	SDSM	SDGAM	SDSM	SDGAM
	Jul	1.609	1.460	1.764	1.625	2.413	1.961	3.974	3.845
	Aug	1.738	1.479	1.727	1.558	3.663	3.338	3.798	3.405
	Sep	1.368	1.234	1.454	1.206	2.582	2.230	2.988	2.525
	Oct	1.339	1.231	1.571	1.740	2.473	2.081	2.809	3.194
	Nov	1.108	0.826	1.410	1.209	2.286	2.006	2.665	1.855
	Dec	1.072	0.916	1.198	1.221	2.092	1.930	2.276	2.020
S6	Jan	1.294	1.205	0.888	0.798	2.476	2.375	1.673	1.581
	Feb	0.968	0.815	0.976	0.820	1.953	1.695	2.024	1.822
	Mar	1.090	0.896	0.959	0.992	1.982	1.498	1.966	2.022
	Apr	1.003	0.894	1.446	1.293	2.008	1.602	2.587	2.380
	May	1.088	0.915	1.131	1.003	1.943	1.652	2.084	1.843
	Jun	1.192	1.019	1.225	1.169	2.746	2.488	2.288	2.164
	Jul	1.560	1.391	1.456	1.329	3.309	2.832	3.057	3.001
	Aug	1.389	1.253	1.669	1.522	2.703	2.164	3.976	3.522
	Sep	1.647	1.347	1.652	1.539	4.416	3.422	4.149	3.550
	Oct	1.184	0.952	1.324	1.163	2.904	2.499	3.190	3.127
	Nov	1.158	1.057	1.257	1.057	2.540	2.398	2.876	2.671
	Dec	1.169	0.927	1.171	1.025	1.869	1.606	2.189	2.014
S7	Jan	1.265	1.121	1.095	1.075	2.724	2.337	2.061	1.974
	Feb	1.059	0.972	1.124	1.084	2.050	1.770	1.982	1.852
	Mar	1.213	1.048	1.233	1.347	2.434	1.911	2.647	2.508
	Apr	1.101	0.956	1.471	1.321	1.937	1.534	2.798	2.517

Indices		Precip-m				Precip-std			
Station	Month	Calibration		Validation		Calibration		Validation	
		SDSM	SDGAM	SDSM	SDGAM	SDSM	SDGAM	SDSM	SDGAM
	May	0.947	0.873	1.227	1.160	1.833	1.555	2.219	2.259
	Jun	1.186	0.898	1.429	1.347	2.386	1.829	3.334	3.447
	Jul	1.519	1.347	1.652	1.520	2.902	2.389	3.530	3.263
	Aug	1.538	1.356	1.610	1.503	3.068	2.767	3.563	2.979
	Sep	1.704	1.498	1.464	1.425	4.218	3.393	2.997	2.999
	Oct	1.259	1.097	1.251	1.208	3.138	2.845	2.750	2.883
	Nov	1.158	1.004	1.308	1.293	2.140	1.718	2.855	2.952
	Dec	1.111	0.972	1.309	1.234	2.108	1.938	1.890	1.608
S8	Jan	1.092	0.994	1.088	1.049	2.166	1.996	2.139	1.997
	Feb	1.359	1.281	1.326	1.196	2.644	2.484	2.129	1.770
	Mar	1.226	1.094	1.228	1.075	2.567	2.182	2.325	1.659
	Apr	1.020	0.865	1.145	0.958	2.017	1.557	2.230	1.807
	May	1.042	0.861	1.381	1.261	2.150	1.948	2.414	2.184
	Jun	1.437	1.299	1.617	1.490	3.326	3.081	3.621	3.389
	Jul	1.427	1.309	1.531	1.364	2.701	2.195	2.909	2.325
	Aug	1.565	1.237	1.789	1.540	3.574	2.989	4.663	4.246
	Sep	1.508	1.270	1.387	1.232	3.324	2.865	3.770	3.617
	Oct	1.085	0.824	1.249	1.207	2.479	1.942	2.974	3.022
	Nov	1.153	0.905	1.246	1.125	2.360	1.977	2.506	2.122
	Dec	1.379	1.190	1.422	1.395	2.384	2.099	2.153	1.696
S9	Jan	1.058	0.915	0.916	0.793	2.049	1.720	1.926	1.559
	Feb	1.047	0.882	1.105	0.862	2.402	2.008	2.725	2.108

Indices		Precip-m				Precip-std			
Station	Month	Calibration		Validation		Calibration		Validation	
		SDSM	SDGAM	SDSM	SDGAM	SDSM	SDGAM	SDSM	SDGAM
	Mar	1.112	0.915	1.102	1.008	2.590	1.961	2.679	2.493
	Apr	0.941	0.865	1.143	0.942	2.470	2.205	2.045	1.704
	May	1.275	1.187	1.544	1.416	2.640	2.471	2.650	2.305
	Jun	1.493	1.354	1.683	1.596	2.770	2.468	3.372	3.660
	Jul	1.780	1.560	2.039	1.652	3.671	3.069	4.451	3.878
	Aug	2.262	2.051	1.734	1.662	5.379	5.092	3.549	2.564
	Sep	1.680	1.507	1.437	1.339	3.892	3.105	3.524	2.938
	Oct	1.169	1.029	1.338	1.148	2.706	2.226	2.782	1.980
	Nov	1.150	0.974	1.315	1.186	2.102	1.742	2.569	2.114
	Dec	0.994	0.895	1.060	0.970	2.238	2.027	2.230	1.880
S10	Jan	1.407	1.260	1.587	1.383	3.302	2.883	3.007	2.169
	Feb	1.525	1.277	1.622	1.381	3.333	2.745	2.914	2.341
	Mar	1.490	1.292	1.603	1.430	3.489	2.966	3.192	2.433
	Apr	1.506	1.356	1.576	1.342	3.506	3.134	3.081	2.452
	May	1.293	1.165	1.325	1.097	2.554	2.145	2.674	2.243
	Jun	1.295	1.174	2.472	2.276	2.911	2.611	4.910	4.918
	Jul	1.576	1.401	1.598	1.427	3.217	2.831	3.165	2.932
	Aug	1.607	1.448	1.629	1.439	3.295	3.125	3.209	2.277
	Sep	1.842	1.712	1.486	1.170	3.861	3.211	3.389	2.401
	Oct	1.387	1.175	1.334	1.235	3.429	3.051	2.723	2.572
	Nov	1.304	1.156	1.514	1.372	2.664	2.188	2.794	2.292
	Dec	1.557	1.408	1.990	1.941	3.177	2.732	3.381	2.618

Table A-2. RMSEs of seasonal Prcp1, SDII, CDD, Prec90p for all stations

Station	Indices	Season	Calibration		Validation	
			SDSM	SDGAM	SDSM	SDGAM
S1	Prcp1 (%)	Spring	8.326	8.011	7.489	7.728
		Summer	6.873	6.533	8.566	7.713
		Fall	6.401	6.020	7.012	6.202
		Winter	7.261	6.902	7.045	6.851
	SDII (mm/wet-day)	Spring	1.958	1.809	1.745	1.878
		Summer	1.693	1.345	1.656	1.379
		Fall	2.490	2.021	2.906	2.202
		Winter	1.687	1.445	1.794	1.456
	CDD (days)	Spring	8.982	8.648	4.391	4.499
		Summer	13.551	13.976	4.928	4.727
		Fall	4.497	4.202	5.945	6.223
		Winter	7.594	7.608	3.749	3.618
	Prec90p (mm/day)	Spring	7.799	7.836	4.500	4.336
		Summer	4.700	4.642	5.671	6.030
		Fall	8.948	9.158	8.587	7.048
		Winter	5.368	5.409	4.552	3.596
S3	Prcp1 (%)	Spring	8.652	8.086	7.596	7.546
		Summer	6.611	6.295	9.043	9.118
		Fall	11.281	11.281	6.048	5.429
		Winter	6.644	6.637	6.854	6.601
	SDII (mm/wet-day)	Spring	2.262	2.101	1.826	1.763
		Summer	1.592	1.431	2.807	2.729

Station	Indices	Season	Calibration		Validation	
			SDSM	SDGAM	SDSM	SDGAM
		Fall	2.812	2.505	2.348	2.150
		Winter	1.304	1.123	2.090	2.062
	CDD (days)	Spring	9.712	9.194	4.011	3.778
		Summer	9.456	9.592	14.958	15.223
		Fall	20.873	21.038	4.492	4.419
		Winter	8.574	8.532	3.173	2.884
	Prec90p (mm/day)	Spring	9.275	9.420	6.166	6.926
		Summer	6.036	6.061	8.736	9.445
		Fall	8.734	8.816	6.836	6.477
		Winter	4.649	4.287	6.318	7.304
S4	Prcp1 (%)	Spring	8.145	7.178	11.323	11.281
		Summer	5.635	5.189	8.195	6.688
		Fall	6.880	6.651	9.177	9.186
		Winter	7.281	6.107	11.406	11.398
	SDII (mm/wet-day)	Spring	1.565	1.418	2.444	2.078
		Summer	1.663	1.323	1.703	1.384
		Fall	2.271	1.983	3.339	2.965
		Winter	1.307	1.156	2.168	2.063
	CDD (days)	Spring	5.006	4.841	18.884	18.890
		Summer	4.430	4.673	4.289	4.172
		Fall	4.366	4.046	17.296	17.321
		Winter	4.177	4.031	22.294	22.267
		Spring	4.547	5.001	5.534	4.954

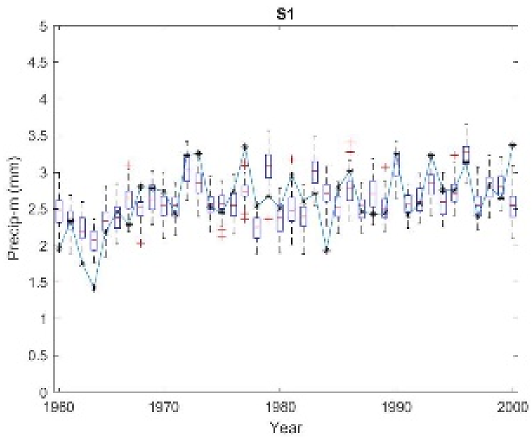
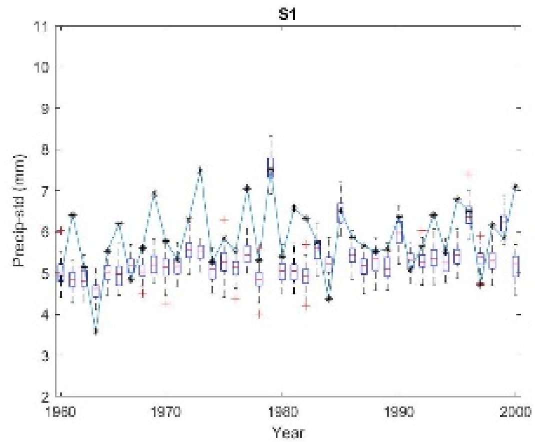
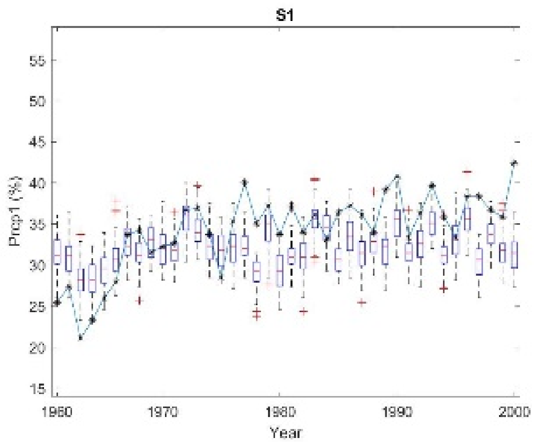
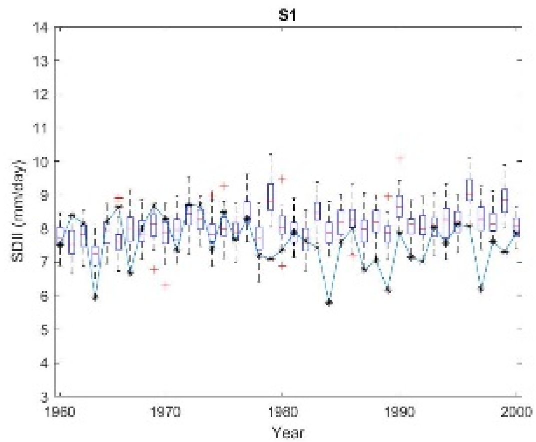
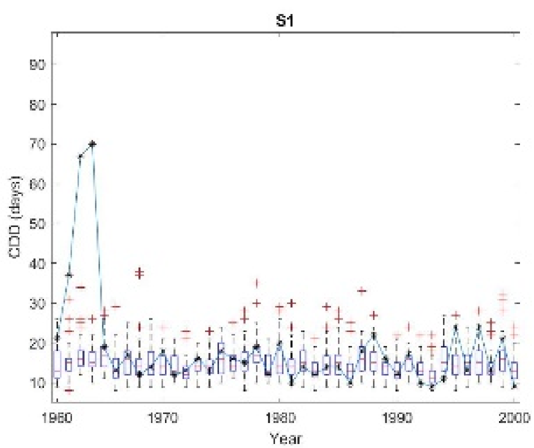
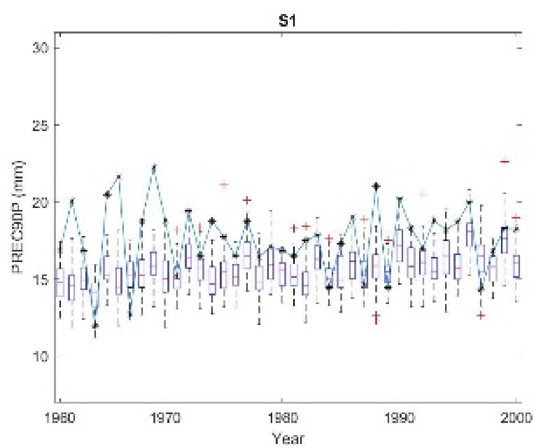
Station	Indices	Season	Calibration		Validation	
			SDSM	SDGAM	SDSM	SDGAM
	Prec90p (mm/day)	Summer	5.253	4.830	4.820	4.284
		Fall	9.086	9.140	17.024	17.815
		Winter	3.552	3.596	6.609	7.263
S5	Prcp1 (%)	Spring	6.809	5.966	7.457	7.873
		Summer	5.538	5.403	8.251	7.957
		Fall	6.024	5.588	6.398	6.010
		Winter	5.985	5.591	10.214	10.659
	SDII (mm/wet-day)	Spring	1.374	1.219	1.726	1.456
		Summer	1.838	1.607	2.353	2.044
		Fall	2.100	1.556	2.380	1.887
		Winter	1.562	1.328	1.758	1.676
	CDD (days)	Spring	6.887	6.643	7.214	7.210
		Summer	4.205	4.106	4.062	4.101
		Fall	3.845	3.637	8.597	8.482
		Winter	3.900	3.888	18.388	18.624
	Prec90p (mm/day)	Spring	4.765	5.194	4.422	3.692
		Summer	5.754	6.114	5.802	4.533
		Fall	6.188	6.296	7.895	7.593
		Winter	5.580	5.991	5.003	5.313
S6	Prcp1 (%)	Spring	6.417	6.168	8.339	7.941
		Summer	5.527	5.135	7.928	7.123
		Fall	6.542	6.389	6.997	6.204
		Winter	6.453	5.689	6.867	6.019

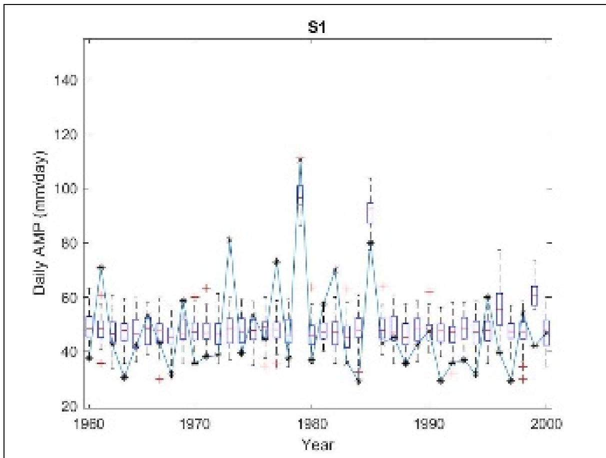
Station	Indices	Season	Calibration		Validation	
			SDSM	SDGAM	SDSM	SDGAM
	SDII (mm/wet-day)	Spring	1.759	1.565	1.483	1.235
		Summer	1.668	1.331	1.592	1.343
		Fall	2.247	1.996	2.759	2.403
		Winter	1.609	1.282	1.679	1.241
	CDD (days)	Spring	4.867	4.676	4.388	4.115
		Summer	5.277	4.944	4.998	5.032
		Fall	4.978	4.597	4.646	4.495
		Winter	4.398	4.348	3.862	3.785
	Prec90p (mm/day)	Spring	5.013	4.662	4.999	4.788
		Summer	5.495	5.777	5.009	5.349
		Fall	8.991	10.225	8.594	8.447
		Winter	4.707	4.435	4.930	4.849
S7	Prcp1 (%)	Spring	8.264	7.975	6.172	6.281
		Summer	5.729	6.047	8.112	7.570
		Fall	6.591	6.666	5.722	5.414
		Winter	6.305	6.119	7.066	6.827
	SDII (mm/wet-day)	Spring	2.320	2.356	2.067	1.833
		Summer	1.569	1.242	1.721	1.574
		Fall	2.498	2.154	2.375	2.113
		Winter	1.345	1.017	1.451	1.203
	CDD (days)	Spring	8.992	8.955	4.404	4.495
		Summer	5.502	5.810	4.667	4.626
		Fall	6.151	6.080	5.376	5.705

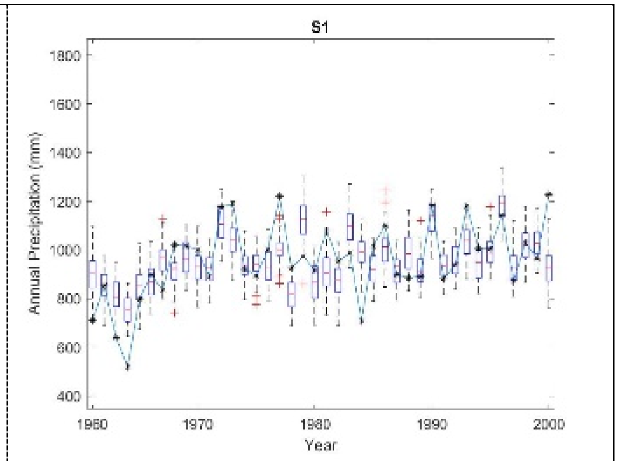
Station	Indices	Season	Calibration		Validation	
			SDSM	SDGAM	SDSM	SDGAM
		Winter	6.595	6.737	4.064	3.729
	Prec90p (mm/day)	Spring	8.175	8.252	6.226	4.985
		Summer	4.988	4.154	6.512	7.295
		Fall	7.603	7.832	6.772	5.702
		Winter	4.694	4.587	4.906	4.476
S8	Prcp1 (%)	Spring	6.357	6.515	7.280	7.075
		Summer	5.857	5.657	6.596	6.762
		Fall	6.244	5.973	6.523	6.227
		Winter	5.507	5.738	7.150	6.771
	SDII (mm/wet-day)	Spring	1.409	1.167	1.739	1.570
		Summer	1.700	1.433	1.896	1.647
		Fall	2.111	1.755	2.727	2.347
		Winter	1.493	1.212	1.715	1.512
	CDD (days)	Spring	7.217	7.496	3.901	4.024
		Summer	4.775	4.803	5.185	5.155
		Fall	4.789	4.470	5.687	5.625
		Winter	3.503	3.442	3.496	3.368
	Prec90p (mm/day)	Spring	4.084	3.590	5.948	5.525
		Summer	5.763	5.477	6.815	7.512
		Fall	8.101	8.544	8.323	8.852
		Winter	4.822	4.648	5.632	5.913
S9	Prcp1 (%)	Spring	7.091	7.030	8.187	8.740
		Summer	5.659	5.626	6.078	5.749

Station	Indices	Season	Calibration		Validation	
			SDSM	SDGAM	SDSM	SDGAM
		Fall	6.901	7.070	5.597	5.143
		Winter	6.127	6.473	7.209	7.190
	SDII (mm/wet-day)	Spring	1.635	1.358	2.060	1.854
		Summer	2.279	2.100	2.019	1.977
		Fall	3.062	2.527	2.171	1.874
		Winter	1.560	1.467	1.728	1.458
	CDD (days)	Spring	7.960	8.179	4.653	4.648
		Summer	4.376	4.236	5.080	4.832
		Fall	3.803	3.664	4.286	4.241
		Winter	4.011	4.088	4.675	4.702
	Prec90p (mm/day)	Spring	5.726	5.254	7.244	7.090
		Summer	7.106	6.990	6.707	7.439
		Fall	9.734	9.228	7.627	6.675
		Winter	5.473	5.340	6.053	5.366
S10	Prcp1 (%)	Spring	5.493	5.746	7.544	7.458
		Summer	7.595	7.416	8.294	7.437
		Fall	6.182	6.307	5.378	5.119
		Winter	5.060	4.947	6.512	6.681
	SDII (mm/wet-day)	Spring	2.632	2.254	2.015	1.828
		Summer	2.810	2.408	2.455	2.211
		Fall	2.289	1.969	2.168	1.790
		Winter	1.866	1.540	2.045	1.699
	CDD	Spring	6.171	6.409	7.718	7.967

Station	Indices	Season	Calibration		Validation	
			SDSM	SDGAM	SDSM	SDGAM
	(days)	Summer	18.693	18.643	4.786	4.051
		Fall	5.981	5.960	4.034	3.877
		Winter	4.489	4.390	4.308	4.400
	Prec90p (mm/day)	Spring	9.173	9.251	5.848	5.240
		Summer	7.976	5.985	9.001	9.715
		Fall	9.874	10.588	6.598	6.111
		Winter	6.174	5.692	6.266	5.770

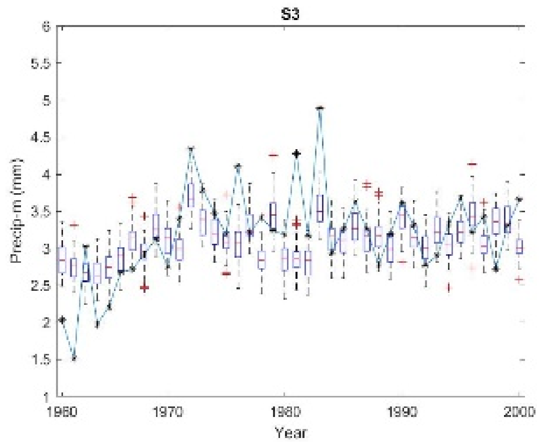
	
Precip_m	Precip_std
	
Prcp1	SDII
	
CDD	Prec90p



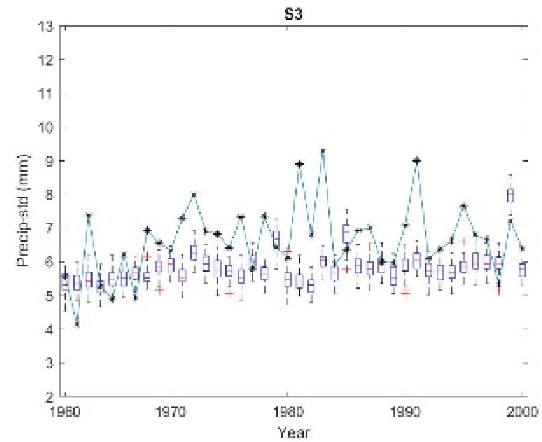


AMS	TAP

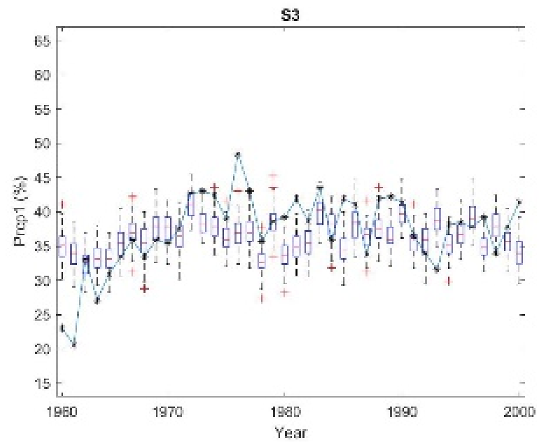
Figure A-3. Boxplots of annual statistics and indices of SDGAM model: Precip_m, Precip_std, Prcp1, SDII, CDD, Prec90p, AMS and TAP for 1961-2000 period at S1



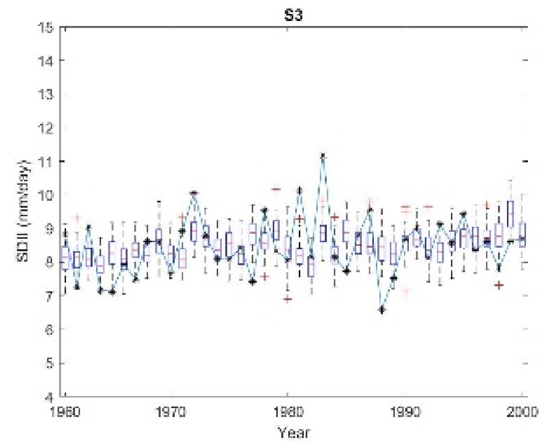
Precip_m



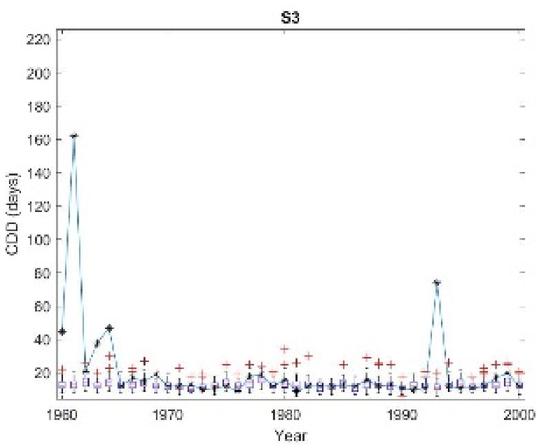
Precip_std



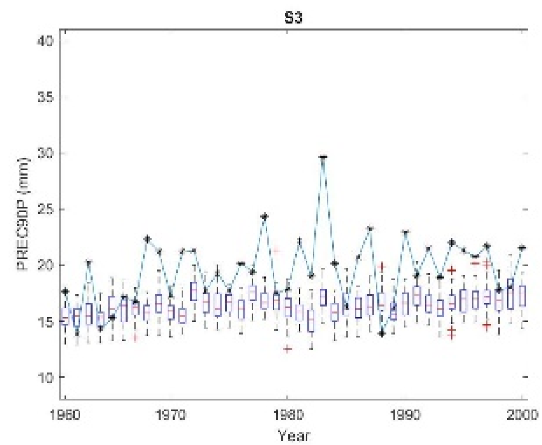
Prcp1



SDII



CDD



Prec90p

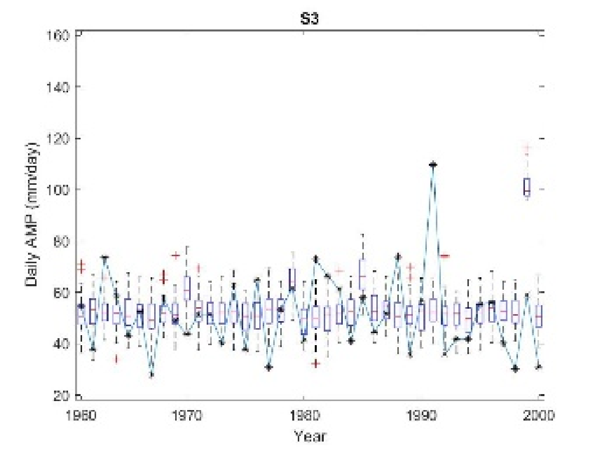
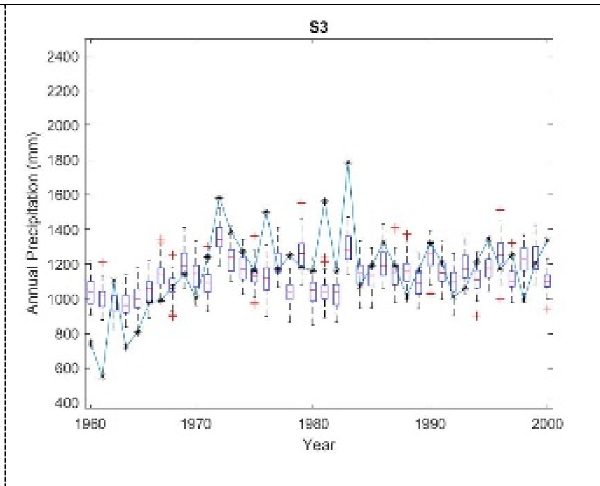
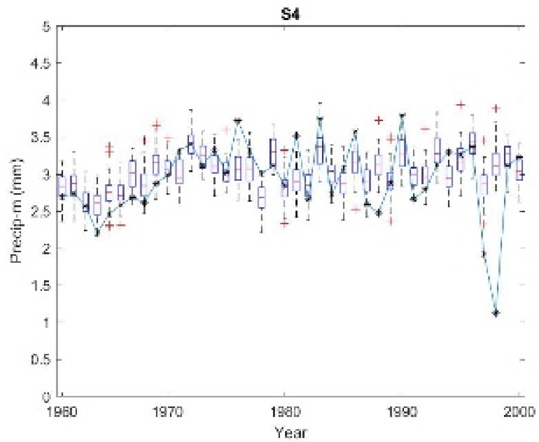
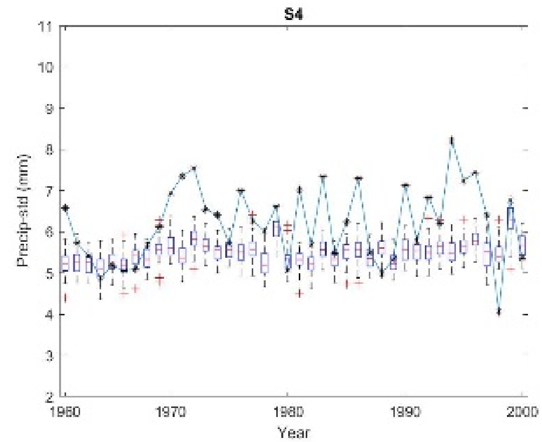
	
AMS	TAP

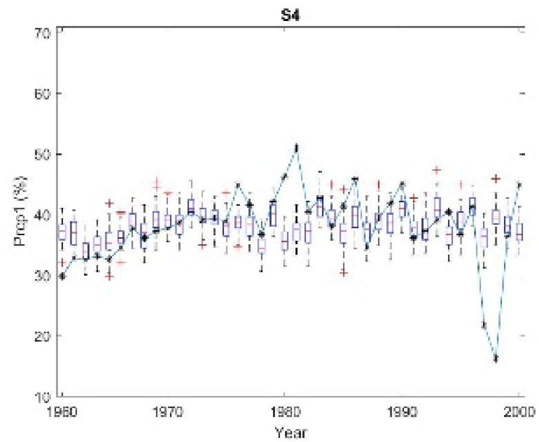
Figure A-4. Boxplots of annual statistics and indices of SDGAM model: Precip_m, Precip_std, Prcp1, SDII, CDD, Prec90p, AMS and TAP for 1961-2000 period at S3



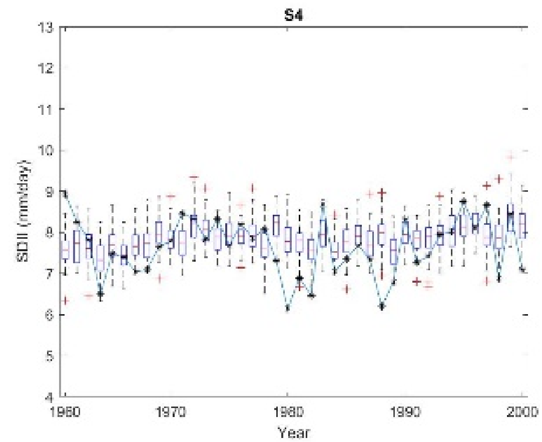
Precip_m



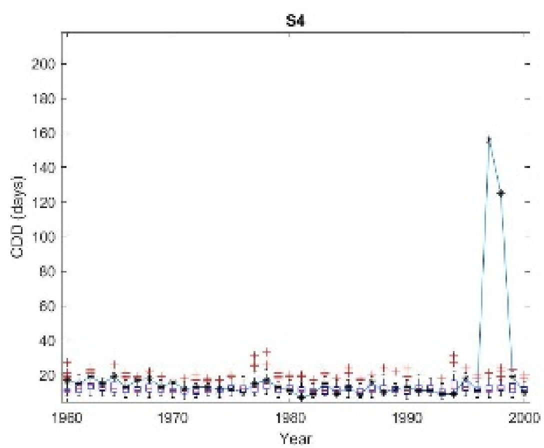
Precip_std



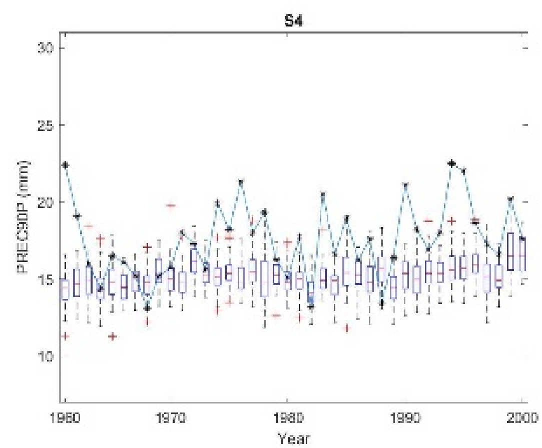
Prcp1



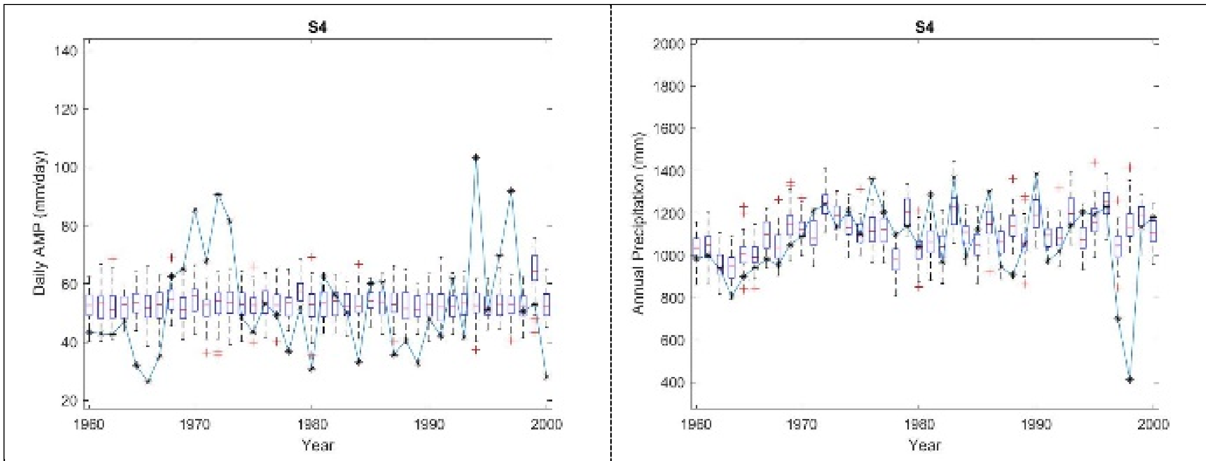
SDII



CDD

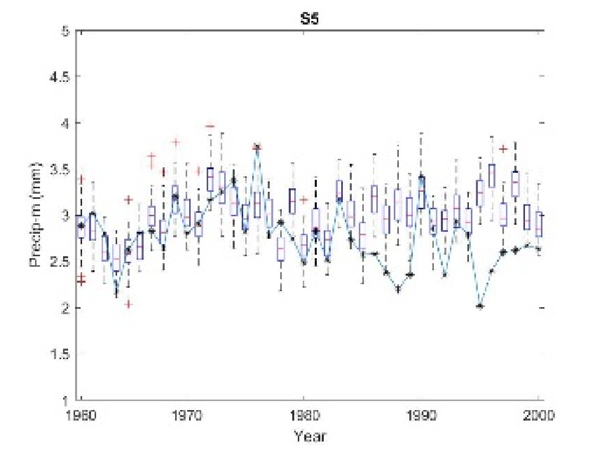


Prec90p

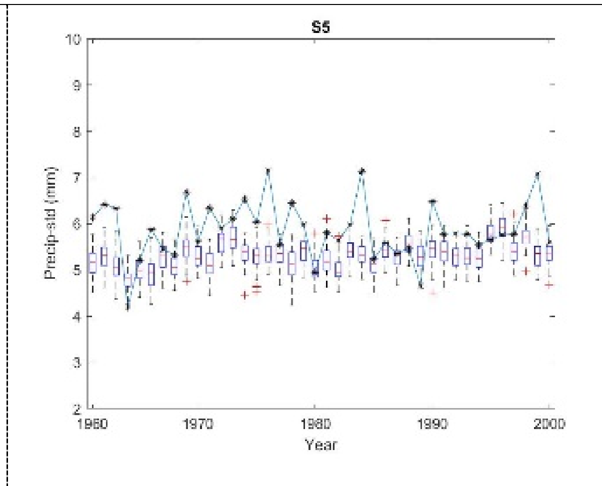


AMS	TAP

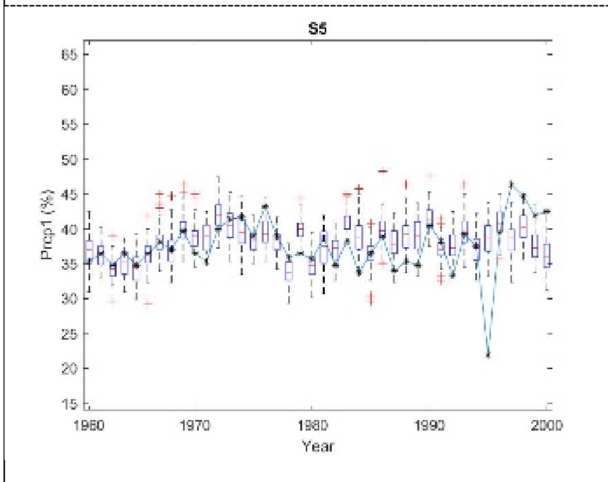
Figure A-5. Boxplots of annual statistics and indices of SDGAM model: Precip_m, Precip_std, Prcp1, SDII, CDD, Prec90p, AMS and TAP for 1961-2000 period at S4



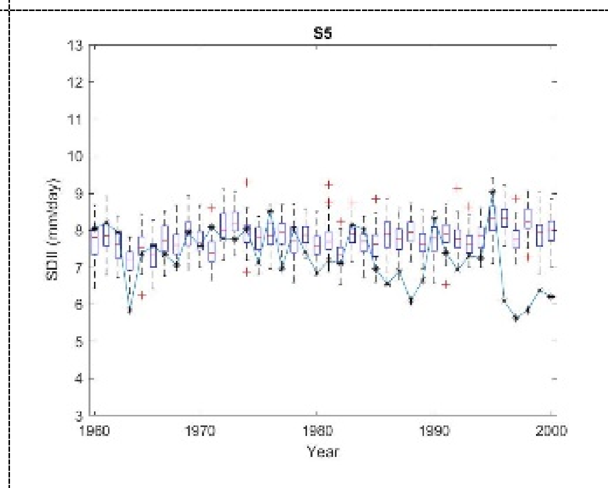
Precip_m



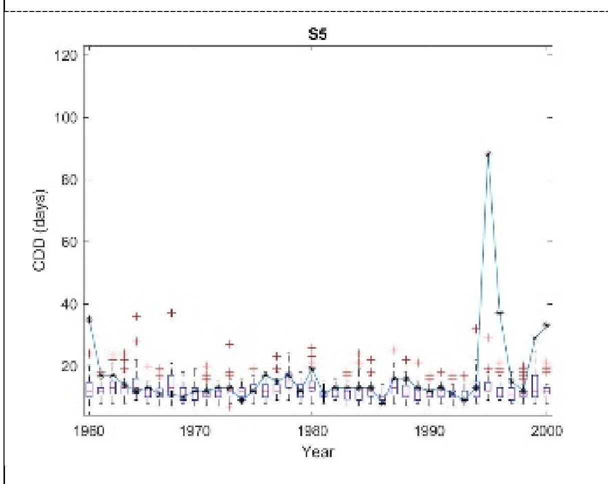
Precip_std



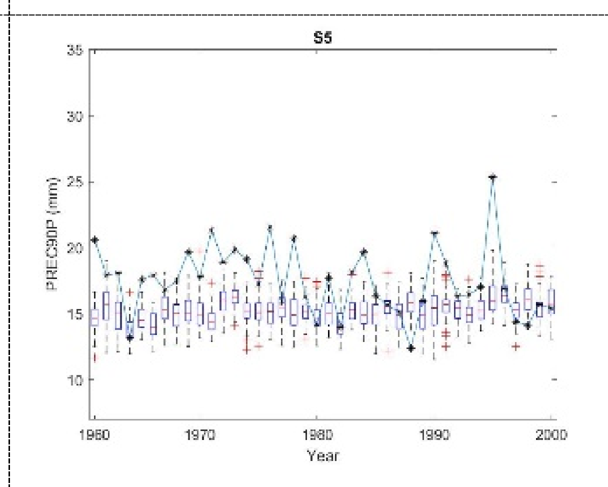
Prcp1



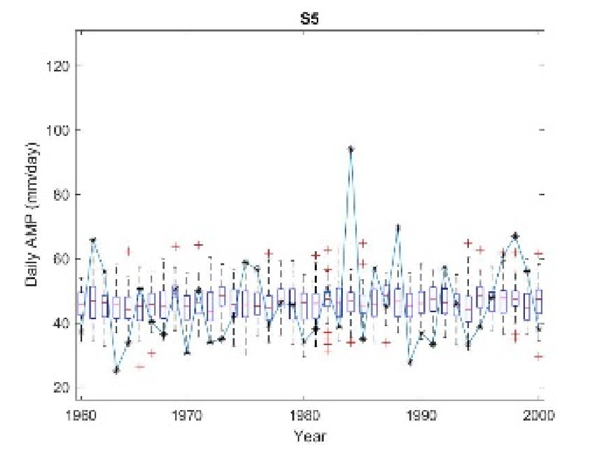
SDII



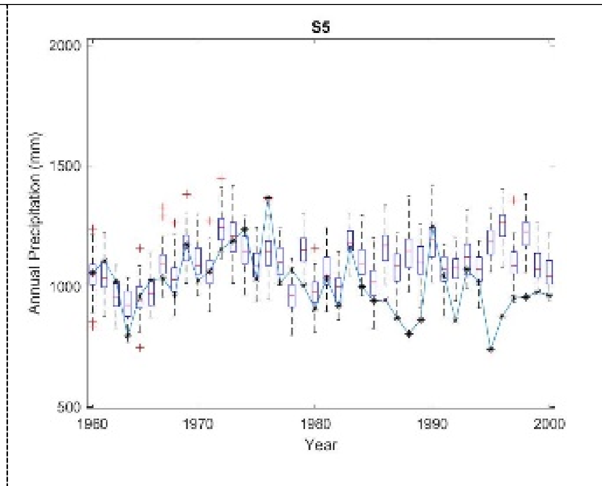
CDD



Prec90p

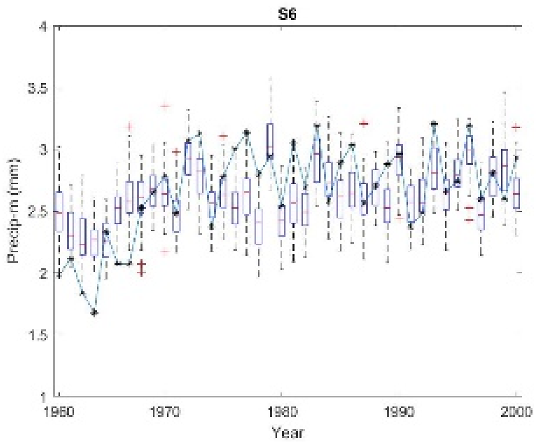


AMS

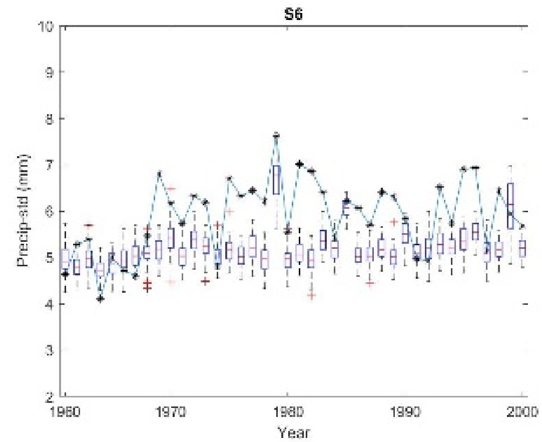


TAP

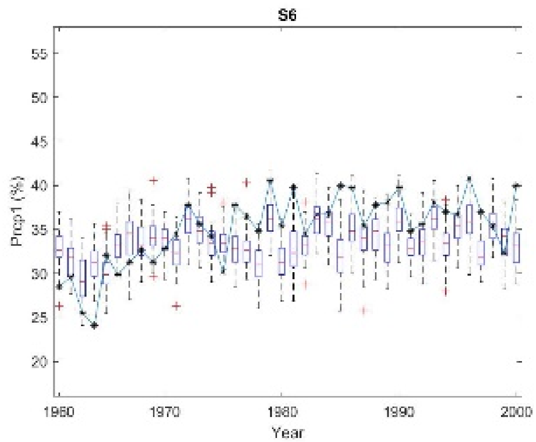
Figure A-6. Boxplots of annual statistics and indices of SDGAM model: Precip_m, Precip_std, Prcp1, SDII, CDD, Prec90p, AMS and TAP for 1961-2000 period at S5



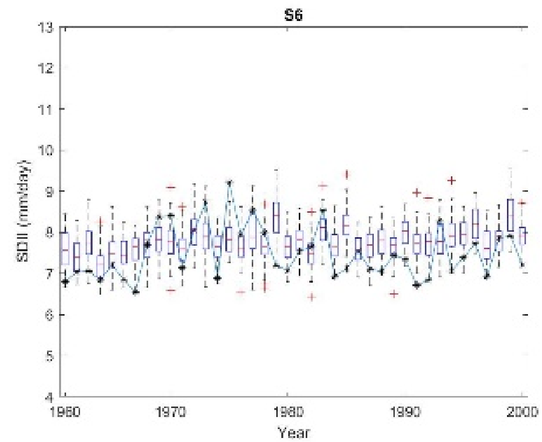
Precip_m



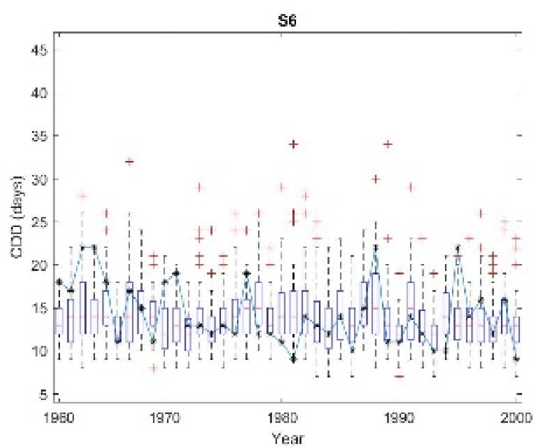
Precip_std



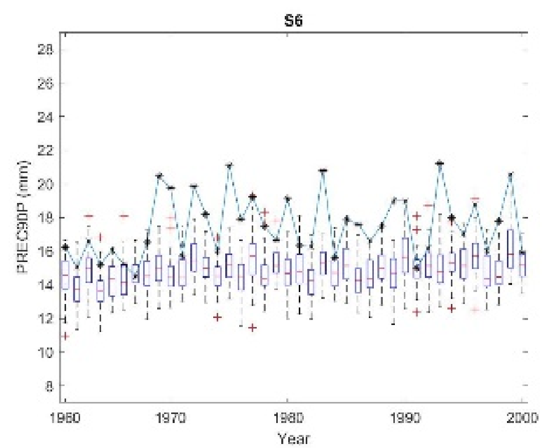
Prcp1



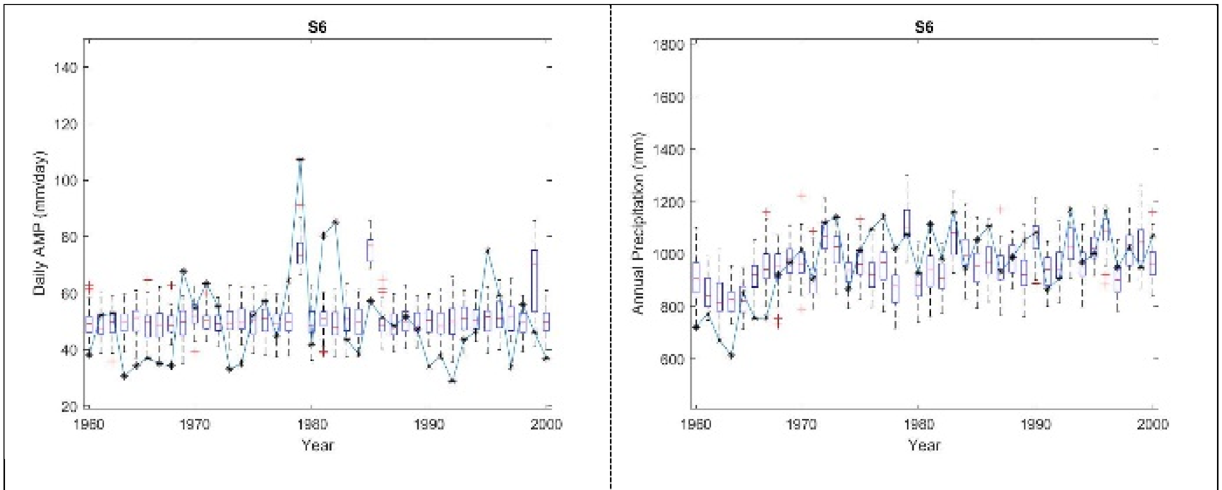
SDII



CDD

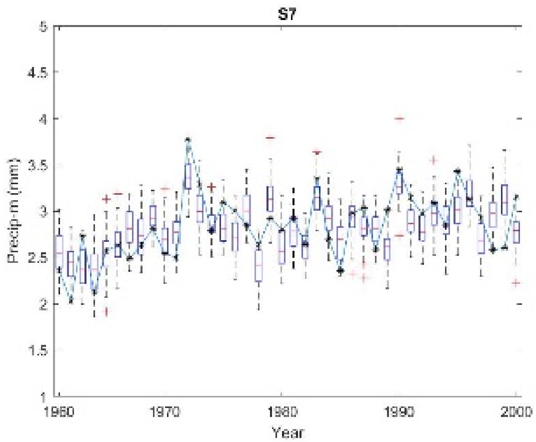


Prec90p

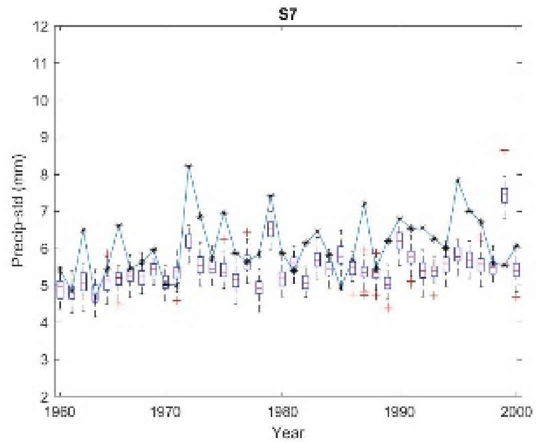


AMS	TAP

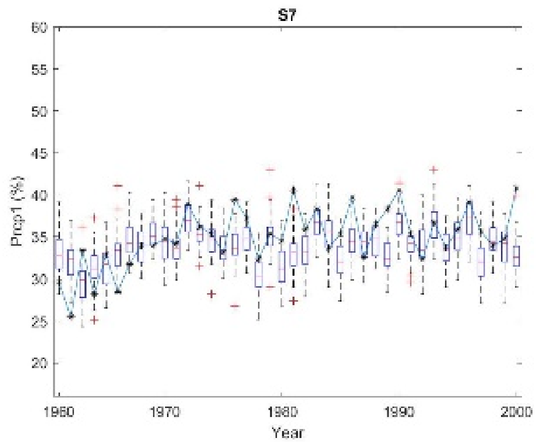
Figure A-7. Boxplots of annual statistics and indices of SDGAM model: Precip_m, Precip_std, Prcp1, SDII, CDD, Prec90p, AMS and TAP for 1961-2000 period at S6



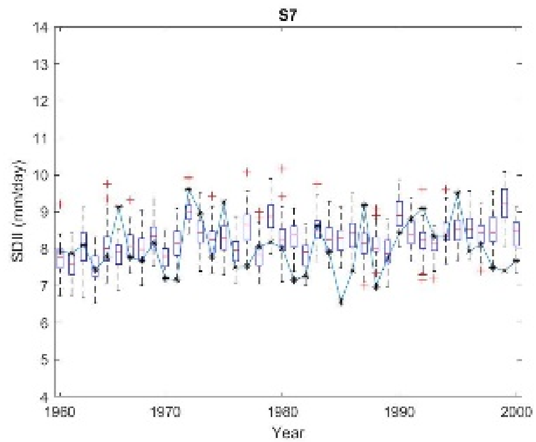
Precip_m



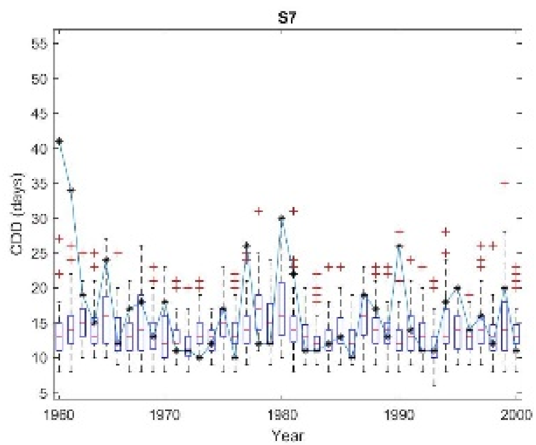
Precip_std



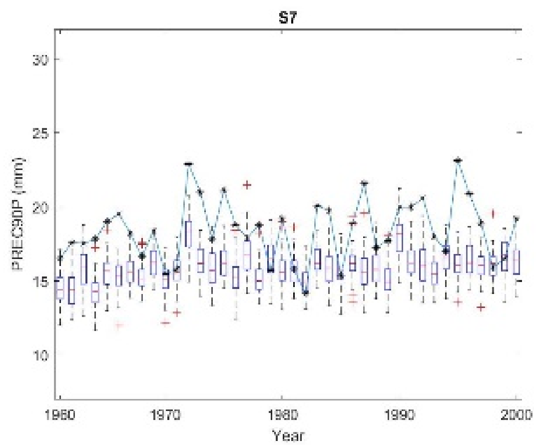
Prcp1



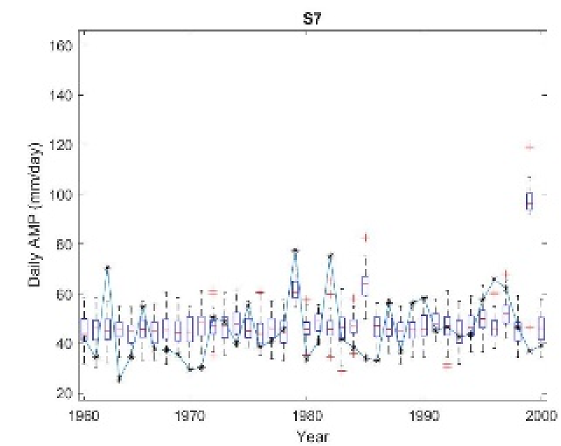
SDII

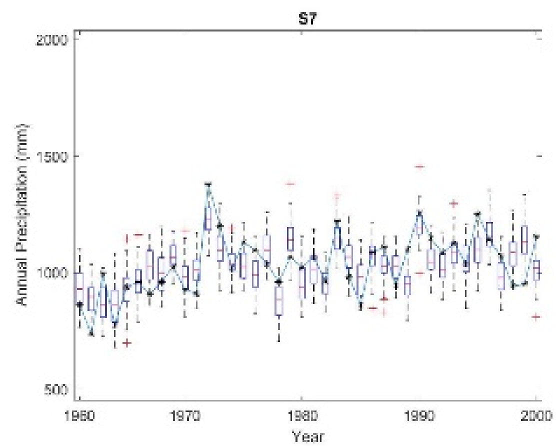


CDD



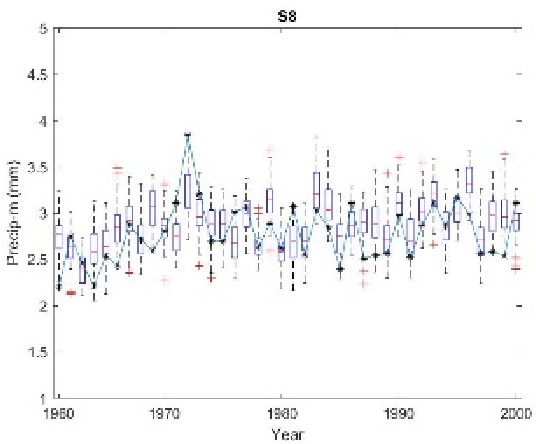
Prec90p



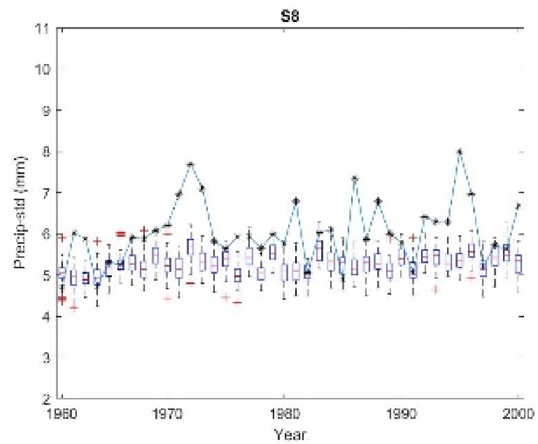


AMS	TAP

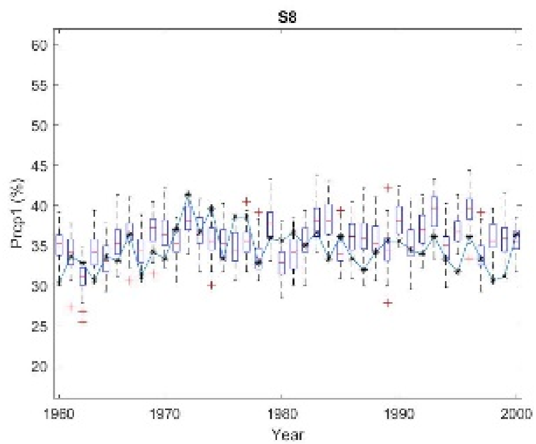
Figure A-8. Boxplots of annual statistics and indices of SDGAM model: Precip_m, Precip_std, Prcp1, SDII, CDD, Prec90p, AMS and TAP for 1961-2000 period at Dorval station (S7)



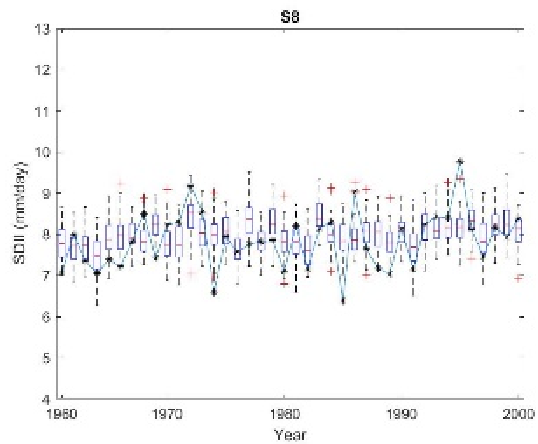
Precip_m



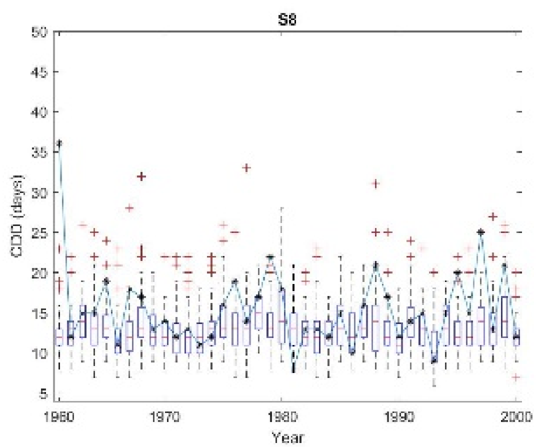
Precip_std



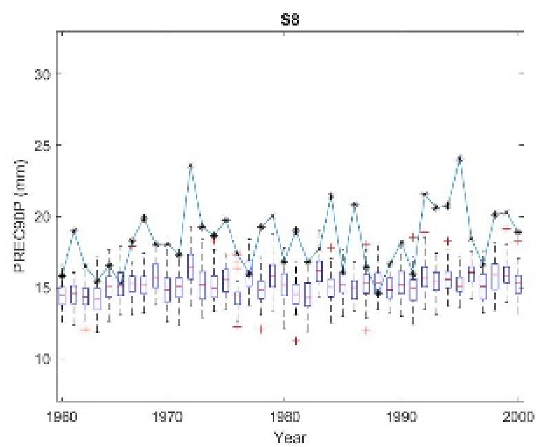
Prcp1



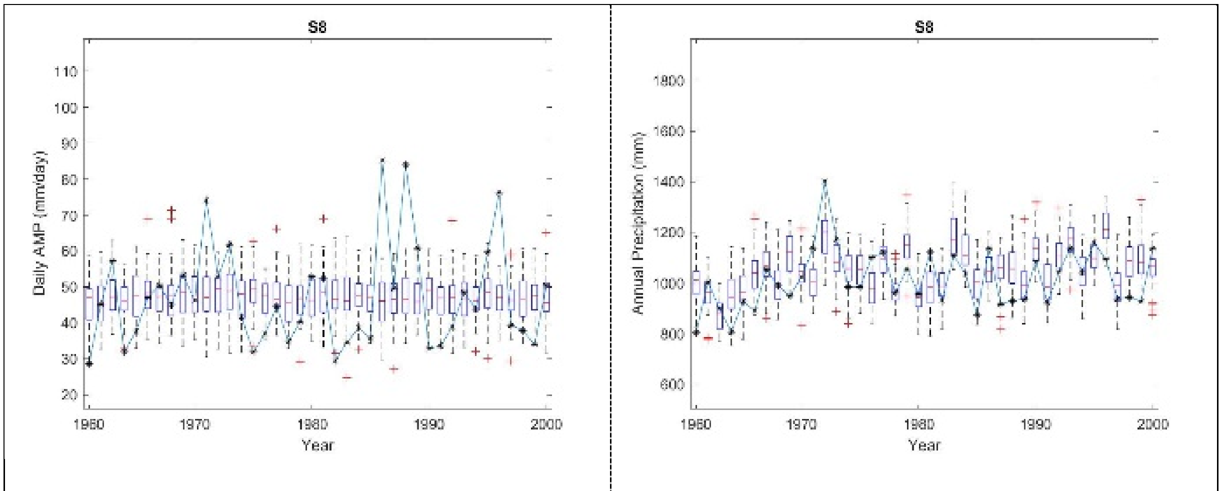
SDII



CDD

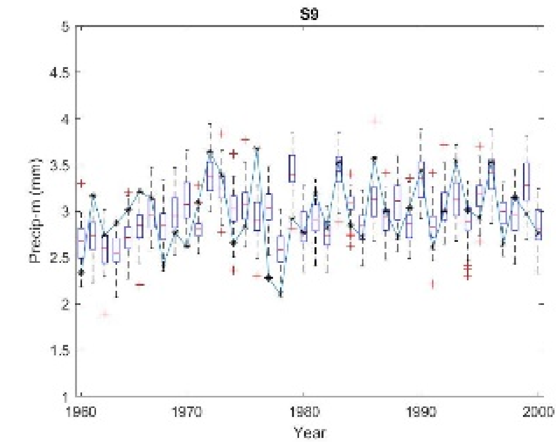


Prec90p

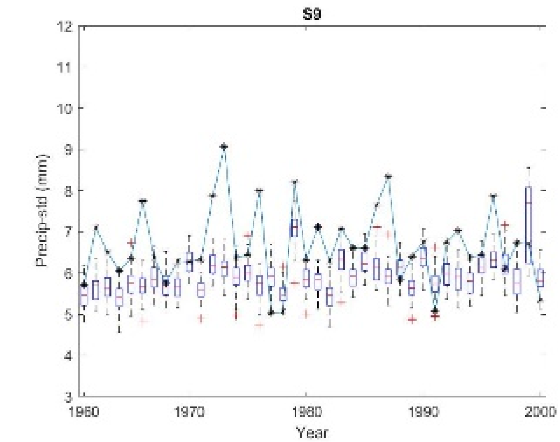


AMS TAP

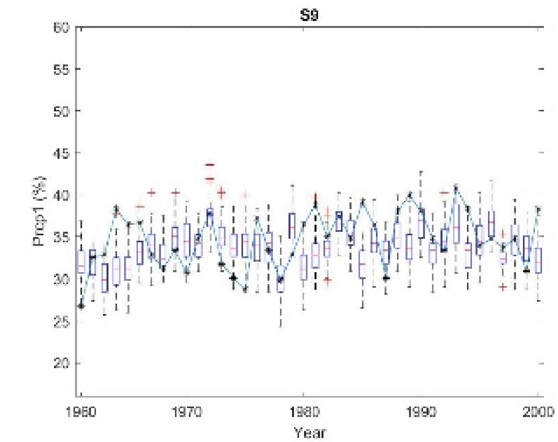
Figure A-9. Boxplots of annual statistics and indices of SDGAM model: Precip_m, Precip_std, Prcp1, SDII, CDD, Prec90p, AMS and TAP for 1961-2000 period at S8



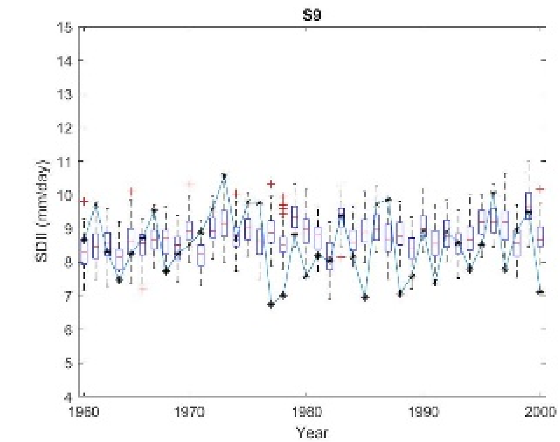
Precip_m



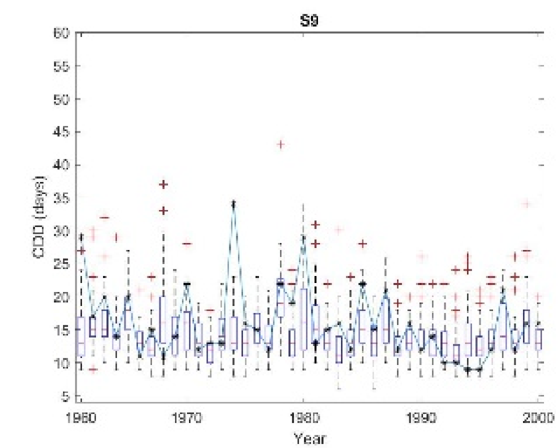
Precip_std



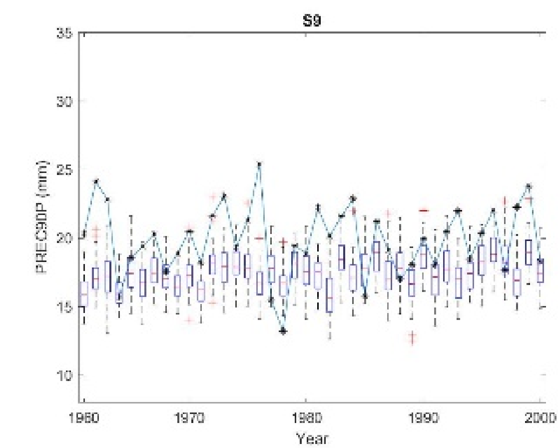
Prcp1



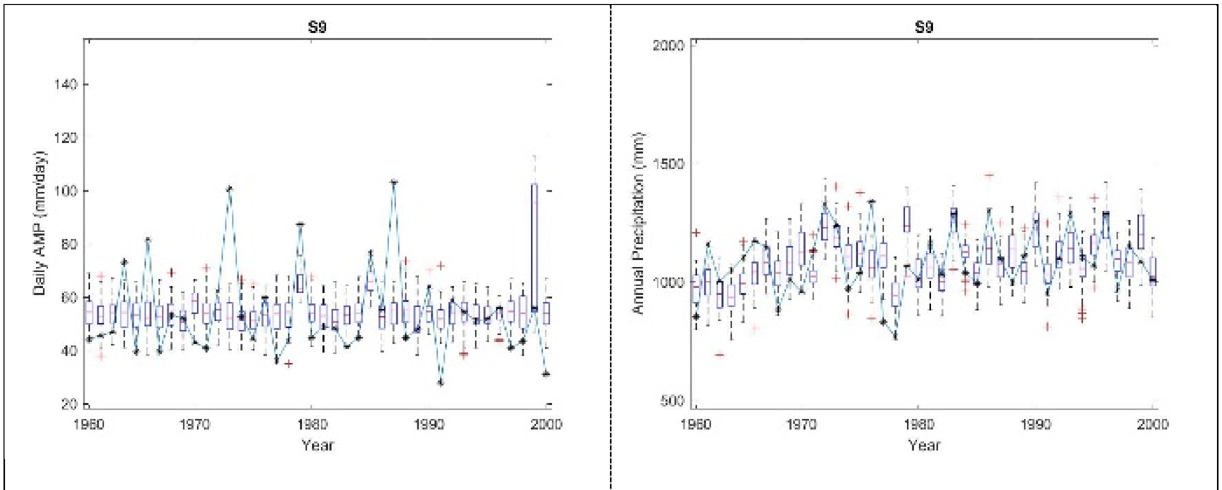
SDII



CDD

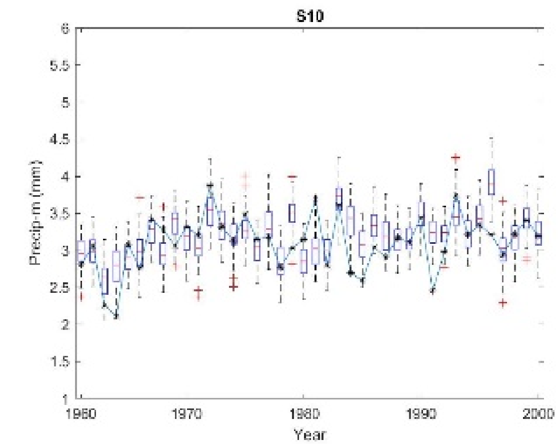


Prec90p

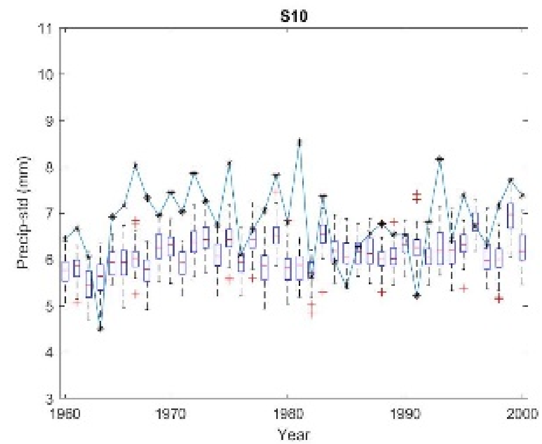


AMS	TAP

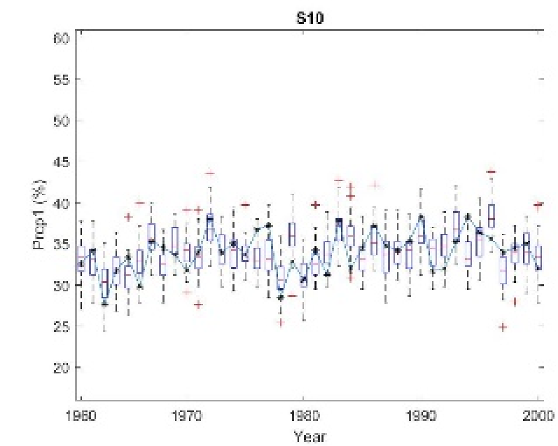
Figure A-10. Boxplots of annual statistics and indices of SDGAM model: Precip_m, Precip_std, Prcp1, SDII, CDD, Prec90p, AMS and TAP for 1961-2000 period at S9



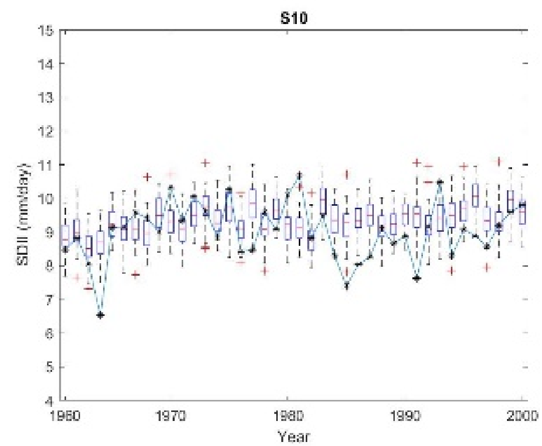
Precip_m



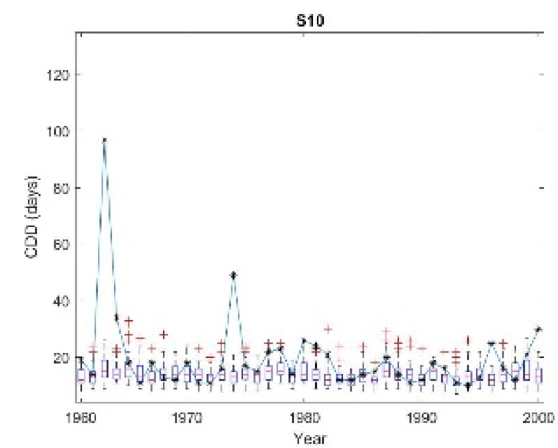
Precip_std



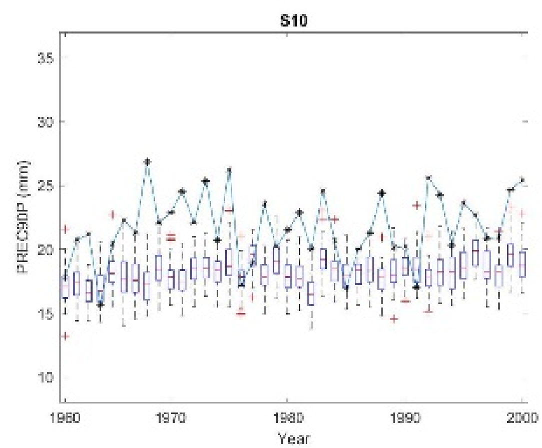
Prcp1



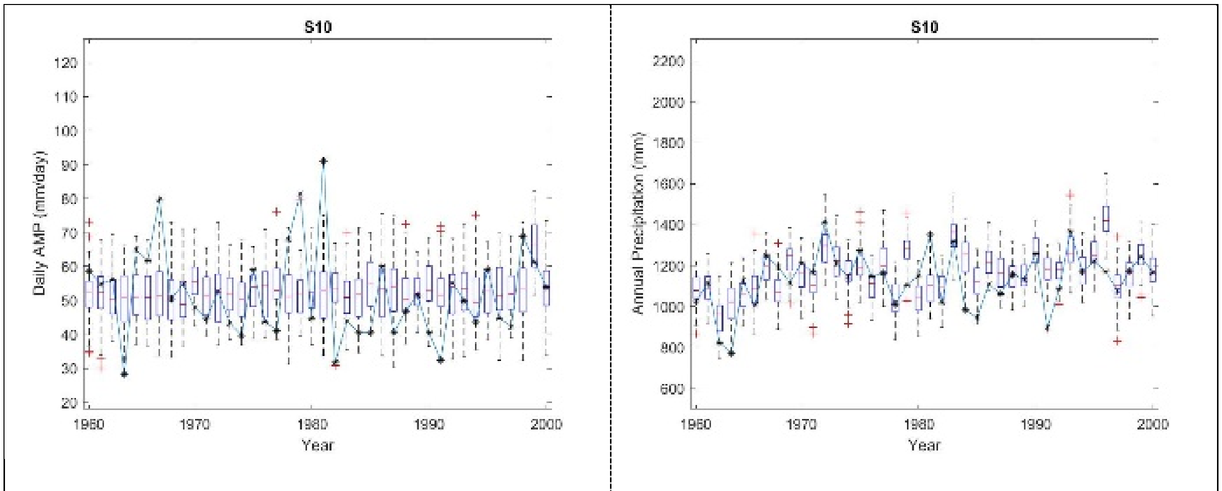
SDII



CDD



Prec90p



AMS	TAP

Figure A-11. Boxplots of annual statistics and indices of SDGAM model: Precip_m, Precip_std, Prcp1, SDII, CDD, Prec90p, AMS and TAP for 1961-2000 period at S10

Appendix B: Supplementary materials for chapter 3

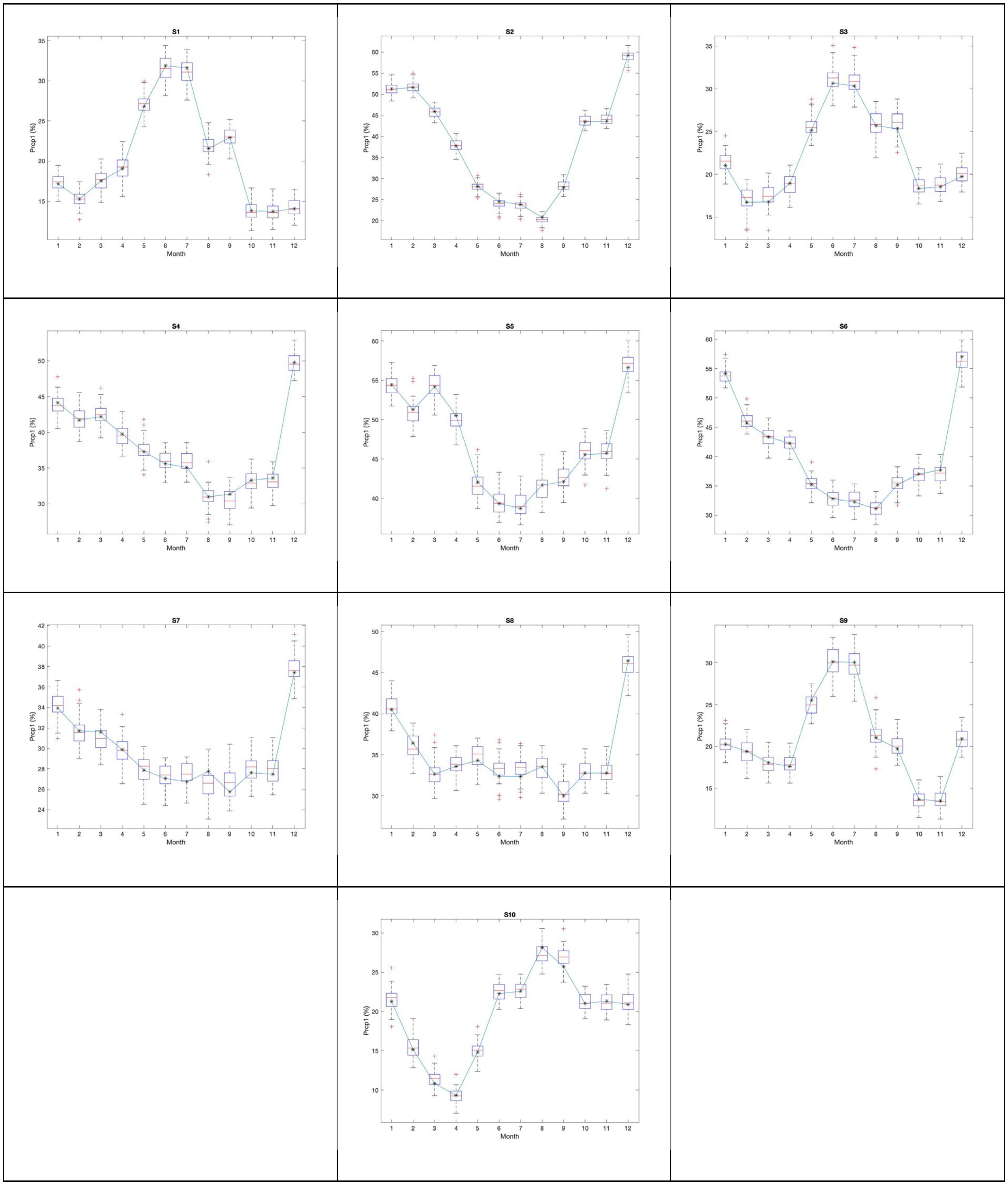


Figure B-1. Boxplots of monthly mean of percentage of wet-day

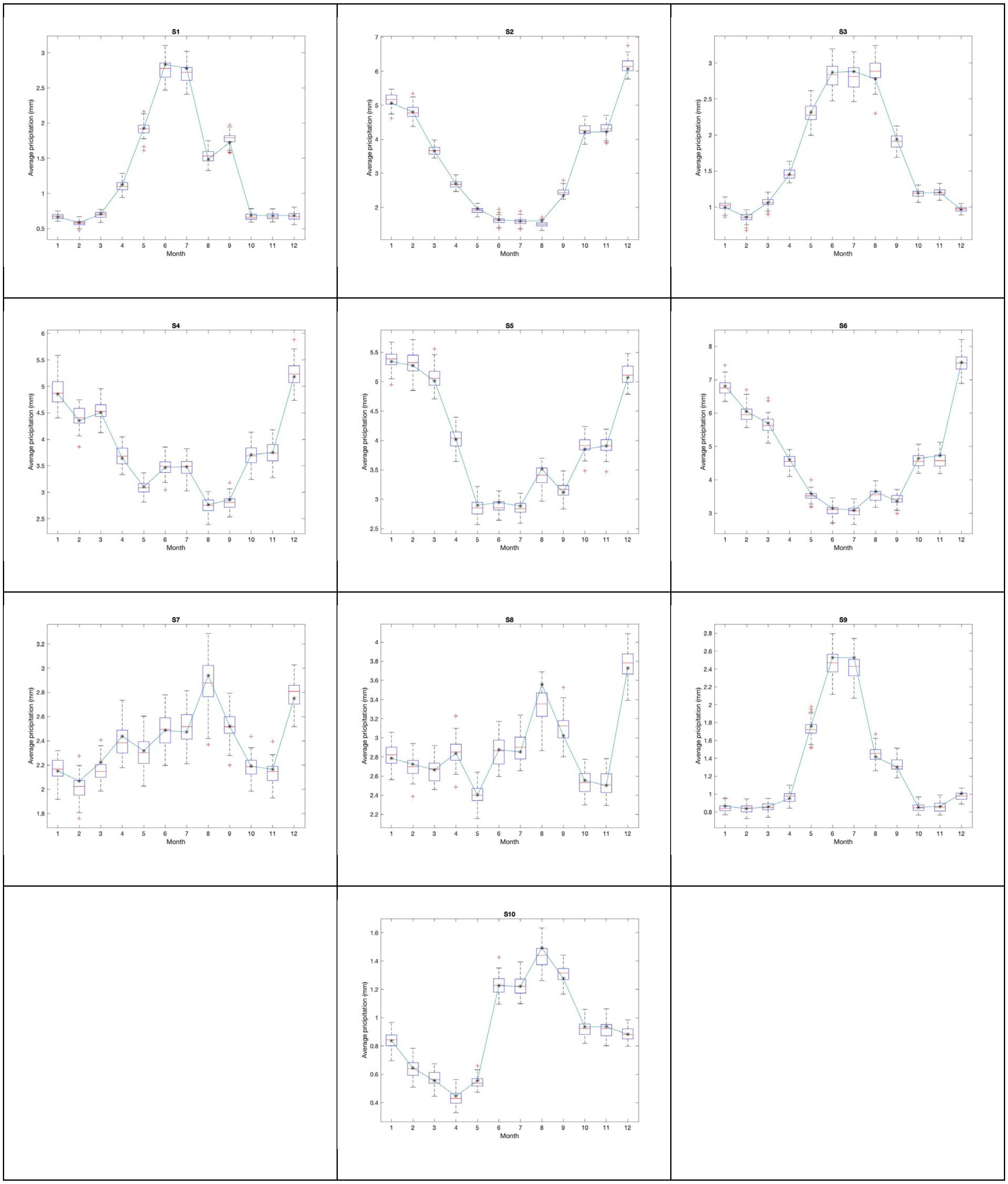


Figure B-2. Boxplots of monthly mean of precipitation

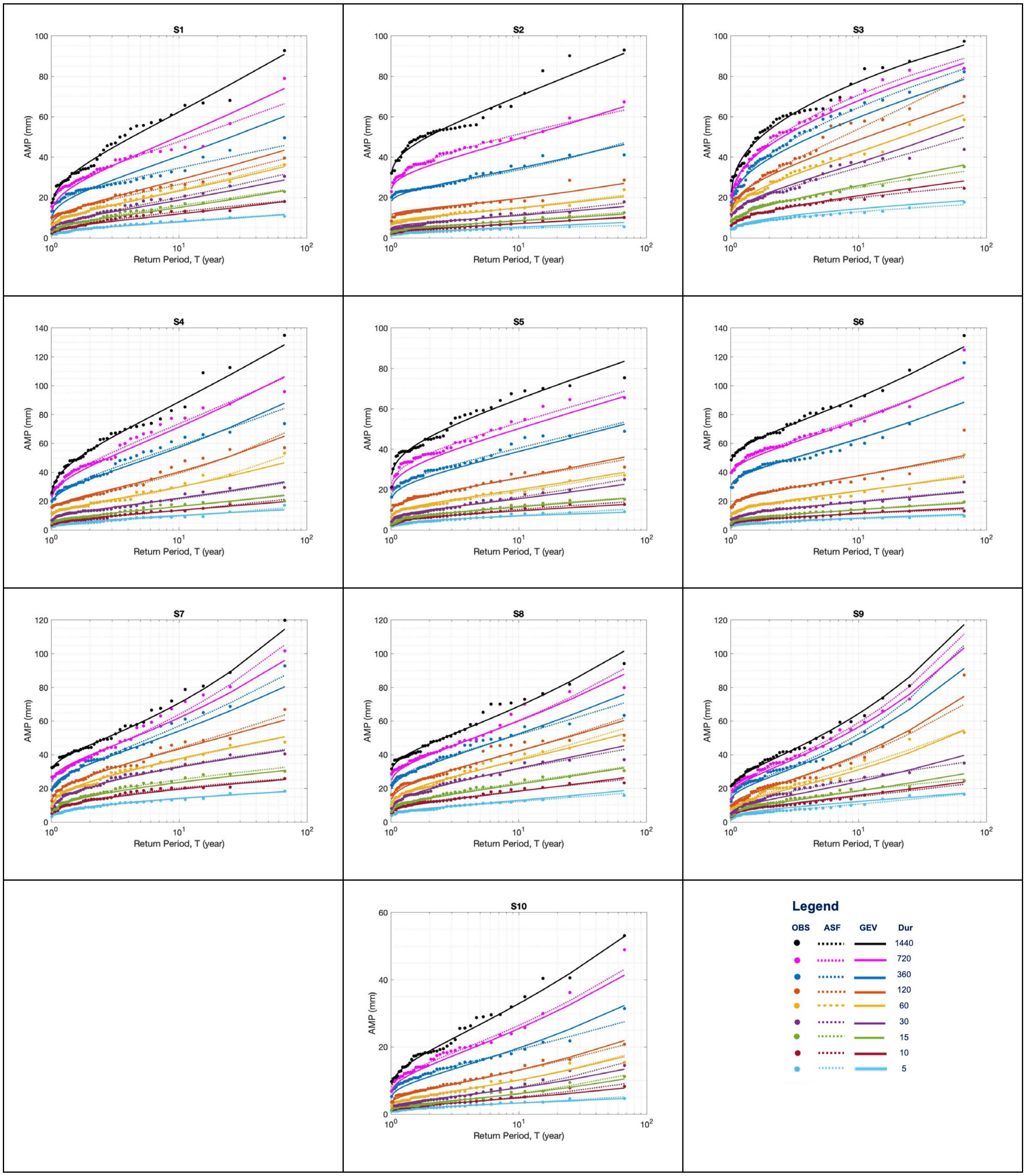


Figure B-3. Probability Plot

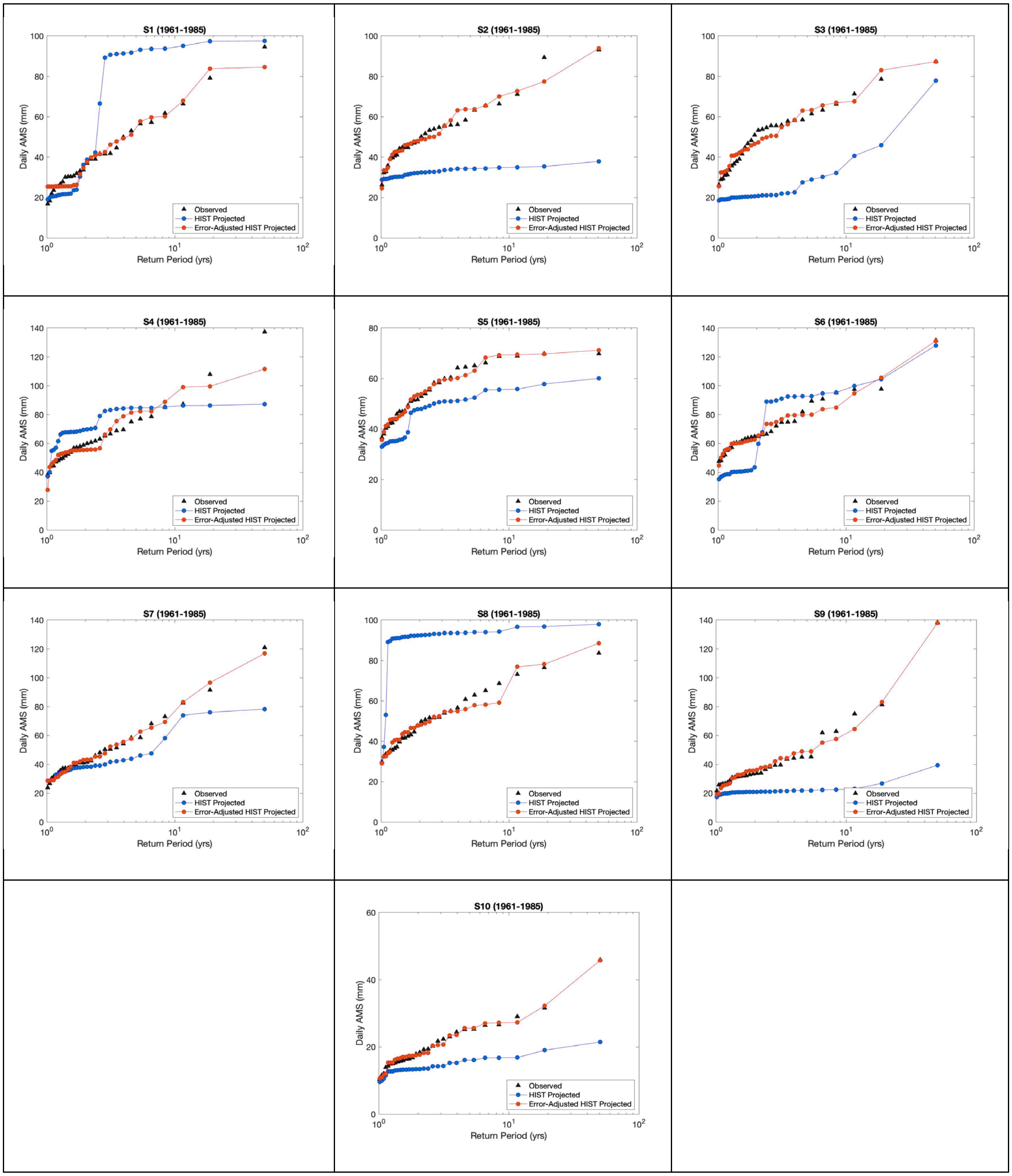


Figure B-4. Bias correction - Calibration period

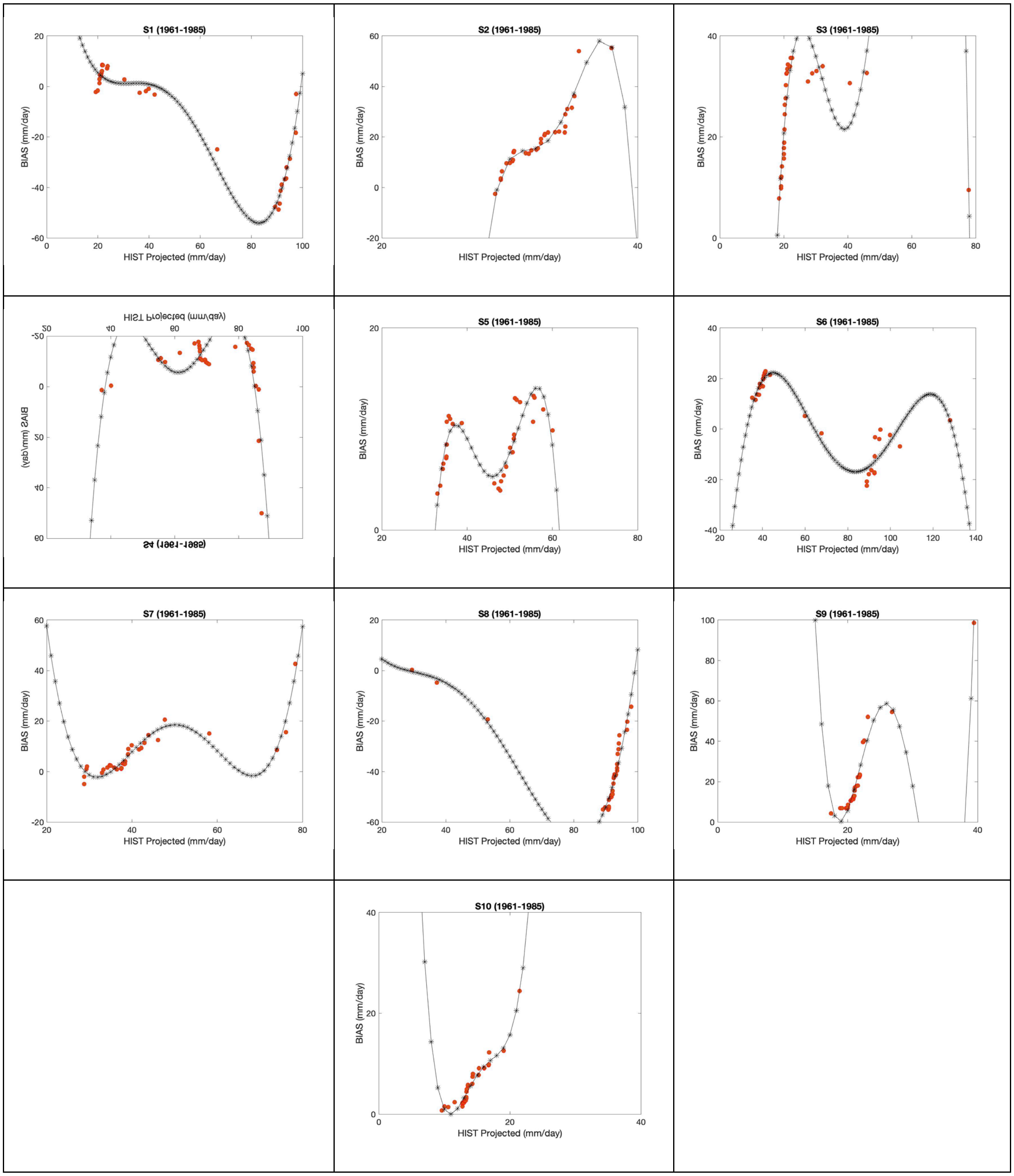


Figure B-5. Bias correction Functions

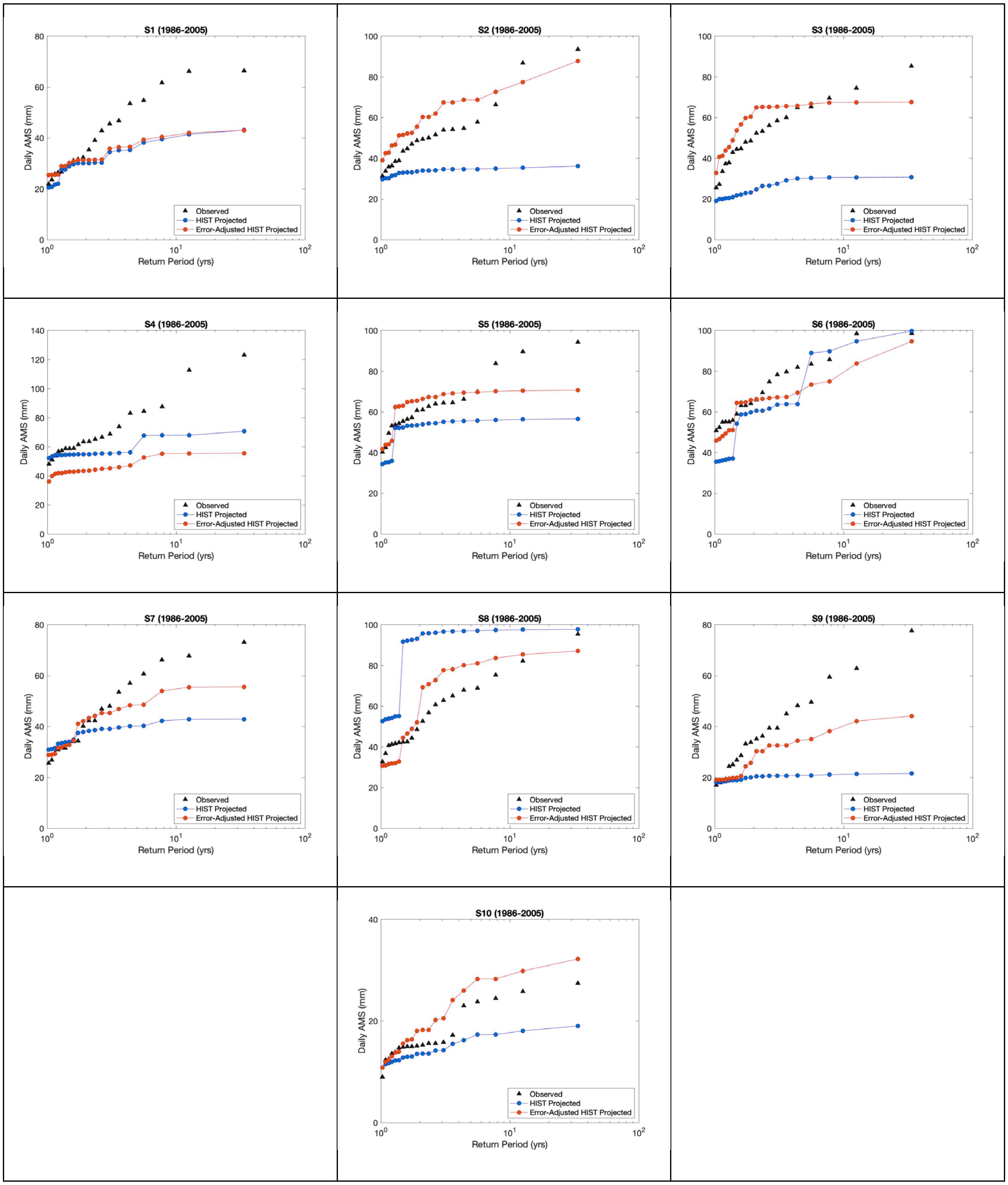


Figure B-6. Bias correction - Validation period

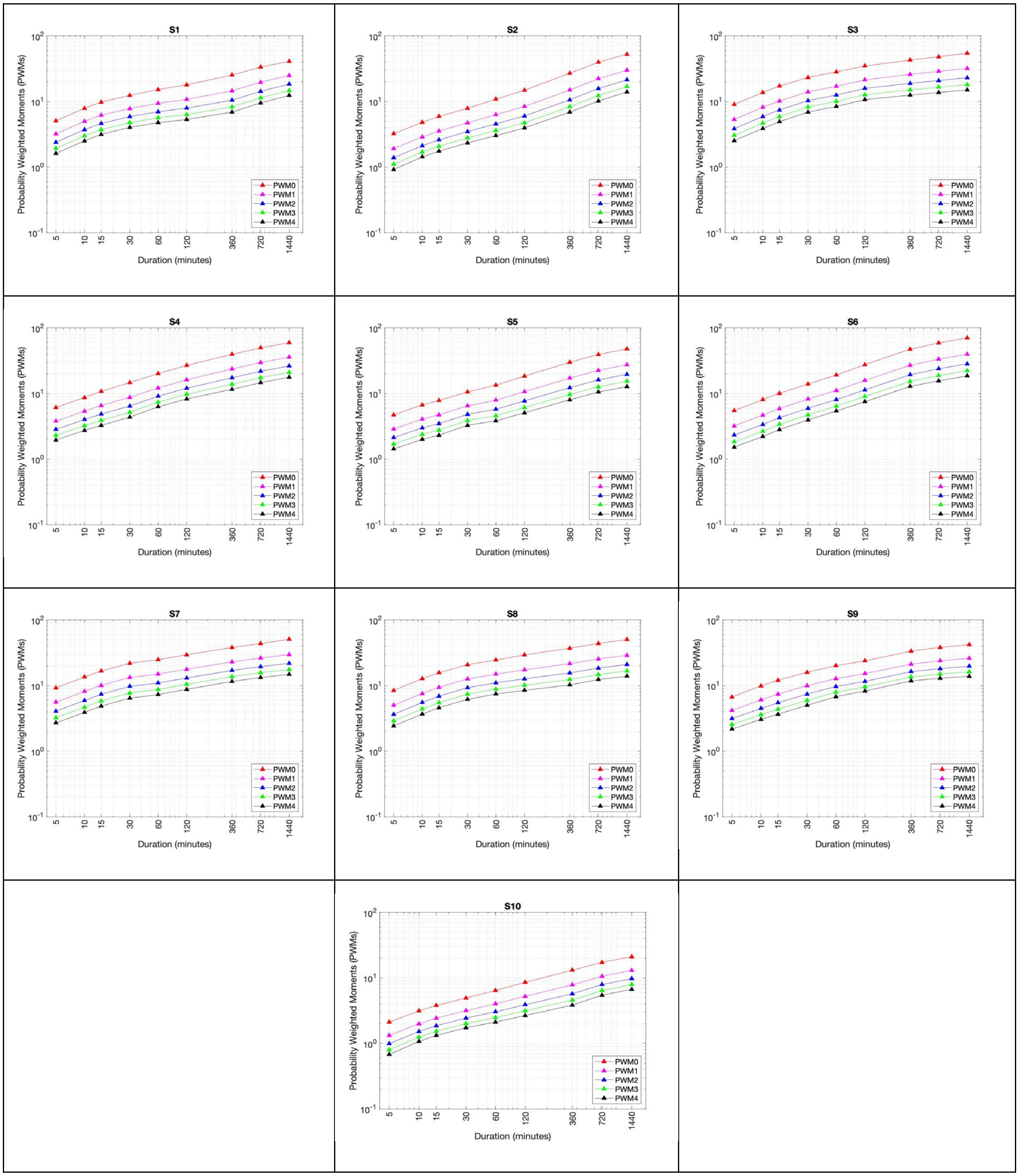


Figure B-7. Log-log plots of the PWMs versus durations

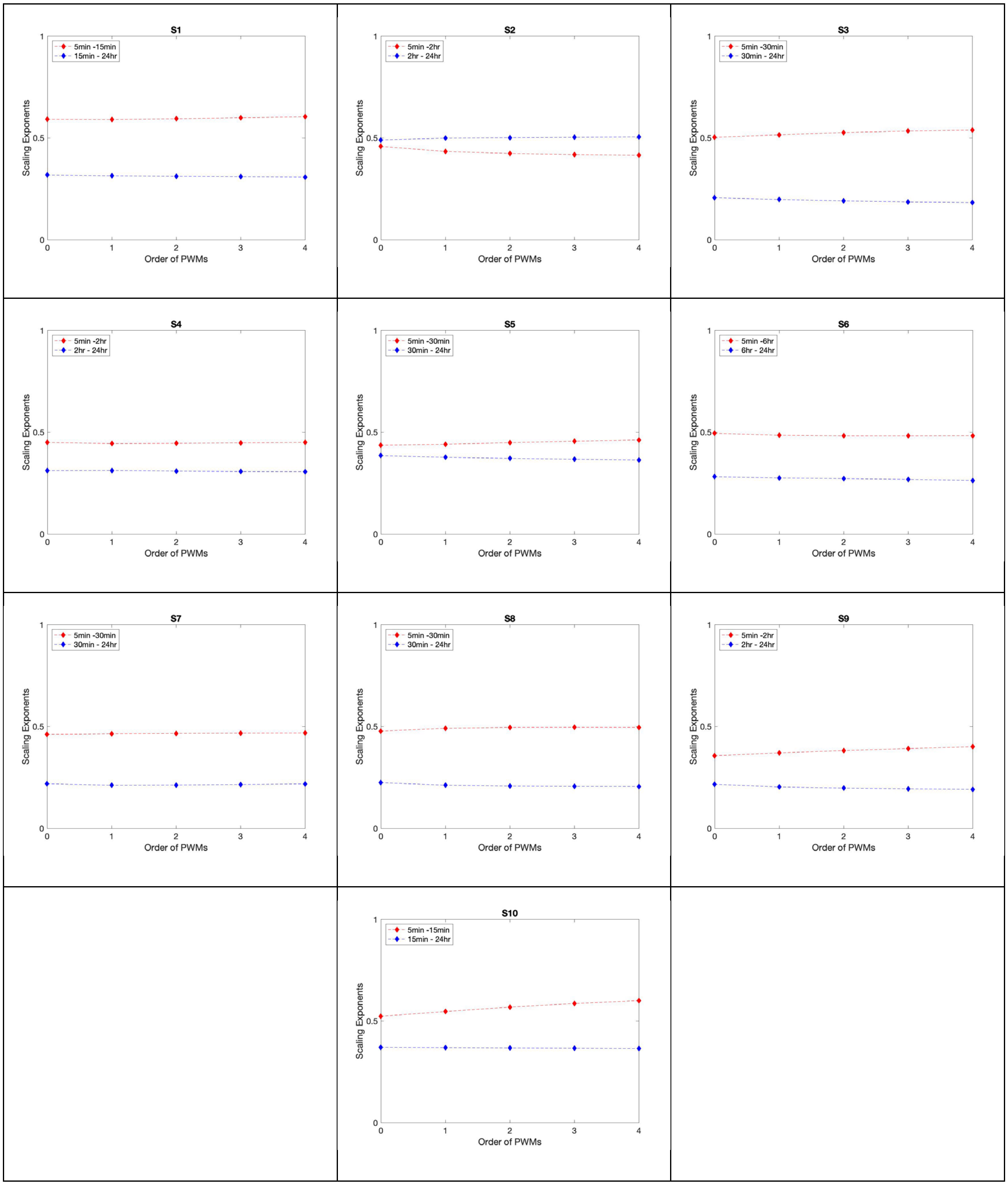


Figure B-8. Scaling exponents plotted against the order of PWMs

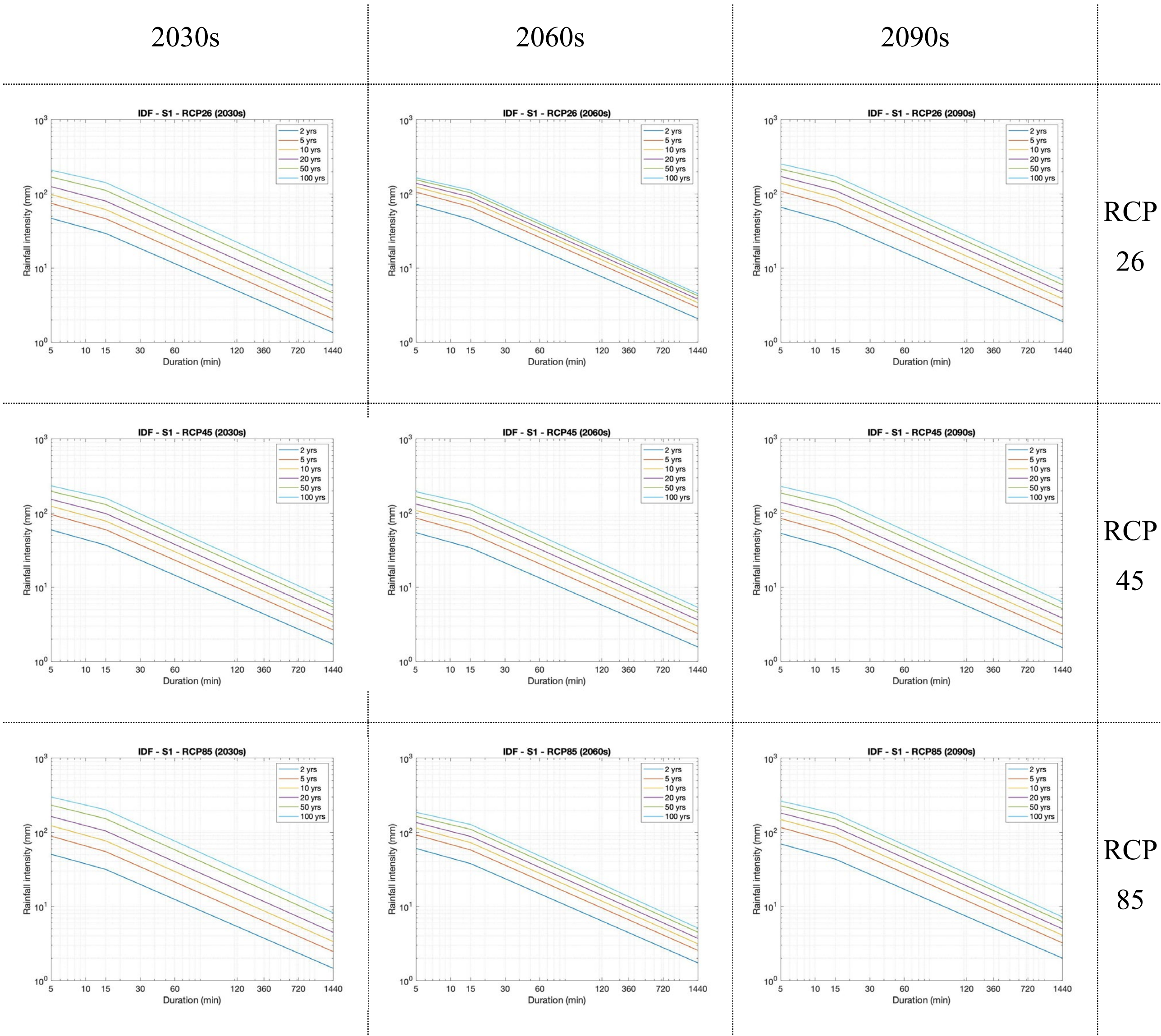


Figure B-9. IDF curves for future periods with different RCPs at S1

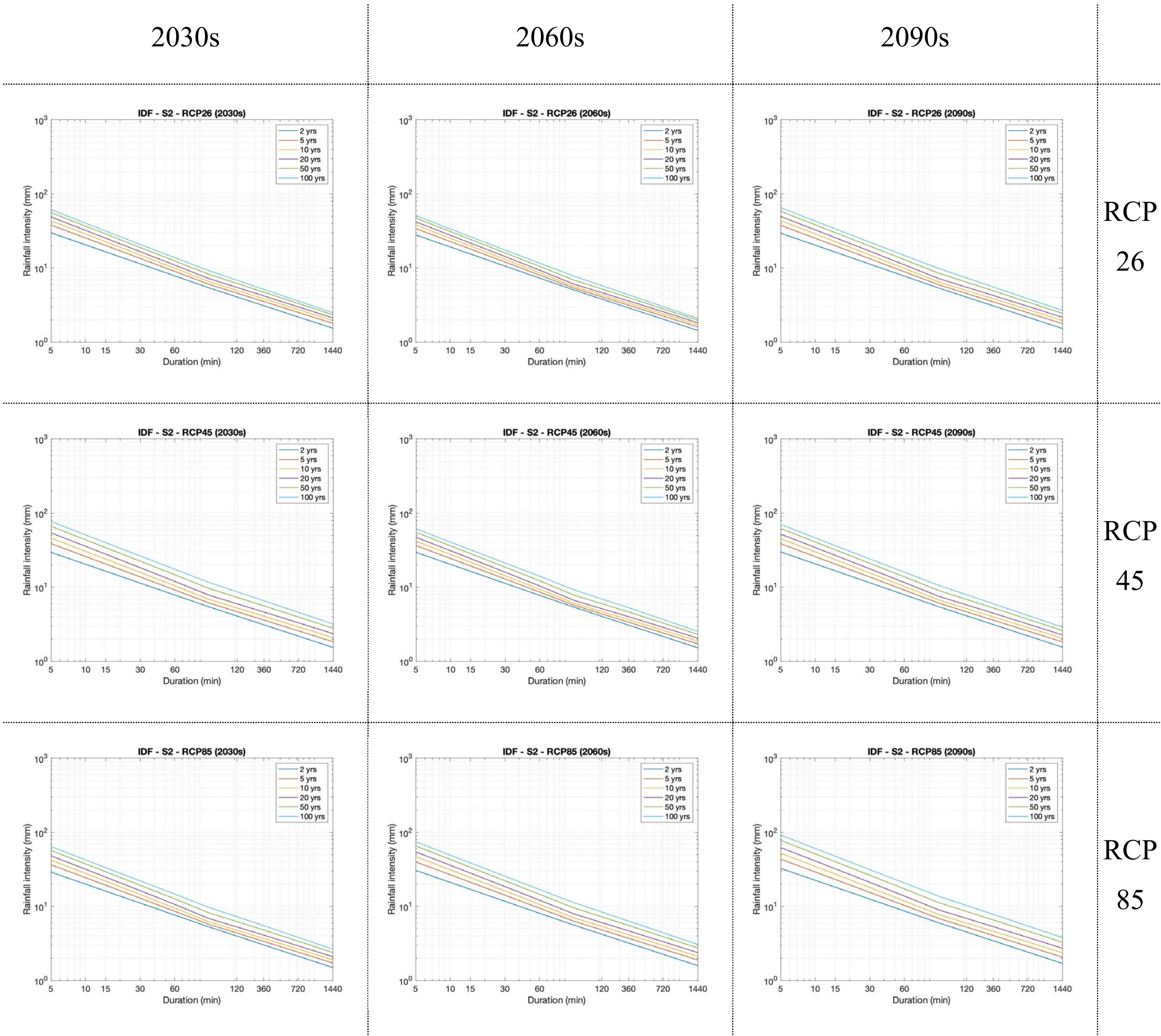


Figure B-10. IDF curves for future periods with different RCPs at S2

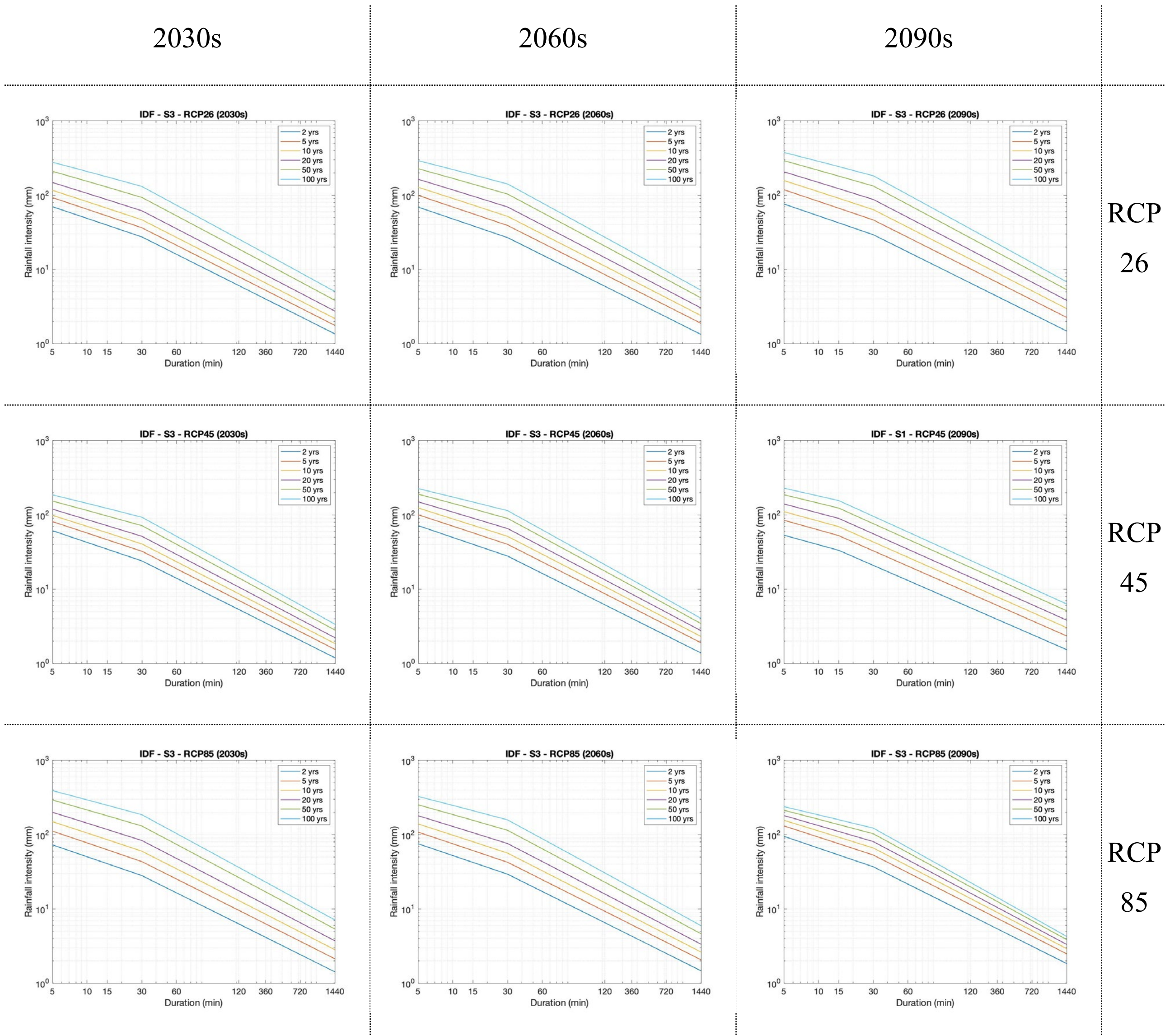
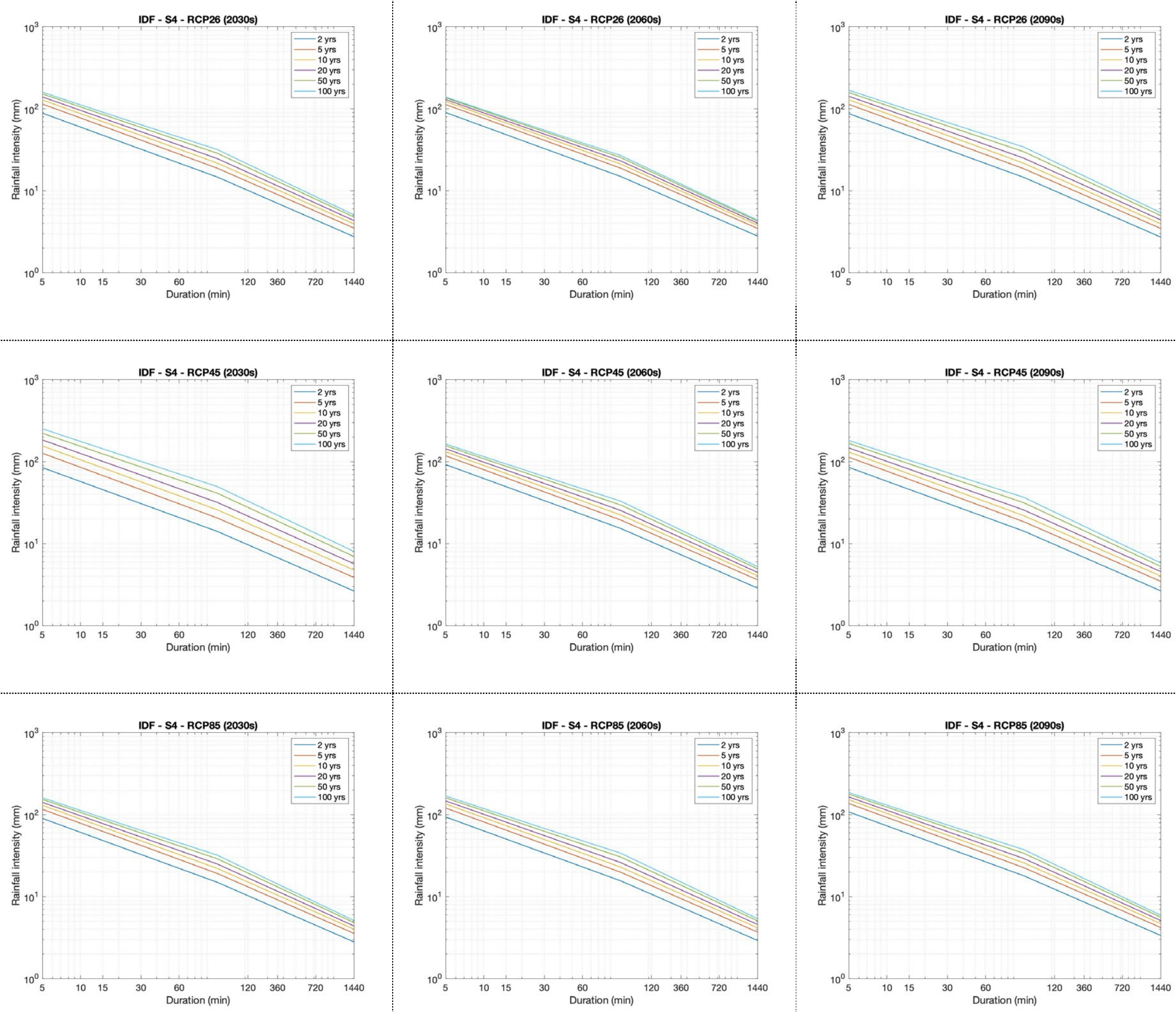


Figure B-11. IDF curves for future periods with different RCPs at S3

2030s	2060s	2090s	



RCP 26

RCP 45

RCP 85

Figure B-12. IDF curves for future periods with different RCPs at S4

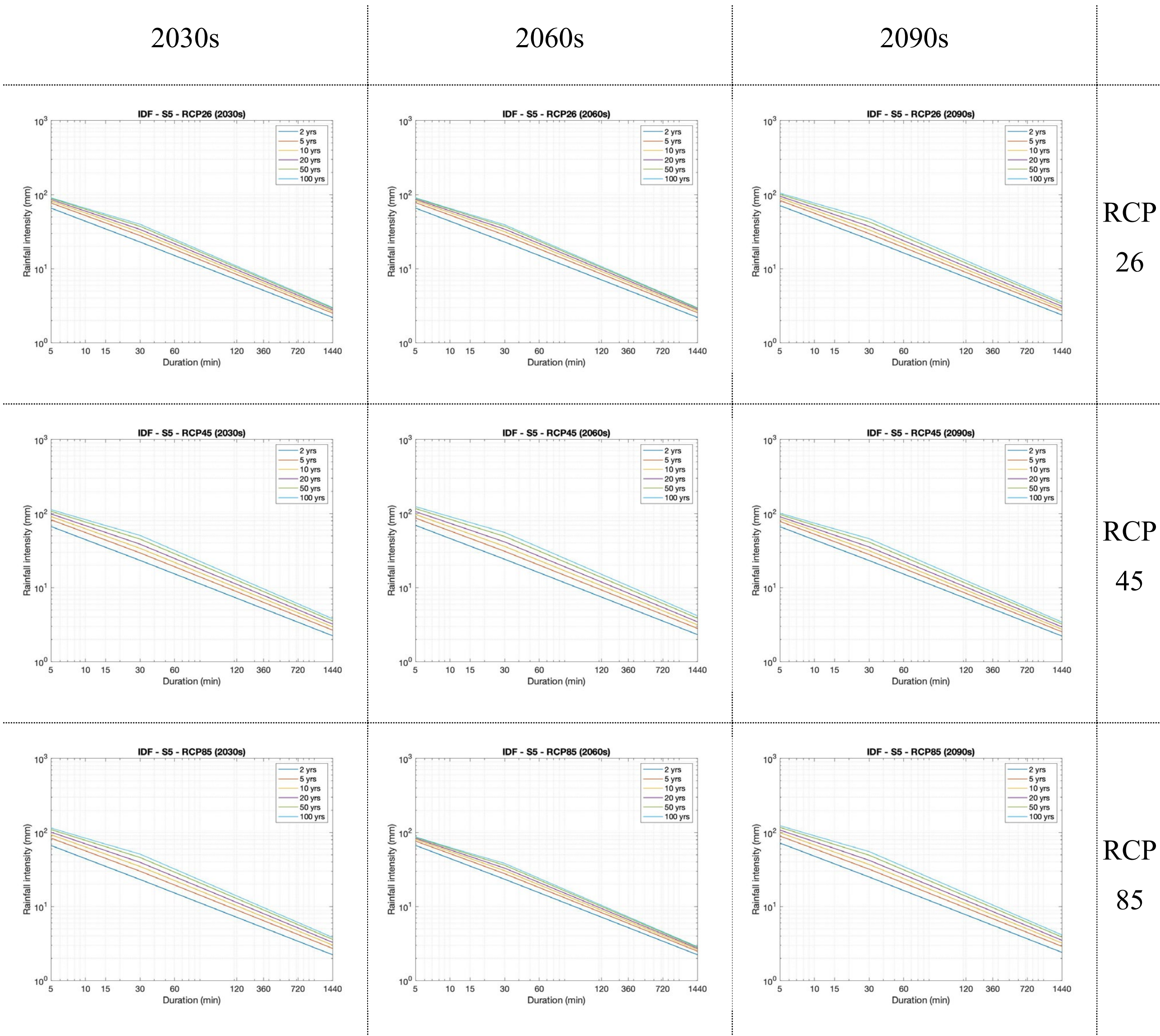


Figure B-13. IDF curves for future periods with different RCPs at S5

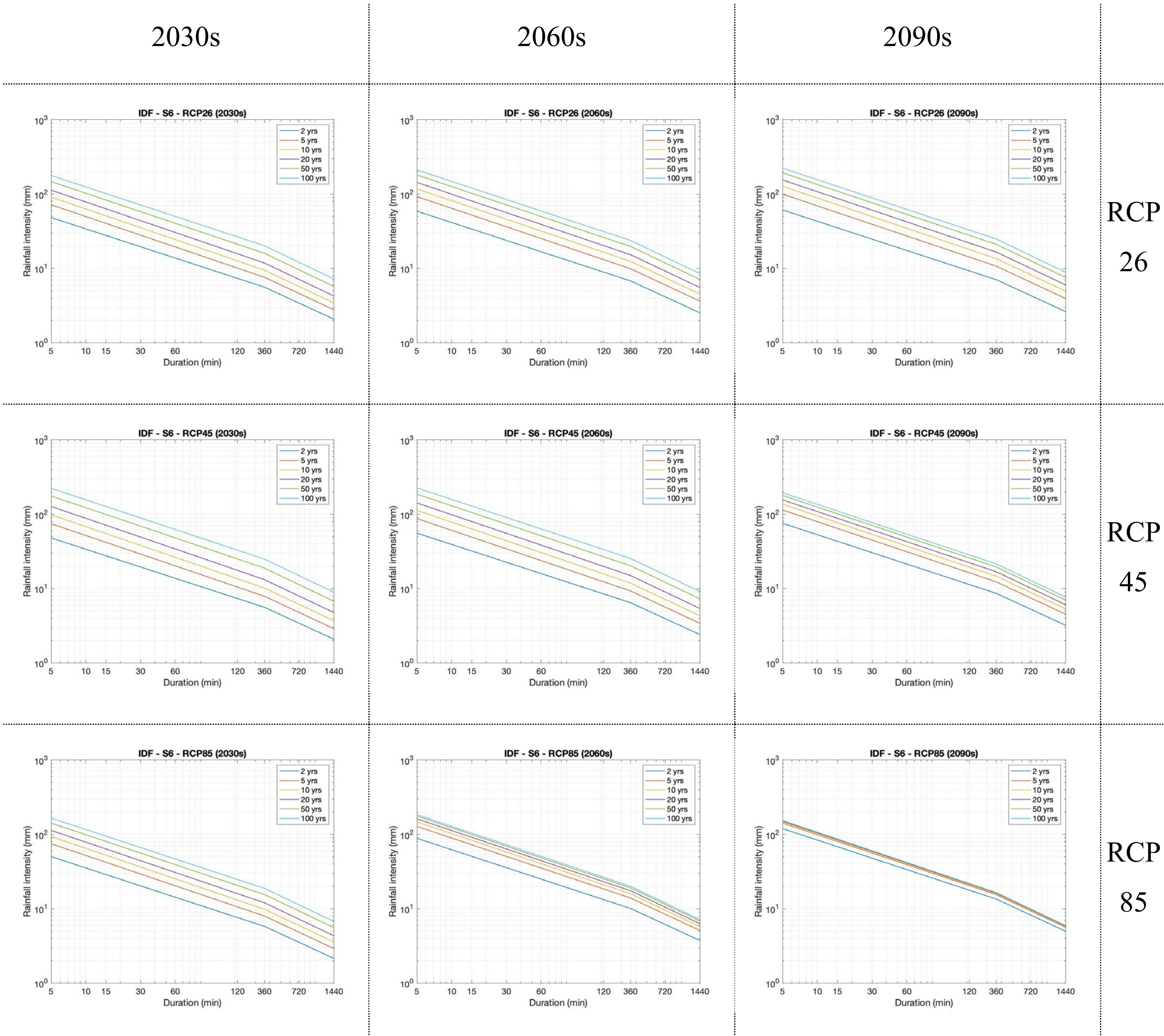


Figure B-14. IDF curves for future periods with different RCPs at S6

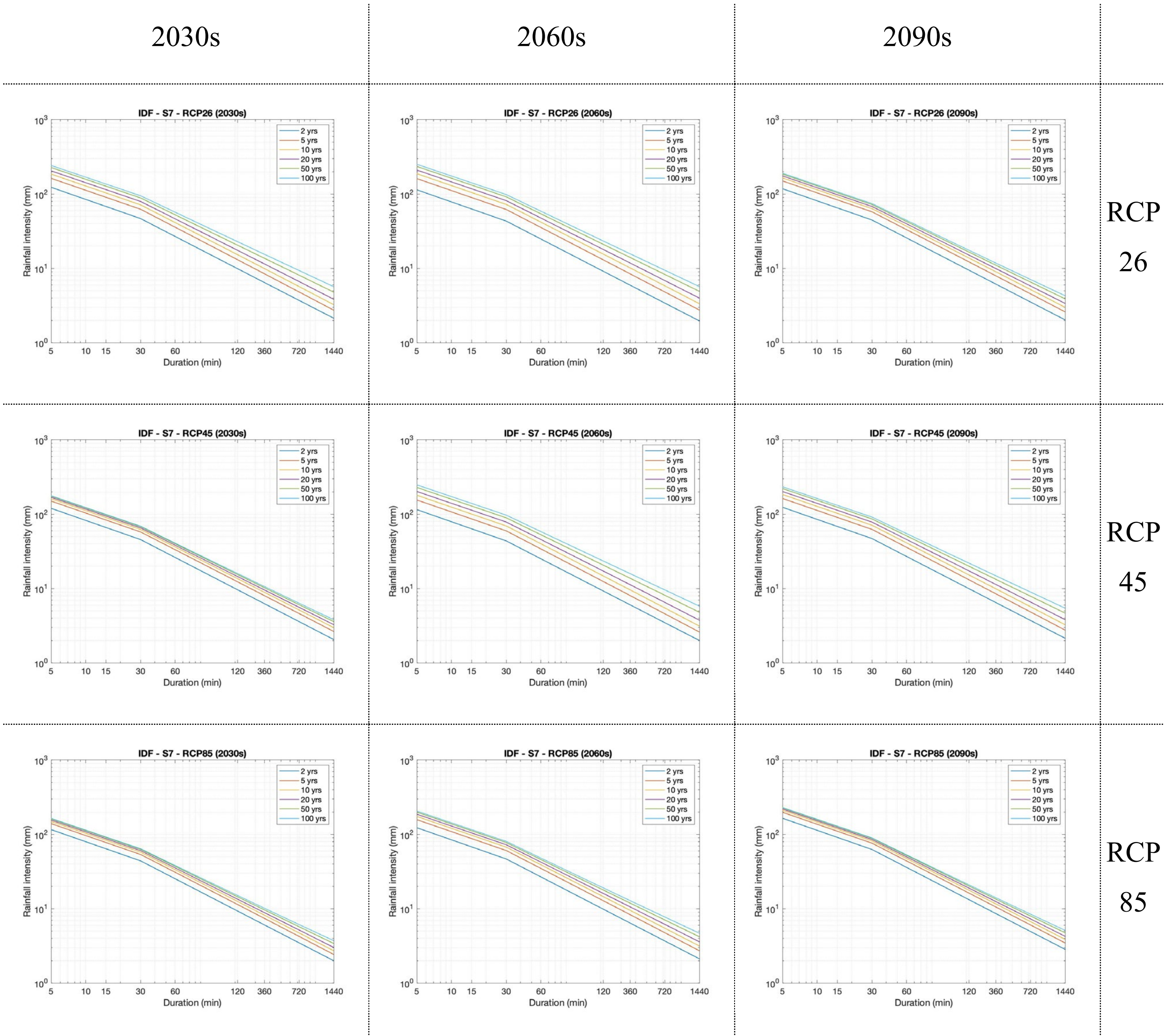


Figure B-15. IDF curves for future periods with different RCPs at S7

	2030s	2060s	2090s	
				RCP 26
				RCP 45
				RCP 85

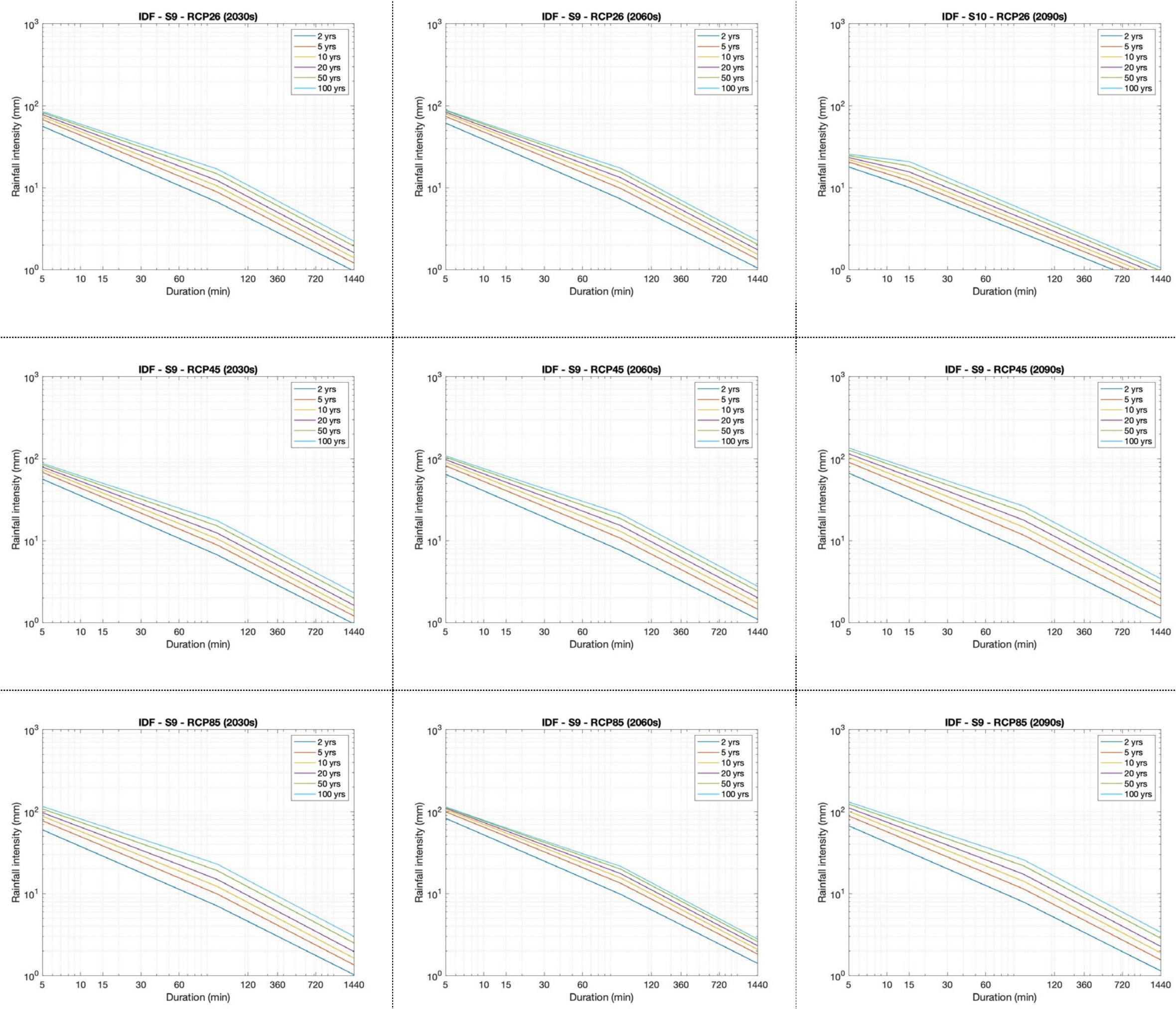


Figure B-16. IDF curves for future periods with different RCPs at S9

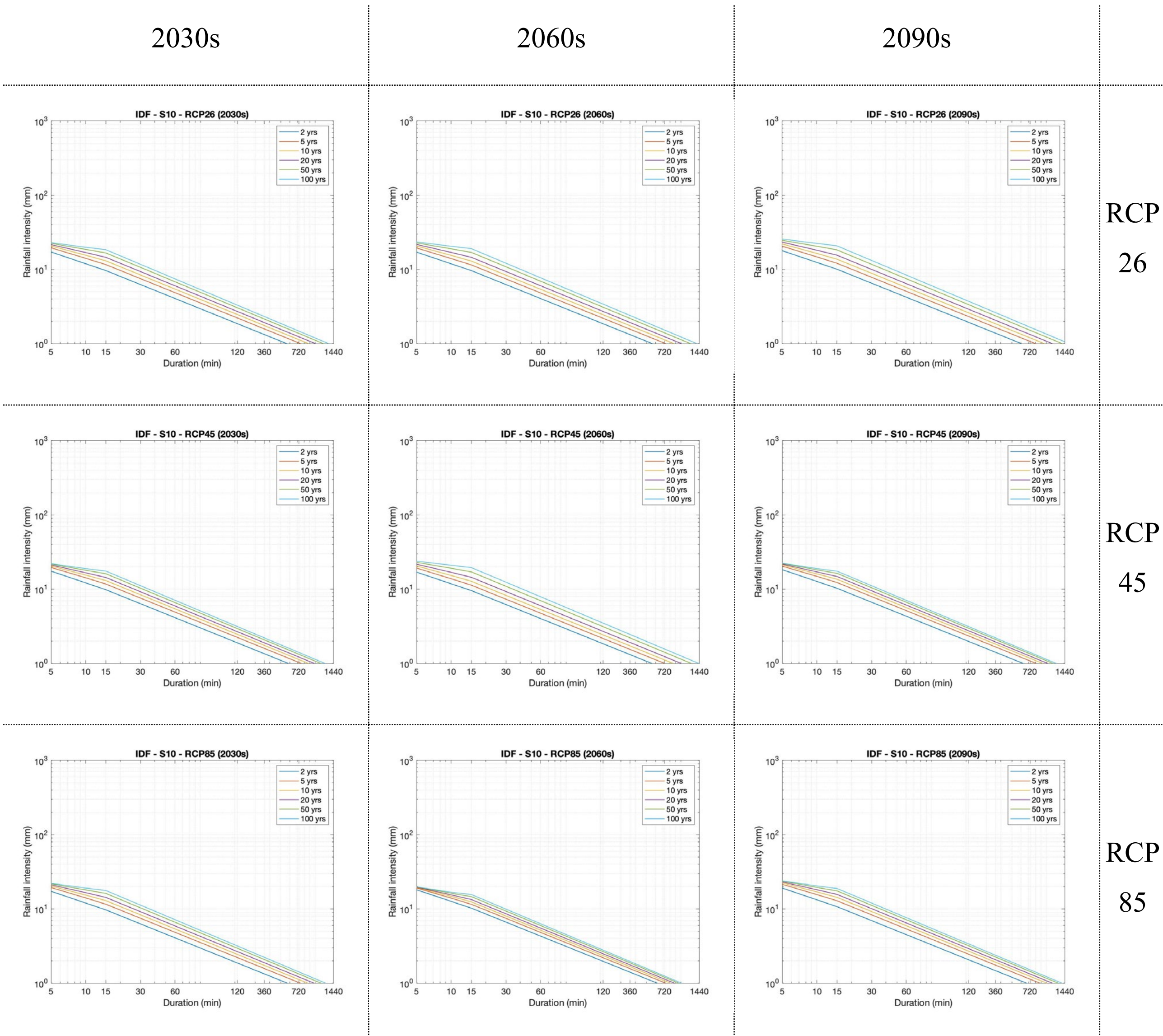


Figure B-17. IDF curves for future periods with different RCPs at S10

Appendix C: Supplementary materials for chapter 4

Table C-1. Information of selected stations in Vietnam

No.	Code	Station name	Province	Lon	Lat	RL (years)
1	002	DienBien	Lai Châu	103	21.22	32
2	003	LaiChau	Lai Châu	103.09	22.04	32
3	005	MuongTe	Lai Châu	102.5	22.22	32
4	006	PhaDin	Lai Châu	103.31	21.34	32
5	008	SinHo	Lai Châu	103.14	22.22	32
6	009	TamDuong	Lai Châu	103.29	22.25	32
7	011	TuanGiao	Lai Châu	103.25	21.35	32
8	012	BacYen	Sơn La	104.25	21.15	32
9	013	CoNoi	Sơn La	104.09	21.08	32
10	014	MocChau	Sơn La	104.41	20.5	32
11	016	PhuYen	Sơn La	104.38	21.16	32
12	017	QuynhNhai	Sơn La	103.34	21.51	32
13	018	SonLa	Sơn La	103.54	21.2	32
14	019	SongMa	Sơn La	103.44	21.04	32
15	023	YenChau	Sơn La	104.18	21.03	32
16	024	ChiNe	Hoà Bình	105.47	20.29	32
17	026	HoaBinh	Hoà Bình	105.2	20.49	32
18	027	KimBoi	Hoà Bình	105.32	20.4	32

No.	Code	Station name	Province	Lon	Lat	RL (years)
19	028	LacSon	Hoà Bình	105.27	20.27	32
20	029	MaiChau	Hoà Bình	105.03	20.39	32
21	030	BacMe	Hà Giang	105.22	22.44	32
22	031	BacQuang	Hà Giang	104.52	22.3	32
23	032	HoangSuPhi	Hà Giang	104.41	22.45	32
24	033	HaGiang	Hà Giang	104.58	22.49	32
25	036	BacHa	Lào Cai	104.17	22.32	32
26	040	PhoRang	Lào Cai	104.28	22.14	32
27	041	SaPa	Lào Cai	103.49	22.21	32
28	045	LucYen	Yên Bái	104.43	22.06	32
29	046	MuCangChai	Yên Bái	104.03	21.52	32
30	048	VanChan	Yên Bái	104.31	21.35	32
31	049	YenBai	Yên Bái	104.52	21.42	32
32	050	ChiemHoa	Tuyên Quang	105.16	22.09	32
33	051	HamYen	Tuyên Quang	105.02	22.04	32
34	053	TuyenQuang	Tuyên Quang	105.13	21.49	32
35	054	BacKan	Bắc Cạn	105.5	22.09	32
36	055	ChoRa	Bắc Cạn	105.43	22.27	32
37	056	NganSon	Bắc Cạn	105.59	22.26	32
38	059	DinhHoa	Thái Nguyên	105.38	21.55	32
39	061	ThaiNguyen	Thái Nguyên	105.5	21.36	32
40	063	MinhDai	Phú Thọ	105.03	21.1	32

No.	Code	Station name	Province	Lon	Lat	RL (years)
41	064	PhuHo	Phú Thọ	105.14	21.27	32
42	067	VietTri	Phú Thọ	105.25	21.18	32
43	069	TamDao	Vĩnh Phúc	105.39	21.28	32
44	071	VinhYen	Vĩnh Phúc	105.36	21.19	32
45	072	BaoLac	Cao Bằng	105.4	22.57	32
46	073	CaoBang	Cao Bằng	106.15	22.4	31
47	075	NguyenBinh	Cao Bằng	105.57	22.39	32
48	077	TrungKhanh	Cao Bằng	106.31	22.5	32
49	078	BacSon	Lạng Sơn	106.19	21.54	32
50	079	DinhLap	Lạng Sơn	107.06	21.32	32
51	080	HuuLung	Lạng Sơn	106.21	21.3	32
52	081	LangSon	Lạng Sơn	106.46	21.5	32
53	084	ThatKhe	Lạng Sơn	106.28	22.15	32
54	085	BacGiang	Bắc Giang	106.13	22.18	32
55	087	HiepHoa	Bắc Giang	105.58	21.21	32
56	088	LucNgan	Bắc Giang	106.33	21.23	32
57	089	SonDong	Bắc Giang	106.51	21.2	32
58	092	BaiChay	Quảng Ninh	107.04	20.58	32
59	093	CoTo	Quảng Ninh	107.46	20.59	32
60	094	CuaOng	Quảng Ninh	107.21	21.01	32
61	097	QuangHa	Quảng Ninh	107.45	21.27	28
62	098	TienYen	Quảng Ninh	107.24	21.2	32

No.	Code	Station name	Province	Lon	Lat	RL (years)
63	099	UongBi	Quảng Ninh	106.45	21.02	32
64	100	BachLongVy	Hải Phòng	107.43	20.08	32
65	106	HonDau	Hải Phòng	106.48	20.4	32
66	107	PhuLien	Hải Phòng	106.38	20.48	32
67	109	BaVi	Hà Tây	105.25	21.09	32
68	111	HaDong	Hà Tây	105.45	20.58	32
69	113	SonTay	Hà Tây	105.3	21.08	32
70	119	Lang	Hà Nội	105.51	21.02	32
71	121	ChiLinh	Hải Dương	106.23	21.05	31
72	122	HaiDuong	Hải Dương	106.18	20.56	32
73	123	HungYen	Hưng Yên	106.03	20.39	32
74	127	HaNam	Hà Nam	105.55	20.33	32
75	129	NamDinh	Nam Định	106.09	20.24	32
76	130	VanLy	Nam Định	106.18	20.07	32
77	131	ThaiBinh	Thái Bình	106.21	20.27	32
78	138	NhoQuan	Ninh Bình	105.44	20.2	32
79	139	NinhBinh	Ninh Bình	105.58	20.14	32
80	140	BaiThuong	Thanh Hóa	105.23	19.54	32
81	142	HoiXuan	Thanh Hóa	105.07	20.22	32
82	146	NhuXuan	Thanh Hóa	105.34	19.38	32
83	150	ThanhHoa	Thanh Hóa	105.47	19.45	32
84	151	TinhGia	Thanh Hóa	105.47	19.27	32

No.	Code	Station name	Province	Lon	Lat	RL (years)
85	153	YenDinh	Thanh Hóa	105.4	19.59	32
86	155	ConCuong	Nghệ An	104.53	19.03	32
87	156	DoLuong	Nghệ An	105.18	18.54	32
88	158	HonNgu	Nghệ An	105.46	18.48	32
89	160	QuyChau	Nghệ An	105.07	19.34	32
90	161	QuyHop	Nghệ An	105.09	19.19	32
91	162	QuynhLuu	Nghệ An	105.38	19.1	32
92	163	TayHieu	Nghệ An	105.24	19.19	32
93	164	TuongDuong	Nghệ An	104.26	19.17	32
94	165	Vinh	Nghệ An	105.4	18.4	32
95	166	HaTinh	Hà Tĩnh	105.54	18.21	32
96	167	HuongKhe	Hà Tĩnh	105.43	18.11	32
97	168	HuongSon	Hà Tĩnh	105.26	18.31	32
98	169	KyAnh	Hà Tĩnh	106.17	18.05	32
99	170	BaDon	Quảng Bình	106.25	17.45	32
100	172	DongHoi	Quảng Bình	106.37	17.29	32
101	175	TuyenHoa	Quảng Bình	106.01	17.53	32
102	176	ConCo	Quảng Trị	107.2	17.1	32
103	177	DongHa	Quảng Trị	107.05	16.51	32
104	178	KheSanh	Quảng Trị	107.44	16.38	32
105	180	ALuoi	Thừa Thiên - Huế	107.17	16.13	31
106	181	Hue	Thừa Thiên - Huế	107.35	16.26	31

No.	Code	Station name	Province	Lon	Lat	RL (years)
107	182	NamDong	Thừa Thiên - Huế	107.43	16.1	32
108	184	DaNang	Đà Nẵng	108.12	16.02	31
109	186	TamKy	Quảng Nam	108.28	15.34	28
110	187	TraMy	Quảng Nam	108.15	15.2	29
111	188	BaTo	Quảng Ngãi	108.44	14.46	27
112	189	LySon	Quảng Ngãi	109.09	15.23	22
113	190	QuangNgai	Quảng Ngãi	108.48	15.07	31
114	191	HoaiNhon	Bình Định	109.02	14.31	29
115	192	QuyNhon	Bình Định	109.13	13.46	32
116	194	SonHoa	Phú Yên	108.59	13.03	30
117	195	TuyHoa	Phú Yên	109.17	13.05	31
118	196	CamRanh	Khánh Hòa	109.09	11.55	29
119	197	NhaTrang	Khánh Hòa	109.12	12.13	31
120	199	TruongSa	Khánh Hòa	111.55	8.39	30
121	201	PhanRang	Ninh Thuận	108.59	11.35	28
122	202	HamTan	Bình Thuận	107.46	10.41	29
123	203	PhanThiet	Bình Thuận	107.06	11.56	29
124	204	PhuQuy	Bình Thuận	107.56	11.31	28
125	205	DacTo	Kon Tum	107.5	14.39	26
126	206	KonTum	Kon Tum	108	14.3	31
127	207	AnKhe	Gia Lai	108.39	13.57	30
128	208	AyunPa	Gia Lai	108.27	13.23	28

No.	Code	Station name	Province	Lon	Lat	RL (years)
129	209	Pleiku	Gia Lai	108.01	13.58	31
130	211	BuonMeThuot	Đắc Lắc	108.03	12.4	30
131	212	BuonHo	Đắc Lắc	108.16	12.55	29
132	213	DacNong	Đắc Lắc	107.41	12	29
133	216	Eakmat	Đắc Lắc	108.08	12.41	27
134	218	MDRac	Đắc Lắc	108.46	12.44	29
135	219	BaoLoc	Lâm Đồng	107.49	11.32	28
136	220	DaLat	Lâm Đồng	108.27	11.57	28
137	221	LienKhuong	Lâm Đồng	108.23	11.45	26
138	224	TriAn	Đồng Nai	107.04	11.05	29
139	226	DongPhu	Bình Phước	106.54	11.32	28
140	228	PhuocLong	Bình Phước	106.59	11.5	29
141	230	TayNinh	Tây Ninh	106.07	11.2	28
142	231	ConDao	Bà Rịa-V.Tàu	106.36	8.41	29
143	233	VungTau	Bà Rịa-V.Tàu	107.05	10.52	28
144	235	MocHoa	Long An	105.56	10.47	28
145	237	MyTho	Tiền Giang	106.24	10.21	27
146	238	CaoLanh	Đồng Tháp	105.38	10.28	28
147	239	BaTri	Bến Tre	106.36	10.03	28
148	242	CangLong	Trà Vinh	106.12	9.59	29
149	243	ChauDoc	An Giang	105.08	10.42	28
150	244	CanTho	Cần Thơ	105.46	10.02	29

No.	Code	Station name	Province	Lon	Lat	RL (years)
151	245	SocTrang	Sóc Trăng	105.58	9.36	29
152	246	PhuQuoc	Kiên Giang	104.08	10.13	28
153	247	RachGia	Kiên Giang	105.04	10	28
154	249	BacLieu	Bạc Liêu	105.43	9.17	27
155	250	CaMau	Ca Mau	105.09	9.11	28

Appendix D: Supplementary materials for chapter 5

Table D-1. Information of selected stations across Canada

No.	Province name	Climate ID	Station name	Lat	Lon	Elevation (m)	Record period	RL (year)
1	NL	8401705	GANDER_AIRPORT_CS	48.95	-54.57	151	1939-2017	70
2	NL	8501900	GOOSE_A	53.32	-60.42	48	1961-2016	53
3	NL	8403820	STEPHENVILLE_RCS	48.57	-58.57	58	1967-2017	48
4	NL	8401501	DEER_LAKE_A	49.22	-57.40	21	1966-2002	36
5	NL	8403506	ST_JOHN_S_A	47.62	-52.73	140	1949-1996	35
6	NL	8403619	ST_LAWRENCE	46.92	-55.38	48	1969-2013	35
7	NL	8400801	BURGEO_NL	47.62	-57.62	10	1967-2013	34
8	PE	8300301	CHARLOTTETOWN_A	46.28	-63.12	48	1967-2016	31
9	PE	8300596	SUMMERSIDE	46.43	-63.83	12	1964-2013	37
10	NS	8205092	SHEARWATER_RCS	44.63	-63.52	24	1955-2016	59
11	NS	8205702	SYDNEY_CS	46.17	-60.03	62	1961-2016	53
12	NS	8202000	GREENWOOD_A	44.98	-64.92	28	1964-2016	44
13	NS	8206495	YARMOUTH_A	43.83	-66.08	42	1971-2016	43
14	NS	8202810	KENTVILLE_CDA_CS	45.07	-64.48	48	1960-2013	37
15	NB	8103201	MONCTON_INTL_A	46.12	-64.68	70	1946-2016	67
16	NB	8100885	CHARLO_AUTO	47.98	-66.33	42	1959-2013	51
17	NB	8101605	FREDERICTON_CDA_CS	45.92	-66.62	35	1959-2015	47
18	NB	8104900	SAINT_JOHN_A	45.32	-65.88	108	1958-2002	40
19	NB	8100989	MIRAMICHI_RCS	47.02	-65.47	33	1964-2015	36
20	QC	702S006	MONTREAL_PIERRE_ELLIOTT _TRUDEAU_INTL	45.47	-73.73	32	1943-2014	61
21	QC	701S001	QUEBEC_JEAN_LESAGE_INTL	46.80	-71.38	60	1961-2015	46
22	QC	7014160	L_ASSOMPTION	45.82	-73.43	21	1963-2017	45
23	QC	7024280	LENNOXVILLE	45.37	-71.82	181	1960-2017	45
24	QC	7060400	BAGOTVILLE_A	48.33	-71.00	159	1961-2017	45
25	QC	7018001	SHAWINIGAN	46.57	-72.73	110	1968-2017	41
26	QC	7055121	MONT_JOLI_A	48.62	-68.22	52	1968-2015	38
27	QC	7047914	SEPT-ILES	50.22	-66.25	52	1969-2014	35
28	QC	7098600	VAL-D_OR_A	48.07	-77.78	337	1961-1995	34

No.	Province name	Climate ID	Station name	Lat	Lon	Elevation (m)	Record period	RL (year)
29	QC	7042395	FORET_MONTMORENCY_RCS	47.32	-71.15	672	1967-2014	33
30	QC	7025280	MONTREAL_MCGILL	45.50	-73.58	56	1906-1992	32
31	QC	7028124	SHERBROOKE_A	45.43	-71.68	241	1962-1994	32
32	QC	7113534	KUUJJUAQ_A	58.10	-68.42	39	1970-2013	32
33	QC	7025745	ORMSTOWN	45.12	-74.05	45	1963-1998	31
34	QC	7028441	THETFORD_MINES	46.10	-71.35	381	1967-1999	31
35	QC	7057287	STE_GERMAINE	46.42	-70.47	510	1966-1999	31
36	QC	7066685	ROBERVAL_A	48.52	-72.27	178	1970-2012	31
37	QC	7103536	KUUJJUARAPIK_A	55.28	-77.75	12	1969-2013	31
38	QC	7020305	ARTHABASKA	46.02	-71.95	140	1963-1998	30
39	QC	7022720	GEORGEVILLE	45.13	-72.23	266	1968-1998	30
40	QC	7024320	LINGWICK	45.63	-71.37	266	1968-1999	30
41	QC	7027200	ST_EPHREM	46.07	-70.97	312	1966-1999	30
42	QC	7027802	SAWYERVILLE_NORD	45.37	-71.53	345	1966-1999	30
43	QC	7028676	VALLEE_JONCTION	46.38	-70.93	152	1966-1999	30
44	QC	7054095	LA_POCATIERE_CDA	47.35	-70.03	30	1962-1995	30
45	ON	6137362	ST_THOMAS_WPCP	42.77	-81.22	209	1926-2007	75
46	ON	6158355	TORONTO_CITY	43.67	-79.40	112	1940-2017	67
47	ON	6144478	LONDON_CS	43.03	-81.15	278	1943-2016	65
48	ON	6158731	TORONTO_INTL_A	43.68	-79.63	173	1950-2017	64
49	ON	6104175	KINGSTON_PUMPING_STATIO N	44.23	-76.48	76	1914-2007	63
50	ON	6139525	WINDSOR_A	42.28	-82.97	189	1946-2007	60
51	ON	6105978	OTTAWA_CDA_RCS	45.38	-75.72	79	1905-2011	54
52	ON	6048268	THUNDER_BAY_CS	48.37	-89.33	199	1952-2012	53
53	ON	6143089	GUELPH_TURFGRASS	43.55	-80.22	325	1954-2017	52
54	ON	6153301	HAMILTON_RBG_CS	43.28	-79.92	102	1962-2016	52
55	ON	6012199	EAR_FALLS_(AUT)	50.63	-93.22	362	1952-2007	50
56	ON	6042716	GERALDTON_A	49.78	-86.93	348	1952-2007	50
57	ON	6131983	DELHI_CS	42.87	-80.55	231	1962-2015	50
58	ON	6127519	SARNIA_CLIMATE	43.00	-82.30	181	1962-2016	49
59	ON	6078285	TIMMINS_VICTOR_POWER_A	48.57	-81.38	294	1952-2007	48
60	ON	6057592	SAULT_STE_MARIE_A	46.48	-84.52	192	1962-2007	46
61	ON	6158875	TRENTON_A	44.12	-77.53	86	1965-2017	46
62	ON	6034073	KENORA_RCS	49.78	-94.38	412	1966-2011	44

No.	Province name	Climate ID	Station name	Lat	Lon	Elevation (m)	Record period	RL (year)
63	ON	6016525	PICKLE_LAKE_(AUT)	51.45	-90.22	390	1953-2007	42
64	ON	6073980	KAPUSKASING_CDA_ON	49.42	-82.43	218	1966-2013	42
65	ON	6085700	NORTH_BAY_A	46.37	-79.42	370	1964-2006	41
66	ON	6037775	SIOUX_LOOKOUT_A	50.12	-91.90	383	1963-2007	40
67	ON	6131415	CHATHAM_WPCP	42.38	-82.22	180	1966-2007	40
68	ON	6106000	OTTAWA_MACDONALD-CARTIER_INT_L_A	45.32	-75.67	114	1967-2007	39
69	ON	6142286	ELORA_RCS	43.65	-80.42	376	1970-2017	39
70	ON	6137154	RIDGETOWN_RCS	42.45	-81.88	205	1959-2016	38
71	ON	6145504	MOUNT_FOREST_(AUT)	43.98	-80.75	414	1962-2016	38
72	ON	6116132	OWEN_SOUND_MOE	44.58	-80.93	178	1965-2006	37
73	ON	6136606	PORT_COLBORNE	42.88	-79.25	175	1964-2007	37
74	ON	6142400	FERGUS_SHAND_DAM	43.73	-80.33	417	1961-2007	37
75	ON	6150689	BELLEVILLE	44.15	-77.38	76	1960-2006	37
76	ON	6068150	SUDBURY_A	46.63	-80.80	348	1971-2007	36
77	ON	6140954	BRANTFORD_MOE	43.13	-80.23	196	1961-2001	36
78	ON	6148105	STRATFORD_WWTP	43.37	-81.00	345	1966-2004	36
79	ON	6100971	BROCKVILLE_PCC	44.60	-75.67	96	1967-2005	35
80	ON	6115811	ORILLIA_BRAIN	44.60	-79.43	250	1965-2004	35
81	ON	6104027	KEMPTVILLE_CS	45.00	-75.63	99	1970-2007	34
82	ON	6074211	KIRKLAND_LAKE_CS	48.15	-80.00	324	1980-2015	33
83	ON	6101901	CORNWALL_ONT_HYDRO	45.03	-74.80	76	1957-1992	33
84	ON	6106400	PETAWAWA_NAT_FORESTRY	45.98	-77.43	183	1961-1994	33
85	ON	6119500	WIARTON_A	44.75	-81.12	222	1973-2007	33
86	ON	6137287	ST_CATHARINES_A	43.20	-79.17	97	1954-2005	33
87	ON	6149387	WATERLOO_WELLINGTON_A	43.45	-80.38	317	1971-2007	33
88	ON	6153194	HAMILTON_A	43.17	-79.93	237	1971-2003	33
89	ON	6166418	PETERBOROUGH_A	44.23	-78.37	191	1971-2006	33
90	ON	6155878	OSHAWA_WPCP	43.87	-78.83	83	1970-2006	32
91	ON	6075435	MOOSONEE_RCS	51.28	-80.62	9	1968-2007	31
92	ON	6122847	GODERICH	43.77	-81.72	213	1970-2016	31
93	ON	6150830	BOWMANVILLE_MOSTERT	43.92	-78.67	99	1968-2001	31
94	ON	6151042	BURKETON_MCLAUGHLIN	44.03	-78.80	312	1969-2001	31
95	MB	502S001	WINNIPEG_A_CS	49.92	-97.25	238	1944-2016	57
96	MB	5062921	THOMPSON_A	55.80	-97.87	224	1971-2017	43

No.	Province name	Climate ID	Station name	Lat	Lon	Elevation (m)	Record period	RL (year)
97	MB	5050919	FLIN_FLON	54.68	-101.68	303	1970-2017	42
98	MB	5012324	PORTAGE_SOUTHPORT	49.90	-98.28	272	1964-2017	40
99	MB	5040681	DAUPHIN_CS	51.10	-100.07	304	1954-2016	40
100	MB	5052890	THE_PAS_CLIMATE	53.97	-101.10	274	1971-2011	39
101	MB	5060608	CHURCHILL_CLIMATE	58.73	-94.07	28	1963-2015	39
102	MB	5060999	GILLAM	56.37	-94.70	145	1972-2017	37
103	MB	5061376	ISLAND_LAKE_A	53.85	-94.65	235	1971-2013	36
104	MB	5010490	BRANDON_RCS	49.90	-99.95	409	1970-2014	35
105	MB	5020725	DEERWOOD_RCS	49.40	-98.32	341	1964-2014	35
106	MB	5061646	LYNN_LAKE_A	56.87	-101.08	356	1969-2005	34
107	MB	5021054	GLENLEA	49.65	-97.12	234	1967-2000	32
108	SK	401HP5R	WEYBURN	49.70	-103.80	588	1962-2017	43
109	SK	4012410	ESTEVAN	49.22	-102.97	580	1964-2016	52
110	SK	4015322	MOOSE_JAW_CS	50.33	-105.53	577	1960-2014	49
111	SK	4016560	REGINA_INT_L_A	50.43	-104.67	577	1941-1995	52
112	SK	4043901	KINDERSLEY_A	51.52	-109.18	693	1966-2016	50
113	SK	4057165	SASKATOON_RCS	52.17	-106.72	504	1960-2017	40
114	SK	4060983	BUFFALO_NARROWS_(AUT)	55.83	-108.42	440	1968-2017	41
115	AB	3012209	EDMONTON_BLATCHFORD	53.57	-113.52	671	1914-2015	69
116	AB	3031094	CALGARY_INT_L_CS	51.12	-114.00	1081	1947-2015	61
117	AB	3012206	EDMONTON_INTERNATIONAL_CS	53.32	-113.62	715	1961-2017	52
118	AB	3025481	RED_DEER_REGIONAL_A	52.18	-113.88	904	1959-2014	49
119	AB	3081680	COLD_LAKE_A	54.42	-110.28	541	1966-2017	49
120	AB	3033890	LETHBRIDGE_CDA	49.70	-112.77	910	1960-2017	47
121	AB	3034485	MEDICINE_HAT_RCS	50.03	-110.72	715	1971-2017	42
122	AB	3062696	FORT_MCMURRAY_CS	56.65	-111.22	368	1966-2017	39
123	AB	3075040	PEACE_RIVER_A	56.23	-117.45	570	1966-2011	39
124	AB	3015523	ROCKY_MTN_HOUSE_(AUT)	52.42	-114.92	988	1964-2017	37
125	AB	3023722	LACOMBE_CDA_2	52.45	-113.77	860	1970-2017	34
126	AB	3030QLP	BROOKS	50.57	-111.85	747	1965-2017	32
127	AB	3073146	HIGH_LEVEL_A	58.62	-117.17	338	1971-2011	32
128	AB	3036681	VAUXHALL_CDA	50.05	-112.13	778	1956-1987	31
129	AB	3053520	JASPER	52.88	-118.07	1062	1963-1994	31
130	AB	3070560	BEAVERLODGE_CDA	55.20	-119.40	744	1961-1994	31

No.	Province name	Climate ID	Station name	Lat	Lon	Elevation (m)	Record period	RL (year)
131	BC	1018611	VICTORIA_GONZALES_CS	48.42	-123.32	61	1925-2017	65
132	BC	1108395	VANCOUVER_INTL_A	49.18	-123.18	4	1953-2017	63
133	BC	1105192	MISSION_WEST_ABBEY	49.15	-122.27	197	1963-2017	54
134	BC	1018621	VICTORIA_INTL_A	48.65	-123.43	19	1965-2017	50
135	BC	1068131	TERRACE_PCC	54.50	-128.62	67	1968-2017	47
136	BC	1038205	TOFINO_A	49.08	-125.77	24	1970-2017	45
137	BC	1126150	PENTICTON_A	49.47	-119.60	344	1953-2002	45
138	BC	1166R45	SALMON_ARM_A	50.68	-119.23	527	1964-2016	44
139	BC	1160899	BLUE_RIVER_A	52.13	-119.28	690	1970-2016	44
140	BC	1103332	HANEY_UBC_RF_ADMIN	49.27	-122.57	147	1963-2005	42
141	BC	1096450	PRINCE_GEORGE_A	53.88	-122.68	691	1960-2002	41
142	BC	1021830	COMOX_A	49.72	-124.90	25	1963-2006	40
143	BC	1106180	PITT_POLDER	49.27	-122.63	5	1965-2007	40
144	BC	1142574	DUNCAN_LAKE_DAM	50.23	-116.97	548	1969-2013	38
145	BC	1163780	KAMLOOPS_A	50.70	-120.43	345	1965-2002	38
146	BC	1013754	JORDAN_RIVER_DIVERSION	48.50	-124.00	393	1964-2003	37
147	BC	1107873	SURREY_KWANTLEN_PARK	49.18	-122.87	78	1962-1999	37
148	BC	1173210	GOLDEN_A	51.30	-116.98	784	1973-2013	37
149	BC	1054500	LANGARA	54.27	-133.07	42	1982-2017	36
150	BC	112G8L1	SUMMERLAND_CS	49.57	-119.65	454	1955-1994	36
151	BC	1123970	KELOWNA_A	49.97	-119.38	429	1969-2004	34
152	BC	1192940	FORT_NELSON_A	58.83	-122.60	381	1966-2002	34
153	BC	1046391	POWELL_RIVER_A	49.83	-124.50	129	1982-2016	33
154	BC	1057050	SANDSPIT_A	53.25	-131.82	6	1972-2004	33
155	BC	1108487	VANCOUVER_UBC	49.25	-123.25	76	1958-1990	33
156	BC	1126510	PRINCETON_A	49.47	-120.52	701	1979-2017	33
157	BC	1157630	SPARWOOD	49.75	-114.88	1137	1980-2016	33
158	BC	1021990	COURTENAY_PUNTLEDGE_BC HP	49.68	-125.03	24	1964-1995	32
159	BC	1025369	NANAIMO_A	49.05	-123.87	28	1985-2017	32
160	BC	1067742	STEWART_A	55.93	-129.98	7	1978-2015	32
161	BC	1113540	HOPE_A	49.37	-121.50	39	1964-1995	32
162	BC	1148211	TRAIL_BIRCHBANK	49.20	-117.73	594	1965-1997	32
163	BC	1152102	CRANBROOK_A	49.62	-115.78	940	1969-2002	32
164	BC	1077500	SMITHERS_A	54.82	-127.18	521	1971-2002	31

No.	Province name	Climate ID	Station name	Lat	Lon	Elevation (m)	Record period	RL (year)
165	BC	1181508	CHETWYND_A	55.68	-121.63	609	1970-2016	31
166	BC	1060841	BELLA_COOLA_A	52.38	-126.60	35	1983-2015	30
167	BC	1125766	OLIVER_STP	49.18	-119.53	297	1973-2005	30
168	NT	2202102	FORT_SIMPSON_CLIMATE	61.77	-121.23	168	1969-2017	42
169	NT	2202401	HAY_RIVER_A	60.83	-115.78	164	1971-2015	39
170	NT	2202801	NORMAN_WELLS_A	65.28	-126.80	72	1974-2016	35
171	NT	2204100	YELLOWKNIFE_A	62.47	-114.43	205	1963-1996	33
172	NT	2202578	INUVIK_CLIMATE	68.32	-133.52	103	1972-2017	32
173	YT	2101310	WHITEHORSE_AUTO	60.73	-135.10	707	1960-2016	44
174	YT	2100880	PELLY_RANCH	62.83	-137.32	445	1966-2014	41
175	YT	2101102	TESLIN_(AUT)	60.17	-132.73	705	1967-2016	35