Title
Frameworks for Univariate and Multivariate Non-Stationary Analysis of Climatic Extremes

Author
Cheng, Linyin

Publication Date
2014

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Frameworks for Univariate and Multivariate Non-Stationary Analysis of Climatic Extremes

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Civil Engineering

by

Linyin Cheng

Dissertation Committee:
Assistant Professor Amir AghaKouchak, Chair
Associate Professor Kuo-lin Hsu
Professor Soroosh Sorooshian

2014

DEDICATION

To

my parents and friends

Without their constantly support, encouragement and love,
this dissertation would not have been possible.

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

Page

ACKNOWLEDGMENTS

I would like to start with expressing my sincere gratitude to lots of people for their support and advice during my Ph.D. study as the beginning of the dissertation. Without them, the dissertation would not have been possible.

First of all, I would like to express sincere gratitude to my Ph.D. adviser, Professor Amir AghaKouchak. I am much honored and fortunate to be his student. During these three years, he gives me new knowledge of climatology and statistical hydrology; he leads me to discover a world of extreme value analysis with loads of joy; he shares his experience and invaluable advice on my research topics, papers and my career. What I learned from him is much more than scientific knowledge, to be patient, respectful, responsible, creative and so on are indeed my life of wealth. With his support and guidance, I received the scholarship to visit National Center for Atmospheric Research lab (NCAR); received Outstanding Student Paper Awards on the 2013 AGU meeting; and was selected to receive prestigious fellowships to continue doing research. Amir is the supervisor and the best mentor to me. All the honors I hope to share with him. Without his help and contribution, none of them would have been possible.

I would also like to express sincere gratitude to Professor Soroosh Sorooshian, Dr. Eric Gilleland, Dr. Richard Katz, Professor Balaji Rajagopalan, Dr. Martin Hoerling and his group members. Professor Sorooshian taught me the best lectures of Hydrology and Hydrologic System, and supported me to pursue further education. During the visit to NCAR, Eric would like to spend time meeting with me every week and share the knowledge in statistics. Eric and Rick gave me invaluable advice and constructive comments on the collaborative papers and on my career development. Without their support, I would not be selected for the Advance Study Program fellowship; the Non-stationary Extreme Value Analysis (NEVA) package is

incomplete; and my knowledge in statistics would be limited. Special thanks to Professor Rajagopalan, who kindly invited me to give a presentation, meet with his group in the University of Colorado, Boulder. He encouraged me all the way and fully supported me to apply for the postdoctoral fellowship. I am grateful to Dr. Hoerling, who provided me with an interview and a research position in the lab of National Oceanic and Atmospheric Administration (NOAA) at Boulder, Colorado. He also broadened my research interests to the physical climate process. All the scientists in his group, particularly, Dr. Judith Perlwitz, Dr. Xiaowei Quan and Secretary Barbara Herrli gave me great help and encouragement during the visiting.

The dissertation committee has given me useful remarks and suggestions. I would like to thank Professor Kuo-lin Hsu and Professor Soroosh Sorooshian for taking time to participate as my dissertation committee and provide invaluable advice on my dissertation.

I thank Diane Hohnbaum, who helped with the administrative works. I also wish to thank all the students in Hydroclimate Research group and Center for Hydrometeorology and Remote Sensing. Their support and friendship are indeed my lifetime treasure. They make the three years of my doctorate study full of happiness, laughs and remarkable memories.

In the end, I wish to thank my parents and my cousin's family for being constantly supportive and making me feel at home always. Their trust and love encourage me to continue my study and to live my dream.

CURRICULUM VITAE

LINYIN CHENG

Education

- ✧ **Ph.D. in Civil Engineering (Hydroclimate),** University of California, Irvine, CA
- ✧ **M.S. in Civil Engineering (Hydrodynamic),** Clarkson University, NY
- ✧ **B.S. in Civil Engineering (Hydraulic and Hydra-electric Engineering),** Sichuan University, Sichuan, China
- ✧ **B.S. in Law,** Law School, Sichuan University, Sichuan, China

Key Qualifications

- ✧ Multivariate Trend Analysis, Conditional Trend Analysis
- ✧ Univariate, Multivariate Extremes, Copula, survival Copula, Multivariate Return Period
- ✧ Empirical Bayes Conditional Extreme Value Analysis, Empirical Bayes, Conditional Extreme Framework
- ✧ Analysis of Non-Stationary Climatic Extremes across Scales, Extreme Value Theory, Non-Stationarity, Bayesian Inference, Bayesian Hierarchical Model, Markov Chain Monte Carlo
- ✧ Apply Expert Advice Algorithm on Deriving Climate Response and Hurricane Tracking, EA Algorithm, Game Theory, Multimodel Simulations
- ✧ Numerical Study on River Ice Dynamics and the Need for the Ice Sluice Gates, Advanced Fluid Dynamics, Numerical Modeling, Computational Fluid Dynamics

Research Experience

- ✧ **UC-Irvine, CA (9/2011 ~ 9/2014)**
  - hydroclimate research group, research on analysis of non-stationary spatio-temporal climatic extremes
  - Ph.D. dissertation, Frameworks for Univariate and Multivariate Non-Stationary Analysis of Climatic Extremes
- ✧ **National Center for Atmospheric Research, CO (6/2013 ~ 9/2013)**
  - Received the Advanced Study Program support to participate in the Graduate Visitor Program at NCAR: Research on Empirical Bayes Conditional Extreme Value Analysis
- ✧ **Clarkson University, NY (9/2009 ~ 5/2011)**
  - hydrodynamic research group, research on numerical study of river ice dynamics
  - Attend the conference, presentation and field study for New York Power Authority (NYPA) on the study on the upper St. Lawrence River ice dynamics
  - M.S. thesis, A Numerical Study on the upper St. Lawrence River Ice Dynamics and the Need for the Ice Sluice Gates

Professional and Graduate Courses

Merging Data
Hydrologic Systems
Watershed Modeling
Hydrology
Geoscience Model Data
Climate Data Analysis
Shallow Water Hydrodynamics
Hydraulic Engineering in Cold Regions
Hydrodynamic Dispersion
Computational Fluid Dynamics
Advanced Fluid Mechanics
Sediment Transport

Major Honors and Awards

- ✧ Research Associate in Physical Sciences Division of NOAA from 10/2015
- ✧ Received the Cooperative Institute for Research in Environmental Sciences (CIRES) Fellowship in Postdoctoral Program from 10/2014
- ✧ Selected for the Advanced Study Program support in Postdoctoral Program (ASP Fellowship) at the National Center for Atmospheric Research in 2014
- ✧ Received the Outstanding Student Paper Awards (OSPAs) for the poster "Non-stationary Extreme Value Analysis in a Changing Climate: A Software Package" in Hydrology session, AGU Fall meeting, 2013
- ✧ Received the Advanced Study Program support to participate in the Graduate Visitor Program at the National Center for Atmospheric Research from 6/8/2013 - 9/14/2013
- ✧ The Teaching Assistant Scholarship and Graduate Research Scholarship for a Ph.D. degree for academic years 2014, 2013, 2012 and 2011 at UC-Irvine
- ✧ The Research Assistant Scholarship for a M.S. degree for academic year 2010 and 2011 at Clarkson University

Computer Skills

- ✧ Programming & Scripting: MATLAB, R, FORTRAN, C
- ✧ Engineering Software: HEC-HMS, GIS, HBV Model, Microsoft Office (Word, Excel, PowerPoint), LaTeX, Framer, SMS, Tecplot, AutoCAD

Members and Conference Papers

- ✧ Member of AGU
- AGU Fall Meeting, December 9-13, 2013: Non-stationary Extreme Value Analysis in a Changing Climate: A Software Package (AGU OSPAs Awards).
- AGU Fall Meeting, December 3-7, 2012: Deriving Climate Response from CMIP5 Ensemble Climate Projections: Application to Analysis of Temperature and Precipitation Extremes.
- Conference papers/presentations: Tracking and Nowcasting of Hurricanes: a Data Fusion Approach, 3rd World Meteorological Organization (WMO) International Symposium on Nowcasting (WSN12), 6-10 August 2012, Rio de Janeiro, Brazil.

Publications

- ✧ Cheng L., AghaKouchak A., Gilleland E., Katz R., 2013, Non-stationary Extreme Value Analysis in a Changing Climate: A Software Package (Climatic Change under revision)
- ✧ Cheng L., AghaKouchak A., Phillips T., 2014, Non-stationary Return Levels of Monthly Temperature Extremes based on CMIP5 Multi-Model Simulations (Climate Dynamics under revision)
- ✧ Cheng L., AghaKouchak A., 2014, Improved Precipitation Intensity-Duration-Frequency Curves for Infrastructure Design in a Changing Climate (Scientific Report under revision)
- ✧ Cheng L., AghaKouchak A., 2013, A Methodology for Deriving Climate Response from Multimodel Ensemble Climate Simulations (Journal of Hydrology under review)
- ✧ Cheng L., Gilleland E., Heaton M., AghaKouchak A., 2013, Empirical Bayes estimation for the conditional extreme value model (Computational Statistics & Data Analysis under review)
- ✧ Cheng L., AghaKouchak A., 2014, An Empirical Bayes Conditional Extreme Value Model for Detecting Changes in the Hydrological Cycle (To be submitted)

ABSTRACT OF THE DISSERTATION

Frameworks for Univariate and Multivariate Non-Stationary Analysis of Climatic Extremes

by

Linyin Cheng

Doctor of Philosophy in Civil Engineering

University of California, Irvine, 2014

Professor Amir AghaKouchak, Chair

Numerous studies show that climatic extremes have increased substantially in the second half of the 20th century. For this reason, analysis of extremes under a non-stationary assumption has received a great deal of attention. In this dissertation, a methodology is developed for deriving non-stationary return levels, return periods, and climate risk assessment using a Bayesian approach. The methodology is presented in the Non-stationary Extreme Value Analysis (hereafter, NEVA) software package. The methodology offers the confidence intervals and uncertainty bounds of estimated non-stationary return levels using both constant and time varying exceedance probability methods. Both stationary and non-stationary components of NEVA are validated for a number of case studies, and have been validated using empirical return levels. The results show that NEVA reliably describes extremes and their return levels. The methodology has been applied for assessing non-stationary extreme return levels in CMIP5 multi-model simulations. Furthermore, the model has been applied for non-stationary precipitation Intensity-Frequency-Duration (IDF). Beyond univariate non-stationary analysis, a novel framework named empirical Bayes conditional extreme value analysis model has been

developed for modeling concurrent and conditional extremes. The methodology has been used for detecting potential changes in the hydrological cycle, and assessing joint occurrences of extreme events.

CHAPTER 1: Introduction

1. Problem Statement

Climate change and variability are likely to affect physical and hydrometeorological conditions and to interact with, and possibly exacerbate, ongoing environmental change (IPCC 2007; Barnett et al, 2006; Schmidli et al., 2005; Solomon et al., 2007; Frich et al., 2002). Climatic extremes including heavy precipitation events and extreme hot days, have substantially increased in the past few decades (Alexander et al 2006; Vose et al. 2005). A recent study shows that even concurrent extremes (e.g., warm-dry and warm-wet conditions) have increased significantly in the second half of the 20th century (Hao et al., 2013). Therefore, there exists a strong need to study extreme weather and climate events across different spatio-temporal scales and to examine potential changes in their frequency and intensity. In the past decades, numerous methods and models have been developed for the analysis of extremes in a changing climate (AghaKouchak et al., 2013). However, there are still major research gaps including:

(1) modeling non-stationarity processes in space and time (e.g., seasonal, interannual time scales). Engineering and hydrologic design considerations have long relied on analyses of extreme-rainfall return intervals. The fundamental assumption behind design concepts is the so-called stationary assumption that indicates frequency of extremes do not change significantly over time. Recent studies have demonstrated that extremes have changed or are likely to change mainly due to climate change (e.g., Milly et al. 2008). Heavy-rainfall events have become more frequent since the middle of the last century, not only in the United States (Karl and Knight 1998), but also in regions across the globe. Warming of the climate system is unequivocal, as is now evident from observations of increases in global average air and ocean temperatures, widespread melting of snow and ice and rising global average sea level

(e.g., IPCC 2007: Synthesis Report). Of course, ignoring time-variant (non-stationary) behavior of extremes could potentially lead to underestimating extremes and failure of infrastructures and considerable damage to human life and society.

(2) modeling concurrent, consecutive, and conditional extremes and their dependencies. The combination or sequences of climate extreme events may have a significant impact on the ecosystem and society, though the individual events involved may not be severe extremes themselves. Current multivariate extreme value models are designed for modeling the dependence between two sets of extremes. However, an extreme event (e.g., extreme heatwave) can happen concurrently with an another non-extreme event (e.g., moderate drought) whose combination could lead to an extreme climatic condition. Current models cannot assess joint occurrence of an extreme event with an event that is only a moderate departure from the mean.

(3) deriving the climate response of the multi-model projections as well as the uncertainty of climate simulations. Climate models have been widely used to derive projections of climate change, climate extremes and their frequency of occurrence over various time scales and emission scenarios. Several national and international efforts, such as the Intergovernmental Panel on Climate Change (IPCC 2007), provide data sets of plausible changes for the future. However, climate projections are subject to significant uncertainties arising from uncertainties in boundary, initial conditions, parameters and model structure (de Elia and Cote 2010; Deque et al. 2007; Kjellstrom and Ruosteenoja 2007). In order to analysis extreme climate variables from multi-model climate simulation, one needs to derive the climate response of the multi-model projections as well as the uncertainty of climate simulations.

The overarching goal of this study is to address the above mentioned research gaps. This study will lead to a better understanding of hydrologic extremes in a changing climate, and will provide valuable tools and techniques for analyzing extreme events. In the following section, the objectives of the project are discussed in details.

2. Objectives

(1) Non-stationary Extreme Value Analysis in a Changing Climate

This study introduces a framework for estimating stationary and non-stationary return levels, return periods, and risks of climatic extremes using Bayesian inference. This framework is implemented in the Non-stationary Extreme Value Analysis (hereafter, NEVA) software package, explicitly designed to facilitate analysis of extremes in the geosciences. In a Bayesian approach, NEVA estimates the extreme value parameters with a Differential Evolution Markov Chain (DE-MC): a genetic algorithm Differential Evolution (DE) for global optimization over the parameter space with the Markov Chain Monte Carlo (MCMC) approach. NEVA includes confidence intervals and uncertainty bounds of estimated return levels through Bayesian inference, with its inherent advantages in uncertainty quantification. The software presents the results of non-stationary extreme value analysis using various exceedance probability methods. We evaluate both stationary and non-stationary components of the package for a case study consisting of annual temperature maxima for a gridded global temperature dataset. The results show that NEVA can reliably describe extremes and their return levels.

(2) Non-stationary Return Levels of CMIP5 Multi-Model Temperature Extremes using NEVA

The NEVA model discussed above is used to evaluate to what extent the CMIP5 climate model simulations of the climate of the 20th century can represent observed warm monthly temperature extremes under a changing environment. The biases and spatial patterns of 2-, 10-, 25-, 50- and 100-year return levels of the annual maxima of monthly mean temperatures (hereafter, annual temperature maxima) from CMIP5 simulations are compared with those of Climatic Research Unit (CRU) observational data considered under a non-stationary assumption. The results show that CMIP5 climate models collectively underestimate the mean annual maxima over arid and semi-arid regions that are most subject to severe heat waves and droughts. Furthermore, the results indicate that most climate models tend to underestimate the historical annual temperature maxima over the United States and Greenland, while generally disagreeing in their simulations over cold regions. Return level analysis shows that with respect to the spatial patterns of the annual temperature maxima, there are good agreements between the CRU observations and most CMIP5 simulations. However, the magnitudes of the simulated annual temperature maxima differ substantially across individual models. Discrepancies are generally larger over higher latitudes and cold regions.

(3) Non-stationary Precipitation Intensity-Duration-Frequency Curves for Infrastructure Design in a Changing Climate

Extreme climatic events are growing more severe and frequent. This observation calls into question how prepared our infrastructure is to deal with these changes. Current infrastructure design is primarily based on precipitation Intensity-Duration-Frequency (IDF) curves with the so-called stationary assumption, meaning extremes will not vary significantly over time. However, climate change is expected to alter climatic extremes, a concept termed non-

stationarity. Using NEVA (Objective 1), we show that given non-stationarity, current IDF curves substantially underestimate precipitation extremes and thus, they may not be suitable for infrastructure design in a changing climate. We show that a stationary climate assumption may lead to underestimation of extreme precipitation by as much as 60%, which increases the flood risk and failure risk in infrastructure systems. We then use the generalized framework outlined above for estimating non-stationary IDF curves and their uncertainties using Bayesian inference.

(4) Empirical Bayes Estimation for the Conditional Extreme Value Model

A methodology is developed for modeling multivariate extreme values through a conditional distribution framework that does not require a priori knowledge of the dependence structure of the variables. Also, this methodology does not require that two extremes to happen at the same time. This conditional extreme value model can be used to assess one extreme value (e.g., an extreme heatwave) conditioned on another non-extreme event (e.g., moderate drought). The main science contribution of this component is an estimation strategy for modeling multivariate extreme values using an empirical Bayesian approach. The approach is tested on different types of synthetic extreme dependence structures, as well as for two real observations consisting of precipitation conditioned on extreme temperature. The simulated data consist of numerous repeated trials in order to gauge the coverage of credible intervals for the parameters. The presented model can be potentially applied in a wide variety of science fields including finance, earth science, environmental science, and biology.

(5) An Empirical Bayes Conditional Extreme Value Model for Detecting Changes in Hydrological Cycle

Greenhouse gases in the atmosphere have been increasing since the industrial revolution, leading to the warming of the Earth through an increase in downwelling infrared radiation. Warming of

the atmosphere increases its water holding capacity and could intensify the hydrological cycle. Several methods have been developed for evaluating changes in climatic variables. On the other hand, numerous indices have been developed for monitoring changes in climate. Most change detection methods, indices, and trend studies focus on changes in one variable at the time. However, hydrologic variables are dependent, and a change in one variable can alter extreme and non-extreme values of other variables. In this study, a new approach for modeling multivariate extreme values through a conditional distribution framework using the empirical Bayes approach is proposed. This chapter highlights the value of empirical Bayes conditional extreme value analysis as a tool for simulating and assessing conditional extremes (e.g., changes in the distribution of precipitation conditioned on extreme temperature). The model has been applied to several locations across the world. This presentation will summarize the findings on changes in the hydrological cycle over the United States and Australia.

(6) A Methodology for Deriving Ensemble Response from Multi-model Simulations

Multi-model ensembles are widely used to quantify uncertainties of climate model simulations. Previous studies have confirmed that a multi-model ensemble approach increases the skill of model simulations. However, one may need to know which ensemble member is more likely to be true, particularly when the ensemble is spread out over a wide area. Typically, ensemble response (climate response) is derived by taking the mean or median of ensemble members. However, strong similarities exist between models (members of an ensemble), which may cause biased climate response toward models with strong similarities. In this study, a model is proposed for deriving the climate response (ensemble response) of multi-model climate model simulations. The approach is based on the concept of the Expert Advice (EA) algorithm that has been successfully applied to the financial sector. The goal of this methodology is to derive an

ensemble response such that every time step is equal or better (less error) than the best model. The methodology is tested using the CMIP5 historical temperature simulations (1951-2005) and Climatic Research Unit observations, and the results show that the EA algorithm leads to smaller error compared to the ensemble mean.

3. Organization of the Dissertation

This dissertation is organized as follows: Chapter 2 presents a model for non-stationary extreme value analysis (NEVA) in a changing climate. In Chapter 3 and Chapter 4, NEVA is used to study non-stationary return levels of CMIP5 multi-model temperature extremes, and non-stationary precipitation Intensity-Duration-Frequency curves for infrastructure design in a changing climate, respectively. Chapter 5 describes the empirical Bayes estimation for the conditional extreme value analysis. Chapter 6 shows the application of the empirical Bayes conditional extreme value model as a tool for detecting changes in the hydrological cycle. Chapter 7 outlines a methodology for deriving ensemble response from multi-model simulations with applications on CMIP5 multi-model temperature data. The last chapter provides a summary of the findings and offers some avenues for further research.

CHAPTER 2: Non-stationary Extreme Value Analysis in a Changing Climate

2.1 Introduction

The Intergovernmental Panel on Climate Change (IPCC) Special Report on Managing the Risks of Extreme Events and Disasters (Field et al. 2012) stressed that continuation of the observed Earth warming would change the frequency, severity and spatial pattern of climatic extremes. Recently, climatic extremes have been widely studied at a range of spatial and temporal scales (Wehner 2013; Jacob 2013; AghaKouchak et al., 2013; Schubert and Lim 2013; Diffenbaugh and Giorgi, 2012; Kharin et al. 2007; Easterling et al. 2000). Climatic extremes, including heavy precipitation events and extreme hot days, have substantially increased in the past few decades (Alexander et al 2006; Vose et al. 2005). A recent study shows that even concurrent extremes (e.g., warm-dry and warm-wet conditions) have increased significantly in the second half of the 20^{th} century (Hao et al., 2013).

Under the assumption of a stationary climate, the concepts of return level and return period provide critical information for design, decision-making, and assessing the impacts of rare weather and climatic events (Rosbjerg and Madsen 1998). For example, the return level with a T-year return period represents an event that has a $1/T$ chance of occurrence in any given year (Cooley 2007). Infrastructure design concepts have long relied on stationary return levels, which assume no change to the frequency of extremes over time (Klein et al. 2009). However, the frequency of extremes has been changing and is likely to continue changing in the future (Easterling et al. 2013; Milly et al. 2008; IPCC 2007). Therefore, concepts and models that can account for non-stationary analysis of climatic and hydrologic extremes are needed (e.g., Cooley 2013; Salas and Obeysekera 2013; Parey et al. 2010).

Katz et al. 2002 present non-stationarity in extremes in terms of changing quantiles (termed "effective return levels"), which vary as a function of time to keep the occurrence probability of an extremal event constant. Alternatively, Rootzén and Katz 2013 introduced the concept of Design Life Level to quantify the probability of exceeding a fixed threshold during the design life of a project. A recent study describes an R-package developed for analysis of extremes based on the concept of effective return levels (extRemes 2.0). Another available R-package (GEVcdn: Cannon 2011) supplies a framework for a conditional density estimation network, and can be used to perform non-stationary extreme value analysis. However, these packages do not provide any non-stationary generalization of the concepts of return period and return level frequently used in hydrology.

The concept of return period can also be extended to a non-stationary framework (e.g., Rootzén and Katz 2013; Salas and Obeysekera 2013). In this study, we introduce a framework for non-stationary extreme value analysis for practical and effective analysis of climate extremes under both stationary and non-stationary conditions using Bayesian inference. The methods presented are available through a software package called Non-Stationary Extreme Value Analysis (NEVA). Under the non-stationary assumption, NEVA provides three different methods for estimation of return levels: (a) standard return levels (commonly used in hydrologic design) in which the exceedance probability is constant for any given return period during the life of the design (hereafter, design exceedance probability); (b) constant thresholds with time varying exceedance probability; and (c) effective return levels. A unique feature of NEVA is that it offers the associated confidence intervals and uncertainty bounds for the return level estimates under non-stationarity. These features make NEVA a practical and attractive tool for users from across

different fields, especially climatology and hydrology, to analyze extremes under both stationary and non-stationary assumptions.

2.2 Extremes in a Non-Stationary Climate: Theory

Extreme Value Theory (EVT) provides a rigorous framework for analysis of climate extremes and their return levels (Katz et al. 2002; Coles 2001). Under a wide range of conditions, the distribution of the maxima or minima converges to one of the three limiting distributions: Gumbel, Fréchet, or Weibull (Katz et al. 2002; Leadbetter et al. 1983; Gumbel 1958). The combination of these three distributions into one family is referred to as the Generalized Extreme Value (GEV) distribution. A variety of studies apply the GEV to analyze extremes (Katz 2013; Towler et al. 2010; AghaKouchak and Nasrollahi 2010; Beniston et al. 2007; El Adlouni et al. 2007; Frei, et al. 2006; Kharin and Zwiers 2005; Zhang et al. 2001; Smith 2001; Kharin and Zwiers 2000; Gumbel 1942). This technique is often referred to as the block maxima approach (e.g., Coles 2001). Another form of the EVT is known as the peak-over-threshold (POT) approach, in which extreme values above a high threshold are analyzed using a generalized Pareto distribution (Coles 2001; Smith 1987). Both annual maxima and POT are widely applied in studying climatic extreme events (Villarini et al. 2011; Li et al. 2005; Davison and Smith 1990).

The cumulative distribution function of the GEV can be expressed as (Coles 2001):

$$\Psi(x) = \exp\left\{-(1+\xi(\frac{x-\mu}{\sigma}))^{\frac{-1}{\xi}}\right\}, \qquad (1+\xi(\frac{x-\mu}{\sigma})) > 0 \tag{2.1}$$

The GEV distribution is flexible for modeling different behavior of extremes with three distribution parameters $\theta = (\mu, \sigma, \xi)$: (1) the location parameter (μ) specifies the center of the distribution; (2) the scale parameter (σ) determines the size of deviations around the location

parameter; and (3) the shape parameter (ξ) governs the tail behavior of the GEV distribution. The limiting case of $\xi \to 0$ gives the Gumbel distribution, $\xi < 0$ the Weibull distribution and $\xi > 0$ the Fréchet distribution.

The extreme value theory for stationary random sequences has been extensively studied (Papalexiou et al. 2013; Li et al. 2005; Leadbetter et al. 1983). In this study, stationarity is defined as time invariance of extremes' properties (Leadbetter et al. 1983). For a non-stationary process, the parameters of the underlying distribution function are time-dependent (Renard et al. 2013; Gilleland and Katz 2011; Katz 2010; Cooley 2009) and hence, the properties of the distribution would vary with time (Meehl et al. 2000). In NEVA, the location parameter is assumed to be a linear function of time to account for non-stationarity (Equation 2.2), while keeping the scale and shape parameters constant:

$$\mu(t) = \mu_1 t + \mu_0 \tag{2.2}$$

where t is the time (in years), and $\beta = (\mu_1, \mu_0, \sigma, \xi)$ are the parameters. Alternative models may be used, such as polynomial trends, step changes, trends on the scale or the shape parameter, etc. (Renard et al. 2013). The methodology presented in this study can be used with different types of trends in location parameter. In hydrology and climate literature, the linear or log-linear models are usually preferred when searching for trends in the occurrence of extreme events (Beguería et al. 2011). While NEVA allows non-stationary σ and ξ ($\sigma(t) = \sigma_1(t) + \sigma_0$, $\xi(t) = \xi_1(t) + \xi_0$), in this study, only non-stationarity with respect to μ is discussed. The primary reason is that modeling temporal changes in σ and ξ reliably requires long-term observations that are often not available for practical applications.

NEVA detects the presence of trends and non-stationarity in extremes in historical data using the Mann-Kendall trend test (Kendall 1976; Mann 1945) at the user's choice of significance level.

The default significance level is $\alpha = 0.05$, which is widely used in hydrological research (Zhang et al. 2004). This nonparametric rank-based test avoids making an assumption about the underlying distribution function (e.g., assuming the data is normally distributed) of hydrological variables (Kundzewicz and Robson 2004). The null hypothesis of no trend is rejected if the test statistic $|Z_S|$ is larger than the critical value $Z_{\alpha/2}$. The test returns either 0 when $|Z_S| \leq Z_{\alpha/2}$ (the null hypothesis of no trend cannot be rejected) or 1 when $|Z_S| > Z_{\alpha/2}$ (the null hypothesis of no trend is rejected). If the null hypothesis is not rejected, NEVA will perform extreme value analysis under the stationary assumption. Upon detection of a trend at the 5% significance level ($\alpha = 0.05$), the GEV distribution parameters will be estimated under the non-stationary assumption (Equation 2.2). This will allow estimating return values in a more realistic way consistent with the behavior of climatic extremes.

NEVA uses a Bayesian technique to infer the GEV distribution parameters under stationary and non-stationary conditions. The Bayesian-based Markov chain Monte Carlo (MCMC) approach for obtaining the posterior distribution of parameters from an arbitrary distribution has become increasingly popular and used in several studies of extremes (Stephenson and Tawn 2004; Coles and Powell 1996). This approach combines the knowledge brought by a prior distribution (in NEVA, the default priors for location and scale parameters are non-informative normal distributions; the shape parameter is a normal distribution with standard deviation of 0.3 (Renard et al. 2013). The prior distributions for all parameters are assumed independently. Users can change the default to give informative priors or change the choice of distribution function) and the observation vector $\vec{y} = (y_t)_{t=1:N_t}$ (Equations 2.4 and 2.5) into the posterior distribution of parameters $\theta = (\mu, \sigma, \xi)$. Here, N_t denotes the number of observations (e.g., annual maxima) in the observation vector $\vec{y}$. Assuming independence between observations, the Bayes theorem for

estimation of GEV parameters under the non-stationary assumption can be expressed as (Renard et al. 2013, Coles 2001):

$$p(\beta \mid \vec{y}, x) \propto p(\vec{y} \mid \beta, x) p(\beta \mid x) \tag{2.3}$$

$$p(\vec{y} \mid \beta, x) = \prod_{t=1}^{N_t} p(y_t \mid \beta, x(t)) = \prod_{t=1}^{N_t} p(y_t \mid \mu(t), \sigma, \xi) \tag{2.4}$$

where $\beta = (\mu_1, \mu_0, \sigma, \xi)$ are the parameters. The stationarity can be treated as a special case of the above equation without $x(t)$:

$$p(\theta \mid \vec{y}) \propto p(\vec{y} \mid \theta) p(\theta) = \prod_{t=1}^{N_t} p(y_t \mid \theta) p(\theta) \tag{2.5}$$

where $x(t)$ denotes the set of all covariate values under the non-stationary assumption. The resulting posterior distributions $p(\theta \mid \vec{y})$ and $p(\beta \mid \vec{y}, x)$ provide information about parameters under stationarity ($\theta = (\mu, \sigma, \xi)$) or non-stationarity ($\beta = (\mu_1, \mu_0, \sigma, \xi)$). The entire process for inferring distribution parameters in NEVA is summarized in Figure 2.1 and Figure 2.2 for stationary and non-stationary conditions, respectively. NEVA generates a large number of realizations from the parameter joint posterior distribution using the Differential Evolution Markov Chain (DE-MC) (Vrugt et al. 2009; Ter Braak and Vrugt 2008; Ter Braak 2006). The DE-MC utilizes the genetic algorithm Differential Evolution (DE) (Ter Braak 2006) for global optimization over the parameter space with the MCMC approach. The DE-MC's simplicity, speed of calculation, and convergence make it favorable over the conventional MCMC (Ter Braak 2006).

The main motivation for combining DE-MC with Bayesian inference is that one can obtain the confidence interval and uncertainty bounds of estimated return levels taking into account the uncertainty in all model parameters (non-stationary:μ_0, μ_1, σ, ξ; stationary: μ, σ, ξ). It is worth

noting that NEVA assesses convergence of the sampling approach statistically. A method known as the criterion $\hat{R}$, suggested by Gelman and Shirley (2011), is built into NEVA as a convergence check. This method suggests that the $\hat{R}$ values should remain below the critical value of 1.1 (see Gelman and Shirley 2011 for more details on computing $\hat{R}$).

In addition to the Mann-Kendall test, the likelihood-ratio test can be used to compare the fit of the two nested models: the *null* model is the stationary (no trend) case (L_{Null}), whereas the *alternative* is the non-stationary (linear trend) case ($L_{Alternative}$). The log-likelihood ratio can be expressed as (Coles 2001):

$$D = -2 \ln \frac{L_{Null}}{L_{Alternative}} \tag{2.6}$$

This likelihood ratio can then be used to test (e.g., at the $\alpha = 0.05$ significance level) whether to reject the null model in favor of the alternative. Based on the modes of the posterior of modeled simulations, the likelihood test returns either 0 when the non-stationary model does not fit significantly better than the stationary model or 1 when the non-stationary model fits significantly better than the stationary model. Note that the Mann-Kendall and likelihood ratio tests are both testing for trends, but under different assumptions: the Mann-Kendall test allows for non-linear trends in the location parameter and any form of distribution, while the likelihood ratio test assumes GEV distribution and only allows for a linear trend in the location parameter.

In order to further evaluate the fit of the *null* model M_1 (i.e., the stationary case), and the *alternative* model M_2 (i.e., the non-stationary case) based on the posterior distributions of sampled parameters, the Bayes factor is computed as:

$$K = \frac{\Pr(DA|M_1)}{\Pr(DA|M_2)} = \frac{\int \Pr(\theta_1|M_1)\Pr(DA|\theta_1,M_1)d\theta_1}{\int \Pr(\theta_2|M_2)\Pr(DA|\theta_2,M_2)d\theta_2} \tag{2.7}$$

where DA denotes input data, and θ stands for model parameters. The term $\Pr(DA|M)$ can be expressed using a Monte Carlo integration estimation as $\Pr(DA|M) = \{\frac{1}{m}\sum_{i=1}^{m} \Pr(DA|\theta^{(i)}, M)^{-1}\}^{-1}$, where m is the sample size (see Kass and Raftery 1995 for more details). A value of $K < 1$ indicates that the non-stationary model (M_2) fits better than the stationary model (M_1). Having multiple tests to detect stationarity or non-stationarity allows a more rigorous assessment of the goodness-of-fit.

While the original NEVA is designed for analysis of maxima in time series, users can apply NEVA for analysis of time series minima using the following transformation (Coles 2001):

$$\min(X_1,...,X_n) = -\max(-X_1,...,X_n) \tag{2.8}$$

where $X_1,...,X_n$ is a time series of i.i.d. random variables.

Using the GEV distribution, NEVA computes the return periods and return levels of extremes (see Equations 2.9 and 2.10). In this approach, return levels are expressed as a function of the return period T (Cooley 2013):

$$T = \frac{1}{1-p} \tag{2.9}$$

where p is the non-exceedance probability of occurrence in a given year (assumed constant under stationarity). The $p-$return level q_p derived from the GEV distribution can be expressed as (Coles 2001):

$$q_p = ((-\frac{1}{\ln p})^{\xi} - 1) \times \frac{\sigma}{\xi} + \mu, \qquad (\xi \neq 0) \tag{2.10}$$

In NEVA, the time-variant parameter ($\mu(t)$) can be derived using different quantiles from the DE-MC. For example, in this dissertation, $\mu(t)$ is computed as: (a) median of $\mu(t)$ (refers to the effective return level for the year corresponding to the midpoint of the time series), and (b) 95

percentile of the DE-MC sampled $\mu(t)$ values. The latter can be considered a low risk (more conservative) approach for extreme value analysis by taking the 95 percentiles of the $\mu(t)$ values in historical observation to be used for future analysis (e.g., the effective return level for a year near the end of the record). The model parameters will then be used to estimate the non-stationary return levels as follows:

$$\tilde{\mu} = Q_k(\mu_{t1}, \mu_{t2}, ..., \mu_{tn}), \qquad (\mu(t) = \mu_1 t + \mu_0) \tag{2.11}$$

$$q_p = ((-\frac{1}{\ln p})^{\xi} - 1) \times \frac{\sigma}{\xi} + \tilde{\mu}, \qquad (\xi \neq 0) \tag{2.12}$$

where κ = 0.5 returns the median of n location parameters $(\mu_{t1}, \mu_{t2}, ..., \mu_{tn})$, and $\kappa = 0.95$ corresponds to the 95 percentile of location parameters (a high quantile $\tilde{\mu}$ indicating low risk extreme value analysis). In this concept the exceedance probability is constant for any given return period during the life of the design. This concept is termed design exceedance probability in this dissertation.

In a recent study, Salas and Obeysekera 2013, proposed another non-stationary counterpart of stationary return levels. In this approach, the probability that the first extreme event exceeding a given fixed threshold will occur at time $x = 1$ is denoted by q_1, and the probability that it will occur at time $x = 2$ is $(1 - q_1)q_2$, and so forth (exceeding probabilities $q_1, q_2, q_3, ..., q_t$ vary over time). With the time varying exceedance probabilities q_t, a non-stationary concept determining the expected return period of the extreme event is outlined in Salas and Obeysekera 2013. This concept is based on the expected waiting time until the first exceedance of a fixed threshold, with the expected waiting time is calculated for time varying exceedance probabilities. In NEVA, the proposed DE-MC-Bayesian approach is integrated into Salas and Obeysekera 2013 to provide an alternative approach for non-stationary return level-return period analysis with time varying

exceedance probability. The parameter estimation, uncertainty assessment, and sampling approach, as well as the log-likelihood test and Bayes factor computation remain similar in both design exceedance probability and time varying exceedance probability methods.

2.3 Results

In the following, NEVA is used for stationary and non-stationary extreme value analysis of annual temperature maxima from the Climatic Research Unit (CRU, New et al. 2000) monthly temperature data (1901-2009). Figure 2.5 displays the global areas in which temperature block maxima exhibit a significant trend at the 5% level and hence, non-stationary behavior (see dark red pixels in Figure 2.5). The white land areas correspond to locations that do not show a significant trend in the annual temperature maxima. NEVA utilizes the suggested non-stationary extreme value analysis algorithm (Figure 2.2) for the dark red pixels in Figure 2.5, and the stationary algorithm (Figure 2.1) in the rest of the pixels. The appropriate type of GEV (stationary or non-stationary) is fitted to each grid of monthly temperature maxima.

Figure 2.6 shows the global annual temperature maxima return levels for the 5-year (a) and 100-year (b) return periods. As mentioned earlier, NEVA generates an ensemble of estimates based on DE-MC sampling. The median of the ensemble is used as the final return level values shown in Figure 2.6. The uncertainty bounds of the computed return levels can be derived based on 5% and 95% confidence intervals of the ensemble as discussed below.

To further explore NEVA's outputs, two pixels in the central (Latitude 40.02 °N, Longitude 105.27 °W) and western (Latitude 34.05 °N, Longitude 118.24 °W) United States are selected for more detailed analysis (see green stars in Figure 2.4). The two locations are close to urban areas in Boulder, CO and Los Angeles, CA where long-term observation stations have been

available. The goodness-of-fit of the model at the local scale is assessed using the Quantile-Quantile (Q–Q) plots of temperature maxima (see Figure 2.3). In both locations, the Mann-Kendall trend test confirms presence of non-stationarity at the 5% significance level (see Figure 2.2). The initial goodness-of-fit of the GEV model is assessed using Quantile-Quantile (Q–Q) plots of fitted and observed temperature maxima (see Figure 2.3). The plot of the return levels versus the corresponding return periods at the two selected locations under both stationary (ignoring the observed trend) and non-stationary assumptions are displayed in Figure 2.7 and Figure 2.8. In both figures, the top panels (a) show return levels under the stationary assumption, while panels (b) exhibit non-stationary return levels for the observation period (here, 1901-2009). Panels (c) and (d) display non-stationary return levels for 100 years beyond observations (e.g., 2010-2109) using median and 95 percentile of sampled location parameters, respectively (see Equations 2.11 and 2.12). Consequently, panels (d) in Figure 2.7 and Figure 2.8 are more conservative estimates of future extreme return levels, and are termed as low risk (hereafter LR).

In the central U.S. (Figure 2.7a), under the stationary assumption the confidence bounds do not encompass the empirical return levels, which indicates the assumptions for this model are not met. On the other hand, for the non-stationary model (Figure 2.7b), the confidence bounds enclose the empirical return levels, indicating reasonable simulations (see the zoom in Figure 2.7b). The selected point in the western U.S. (Figure 2.8) exhibits a similar behavior. The non-stationary envelope of simulations (Figure 2.8b) encompasses all of the empirical return levels, while the stationary ensemble (Figure 2.8a) does not enclose all the points, including few points at the beginning and the last observation.

In the central U.S., the return levels under the stationary assumption (Figure 2.7a) are much lower than those under the non-stationary assumption (Figure 2.7c and 2.7d). For example, the

return levels corresponding to the 50-year annual temperature maxima (ensemble median - red dashed lines) are 14.3° C, 15.5° C and 16.1° C, under stationary, non-stationary, and LR non-stationary, respectively. This result indicates that an unrepresentative assumption of stationarity would lead to misinterpretation (in this example, underestimation) of extreme climatic conditions. Another example is the pixel in the western U.S., where the positive trend is not as strong as the one in the central U.S. (compare Figure 2.4a and 2.4b). Nonetheless, if the observed linear trend continues in the future, the return levels will be underestimated under the stationary assumption. Considering a 50-year return level (ensemble median - red dashed lines), it is 28.5° C (stationary), 29.1° C (non-stationary), and 29.4° C (LR non-stationary). It should be noted, that the annual maxima is based on mean monthly temperature values and the daily maxima may exceed these values.

Once the parameters are sampled and return levels are simulated, the non-stationarity assumption included in the location parameter is tested using the log-likelihood and Bayes factor approach discussed in Equation 2.6 and 2.7. In both the central and western U.S., the log-likelihood test and Bayes factor confirm that the simulations exhibit non-stationary behavior consistent with the Mann-Kendall test results (p-values smaller than the 0.05 significance level) – See Table 2.1.

As shown in the model flowchart (Figure 2.2), NEVA can provide non-stationary return periods based on a time varying exceedance probability. Figure 2.9 presents return period vs. return level under stationarity and the corresponding non-stationary return periods for the two selected points in the central and western U.S. In this framework, the exceedance probability q_t varies through time. Since temperature extremes exhibit an upward trend at both locations, the exceedance probability q_t will increase over time. The probability distribution of the waiting time for the first extreme event to exceed a given threshold is a generalization of the geometric distribution,

which enables determining the expected return period. For instance, in the western U.S., the 50-year return period under the stationary assumption corresponds to an approximate 30-year return period under a non-stationary condition (see Figure 2.9b). In other words, an exceedance probability of 0.02 will increase to around 0.03 in a non-stationary climate. Similarly, in the central U.S., a 50-year extreme changes to a 22-year event in a changing climate. This framework allows displaying stationary and non-stationary return periods against each other.

As shown in the model flowchart (Figure 2.2), NEVA generates non-stationary return levels based on both the standard definition in hydrology and the concept of effective return level. As an example, Figure 2.10 demonstrates effective return levels for the two selected points in the central and western U.S. The figures show return levels versus the time covariate used in the linear regression (Equation 2.2). In this concept, the return levels vary over time such that the probability of occurrence remains constant. Basically, the effective return level indicates what return level should be used for all years to have the same risk. In the western U.S., the effective return level corresponding to a 50-year (0.02 probability of occurrence) event during 1901-1950 is 28.5° C; whereas the same risk for a 100-year period (1901-2000) would be 28.8° C. Similarly, for another 100-year period (e.g., 2001-2100) the 50-year event would be different (here, 29.3° C). By providing both the standard and the effective return levels, as well as the integrated time-varying exceedance probability non-stationary return periods, NEVA allows the users to use the one that fits their application.

2.4 Conclusions

Substantial evidence shows that the climate is non-stationary, possibly due to anthropogenic climate change. The assumption of stationarity in extreme value analysis is therefore questionable and statistical models that explicitly allow for non-stationarity are much needed.

Specifically, statistical models that can provide estimates of return levels under non-stationary conditions are essential for design and risk assessment purposes. In this study, a practical package named Non-stationary Extreme Value Analysis (NEVA) package is introduced for assessing extremes in a changing climate.

NEVA offers a framework for estimating non-stationary return levels, return periods, and risks of climatic extremes using Bayesian inference. In this approach, the model parameters are estimated using a Differential Evolution Markov Chain (DE-MC) for global optimization over the real parameter space with the Markov Chain Monte Carlo (MCMC) approach. NEVA also provides the confidence interval and uncertainty bounds of estimated return levels by combining DE-MC with Bayesian inference. A unique feature of the model is non-stationary extreme value analysis using both design exceedance probability and time varying exceedance probability methods.

The features and capabilities of NEVA can be summarized as follows: (a) the framework assesses trends in the observations; (b) depending on the trend, it performs stationary or non-stationary analysis of extremes and can test which model describes the data more appropriately based on the model outputs; (c) it provides non-stationary return levels based on three methods including one that resembles the standard approach in hydrology under stationarity, one based on expected waiting time with time varying exceedance probability, and effective return levels; and (d) NEVA includes a sampling framework that leads to uncertainty bounds of simulations. The return level and return period estimates can be used in hydrology and climate studies to assess the risk (probability of occurrence) of extremes.

By providing confidence intervals (e.g., 5% and 95% quantiles), NEVA offers a range of return levels, and the user can select the upper bound (low risk) or the lower bound (high risk)

depending on the application at hand. Users can change the upper and lower bound quantiles of the simulated ensembles and also the significance level of the trend analysis component. Both stationary and non-stationary components of the package are evaluated using Climatic Research Unit (CRU) observations. The results indicate that NEVA simulates GEV-based return levels consistent with empirical observations. While the focus of this dissertation is on climate extreme value analysis, the methodology can potentially be used in different areas (hydrology, ecology, and economics) and with different data sets.

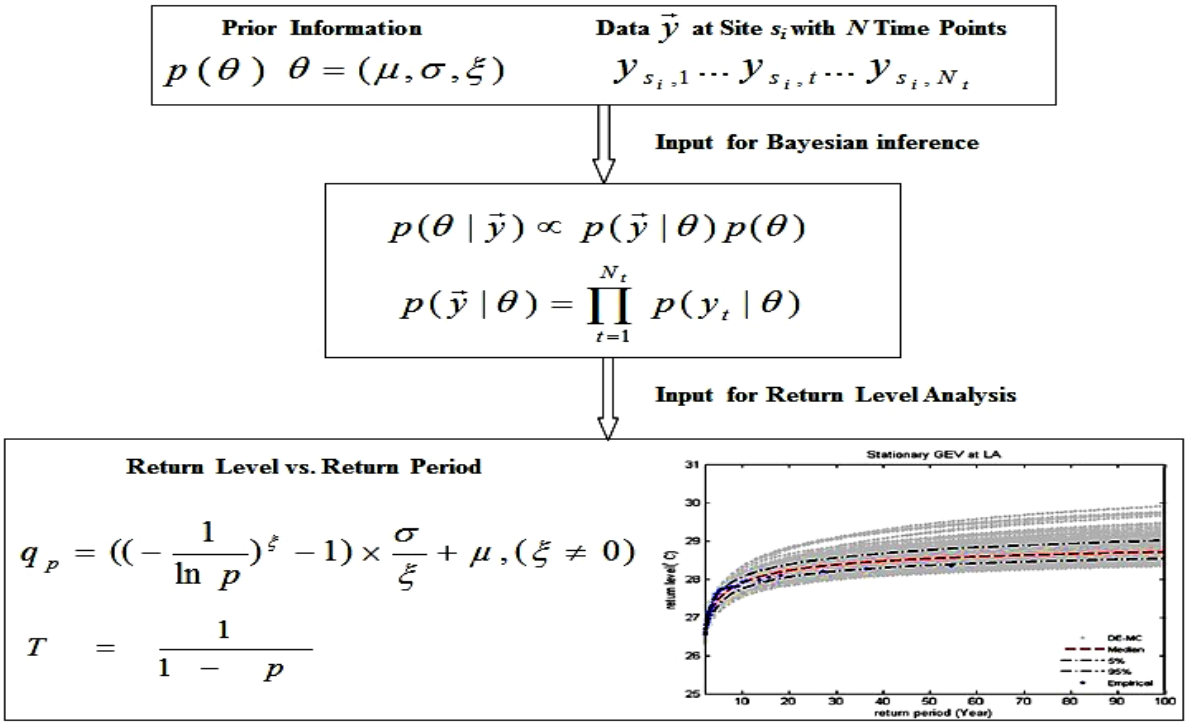

Figure 2.1 NEVA's stationary GEV framework for extreme value analysis. The outputs are return levels versus return periods.

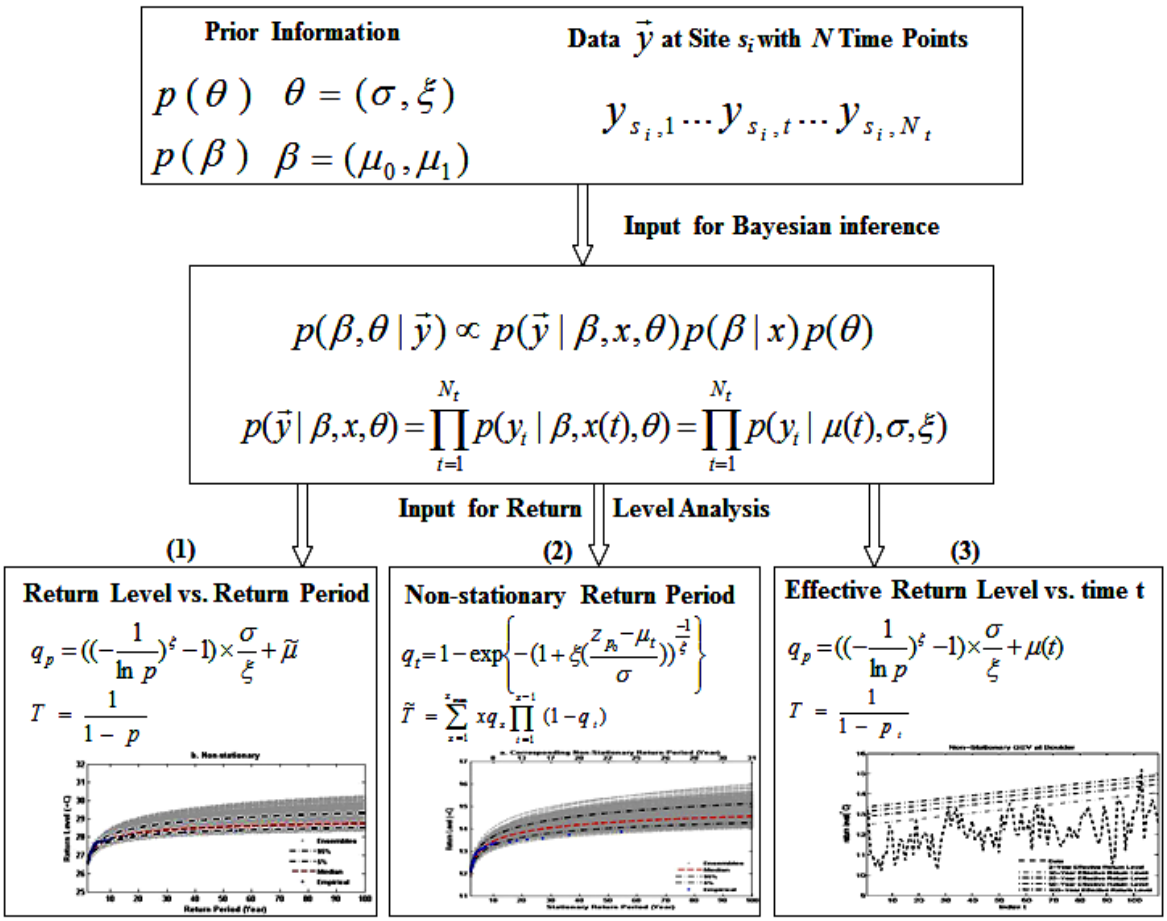

Figure 2.2 NEVA's non-stationary GEV framework for extreme value analysis. The model outputs include: (1) standard return levels with constant exceedance probability; (2) standard return levels with time varying exceedance probability; and (3) effective return levels.

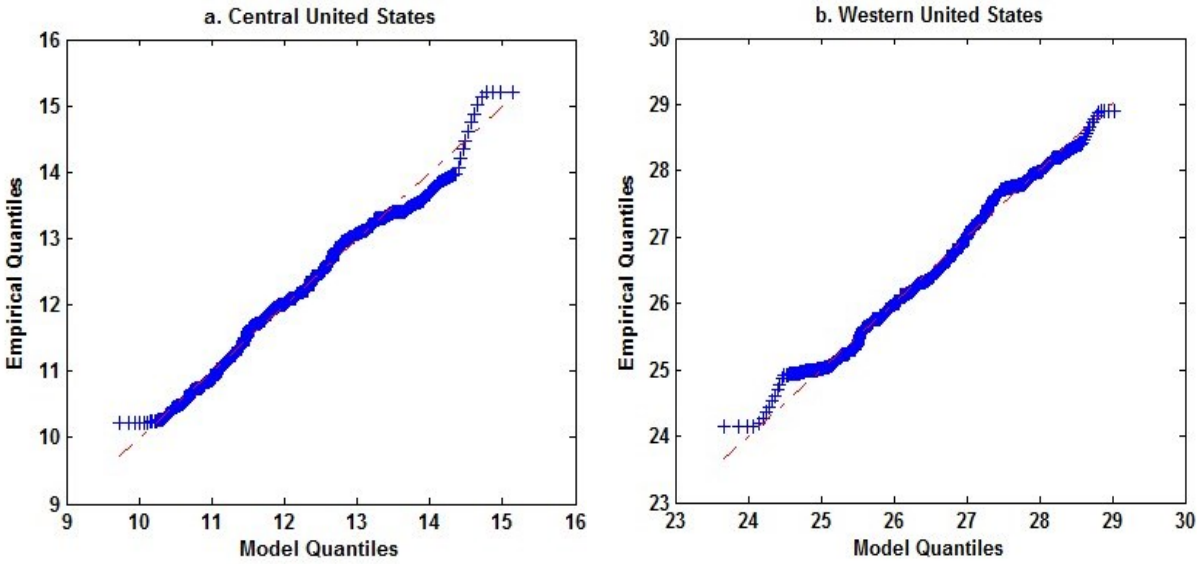

Figure 2.3 Model quantiles vs. empirical quantiles of the annual monthly temperature maxima (^{o}C) in the central (a) and western (b) United States.

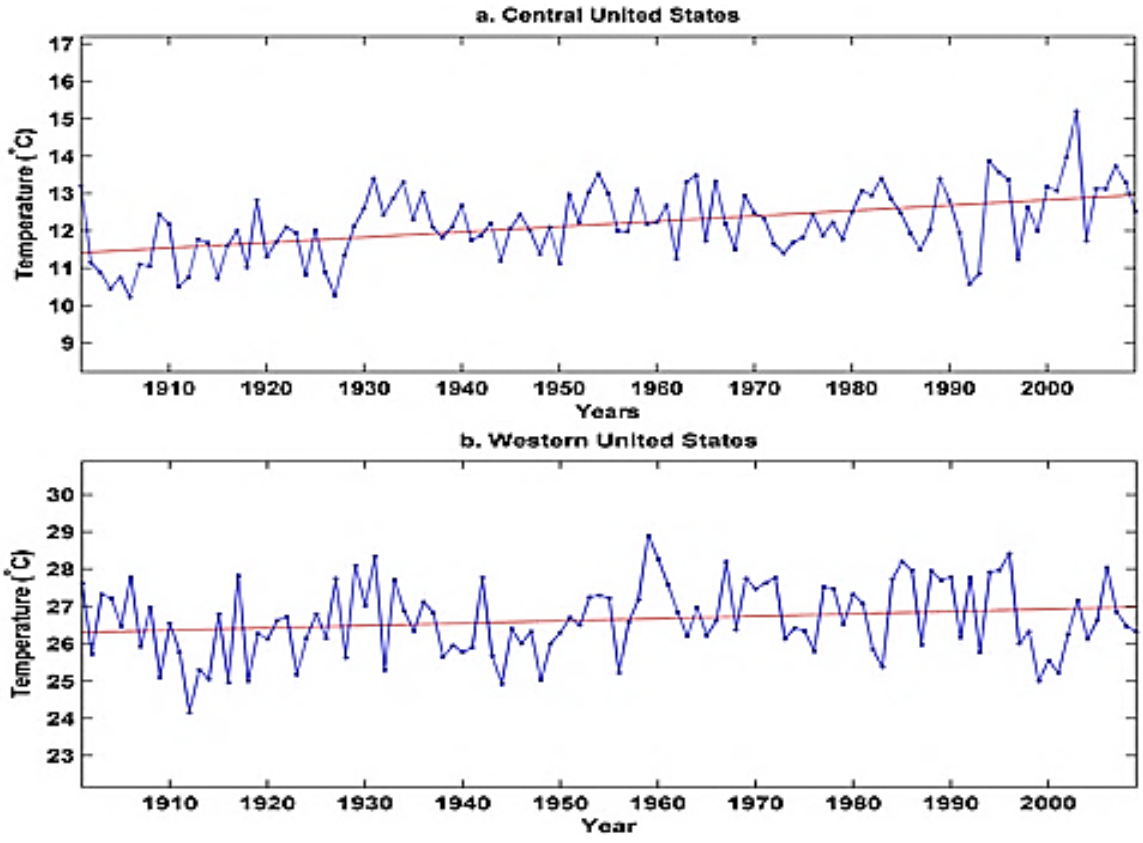

Figure 2.4 Trends in annual monthly temperature maxima in the central (a) and western (b) United States.

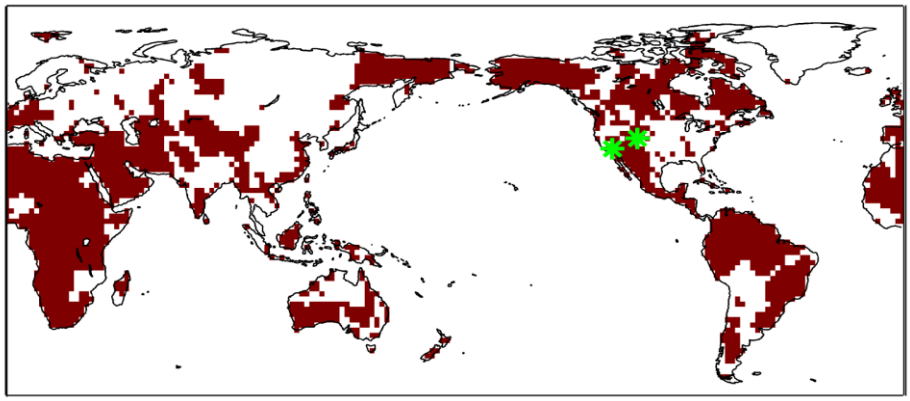

Figure 2.5 Global Mann-Kendall Trend Analysis (Significant trend in red; No significant trend in white). The Star-marked locations are the pixels selected for time series analysis.

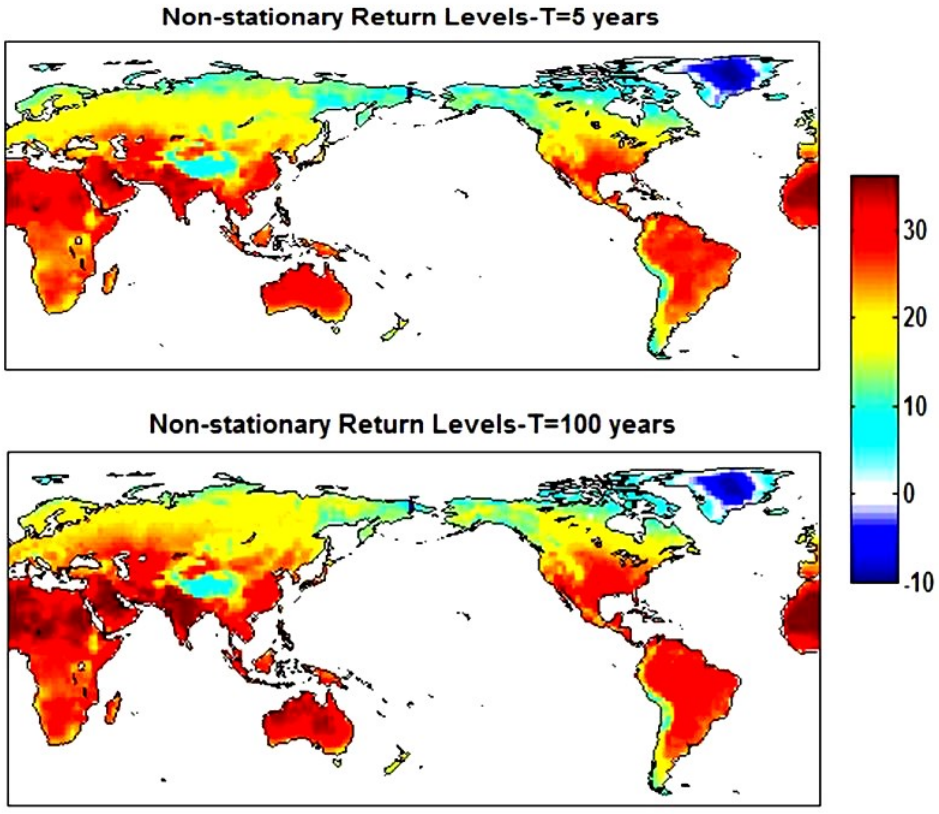

Figure 2.6 5-year (a) and 100-year (b) annual monthly temperature maxima return levels (^{o}C) under the non-stationary assumption, derived using the standard return levels with constant exceedance probability.

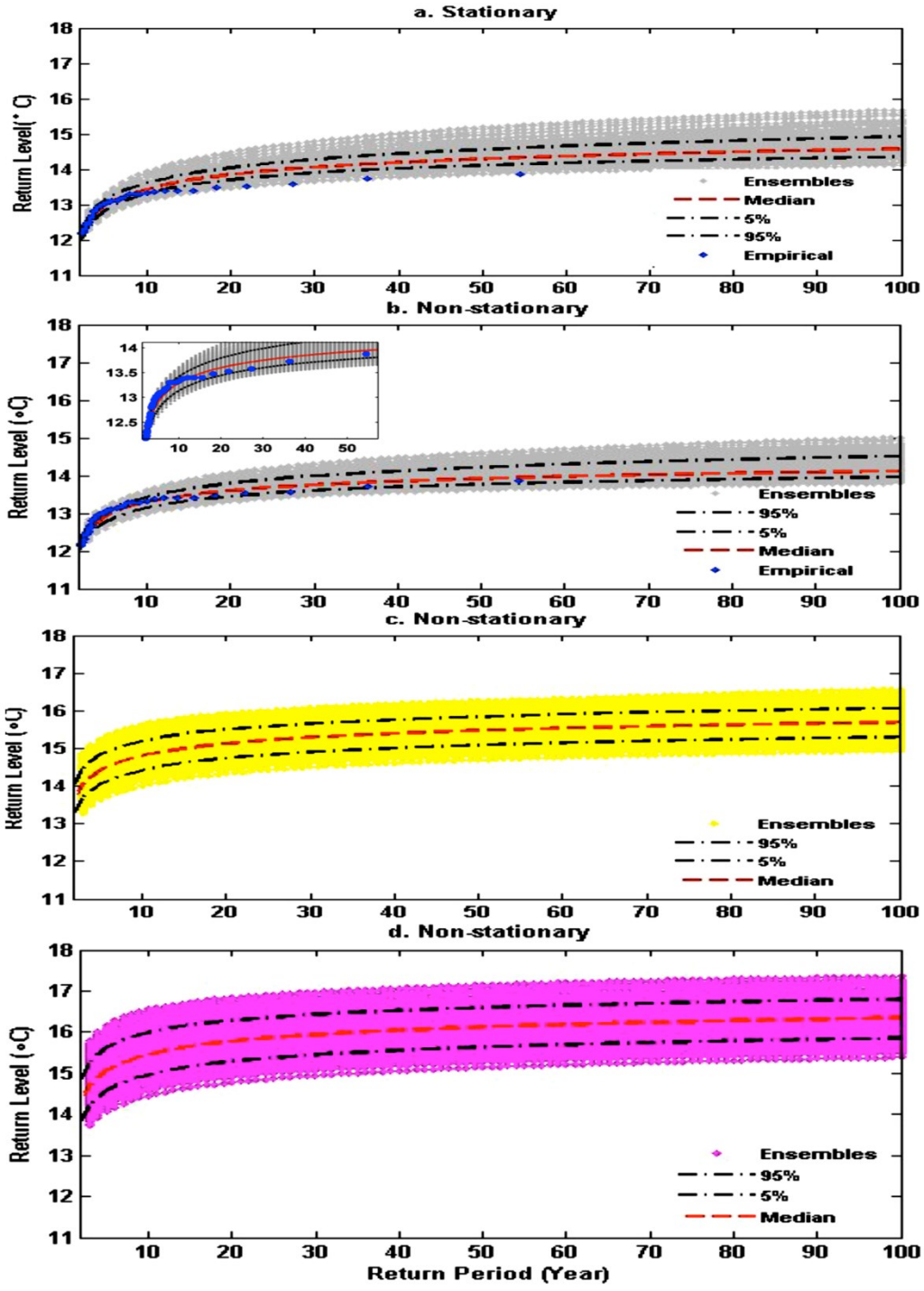

Figure 2.7 Annual monthly temperature maxima return level vs. return period in the selected point in the central U.S. under stationary (a), non-stationary during the period of observations 1901-2009 (b), non-stationary based on median of sampled parameters (c), and non-stationary based on the 95 percentile of the sampled parameters or LR non-stationary (d).

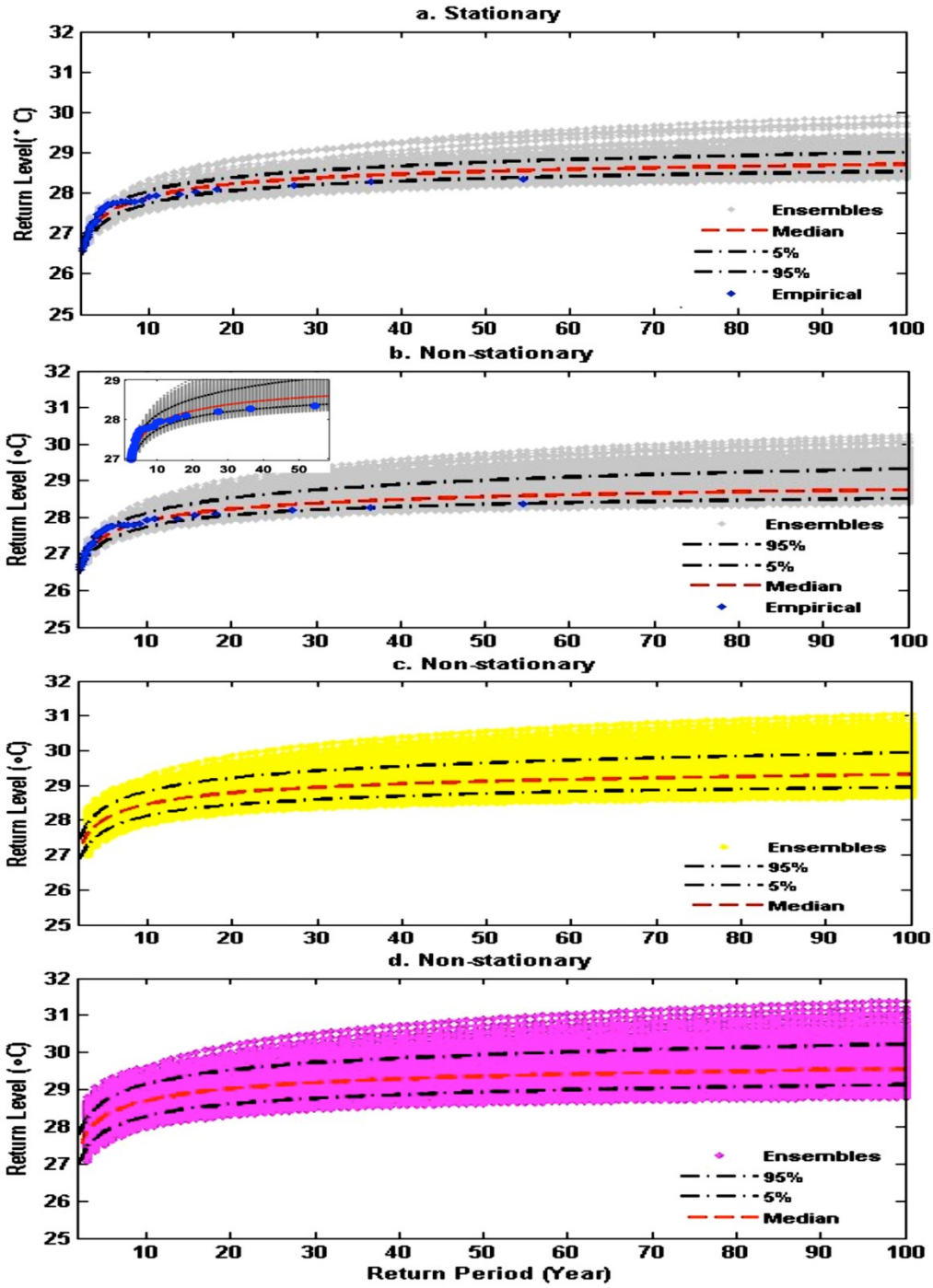

Figure 2.8 Annual monthly temperature maxima return level vs. return period in the selected point in the western U.S. under stationary (a) non-stationary during the period of observations 1901-2009 (b), non-stationary based on median of sampled parameters (c), and non-stationary based on the 95 percentile of the sampled parameters or LR non-stationary (d).

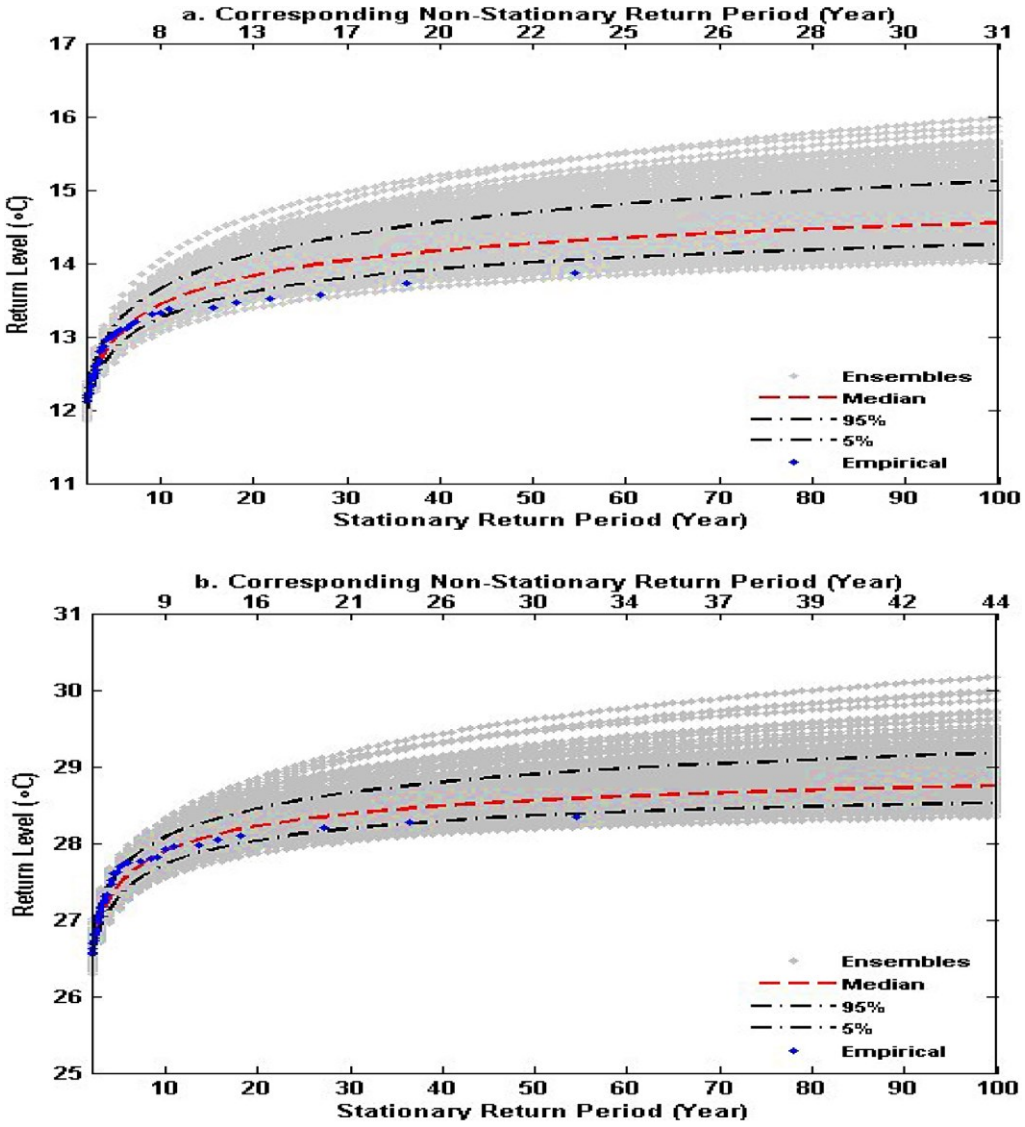

Figure 2.9 Annual monthly temperature maxima return levels vs. return period under stationary (bottom axes) and the corresponding non-stationary (top axes) assumption at the selected points in the central (a) and western (b) United States.

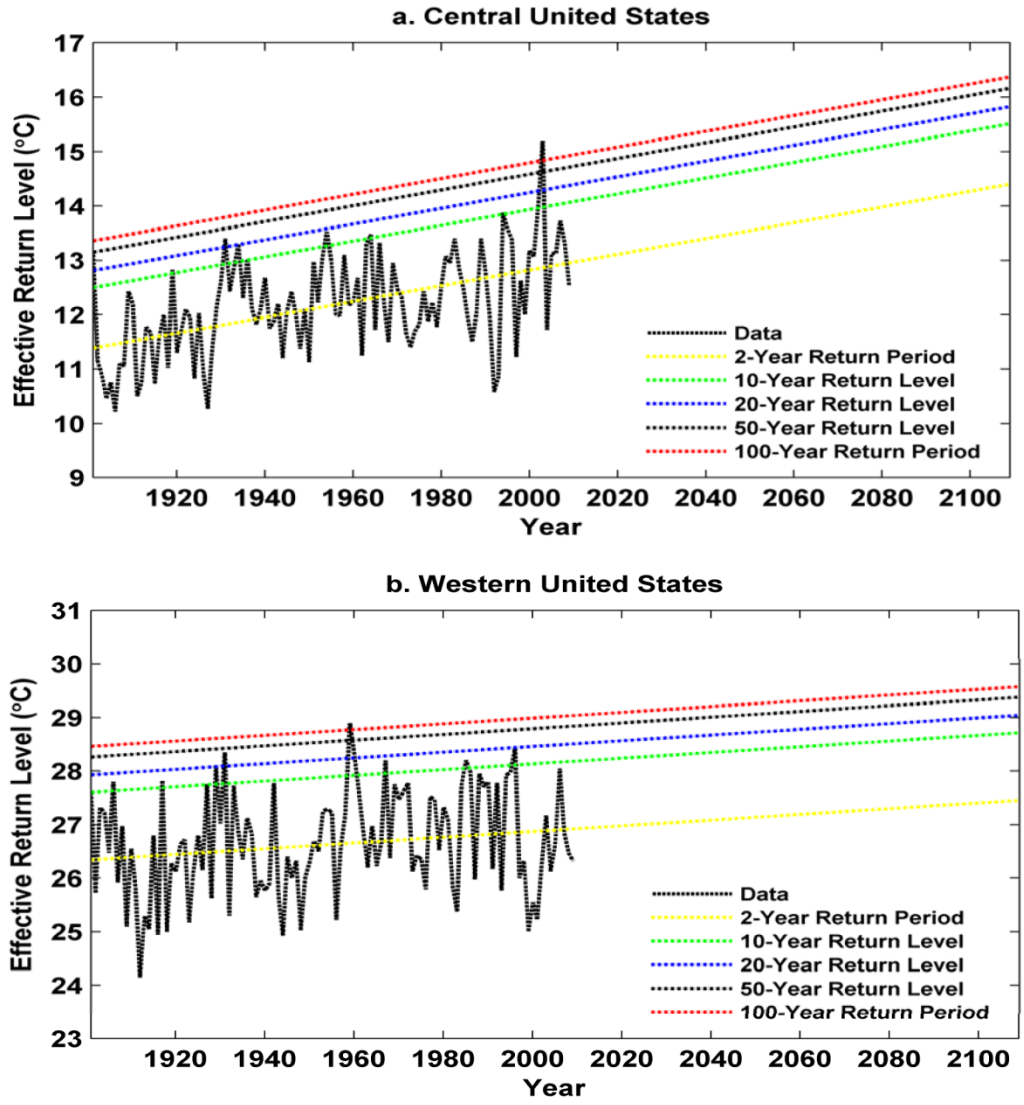

Figure 2.10 Effective return levels under the non-stationary assumption at the selected points in the central (a) and western (b) United States.

Table 2.1 Results of the Likelihood Ratio Test and Bayes Factor at the selected pixels in the central and western United States. H-statistics = 1 indicates that the null hypothesis (i.e., stationary model) is rejected (p-value less than the significance level α=0.05), while H-statistics = 0 means that the null hypothesis cannot be rejected. Bayes Factor K < 1 means the null hypothesis (i.e., stationary model) is not in favor.

Location	Likelihood Ratio Test			
	D	*p* -value	H-statistics	Test Interpretation
Central U.S.	29.06	7.00e-8	1	Reject *Stationary Model*
Western U.S.	5.16	0.02	1	Reject *Stationary Model*
Location	Bayes Factor			
	K	Interpretation		
Central U.S.	1.18e-6	Reject *Stationary Model*		
Western U.S.	0.01	Reject *Stationary Model*		

Table 2.2 NEVA and extRemes 2.0 Comparison.

Stationary	**extRemes 2.0**			**NEVA**		
Western U.S.	μ	σ	ξ	μ	σ	ξ
2.50%	26.06	0.84	-0.42	26.11	0.86	-0.39
Mean	26.3	0.99	-0.29	26.3	0.98	-0.31
97.50%	26.53	1.18	-0.14	26.5	1.12	-0.18
$\hat{R}$	NA			1.001	1.003	1.001
Acceptance Rate	0.66			0.41		
Return Level (˚ C)	2-yr	20-yr	100-yr	2-yr	20-yr	100-yr
2.50%	26.4	28	28.48	26.4	27.61	27.93
Mean	26.65	28.28	28.84	26.64	28.22	28.73
97.50%	26.89	28.67	29.55	26.9	29.09	30.03
Central U.S.	μ	σ	ξ	μ	σ	ξ
2.50%	11.61	0.78	-0.27	11.64	0.8	-0.26
Mean	11.82	0.91	-0.18	11.81	0.89	-0.18
97.50%	12.02	1.06	-0.05	11.98	0.98	-0.08
$\hat{R}$	NA			1.003	1.006	1.007
Acceptance Rate	0.65			0.40		
Return Level (˚ C)	2-yr	20-yr	100-yr	2-yr	20-yr	100-yr
2.50%	11.92	13.59	14.23	11.92	13.3	13.8
Mean	12.14	13.91	14.67	12.12	13.86	14.59
97.50%	13.91	14.36	15.5	12.34	14.58	15.77

Non-Stationary	**extRemes 2.0**				**NEVA**			
Western U.S.	σ	ξ	μ_0	μ_1	σ	ξ	μ_0	μ_1
2.50%	0.84	-0.41	25.58	0.001	0.85	-0.39	25.83	0.003
Mean	0.97	-0.28	25.98	0.006	0.96	-0.28	26	0.005
97.50%	1.16	-0.13	26.31	0.01	1.09	-0.13	26.18	0.008
$\hat{R}$	NA				1.001	1.007	1.001	1.001
Acceptance Rate	0.47				0.43			
Effective Return Level (˚ C)	2-yr	20-yr	100-yr		2-yr	20-yr	100-yr	
t= 50 years (1901~1950)	26.61	28.22	28.77		26.6	28.19	28.72	
Central U.S.	σ	ξ	μ_0	μ_1	σ	ξ	μ_0	μ_1
2.50%	0.69	-0.34	10.73	0.009	0.71	-0.3	10.88	0.011
Mean	0.82	-0.23	11.08	0.014	0.79	-0.22	11.09	0.015
97.50%	0.98	-0.11	11.42	0.020	0.89	-0.13	11.25	0.018
$\hat{R}$	NA				1.02	1.004	1.01	1.008
Acceptance Rate	0.44				0.43			
Effective Return Level (˚ C)	2-yr	20-yr	100-yr		2-yr	20-yr	100-yr	
t= 50 years (1901~1950)	12.09	13.57	14.13		12.09	13.52	14.07	

CHAPTER 3: Non-stationary Return Levels of CMIP5 Multi-Model Temperature Extremes

3.1 Introduction

Losses of life and economic damage due to extreme weather and climate events have been steadily increasing since the 1930's in the United States (Easterling et al., 2000). During the period 1979 - 1992, for example, on average 384 people were killed by excessive heat each year (NOAA 1995; Kilbourne 1997). In fact, in this period over the United States, excessive heat accounted for more reported deaths annually than hurricanes, floods, tornadoes, and lightning combined (NOAA 1995). Furthermore, agriculture products such as wheat, rice, corn and maize can be significantly reduced by extreme high temperatures at key development stages (NOAA 1980). Numerous studies indicate that climate extremes are likely to intensify in the future under different plausible climate scenarios (Alexander et al., 2006; IPCC 2007); AghaKouchak et al. 2013). Therefore, there is a need to study extreme weather and climate events across different spatial and temporal scales.

Currently, some 20 international climate modeling groups are providing Coupled Model Intercomparison Project Phase 5 (CMIP5) historical and projected climate simulations (Taylor et al., 2012). The scope of CMIP5 also is broader than previous model intercomparison projects (e.g. CMIP3), with carbon emission-driven Earth System Model (ESM) experiments now represented along with the typical concentration-driven atmosphere-ocean general circulation model (AOGCM) simulations (Meehl and Bony 2011). More than previous model intercomparisons, CMIP5 also includes AOGCM simulations incorporating aerosol chemistry, as well as time-slice experiments performed with high-resolution (approx. 25 km horizontal grid) atmosphere-only models (Meehl and Bony 2011). Thus, the multi-model gridded CMIP5

datasets provide an unprecedented opportunity to analyze and assess climate variability and change.

In a recent study, Kharin et al. (2013) argued that the global warm temperature extremes in the late 20th century climate are reasonably simulated by the CMIP5 models (differences CMIP5 models and reanalysis data were within a few degrees C over most of the globe). However, the inter-model differences of warm temperature extremes are generally large over land with a standard deviation of around 4 °C (Kharin et al., 2013). Kharin et al. (2013) concluded that upward trends of warm extremes exceed those of cold extremes over tropical and subtropical land regions. Morak et al. (2013) showed that there is a significant increase in the trend in warm temperature extremes during both boreal cold and warm seasons over the second half of the 20th century.

Return periods and return levels (also known as return values) are often used to describe and assess risk (Cooley et al., 2007; AghaKouchak and Nasrollahi 2010; Katz 2010; Cooley 2013). In theory, the return period (T) of an event is the inverse of its probability of occurrence in any given year. That is, the n-year return level corresponds to an exceedance probability (by an annual extreme) of 1/n. In the statistical literature, there are different definitions for return period and return level; for alternative definitions, the interested reader is referred to Bonnin et al. (2004), Mays (2010), and AghaKouchak et al. (2013).

In recent years, Extreme Value Theory (EVT) has been widely used for analysis of climate extremes and their return levels (Zwiers and Kharin, 1998; Clarke 2000; Katz et al., 2000); Kharin and Zwiers, 2005; Kunkel 2013; Parey et al., 2010; Cooley 2013). Fisher and Tippett (1928) introduced the concept of asymptotic theory in extreme value distributions and laid the foundation for a generalized approach to extreme value analysis. Gnedenko (1943)

mathematically proved that three families of extreme value distributions- namely Weibull, Gumbel and Fr´echet can represent the limiting distributions of extremes in random variables. The Generalized Extreme Value (GEV) distribution is essentially a combination of these three distribution families, and has been applied in a variety of studies (Gumbel 1942; Smith 2001; Katz 2013).

Numerous studies indicate that climatic extremes (e.g., hot days, heavy precipitation) have increased significantly, particularly in the second half of the 20th century (Karl and Knight, 1998; Easterling et al., 2000; Vose et al., 2005; Hansen et al., 2010; Villarini et al., 2011; Hao and AghaKouchak 2013; Wehner 2012; Field et al., 2012). In addition to the number, the frequency of extremes has been changing in the past, and it is likely to change in future (Milly et al., 2008; Easterling et al., 2000; IPCC 2007). Evidently, ignoring time-varying (non-stationary) behavior of extremes could lead to underestimation of extremes and considerable damages to human life and society (McMichael, 2003). Therefore, it is necessary to assess non-stationarity in the CMIP5 climate models simulations, and to document to what extend the patterns are consistent with observations.

In this study, the GEV is used to investigate the return levels of annual monthly temperature maxima considering a changing climate. The primary objective of this dissertation is to evaluate to what extent the CMIP5 model simulations of the historical climate of the period 1901-2005 can represent observed warm monthly temperature extremes under the non-stationary assumption. The return levels of temperature maxima estimated from the CMIP5 climate simulations are compared with those of Climatic Research Unit (CRU) temperature observations. The dissertation is organized as follows. After this introduction, the data sets and study area are discussed. Section 3.3 presents the same methodology for non-stationary extreme value analysis

in Chapter 2.2. The results including representation of annual maxima in their return levels are provided in Section 3.4. The last section summarizes the main results and offers concluding remarks.

3.2 Study Area and Data Resources

Monthly observations of temperature provided by the Climatic Research Unit (CRU, New et al., 2000; Mitchell and Jones, 2005), available in a 0.5^{o} spatial resolution, are used as reference data. CRU observations have been validated and used in numerous studies of historical climate variability (e.g., Tanarhte et al., 2012). In this study, 41 CMIP5 historical monthly temperature simulations from 1901 to 2005 are subjected to extreme value analysis, and a subset of 17 of these simulations are investigated in more detail. These data sets represent the most extensive and ambitious multi-model simulations that contribute to the World Climate Research Programme's CMIP multi-model dataset (Meehl and Bony, 2011; Taylor et al., 2012). For this extreme value analysis, the CMIP5 model simulations and CRU observations all are regridded to a common 2 × 2-degree resolution. This regridding entailed use of bilinear interpolation, with special attention given to appropriate use of model-specific land fraction masks so as to minimize data distortions along coastlines. The selected models (Table 3.1) include physical climate models (without a prognostic global carbon cycle), as well as earth system models (with the designation "ESM" appearing in the model title). The former are run in a standard "historical climate" configuration with prescribed historically increasing CO_2 concentrations (i.e. with the prognostic carbon cycle turned "off"), and the latter are run either with prescribed atmospheric CO_2 concentrations or with CO_2 emissions (fluxes corresponding to the prescribed historically increasing CO_2 concentrations–designated as "esm" experiments in Table 3.1).

This study focuses on global land areas (excluding Antarctica) for which the CRU observations are provided. From CRU observations and CMIP5 simulations, pixel-based annual monthly temperature maxima (hereafter, annual temperature maxima) are extracted for estimation of extreme temperature return levels. It should be noted that the Hadley Centre has adopted an unconventional time model for all their CMIP5 outputs, with an endpoint in November rather than December of 2005, and thus HadGEM2 models have one month less data compared to the other models. This issue will not affect the results in the Northern Hemisphere since the annual maxima of monthly data do not often occur in December. However, it might slightly impact the analyses in the Southern Hemisphere.

3.3 Methodology

Refer to the Chapter 2.2.

3.4 Results

3.4.1 Representation of Annual Temperature Maxima

In the first step, the annual maxima of CMIP5 temperature simulations, determined from monthly time series, are compared with those of the CRU observations. Figure 3.1 (top left) displays the mean annual temperature maxima from 1901 to 2005 as represented by CRU observations. The rest of the panels in Figure 3.1 demonstrate the differences between CMIP5 climate simulations and CRU observations (CMIP5 model - observation). In this figure, positive (negative) values indicate overestimation (underestimation) of the annual temperature maxima. Figure 3.1 shows the results for the 17 CMIP5 models listed in Table 3.1. One can see that the climate models individually display different patterns of overestimation and underestimation. In most parts of the globe, the discrepancies between the model simulations and observations are

within ± 1 to 3 °C, indicating a reasonable overall agreement between model simulations and CRU observations. However, local errors for some regions may be as high as ±10 °C (see also the empirical cumulative distribution of the mean error in Figure 3.2).

Figure 3.1 shows that over the United States most models (excepting HadGEM2-ES_esm, CCSM4, CSIRO-ACCESS1-0, CESM1-WACCM, MIROC-ESM and CanESM2) tend to underestimate the mean annual temperature maxima by 1 to 3 °C. Here, CanESM2 instead substantially overestimates the mean annual temperature maxima. Over Australia, on the other hand, several models (e.g. CSIRO-ACCESS1-0, HadGEM2-ES_esm, MPI-ESM-P and CanESM2) demonstrate little or no bias. Over Amazonia, the mean annual temperature maxima are underestimated in most models, except in few models (e.g., GFDL-CM3, CanESM2) where they are overestimated.

The results indicate that model simulations particularly diverge from one another over cold regions (e.g., northern Russia, and Canada) except for Greenland, where most models (but not MIROC-ESM and INMCM4_esm) underestimate the mean annual temperature maxima. Such a consistent underestimation could substantially impact model-based analyses of changes in ice-sheets, and snow/glacier melt studies. Krabill et al. (2004) reported that Greenland is losing coastal ice sheets quite rapidly (see also Ren et al., 2011; Kjær et al., 2012). CMIP5 models' underestimation of annual maxima climatology implies that the ice loss rate in Greenland might be greater than that reported in model-based studies. Similar to the modeling results by Alley et al. (2005) and Reeh (1989), rapid ice-marginal changes may indicate greater ice-sheet sensitivity to warming than has been considered previously. However, over other cold regions that are at most risk of accelerated ice melt (e.g. Alaska, Northern Canada, and Siberia), most models tend

to overestimate the mean annual temperature maxima relative to the CRU reference data (Figure 3.1).

It is also noteworthy that the model simulations collectively underestimate the mean annual maxima over arid and semi-arid regions (e.g., Sahara, southwestern U.S.), that are most subject to severe heat waves and droughts. Considering the magnitudes of deviations from the CRU, there is a better agreement between CMIP5 simulations and observations in such subtropical regions than in high-latitude cold regions. This is consistent with the findings reported in Kharin et al. (2007) based on CMIP3 climate model simulations.

Figure 3.2 displays the ensemble mean (top left), inter-model standard deviation (top right), and range (bottom) of the annual temperature maxima in CMIP5 simulations. The figure shows that the inter-model variability and range of simulations are more variable over Siberia, the western United States, and parts of the Middle East and Sahara compared to other regions.

3.4.2 Return Levels of Temperature Extremes

Using the annual temperature maxima from CMIP5 multi-model simulations and CRU observations, temperature return levels are derived for different return periods by fitting the appropriate type of GEV (stationary/non-stationary) to the block maxima of temperature extremes. Return levels of annual temperature maxima are derived and reported for the return periods T of 2, 10, 25, 50, and 100 years.

As an example, Figure 3.3 shows the 2-year temperature return levels based on CRU observations (top left) and on the selected subset of 17 CMIP5 climate model simulations. In Figure 3.3, the global temperature values range from -11 to 35 °C. Overall, Figure 3.3 indicates that there are good agreements between the observed and CMIP5 simulated spatial patterns of 2-

year annual temperature maxima, but that the magnitudes of 2-year annual temperature maxima represented by the selected CMIP5 models differ substantially.

Figure 3.4 presents the differences in CMIP5 simulated 2-year annual temperature maxima with respect to CRU observations. One can see that there are variations in both the magnitude and sign of the error of 2-year return levels across CMIP5 climate simulations. This implies that CMIP5 climate models capture the spatial patterns of temperature extremes well; however, individual models may be biased with respect to observations. As shown, over most parts of the world, the biases are within ± 4 °C. For a higher return level, one expects the differences in temperature simulations to increase relative to observations. For example, Figure 3.5 presents the differences in 25-year-return annual temperature maxima, as simulated by CMIP5 models with respect to CRU temperature observations. As shown, the patterns of differences remain similar, but the range of differences between simulated and observed annual temperature maxima increases at 25-year return level relative to the 2-year return level.

As another example, Figure 3.6 displays the 100-year return levels for the CRU observations and the selected CMIP5 simulations. One can see that the patterns of annual temperature maxima are similar to those of Figure 3.3, but with higher magnitudes of annual temperature maxima (as expected). The figure shows that the warmest months across the globe typically occur over the Sahara, the Middle East, and Australia. The differences in CMIP5 100-year simulated and observed annual temperature maxima are presented in Figure 3.7. As shown, the biases of the 25-, and 100-year return temperature simulations are larger than those of 2-year-return simulations in Figure 3.4. However, the spatial patterns of temperature extremes are in a good agreement with CRU observations and consistent across different return periods (compare the model simulations with the upper left panels in Figure 3.3 and 3.6). Overall, the regional biases of simulated annual

temperature extremes at high return levels (e.g., 100-year) are consistent with those of the lower return levels (e.g., 2-year).

Not shown here for brevity are the spatial patterns and biases of 10-, 25- and 50-year return levels of extreme temperature simulations by CMIP5 models, which are consistent with the results presented in Figure 3.3 to 3.7. For a quantitative evaluation of the extremal simulation by CMIP5 models, Figure 3.8 (top) summarizes the Mean Error (ME) for all the 41 CMIP5 climate model simulations of 2-, 10-, 25-, 50, and 100-year annual temperature maxima return levels with respect to CRU observations. As anticipated, ME values are larger at higher return levels. One can see that considering the global averages, most models overestimate the simulated return levels of the annual temperature maxima, while fewer models (e.g., FGOALS-g2, INMCM4_esm, NorESM1-ME) underestimate the temperature extremes. Among the models, FGOALS-s2, CanESM2 and MIROC5 exhibit the highest global averages of the ME of the annual temperature maxima. Most models either systematically overestimate or underestimate the extreme return levels, except the BCC model experiments in which the shorter return levels (2- and 10-year) are underestimated and the longer ones are overestimated. Figure 3.8 (bottom) displays boxplots of the differences between CMIP5 simulations and CRU observations. The figure shows medians, 25th and 75th percentiles, and whiskers (variability outside respective percentiles) of differences in Celsius degrees. Figure 3.8 (bottom) indicates that while local differences can be large, most differences (between 25th and 75th percentiles) fall within ± 2 Celsius degrees.

The MCMC component of the DE-MC model used in this study allows the upper and lower bounds and confidence intervals of the temperature return levels to be derived. The uncertainty bounds would be different across either models or space (simulation grid boxes). As an example,

Figure 3.9 shows sample uncertainty bounds, median, and the 5% and 95% confidence bounds of the annual temperature maxima based on the DE-MC for CRU reference data and over grid boxes in two different locations in the western and central United States under the non-stationary assumption. The figure confirms that the inference uncertainty is larger at higher return levels (e.g., because of larger sampling errors). One can see that the uncertainties of the estimated return levels also vary over different regions. It should be noted that this approach provides uncertainties associated with the statistical analysis of extremes, but does not include uncertainties associated with model physics.

3.5 Discussion and Concluding Remarks

The objective of this study is to evaluate to what extent the CMIP5 climate model simulations can represent observed warm monthly temperature extremes under a changing climate. The biases of simulated annual temperature maxima are quantified for the selected CMIP5 models. Furthermore, the 2-, 10-, 25-, 50, and 100-year return levels of the annual temperature maxima from CMIP5 simulations are compared with those derived from CRU observations.

The results show that most, but not all, CMIP5 climate models tend to underestimate the mean annual temperature maxima over the United States and Amazonia. The CMIP5 models particularly disagree with each other over cold regions (e.g., Russia, northern Canada), with the exception of Greenland where most climate models underestimate the mean annual temperature maxima. This underestimation of the annual temperature maxima is likely to affect model-based representations of changes in ice-sheets and snow/glacier melt. In contrast, over Alaska, Northern Canada and Siberia, most CMIP5 simulations overestimate the annual temperature maxima compared to those derived from the CRU reference data.

Over arid and semi-arid regions (e.g., the Sahara, southwestern U.S., and Middle East), most climate models also underestimate the mean annual temperature maxima. Considering the magnitudes of deviations from the CRU, however, there is a better agreement between CMIP5 model simulations and observations in subtropical regions than in high-latitude cold regions.

The return level analyses show that there are good agreements between the observed and CMIP5 simulated spatial patterns of 2-, 10-, 25-, 50- and 100-year annual temperature maxima. While the simulated spatial patterns of the temperature extremes are similar, the magnitudes of the return levels of the annual temperature maxima represented by CMIP5 climate models are biased with respect to CRU observations. In addition, there are variations in both the magnitude and sign of the biases of the annual temperature maxima return levels across the CMIP5 simulations. The results reveal that most CMIP5 simulations overestimate the global averages of the annual temperature maxima at different return periods (see Figure 3.8).

Given the state of the science in climate modeling, one would not expect the coupled Atmosphere/Ocean General Circulation Model (AOGCMs) and earth system models (ESMs) to reproduce the magnitudes of the observed historical extremes very accurately. Rather, one expects the models to reasonably simulate large-scale patterns of change in occurrences of climate extremes (Tebaldi et al., 2006). Overall, the results of this study indicate that the models capture the spatial patterns of temperature extremes well, but that individual models are biased relative to the CRU observations.

Table 3.1 List of 17 climate models whose simulations are displayed in Figures 3.1 to 3.7 and their related institutions and countries

Models	Institution	Country
BCC-CSM1-1_esm	Beijing Climate Center, China Meteorological Administration	China
MIROC-ESM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute, The University of Tokyo National Institute for Environmental Studies	Japan
NorESM1-M	Norwegian Climate Centre	Norway
IPSL-CM5A-LR	Institut Pierre-Simon Laplace	France
GFDL-CM3	NOAA Geophysical Fluid Dynamics Laboratory	USA
CCSM4	National Center for Atmospheric Research	USA
GISS-E2-H	NASA Goddard Institute for Space Studies	USA
INMCM4_esm	Institute for Numerical Mathematics	Russia
HadGEM2-ES_esm	Met Office Hadley Centre	UK
CSIRO-ACCESS1-0	Commonwealth Scientific and Industrial Research Organisation, and Bureau of Meteorology	Australia
MRI-ESM1_esm	Meteorological Research Institute	Japan
MPI-ESM-P	Max Planck Institute for Meteorology	Germany
CanESM2	Canadian Centre for Climate Modelling and Analysis	Canada
FGOALS-g2	Institute of Atmospheric Physics Chinese Academy of Sciences and Tsinghua University	China
CESM1-CAM5	National Science Foundation, Department of Energy, and National Center for Atmospheric Research	USA
CNRM-CM5	Centre National de Recherches Meteorologiques Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique	France
CESM1-WACCM	National Science Foundation, Department of Energy, and National Center for Atmospheric Research	USA

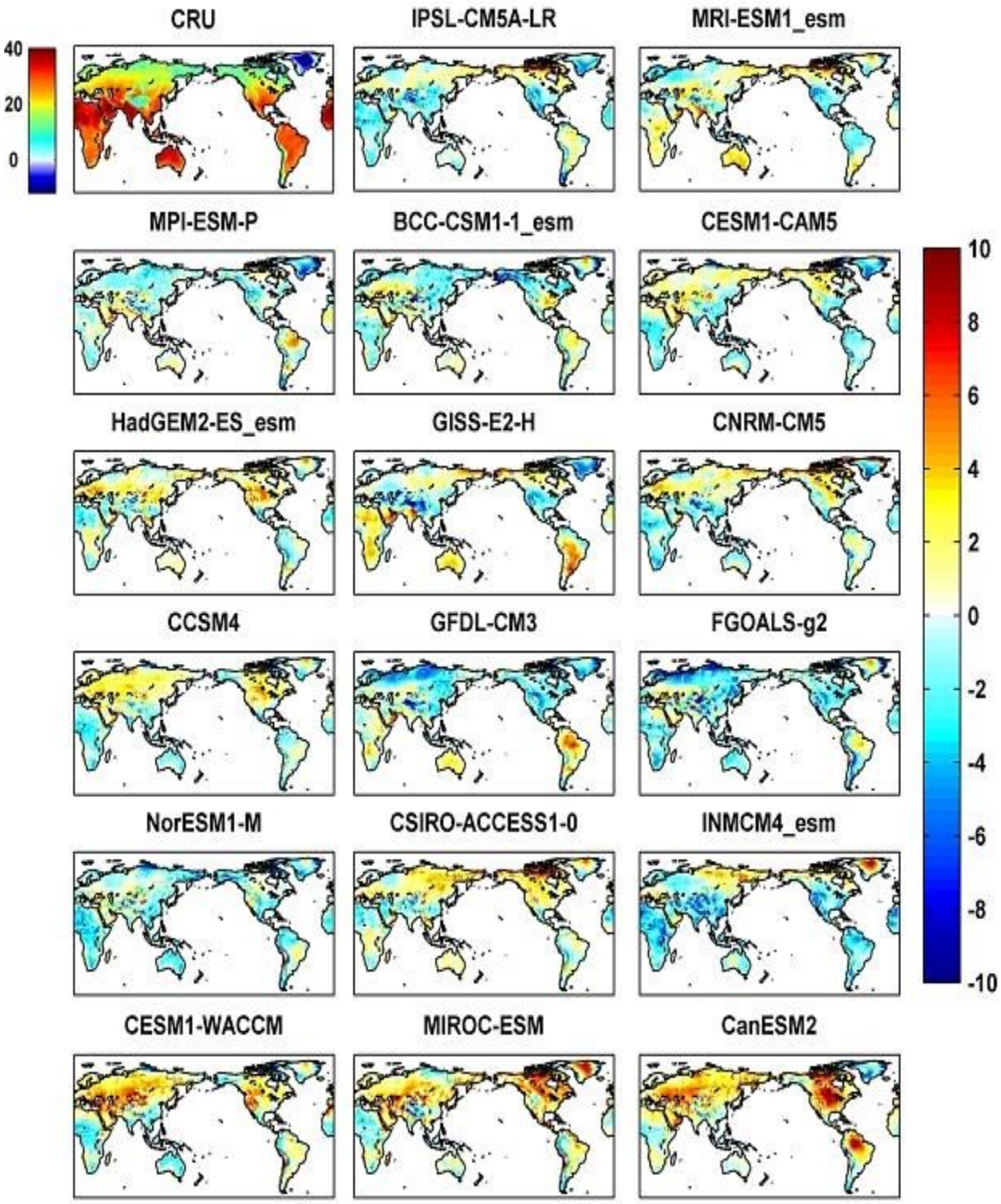

Figure 3.1 Mean annual temperature maxima (in degrees Celsius) based on 1901-2005 Climatic Research Unit (CRU) observations (upper left panel), and the differences between selected CMIP5 climate simulations and CRU reference data (remaining panels).

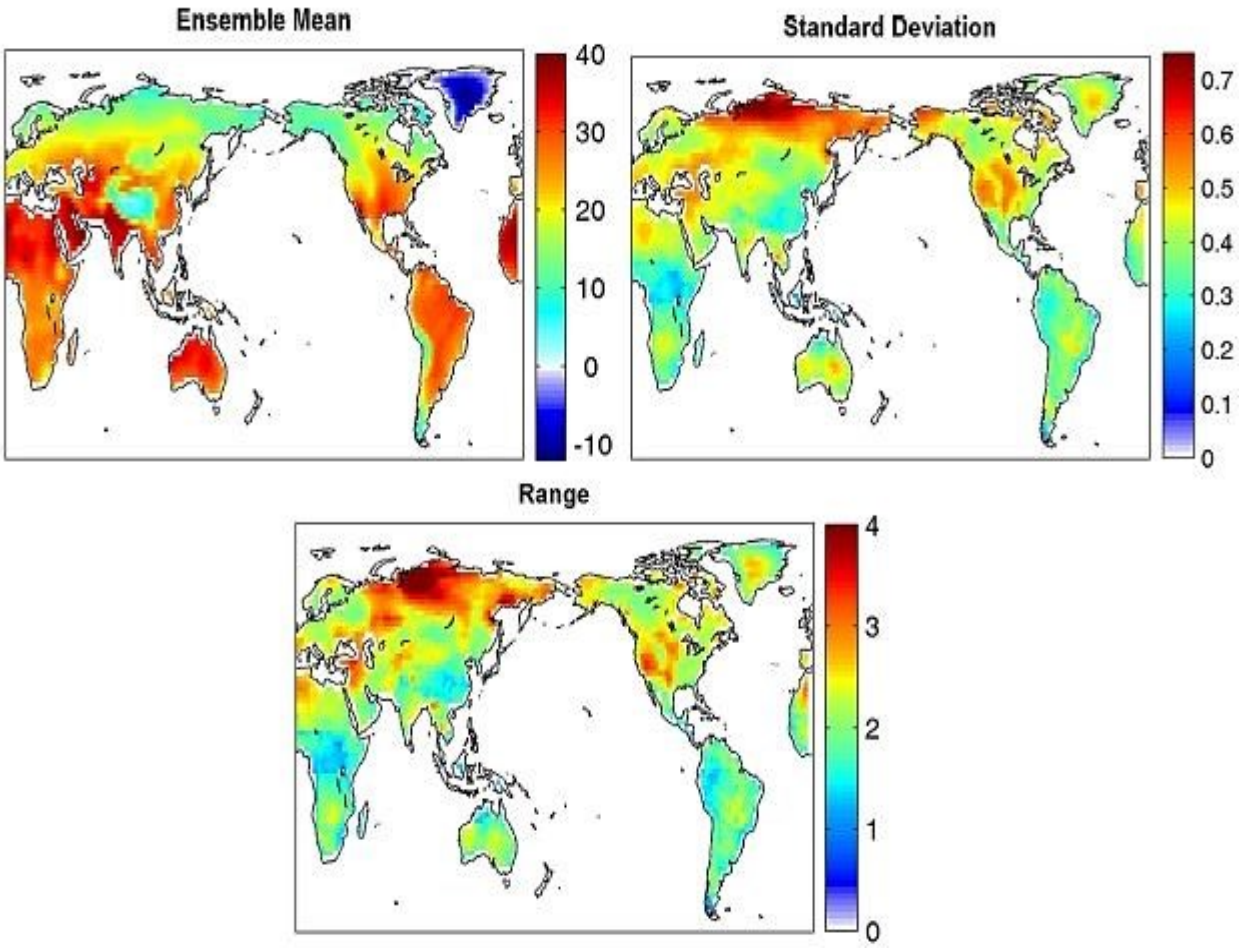

Figure 3.2 Ensemble mean (top left), inter-model standard deviation (top right), and range (bottom) of the annual temperature maxima in CMIP5 simulations.

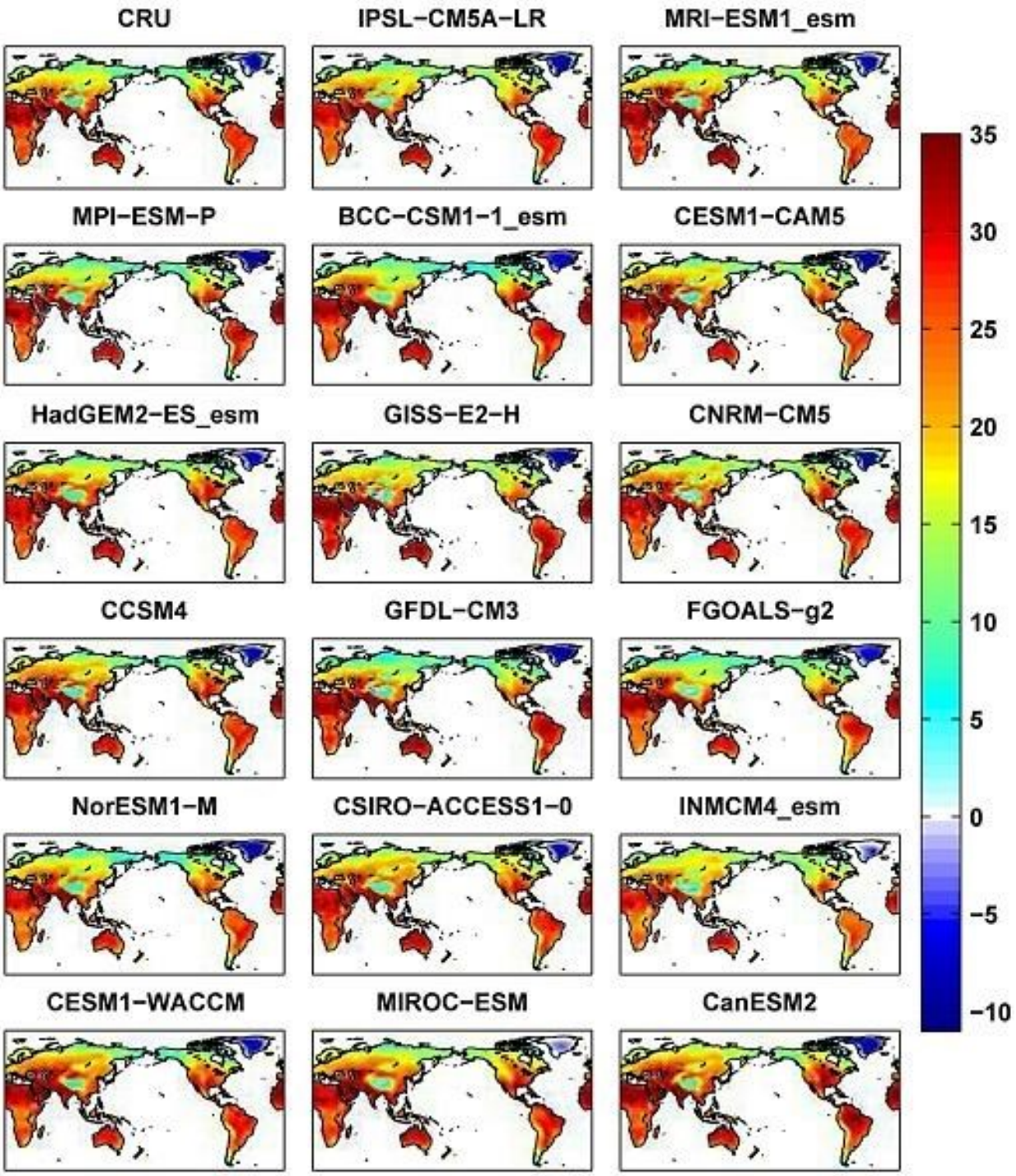

Figure 3.3 2-year return level (in degrees C) of the annual temperature maxima based on the CRU observations (upper left panel), and on selected CMIP5 climate model simulations (remaining panels).

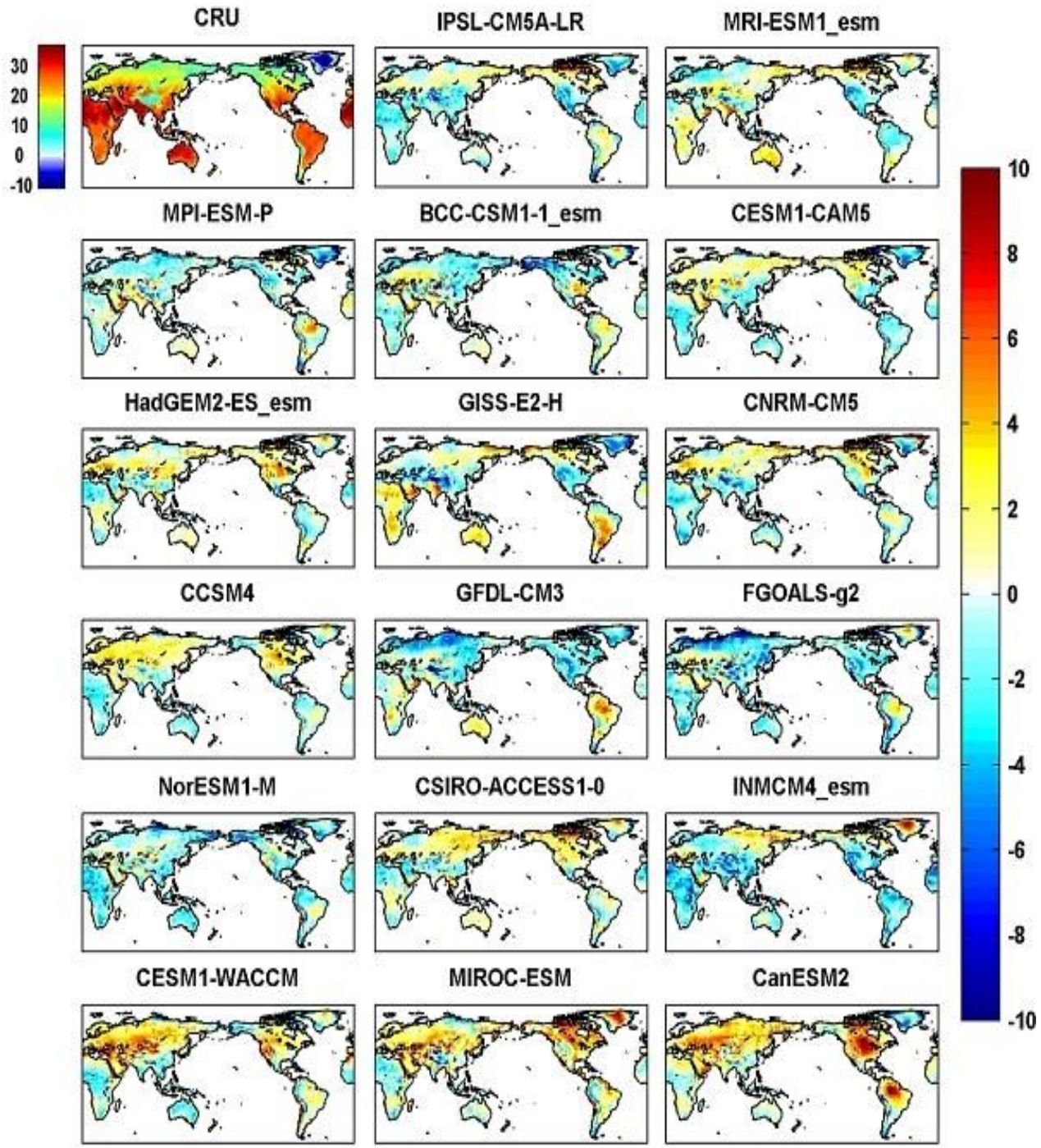

Figure 3.4 2-year return level (in degrees C) of the annual temperature maxima based on the CRU observations (upper left panel), and return-level differences between selected CMIP5 climate simulations and CRU reference data (CMIP5 - CRU; remaining panels).

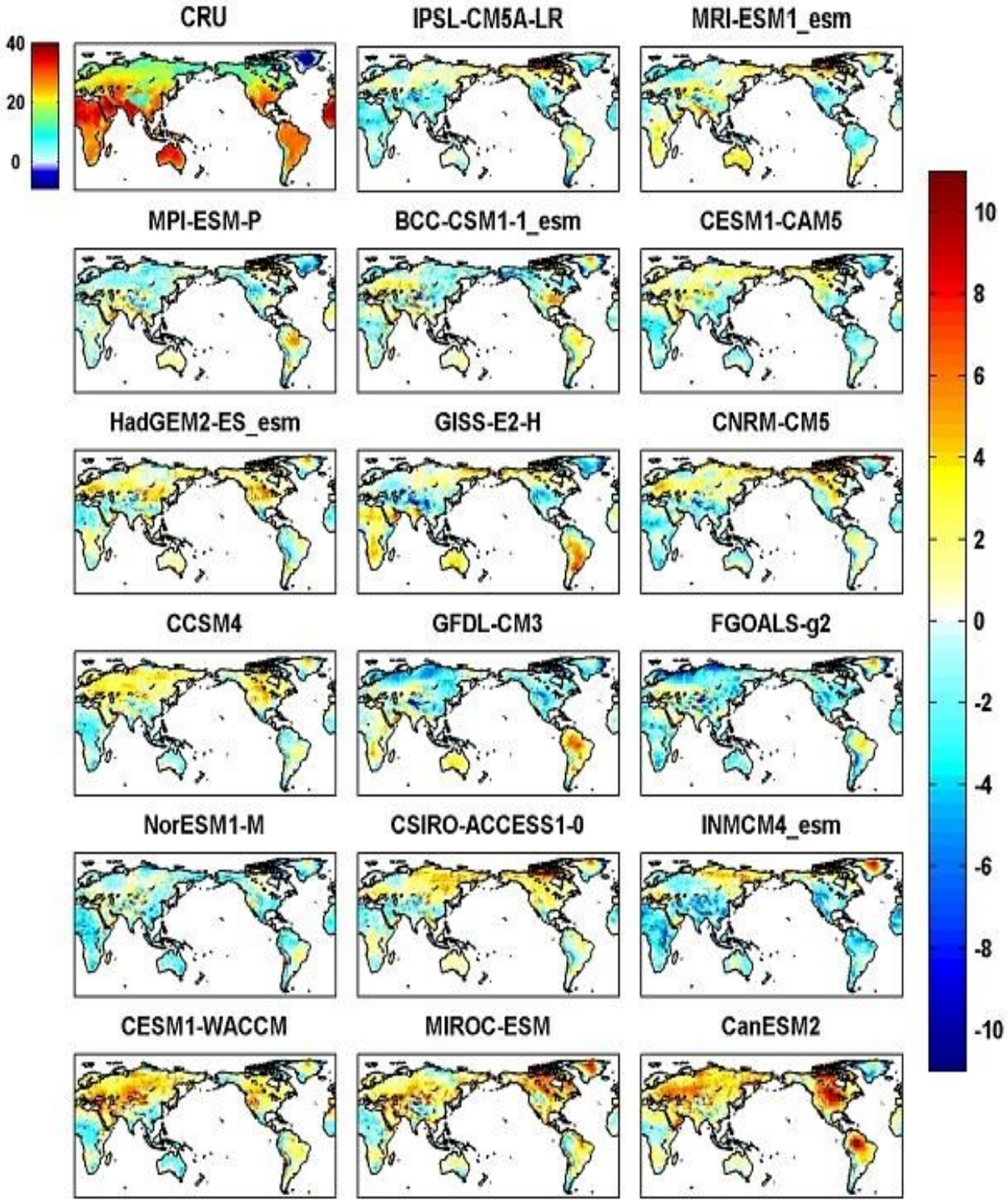

Figure 3.5 25-year return level (in degrees C) of the annual temperature maxima from CRU observations (upper left panel), and return-level differences between selected CMIP5 climate simulations and CRU reference data (CMIP5 - CRU; remaining panels).

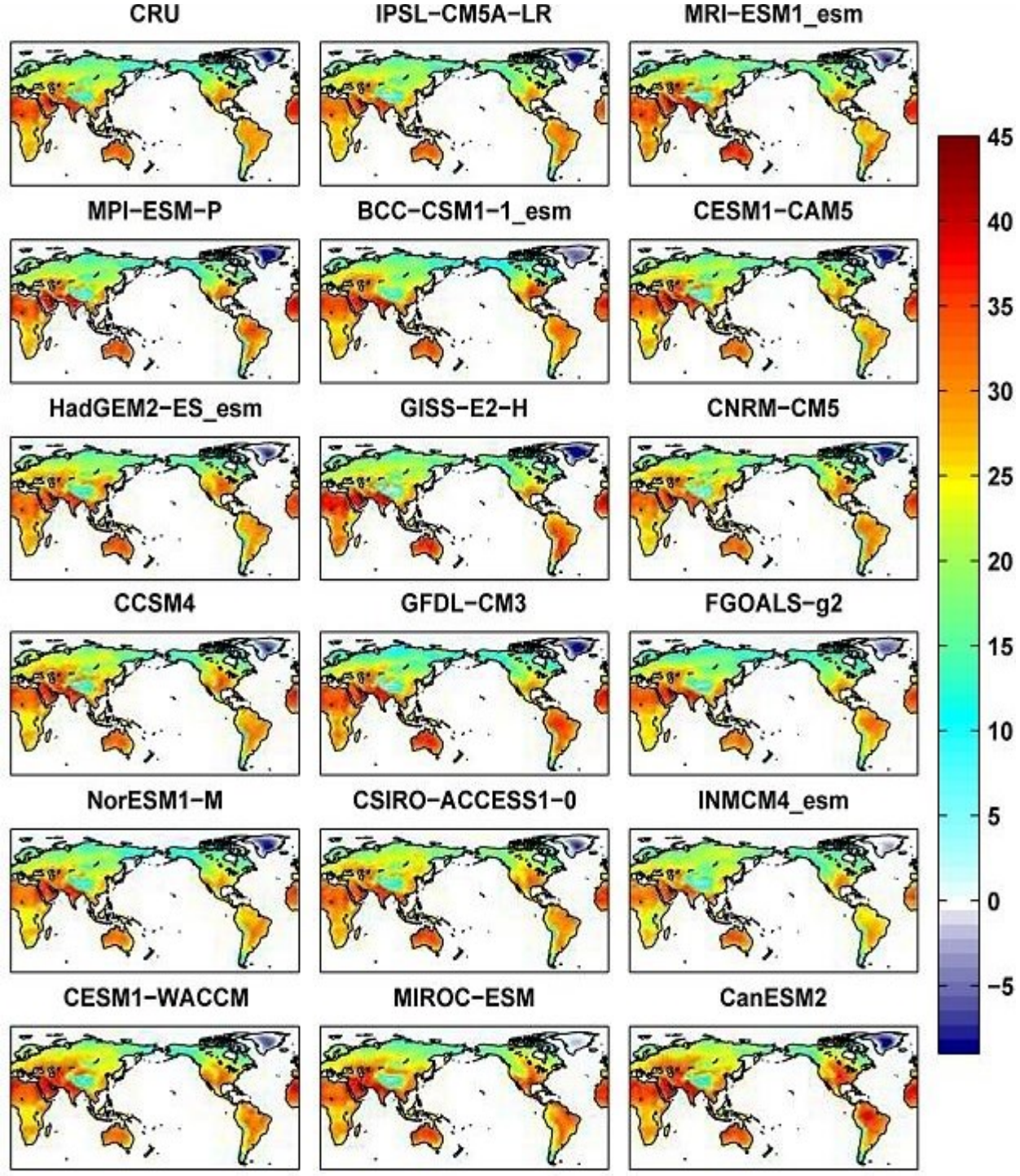

Figure 3.6 100-year return level (in degrees C) of the annual temperature maxima based on the CRU observations (upper left panel), and on selected CMIP5 climate model simulations (remaining panels).

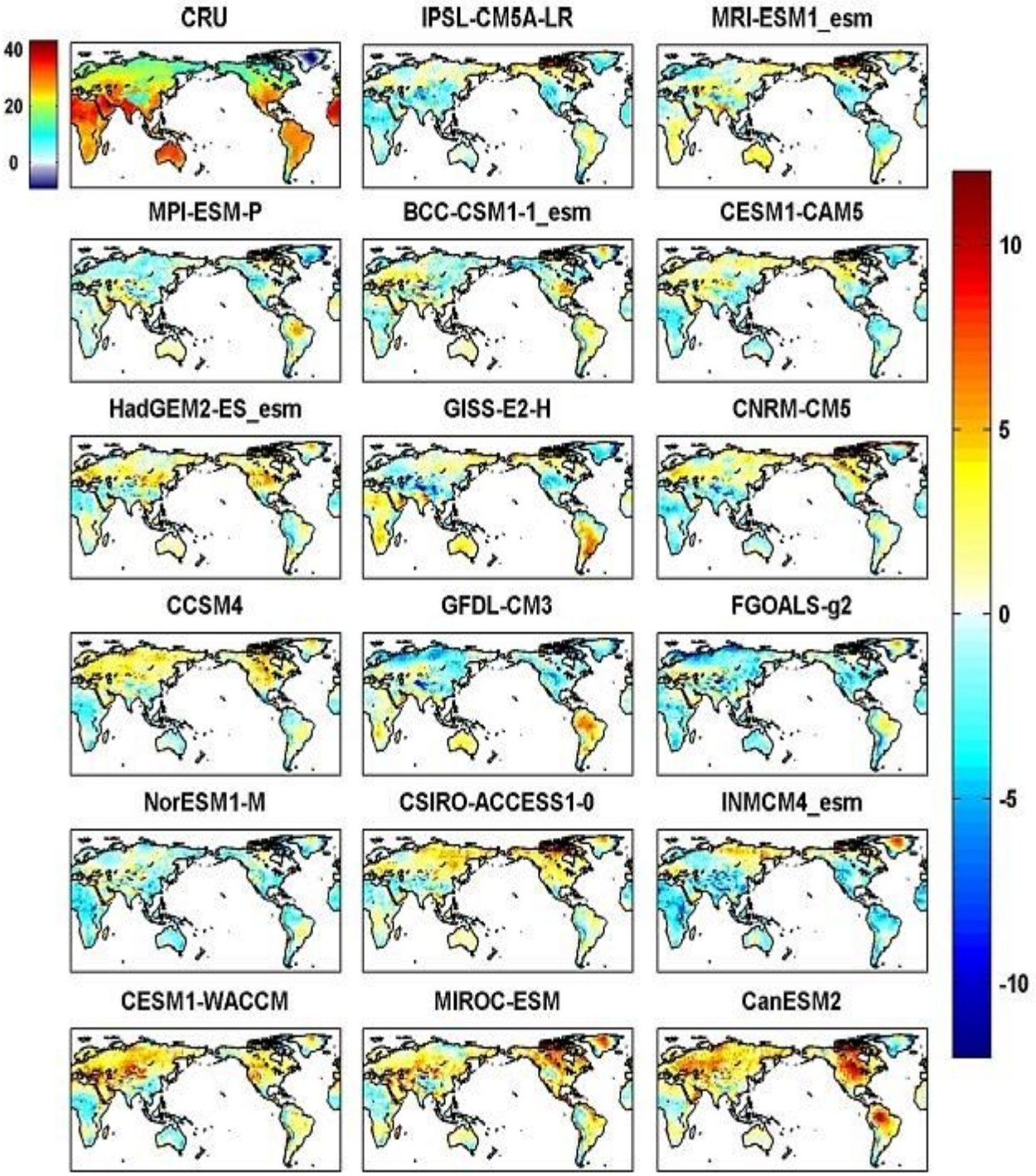

Figure 3.7 100-year return level (in degrees C) of the annual temperature maxima from CRU observations (upper left panel), and return level differences between selected CMIP5 climate simulations and CRU reference data (CMIP5 - CRU; remaining panels).

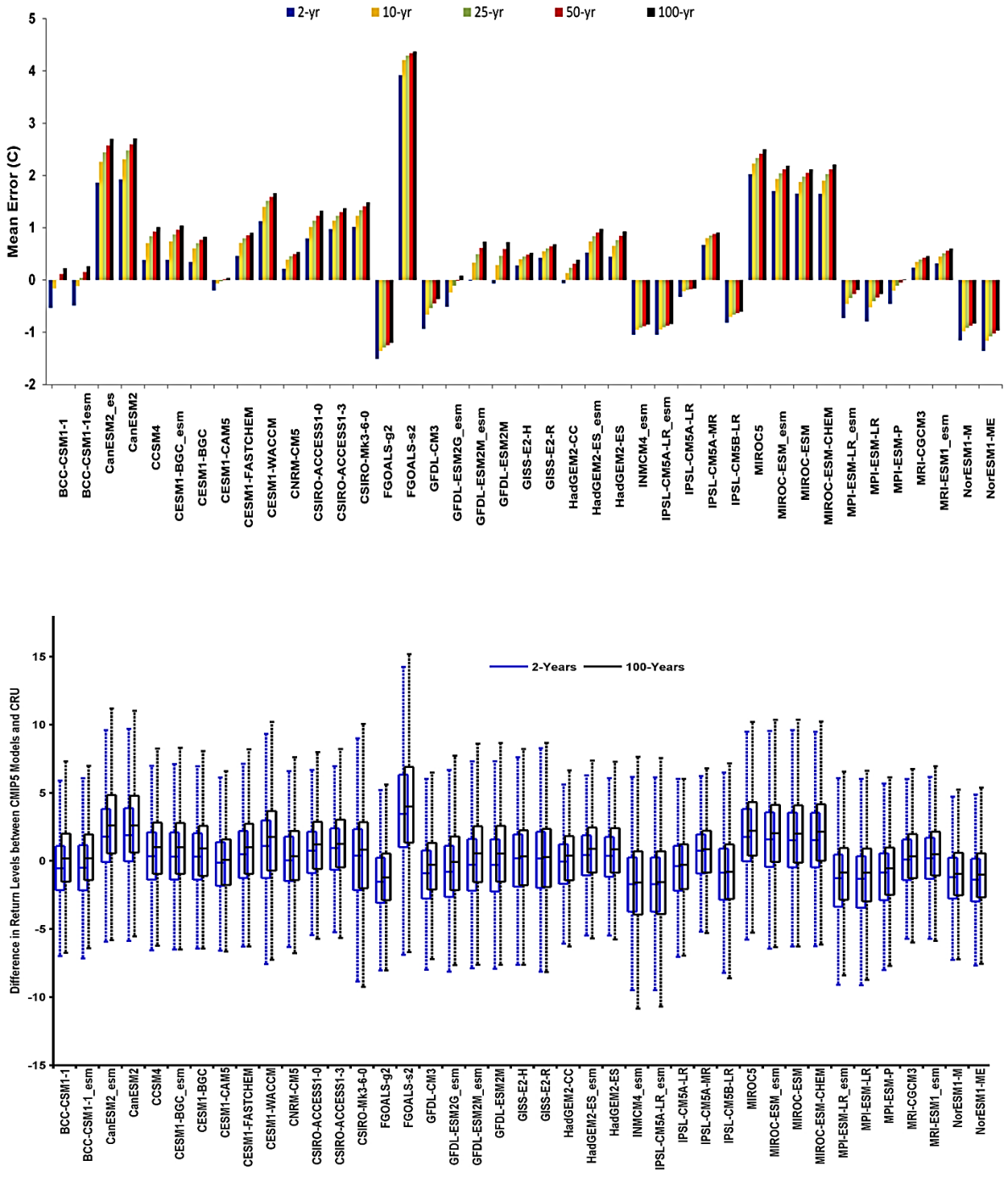

Figure 3.8 (top): Mean Error (ME) of the 2-, 10-, 25-, 50-,and 100-year temperature maxima (Degree Celsius) simulations based on 41 CMIP5 simulations relative to Climatic Research Unit (CRU) observations; (bottom): boxplots of differences (degrees C) between CMIP5 2-and 100-year return levels relative to CRU observations.

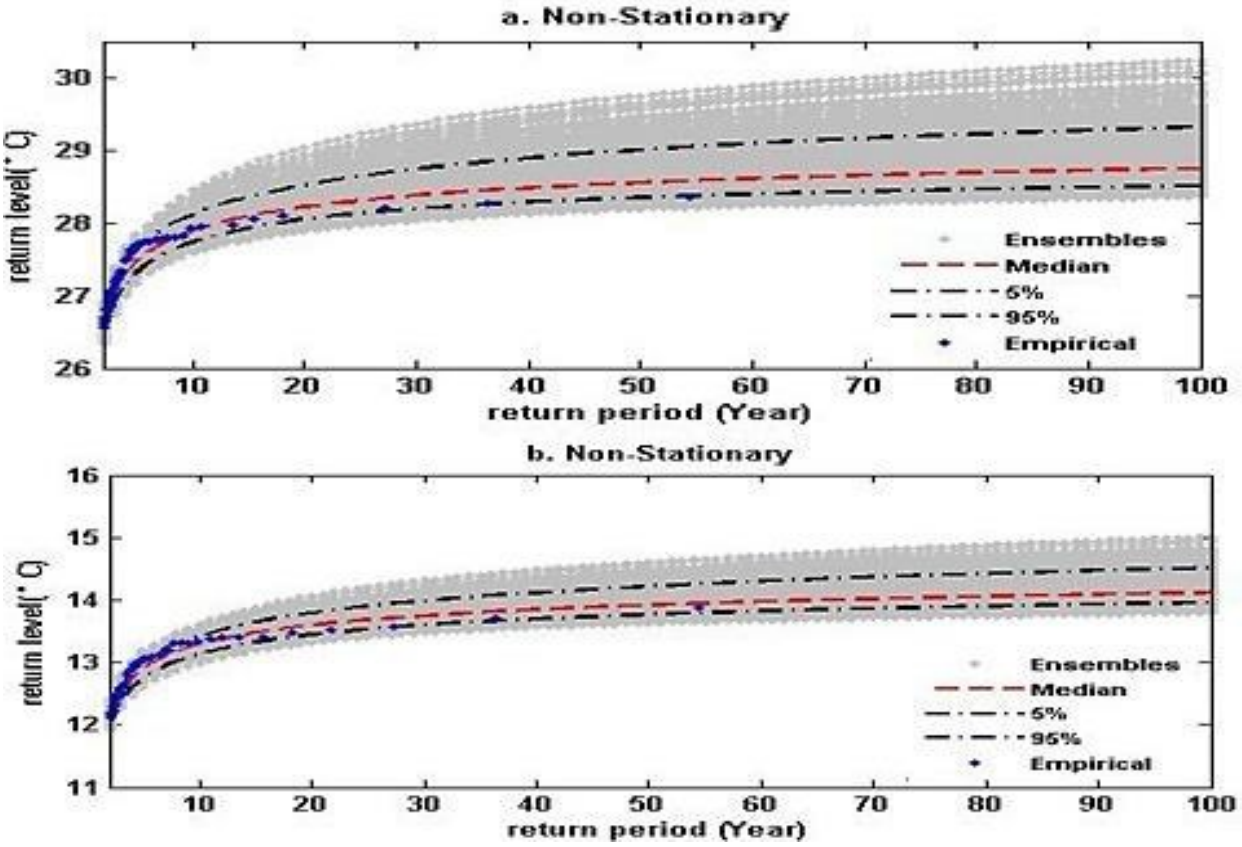

Figure 3.9 Sample uncertainty bounds, median, and the 5% and 95% confidence bounds of the annual temperature maxima based on the DE-MC model for CRU reference and over two pixels in the a. western and b. central United States.

CHAPTER 4: Non-stationary Precipitation Intensity-Duration-Frequency Curves for Infrastructure Design in a Changing Climate

4.1 Introduction

Human activities during the past 50 years have caused an increase in global temperature (IPCC 2013; Karl et al. 2009; Hao et al., 2013). The rising temperatures boost in atmosphere's water holding capacity by about 7% per 1°C warming, thus directly affecting precipitation (Trenberth 2011). Higher atmospheric water vapor creates more intense precipitation events (Kunkel et al. 2013). Consequentially, global warming increases the risk of extreme floods and damage to infrastructure such as dams, roads and sewer and storm water drainage systems (Jongman et al., 2014). Indeed, ground-based observations in the U.S. show an increase in extreme rainfalls by around 20% (Karl et al. 2009). Global-scale studies also show increased precipitation in northern Australia, central Africa, Central America and parts of southwest Asia (Damberg and AghaKouchak, 2013). Due to rising temperatures, subsequent increases in atmospheric moisture content and moisture transport into storm, expected extreme precipitation, or probable maximum precipitation (PMP), may increase globally (Kunkel et al. 2013) and will likely lead to more frequent and severe catastrophic floods (Das et al. 2011; Groisman et al. 2005; Roger Few 2003). Current infrastructure design concepts to deal with flooding and precipitation are based on local rainfall Intensity-Frequency-Duration (IDF) curves (Chen et al. 2013). These curves are widely used in municipal storm water management and other engineering design applications across the world (Endreny and Imbeah 2009). The IDF curves are based on historical rainfall time series data and designed to capture the intensity and frequency of precipitation for different durations. Rainfall intensities corresponding to particular durations (e.g., 1-hr, 2-hr, 6-hr, 24-hr) are obtained by fitting annual extreme rainfall to a theoretical probability distribution. However,

current IDF curves are based on the concept of stationarity, which assumes that the occurrence probability of extreme precipitation events is not expected to change significantly over time (Simonovic and Peck 2009).

Given the observed increase in heavy precipitation events and the expected increase in PMP, we argue that the IDF curves should be updated to account for a changing climate (Simonovic and Peck 2009). Although some studies evaluate changes in precipitation intensity or frequency, the methods for assessing changes in precipitation intensity, duration and frequency and their uncertainty in a non-stationary climate are limited (Chen et al. 2013; Cooley 2013; Salas and Obeysekera 2013; Hassanzadeh et al, 2013; Yilmaz and Perera 2013; Park et al. 2011; Endreny and Imbeah 2009; Zwiers and Kharin 1998).

In this study, we assess the effect of non-stationarity on IDF curves and the occurrence of extremes using NEVA. We also outline a generalized framework for constructing IDF curves under non-stationary conditions. The fundamental concept is based on the Generalized Extreme Value distribution (GEV) combined with Bayesian inference for uncertainty assessment (see Methods Section). Our analyses are based on ground-based observations of precipitation extremes (here, annual maximum) from the United States National Oceanic and Atmospheric Administration Atlas 14, the basis for IDF curves in the United States (Bonnin et al. 2006). Following the NOAA Atlas 14 approach, the annual maxima series is constructed by extracting the highest precipitation amount for a particular duration in each successive water year. Historical rainfall data (1949-2000) from five stations are used to assess IDF curves under non-stationarity (Table 4.1). In the selected stations and based on the Mann-Kendall trend test (Kendall 1976), precipitation extremes exhibit non-stationary behavior over different durations at

95% confidence level (see Figure 4.4). The presence of a statistically significant increasing trend in precipitation extremes violates the basic assumption of stationary IDF curves.

Given detection of a statistically significant trend, the IDF curves should be derived based on the non-stationary model presented in Method Section (Equations 2.1 and 2.2). A unique feature of this modeling framework is that it offers uncertainty bounds of IDF estimates (see Equations 2.3 and 2.4).

4.2 Methodology

Refer to the Chapter 2.2.

4.3 Results

To illustrate, the stationary and non-stationary IDF curves for different return periods (2-, 10-, 25-, 50- and 100-year) and durations for the White Sands National Monument Station in New Mexico are presented in Figure 4.1. As discussed in Method Section, the confidence interval and uncertainty bounds of IDF curves can be obtained simultaneously in the proposed framework. The gray lines show the uncertainty bound of the non-stationary IDF estimates based on the Differential Evolutionary Monte Carlo algorithm built into the generalized extreme value distribution (refer to Chapter 2.2). We found that the stationary assumption delivers IDF curves that substantially underestimate extreme events. If such a stationary IDF curve is used for an infrastructure design, the project may not be able to withstand more extreme events, which are shown by non-stationary estimates for the same return period.

For example, for a 2-year 2-hr storm (i.e., an event with a return period of 2 years and duration of 2 hours), the difference between the non-stationarity (14.7 mm/hr) and stationarity (9.1 mm/hr) extreme precipitation is about 5.6 mm/hr (+61.5%); while for a 10-year 1-hr event, the difference between non-stationarity (35.0 mm/hr) and stationarity (25.9 mm/hr) extreme precipitation is

over 9.1 mm/hr (+35.1%). Even for a small watershed, this extra 9.1 mm/hr (+35.1%) precipitation would lead to a significant increase in flood peak. In other words, a stationary assumption will underestimate the peak flood and as a result the actual flood risk will be higher than what the system or infrastructure is designed for.

The differences between the non-stationary and stationary estimates decrease for longer durations (e.g., 168-hr precipitation). This implies that in this station, shorter precipitation events have been intensified more in the past decades, while longer events have not changed as much. For the same station shown in Figure 4.1, the boxplots of differences between the non-stationary and stationary precipitation extremes are presented in Figure 4.2. The figure shows that for all durations and return periods, the quantile boxes are above zero, indicating underestimation of extremes in a stationary assumption. In all durations, the uncertainty increases as the return period increases. Consequently, the uncertainties in the stationary and non-stationary precipitation differences increase at higher return periods.

By examining storm durations, we found that the shorter the duration the larger the differences between the non-stationary and stationary extremes. As an example, for the 100-year return period, the differences between non-stationary and stationary IDF curves of 1-hr and 2-hr events reduce from 4.7-15.6 mm/hr to 1.4~7.3 mm/hr, while for a 168-hr storm, the difference approaches zero (see Figure 4.2). Similar behavior is observed in the other stations and as a result we have focused on shorter durations. For the other stations in Nevada (NV), California (CA) and North Carolina (NC), Figure 4.3 summarizes the differences between the stationary and non-stationary precipitation extremes for different return periods and durations (upper panels: 1-hr duration; lower panels: 2-hr duration). The boxplots show the median (center mark), and the 25th (lower edge) and 75th (upper edge) percentiles of the differences between stationary and

non-stationary estimates. Similar behavior emerges in different stations, indicating that non-stationary estimates are larger than their corresponding stationary values. Such difference in underestimation raises the risk of extreme floods and damage to infrastructure if non-stationarity is ignored.

4.4 Summary and Conclusion

Infrastructure health and safety during precipitation extremes is closely related to human health and security, particularly downstream of major structures (e.g., dams, spillways, reservoirs). For this reason, methods that can account for changing precipitation extremes are essential for updating engineering standards and design codes. Potential non-uniform and climate-induced changes on heavy rainfall events calls into question the accuracy and adequacy of current infrastructure design concepts, which rely on an assumption of climate stationarity. We show that ignoring the stationary assumption could lead to substantial underestimation of extremes, especially at sub-daily durations (e.g., 1-hr, 2-hr). We also outline a novel framework to create the next generation of IDF curves to be incorporated into infrastructure design. However, infrastructure design and construction requires substantial investment over a long period of time and effective integration of this methodology as well as development of adaptive design frameworks will require collaborative and interdisciplinary research with engineers, policy makers, economists, climate scientists and decision makers.

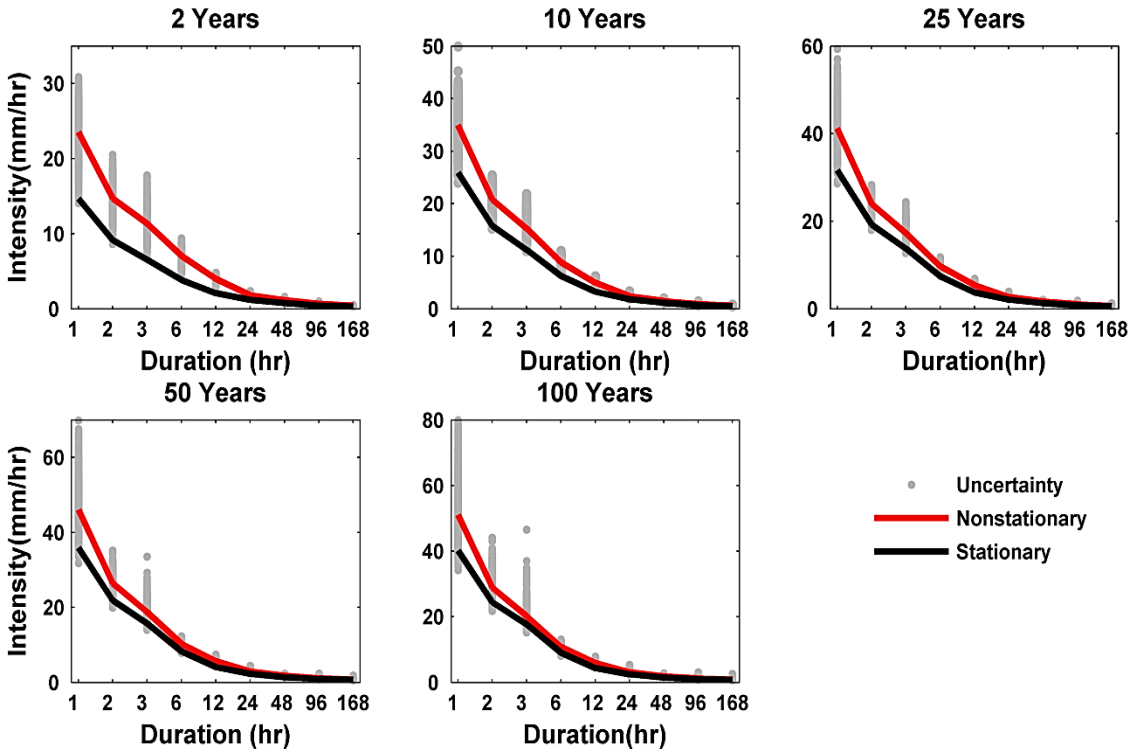

Figure 4.1 Non-stationary vs. stationary IDF Curves for different return periods and durations at the selected station in White Sands National Monument Station, New Mexico. The stationary assumption consistently underestimates the IDF curves over different durations. The gray area shows the uncertainty bound of non-stationary IDF estimates.

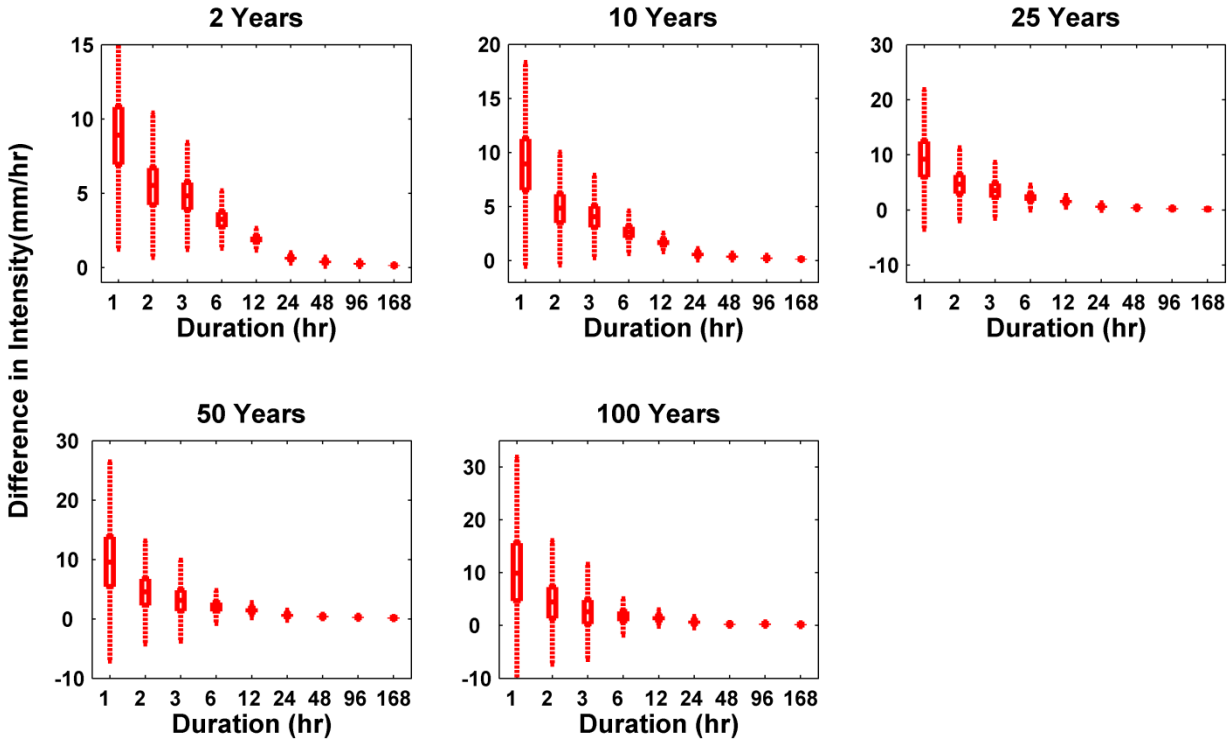

Figure 4.2 Differences between the non-stationary and stationary precipitation extremes for different return periods and durations in White Sands National Monument Station, New Mexico. The boxplots show the median (center mark), and the 25th (lower edge) and 75th (upper edge) percentiles.

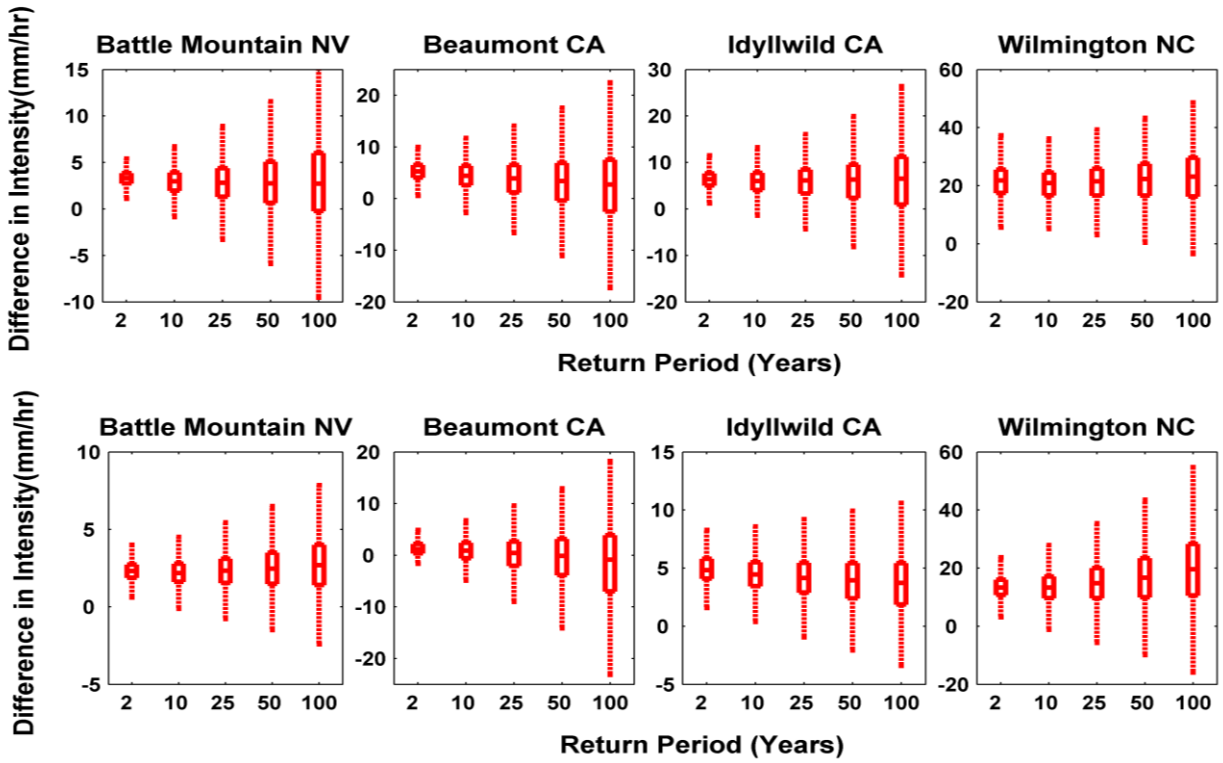

Figure 4.3 Difference between the non-stationary and stationary precipitation extremes for different return periods and durations (upper panels: 1-hr duration; lower panels: 2-hr duration). The boxplots show the median (center mark), and the 25th (lower edge) and 75th (upper edge) percentiles.

Table 4.1 Selected stations for analysis of Intensity-Duration-Frequency curve analysis under stationary and non-stationary assumptions.

Station Name	State	Latitude	Longitude
White Sands National Monument	**NM**	**32.7817**	**106.1747**
Battle Mountain	**NV**	**40.6167**	**116.8667**
Beaumont	**CA**	**33.9292**	**116.9750**
Idyllwild Fire Dept	**CA**	**33.7472**	**116.7144**
Wilmington WSO Airport	**NC**	**34.2683**	**77.9061**

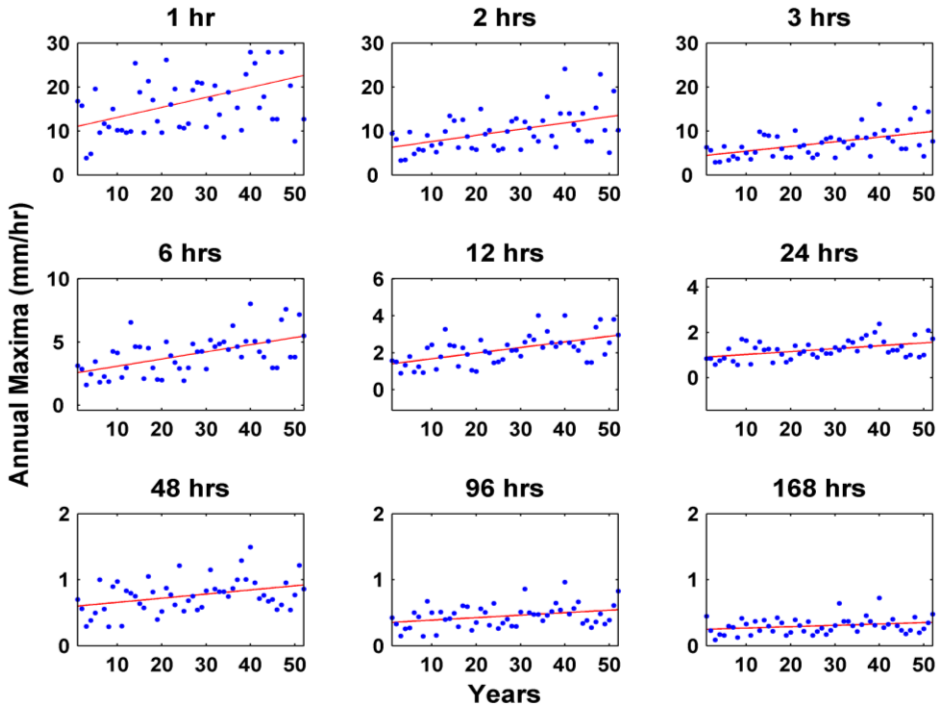

Figure 4.4 Significant trends in annual maxima precipitation over different durations in the selected station in White Sands National Monument, New Mexico (latitude 40.62°, longitude 116.87°).

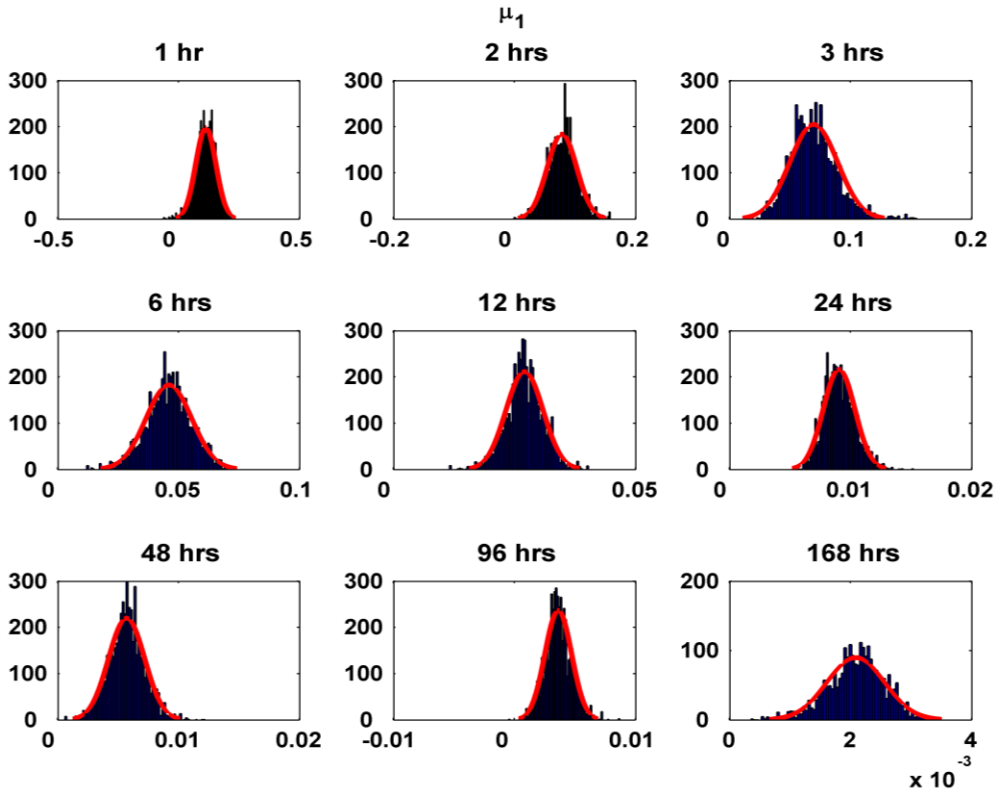

Figure 4.5 Posterior distribution of the regression parameter μ_1 in the selected station in White Sands National Monument Station, New Mexico (latitude 40.62°, longitude 116.87°).

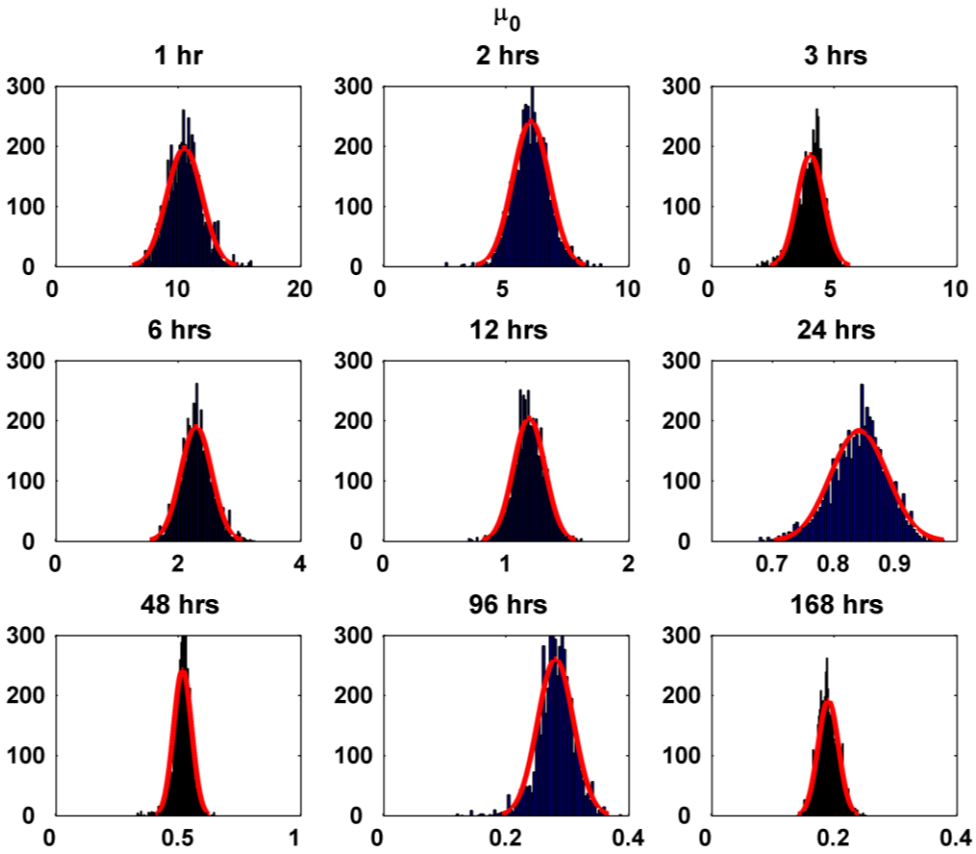

Figure 4.6 Posterior distribution of the regression parameter μ_0 in the selected station in White Sands National Monument Station, New Mexico (latitude 40.62°, longitude 116.87°).

CHAPTER 5: Empirical Bayes Estimation for the Conditional Extreme Value Model

5.1 Introduction

Heffernan and Tawn (2004) introduced new methodology for modeling multivariate extreme values through a conditional distribution framework that does not require a priori knowledge of the dependence structure nor that the variables be simultaneously extreme. For positively associated random variables, the model is quite simple, and has primarily two parameters α and β that describe the dependence. Heuristically, α describes the strength of dependence and β the variability in the dependence (e.g., Keef et al., 2009a); although special cases make such a precise description less straightforward. For negatively associated parameters, the form is more complicated, but can be avoided using an appropriate transformation so that the simpler form will hold (Keef et al., 2013a).

A drawback to the approach concerns the estimation procedure for the model as no simple, closed-form distribution exists in general (i.e., without assuming a specific dependence structure) for the conditional distribution. To quote Richard L. Smith in his comment to Heffernan and Tawn (2004):

"The authors use maximum likelihood for estimating the generalized Pareto margins, Gaussian estimation for the conditional means and standard deviations and pseudolikelihood estimation for combining the various conditional distributions into a multivariate family, a veritable witches' soup of estimation methods, all nicely stirred up with the bootstrap as seemingly the only means of keeping control of all the estimation errors. Although I applaud the authors' eclecticism, would it not be better to have a more coherent estimation strategy?"

The aim of this dissertation is to provide a "more coherent estimation strategy" by using a Bayesian estimation approach.

The most difficult hurdle in developing an estimation strategy for the Heffernan and Tawn (2004) model concerns the lack of a distribution function for the conditional likelihood. To circumvent this issue, we employ an empirical estimation strategy. Owen (1988) developed empirical likelihood techniques as a robust alternative to classical likelihood approaches or the bootstrap (Owen, 1990; Mengersen et al., 2013). He demonstrated that, for some categories of statistical models, when the likelihood function is numerically unavailable or not entirely known, empirical likelihood methods can be used to bypass simulations from the model while converging in the number of observations. Empirical likelihood has been shown in a wide range of situations to have properties analogous to a real likelihood (Li, 1995; Jing, 1995; Chen, 1994; Qin and Lawless, 1994; Hall and La Scala, 1990). Although further investigation of this methodology is needed, it appears to be a valuable approach in distribution-free contexts. Therefore, empirical likelihood provides another route in this study to tackle the difficulty in the conditional likelihood estimation.

5.2 Model and Estimation Methodology

5.2.1 The Conditional Extreme Value Model

We describe the Heffernan and Tawn (2004) conditional extreme value model (henceforth, the HT model) using the approach of Heffernan and Resnick (2007). That is, suppose $X$and $Y$are two suitably transformed random variables, and assume that

$$\Pr\left\{Y - u > y, \frac{X - a(Y)}{b(Y)} \leq z \middle| Y > u\right\} \to \exp(-y)\, G(z)\ as\ u \to \infty \tag{5.1}$$

where G is a non-degenerate distribution and $a(Y)$ and $b(Y)$ are normalizing functions for $y > 0$. Through examing a wide class of copula dependence models, using Gumbel margins for X and Y, Heffernan and Tawn (2004) found that the forms for $a(X)$ and $b(Y)$ fell into the simple class:

$$a(Y) = \alpha y \; and \; b(Y) = y^{\beta} \tag{5.2}$$

For positively associated variables X and Y, $\alpha \in [0,1]$ and $\beta \in (-\infty, 1)$. For the more complicated form of negatively associated X and Y, Keef et al. (2013a) used a Laplace transformation to ensure Equation 5.2 was valid and α $[-1,1]$. The parameters α and β are interdependent and control the dependence between the variables X and Y. With the Laplace transformation, $\alpha < 0$ implies negative dependence, and $\alpha > 0$ implies positive dependence. Weakly dependent X and Y are possible as $\alpha \to 0$, but it is also possible that strong dependence exists even when $\alpha = 0$ (Gilleland et al., 2013; Heffernan and Tawn, 2004). The parameter β measures the variability of the dependence with highly negative values indicating lower variability. Let,

$$Z = (X - a(Y))/b(Y) \tag{5.3}$$

Equation 5.1 implies conditional independence between Z and Y given $Y > u$. As mentioned, no simple, closed-form expression exists for G (the most difficult hurdle in developing Bayesian statistical inference). The primary contribution of this article is to apply the empirical likelihood approach to circumvent this issue. From Equations 5.1, 5.2 and 5.3, define

$$X_{|Y>u} = \alpha Y + Y^{\beta} Z_{|Y>u} \tag{5.4}$$

Subsequently, the key role in estimating the joint distribution function (df) of X and Y, conditional on $Y > u$ for u large, is to know the parameters α and β and the df G. Estimation for the parameters α and β is an active area of research (Keef et al., 2009a,b; Lamb et al., 2010; Jonathan et al., 2013), and estimation of G is performed through resampling from the empirical df of the "residual" vectors Z in the HT model after achieving reasonable estimates for α and β. The estimation method from the HT model is semi-parametric and involves the following steps:

1. Estimate the marginal df's for each variable separately.

2. Transform each variable in order that they each follow a Laplace df (following Keef et al., 2013a).

3. Estimate the parameters of the parametric model conditional on large values of the conditioning variable.

4. Information about G (e.g., functions such as the mean and variance, etc.) can be simulated using the empirical df of the estimated standardized residuals. Back transformation can be used to put these estimates onto the original scale.

Heffernan and Tawn (2004) suggested using a hybrid, semi-parametric, model for step 1 of the following form that accounts for both the extreme and non-extreme values (Coles and Tawn, 1991)

$$\hat{F}_{X_i}(x) = \begin{cases} 1 - (1 - \tilde{F}_{X_i}(x))(1 + \xi_i(x - u_{X_i})/\sigma_i)_+^{-1/\xi_i} & for \;\; x > u_{X_i} \\ \tilde{F}_{X_i}(x) & for \;\; x \leq u_{X_i} \end{cases} \quad (5.5)$$

where $\tilde{F}_{X_i}(x)$ is the empirical df of the X_i values; ξ_i is the shape parameter and $\sigma_i > 0$ is the scale parameter for each individual variable, fitted with the generalized Pareto (GP) df as a model for the upper tail of the univariate extremal exceedances over a high threshold (i.e., u_{X_i} in Equation 5.5). The threshold excess probability is

$$Pr\{Y - u \leq y | Y > u\}, \; for \; u \; large \quad (5.6)$$

which is approximately equivalent to the GP df for a sufficiently large threshold, u, given by

$$H(y) = 1 - (1 + \frac{\xi y}{\sigma})_+^{-1/\xi} \quad (5.7)$$

with $z_+ = max(z, 0)$. In particular, it is $1 - H(y) = Pr\{Y - u > y | Y > u\}$ for u large, that is of primary concern. Note that the limit $\xi \to 0$ yields the exponential df, $exp(-y/\sigma)$. In the HT model, Y is assumed to follow a standard exponential df (i.e., $exp(-y)$ in Equation 5.1) without loss of generality because it can be obtained through a simple transformation.

In step 2, the original data are first transformed to X^* using the Laplace transformation with the estimated df described by Equation 5.5 in order to maintain both the marginal and dependence features of the multivariate data. Specifically,

$$X^* = \begin{cases} \log\{2\hat{F}_{X_i}(X_i)\} & for\ X_i < \hat{F}_i^{-1}\left(\frac{1}{2}\right) \\ -\log\left\{2\left(1-\hat{F}_{X_i}(X_i)\right)\right\} & for\ X_i \geq \hat{F}_i^{-1}\left(\frac{1}{2}\right) \end{cases} \tag{5.8}$$

where $\hat{F}_i$ is estimated according to Equation 5.5 using maximum likelihood estimation for the GP portion and empirical estimation for $\tilde{F}_i$.

Heffernan and Tawn (2004) used non-linear least squares estimation in Equation 5.4 to estimate α_i and β_i for each X_i under the working assumption that Z follows a normal df. Obviously, the assumption of a normal df for Z is inappropriate as it implies that $X_{|Y>u}$ is also normally distributed, which generally may not be the case. To counteract the inherent estimation bias from this approach, Keef et al. (2013a) imposed joint constraint on the dependence parameters (α, β) in order to limit the upper quantiles of $X_{|Y>u}$ to be less than or equal to x_F, the value that would be observed under asymptotic dependence. From these estimates in step 3, Heffernan and Tawn (2004) obtain new estimates $\hat{Z}_i = (X_{i|Y>u} - \hat{a}_i(Y_i))/\hat{b}_i(Y_i)$ from which simulations from $\hat{G}$ are obtained.

To incorporate the uncertainty inference at each stage of the estimation procedure, a bootstrap procedure is proposed by Heffernan and Tawn (2004). Note that this approach differs from that of incorporating covariates into the parameters of a univariate extreme value df in that a distribution for values of one variate is conditional on only the extreme values of another variable. Therefore, the dependence is on the processes themselves rather than indirectly through distributional parameters (Gilleland et al., 2013; Jonathan et al., 2012).

5.2.2 Empirical Bayes Estimation

In order to use empirical Bayes estimation, we need to infer $\hat{\alpha}$, $\hat{\beta}$ and $\hat{G}$.

However, because G does not have a simple, closed-form expression and we only need to simulate from G, only the estimates $\hat{\alpha}$ and $\hat{\beta}$ are required. That is, $p(\alpha,\beta) = [\alpha,\beta|Z,Y > u]$ is what we are, ultimately, seeking to estimate. The Bayesian approach for simulating data from an arbitrary distribution has become increasingly popular and used in numerous studies (Stephenson and Tawn, 2004; Coles and Powell, 1996). From Equation 5.1 and Equation 5.3, we have that

$$[Z \,|\, \alpha,\beta,X^*,Y > u][X^*|Y > u,\theta_X\,][Y^*\,|Y > u,\theta_Y\,] \tag{5.9}$$

where θ_X and θ_Y are parameters in the marginal in Equation 5.5 and $p(\alpha,\beta)$ is proportional to $[Z \,|\, \alpha,\beta,X^*,Y > u]$ times a prior distribution. More specifically,

$$\begin{aligned} p(\alpha,\beta) &\propto [Z \,|\, \alpha,\beta,X^*,Y > u][\alpha,\beta,\theta_X\,,\theta_Y\,|\eta][\eta] \\ &= [Z|\alpha,\beta,\theta,X^*,Y > u][X^*|Y > u,\theta_X\,][Y^*\,|Y > u,\theta_Y\,][\alpha,\beta,\theta_X\,,\theta_Y\,|\eta][\eta] \\ &= [Z|\alpha,\beta,\theta,X^*,Y > u][X^*|Y > u,\sigma_X,\xi_X\,][Y^*\,|Y > u,\sigma_Y,\xi_Y\,] \\ &\quad [\alpha,\beta|\eta][\sigma_X\,,\xi_X|\eta][\sigma_Y,\xi_Y\,|\eta][\eta] \end{aligned} \tag{5.10}$$

where $\theta_X = (\sigma_X,\xi_X)$ and $\theta_Y = (\sigma_Y,\xi_Y)$ are the marginal GP df parameters for X and Y, respectively, and η represents any additional hyper parameters pertaining to the prior df for α, β and θ. For simplicity, let $A \equiv [Z|\alpha,\beta,\theta,X^*,Y > u], B \equiv [X^*|Y > u,\sigma_X\,,\xi_X\,]\,[Y\,|Y > u,\sigma_Y\,,\xi_Y\,], C \equiv [\alpha,\beta|\eta], D \equiv [\sigma_X\,,\xi_X\,|\eta], and\ E \equiv [\sigma_Y\,,\xi_Y\,|\eta]$.

Reasonable choices for B, D and E are available by fitting the GP df (see Equation 5.5 and Equation 5.7), and additional hyper parameters η are not of great concern. It may be difficult to obtain a joint df for C, but we can make the reasonable assumption of conditional independence between α and β and follow the method introduced by Keef et al. (2013b), which is a fast estimate for α that does not rely on knowing anything about G, to help inform the priors. In

theory, the prior knowledge on parameters does not depend on the observations Y, and should therefore be specified without using observations, but rather using any external source of knowledge (Renard et al., 2013). Thus, we suppose that prior information of parameters should not produce much effect on the simulation results. However, we found that in this specific problem, prior information does matter. Therefore, relatively informative priors are preferred here. Particularly, the informative initials are estimated by looking into different qth quantiles of $X^*|Y^*$ when Y^* values are within two different intervals as $Y_i^* \in (u^* - \delta_u, u^* + s_u)$ and $Y_i^* \in (v^* - \delta_v, v^* + s_v)$, where the intervals are around the conditional threshold u, and $v > u$ within the range of the observation of Y_i^*. Then take the median of the estimated α^*s as the initial point $\hat{\alpha}$. This initial estimation approach is systematically introduced by Keef et al. (2013b) with more details. An initial estimate for β is achieved through a linear regression on the log-transformed Equation 5.3; namely,

$$\beta log(Y^*) + log(Z) = log(X^* - \hat{\alpha}(Y^*)) \quad given\ Y^* > u^* \tag{5.11}$$

with Z being some random residuals, the informative initial of β is obtained and subsequently, we have some knowledge about C. Given prior distributions specified in the above manner, the remaining problem, of course, is to estimate A. This is where an empirical estimation strategy is employed, by assuming A as a prior df for (α, β) in the following empirical likelihood estimation. Empirical likelihood provides likelihood ratio statistics for parameters by profiling a nonparametric likelihood; the approach is analogous to that used for parametric models (Qin and Lawless, 1994). Owen (1990) showed that for d-variate independent and identically distributed (henceforth, i.i.d.) random variables Y (each variate as $y_1, \ldots y_n$), with an unknown distribution density f, mean μ_θ and variance ${\sigma_\theta}^2$, the approach applies to quite general parameters $\theta(f)$, where θ is the parameter associated with f. Rather than defining the likelihood from the density

f as usual, the empirical likelihood method starts by defining parameters of interest θ as functionals of f, for instance as moments of f, and then profiles a nonparametric likelihood (Mengersen et al., 2013). More precisely, given a set of constraints of the form

$$E_f[h(Y,\theta)] = 0 \tag{5.12}$$

where the dimension of h sets the number of constraints unequivocally defining θ, the empirical likelihood is defined as

$$L(\theta|y) = \max_p \prod_{i=1}^{n} p_i \tag{5.13}$$

for p in the set

$$p \in [0,1]^n,\ \sum p_i = 1, \sum_i p_i\, h(y_i,\theta) = 0 \tag{5.14}$$

where $p_1, \dots p_n$ are nonnegative real numbers summing to unity. The validation of the empirical likelihood approximation is also provided by Owen (1988, 1990). He has proved under mild conditions, if θ satisfies Equation 5.12, then $-2\log(\frac{L(\theta|y)}{n^{-n}}) \to \chi_d^2$ in distribution when $n \to \infty$ and note that n^{-n} is the maximum of $L(\theta|y)$.

In general, the basic idea in this approach is to maximize the empirical likelihood (see Equation 5.13) subject to constraints provided by Equation (5.12) which reflect the characteristics of the quantity of interest. For instance, in the one-dimensional case when $\theta = E_f[Y]$, the empirical likelihood in θ is the maximum of the product $(p_1, \dots p_n)$ under the constraint $p_1 y_1 + \cdots + p_n y_n = \theta$. Solving Equation 5.13 is based on the Newton-Lagrange algorithm and more are derived with details in Mengersen et al. (2013), Qin and Lawless (1994) and Owen (1990, 1988). Due to its ability to conduct a nonparametric inference without knowledge of higher order moments of the distribution while implicitly taking them into consideration (according to Chen and Cui, 2003), when applying to the conditional likelihood estimation in this study, the first,

$$E[Y - \mu_\theta] = 0 \tag{5.15}$$

and second,

$$E[(Y - \mu_\theta)^2 - \sigma_\theta^2] = 0 \quad (5.16)$$

statistical moments of "residual" vectors Z (in Equation 5.3) and conditional vectors of $X|Y$ in the HT model are used as sufficient constraints to estimate the empirical likelihood of A. And if Z has the mean vector μ and vector of standard deviation σ, the respective conditional mean and standard deviation vectors of $X^*|Y^*$, for $Y^* > u^*$, are $\alpha Y^* + \mu Y^{*\beta}$ and $\sigma Y^{*\beta}$, respectively (see Keef et al., 2013a). Now we have all the components, A, B, C, D and E in the Bayesian framework (see Equation 5.10), in order to derive the empirical Bayes estimation are clarified and put together.

To estimate the parameters using Bayesian inference, a large number of realizations is generated from the parameters' posterior distributions, using the Differential Evolution Markov Chain (DE-MC; Ter Braak, 2006). The DE-MC utilizes the genetic algorithm Differential Evolution (DE; Ter Braak, 2006; Storn and Price, 1997) for global optimization over real parameter space with Markov Chain Monte Carlo (MCMC) approach (Ter Braak, 2006; Gilks et al., 1996). The advantages of simplicity, speed of calculation and convergence makes DE-MC favorable over the conventional MCMC (Ter Braak, 2006). In this model, for example, five Markov Chains are constructed in parallel, and are allowed to learn from each other by generating candidate draws based on two random parent Markov Chains (rather than to run independently - (see ter Braak, 2004; Gelman and Shirley, 2011)) such that the equilibrium distribution is the target posterior distribution. Meanwhile, the uncertainty of each parameter in Equation 5.10 is estimated.

5.3 Simulation Experiment

In this study, the results of the conditional extreme value analysis (henceforth, conditional EVA) simulated by the proposed empirical Bayes estimation approach are compared with the HT

model using the R (R Core Team, 2013) package *texmex* developed by Southworth and Heffernan (2010), which allows for the constrained estimation of the dependence parameters with the Laplace transformation on the marginal variables following Keef et al. (2013a).

5.3.1 Background

Ledford and Tawn (1996) identified four classes of extremal dependence. The first class is that of asymptotically dependent distributions. The other three classes comprise distributions with asymptotically independent dependence structures exhibiting positive extremal dependence, near extremal independence and negative extremal dependence for a d-dimensional variable Y. These three classes correspond respectively to joint extremes of Y occurring more often than, approximately as often as or less often than joint extremes if all components of the variable were independent (see Ledford and Tawn, 1996, for more details). In this study, the focus is to clarify the performance of interpreting dependence structure (see Equation 5.2) by the empirical Bayes estimation approach for different types of dependence derived in detail in Section 8 of Heffernan and Tawn (2004). We choose the extremal dependent types below to be simulated, which are also described in Keef et al. (2013a).

Independence. Here $\alpha = \beta = 0$ and G factorizes into Laplace distribution functions.

Asymptotic dependence. Here $\alpha = \pm 1$ (with +1 indicating positive dependence and -1 implying negative dependence) and $\beta = 0$ and G takes a range of forms.

Asymptotic independence. Variable X is asymptotically independent of variable Y if $-1 < \alpha < 1$.

The simulated data for analyzing these three types are randomly generated from bivariate extreme value distributions using the R (R Core Team, 2013) package POT (Ribatet, 2006). Mainly two types of models are used as listed in Table 5.1. The experiment is particularly

designed to see the performance in estimating dependence parameters α and β with different exceedance sizes based on the proposed empirical Bayesian approach. Initially, sample sizes of 1000, 3000, 5000, 10000 are generated and the high threshold of 99% quantile is selected for all cases to keep independence between Z and Y , thus exceedence sizes of 10, 30, 50 and 100 are left to be experimented on (see an example in Figure 5.1 and Laplace-transformed results in Figure 5.2). All simulation cases are repeated for 800 trials, to see the percentage of times (over 800 trials) that the true parameter(s) fell within the estimated 95% credible interval (henceforth, 95% CI), and to explore how the sample sizes (i.e., 10, 30, 50 and 100) might affect estimation inference. Ideally, the percentage of fall-in times for the true parameter(s) should be approximately 0.95 when considering a 95% CI. Another aspect of this experiment also compares simulations given vague priors (uniform distributions with wide support and random initials) for dependence parameters with relative informative ones to check the consequential effect from initials and priors.

5.3.2 Simulation Results

The selected three different forms of dependence structures described in section 5.3.1 (Figure 5.1 and 5.2) are tested with the proposed empirical Bayes estimation approach. For the simulation experiment, with exceedance sizes of 10, 30, 50 and 100, the percentage of times that the true parameter(s) fell within the estimated 95% CI are shown for parameters α and β individually, as well as for when both parameters fell within the bounds simultaneously in Table 5.2 and 5.3. Table 5.2 summarizes the simulation results associated with vague priors for dependence parameters, while Table 5.3 displays the results having used informative priors. In both tables, it is clear that results improve, if only slightly, with increasing sample sizes, except for the asymptotic dependence case, where the estimation performs relatively poorly for the parameter,

α. This inefficiency may be caused by the fact that it is a special case of estimating a single point in a continuous parameter space (at least for one type of exact dependence). By comparing Table 5.2 and Table 5.3, with Informative Priors, the performance of hit percentage (the true parameter(s) fell within the estimated 95% CI over 800 trials) for either α or β, or both are much better. For example, in the Independence case, with exceedances of 50, Table 5.3 shows that the individual parameter as well as both have the hit percentage over 0.9, while in Table 5.2 the inference performance is 0.5 on average. That is, over 800 trials, the true parameter(s) fell within the estimated 95% CI with Informative Priors over 720 times, while it fell in the interval approximately only 400 times, on average, when using Vague Priors. In some cases, such as in the asymptotic independence case with exceedances over 10 (e.g., exceedance size of 50 and 100, then the percentage for both is around 0.914 and 0.915), or in the independence simulation with larger sample size (e.g., same exceedance sizes, the percentage for α is about 0.961 and for β is approximately 0.958), the percentage even reaches over the ideal situation which is around 0.95. The identifiability issue with β may be because the model at some point is actually $X^* = \alpha Y + Y^{\beta}(Z - \mu)/\sigma$, where μ and σ are the mean and standard deviation vectors of Z, respectively, and as noticed, it is not possible to differentiate β from μ and σ.

We feel that it is reasonable to obtain prior information for β so that this issue is not a major concern. Overall, simulation results show that the proposed approach performs fairly well for most types of dependence structures, but strong and reasonable prior information for β is generally necessary. And the benefit from increasing samples is not so obvious as using Informative Priors.

5.4 Temperature and Precipitation Test Case

5.4.1 Background

In this study, the proposed empirical Bayes estimation approach is further applied on two real data cases, and simulated results are compared with the (slightly modified) HT estimation strategy as implemented by the R (R Core Team, 2013) package *texmex* (Southworth and Heffernan, 2010). Monthly observations of precipitation and temperature from the Climatic Research Unit (CRU; New et al., 2000; Mitchell and Jones, 2005) regridded in a common 2 × 2-degree spatial resolution, are used to provide the test case data. Historical monthly precipitation and temperature records are available from 1901 to 2009. CRU observations have been validated and used in numerous studies of historical climate variability (e.g. Tanarhte et al., 2012; Hao et al., 2013). To identify extreme conditions, two grid points in the central (Latitude 40.02 °N, Longitude 105.27 °W) and western (Latitude 34.05 °N, Longitude 118.24 °W) United States are selected for conditional extreme dependence structure analysis in Figure 5.3. The two locations are close to urban areas in Boulder, Colorado and Los Angeles, California where long-term observation stations have been available. In both cases, we consider the precipitation conditional on the temperature's being extreme, i.e., precip | temp $>$ u, for u large. The marginal threshold level (i.e., u) of temperature at the two locations is corresponding to the 0.97 quantile. For the upper tail of precipitation data, the threshold level Los Angeles is taken to be the 0.97 quantile, and for Boulder, the 0.99 quantile. The proposed empirical Bayes approach is applied to infer the dependence structure parameters, α describing the dependence strength between precipitation and extremal temperature and β outlining the dependence variability, along with the scale and shape parameters for the upper tail of precipitation and temperature distributions.

5.4.2 Test Case Results

Using the CRU precipitation and temperature monthly data (see Figure 5.3), analyzing precip | temp $> u$, for u large, the dependence structure controlling parameters, α and β, and distribution parameters of σ and ξ fitted with the GP df for each variable at two locations (Boulder and Los Angeles) are presented in Table 5.4. The parameters derived by the empirical Bayes estimation approach and the HT model are simulated with informative priors for the α and β. In both locations, the dependence relationship of precipitation given extreme temperature tends to show asymptotic independence, and in Los Angeles, it is near extremal independence, while in Boulder, it is towards negative extremal dependence. Looking at the parameter β, it appears that the variability of the dependence is relatively lower in Los Angeles than that in Boulder. Parameters σ_1 and ξ_1 describe the scale and shape of temperature data which indicates a bounded upper tail distribution, while σ_2 and ξ_2 stand for the precipitation distribution. In Los Angeles, the precipitation distribution shows a heavy tail property (indicated by $\xi_2 > 0$), while in Boulder, it shows a bounded upper tail (see $\xi_2 < 0$). Table 5.4 also compares results using the empirical Bayes estimation approach and the R (R Core Team, 2013) package *texmex* for the HT method. In general, from the table, we can see that all the parameters, including dependence structure parameters α and β, and GP df parameters σ and ξ, inferred by the two approaches are consistent with each other.

5.5 Summary, Conclusions and Discussion

The conditional EVA approach introduced by Heffernan and Tawn (2004) is an important new methodology for modeling multivariate extreme values through a conditional distribution framework. Although this approach does not require a priori knowledge of the dependence structure nor that the variables be simultaneously extreme, a difficulty for estimating the

parameters is that no simple, closed-form distribution exists in general for G. Therefore in the original approach, several estimation methods and constraints are mixed together to evaluate the G df and counteract the inherent bias. This disadvantage motivates the development of the empirical Bayes estimation approach proposed in this study.

Simulations are employed to reproduce known dependence structures with 800 repeated trials for each of three types of dependence and sample sizes, and to test how well the estimation procedure performs. Additionally, precipitation data conditional on having extreme temperature is also analyzed and compared to the (slightly modified) estimation strategy proposed by Heffernan and Tawn (2004). Simulations show generally good coverage of credible intervals. However, the parameter β is relatively hard to infer precisely, and sometimes it is not unique, so strong prior information for β is generally necessary, which is the primary hold-back in this approach and might require further refinement.

The identifiability issue with β may result from the model at some point falling into the function of $X^* = \alpha Y + Y^\beta (Z - \mu)/\sigma$, where μ and σ are the mean and standard deviation vectors of Z, respective, and as noticed, it is difficult to differentiate β from μ and σ. To possibly solve this issue, one may include μ and σ as parameters of interest in the empirical likelihood estimation by imposing other prior knowledge for those parameters, but still, the identifiability problem might not disappear. Another possibility is that one may explore and include higher order moments (other than first and second moments in this study) of the conditional distribution as empirical likelihood constraints, which may also be difficult to identify without any additional assumption because the conditional distribution is generally, numerically unknown. As for the inefficiency to estimate a single point in a continuous parameter space (at least for one type of exact dependence), a possible extension for the estimation model would be to use a reversible jump

Markov Chain to include probability terms for special cases (i.e., $\alpha = 0, \beta = 0, \alpha = \pm 1, \beta = 1$) analogously as proposed in the univariate setting by Stephenson and Tawn (2004) for the shape parameter, $\xi = 0$ versus $\xi \neq 0$. Such a scheme might help with the estimation in the asymptotic dependence case and will be explored in a future study.

The presented model can be potentially applied in a wide variety of science fields including finance, earth science, environmental science, and biology. Particularly, this model can be used for assessing spatial climatic extremes (see Gilleland et al., 2013). A myriad of papers show climatic extremes have been changing and are projected to change in the future (e.g. Wehner, 2013; AghaKouchak et al., 2013; Field et al., 2012; Schubert and Lim, 2013; Easterling et al., 2000; Alexander et al., 2006). Even concurrent extremes (e.g., joint precipitation and temperature extremes) have been reported to have increased/changed over time (Hao et al., 2013). The proposed methodology allows assessing one extreme variable conditioned on another and hence, we expect it to be a useful tool for conditional extreme value analysis.

Table 5.1 Dependence models for bivariate extreme value distributions used in this study

Dependence Models	**Negative Logistic (nlog)**	**Logistic (log)**
Formula	$V(x,y) = \frac{1}{x} + \frac{1}{y} - (x^{\gamma} + y^{\gamma})^{-\frac{1}{\gamma}}$	$V(x,y) = (x^{-\frac{1}{\gamma}} + y^{-\frac{1}{\gamma}})^{\gamma}$
Pickands' Dependence	$A(\omega) = 1 - [(1-\omega)^{-\gamma} + (\omega)^{-\gamma}]^{-\frac{1}{\gamma}}$	$A: [0,1] \rightarrow [0,1];$ $\omega \mapsto [(1-\omega)^{\frac{1}{\gamma}} + (\omega)^{\frac{1}{\gamma}}]^{\gamma}$
Independence	$\gamma \rightarrow 0$	$\gamma = 1$
Total Dependence	$\gamma \rightarrow +\infty$	$\gamma \rightarrow 0$

Table 5.2 Results from fitting the conditional extreme value model to simulated data using the empirical Bayes estimation approach with *Vague Priors* proposed here. For each (exceedance) sample size, the percentage (of 800 trials) given is the percentage of times that the *true* parameter(s) fell within the estimated 95% CI's. Results are shown for parameters individually, as well as for when both parameters fell within the bounds simultaneously.

	Exceedance Size	10	30	50	100
Asymptotic Independence	α	0.796	0.666	0.745	0.701
	β	0.821	0.883	0.938	0.874
	both	0.661	0.588	0.713	0.630
Independence	α	0.800	0.639	0.773	0.695
	β	0.499	0.409	0.396	0.349
	both	0.433	0.301	0.336	0.296
Asymptotic Dependence	α	0.674	0.518	0.585	0.278
	β	0.685	0.674	0.633	0.484
	both	0.439	0.331	0.345	0.135

Table 5.3 Results from fitting the conditional extreme value model to simulated data using the empirical Bayes estimation approach with *Informative Priors* proposed here. For each (exceedance) sample size, the percentage (of 800 trials) given is the percentage of times that the *true* parameter(s) fell within the estimated 95% CI's. Results are shown for parameters individually, as well as for when both parameters fell within the bounds simultaneously.

	Exceedance Size	10	30	50	100
Asymptotic Independence	α	0.941	0.708	0.946	0.923
	β	0.941	0.991	0.960	0.993
	both	0.890	0.704	0.914	0.915
Independence	α	0.965	0.975	0.961	0.940
	β	0.781	0.786	0.934	0.958
	both	0.764	0.768	0.901	0.905
Asymptotic Dependence	α	0.844	0.585	0.741	0.995
	β	0.956	0.929	0.980	0.636
	both	0.823	0.560	0.730	0.634

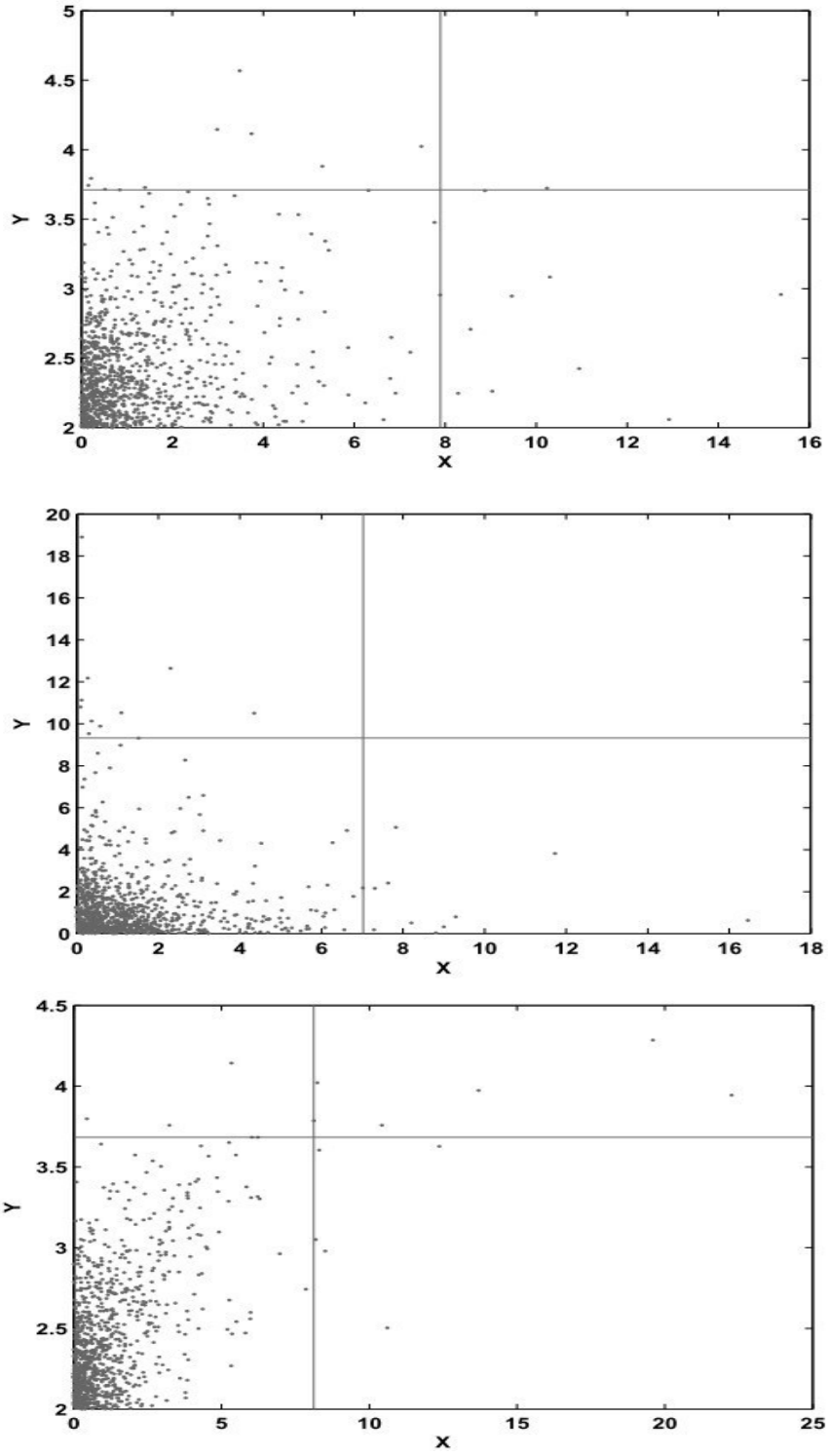

Figure 5.1 Scatter plots of randomly generated data of sample size 1000. The marginal threshold level, corresponding to the 0.99 quantile is shown in grey lines. Data are shown on original scales. Asymptotic Independence (upper), Independence (middle), Asymptotic dependence (lower).

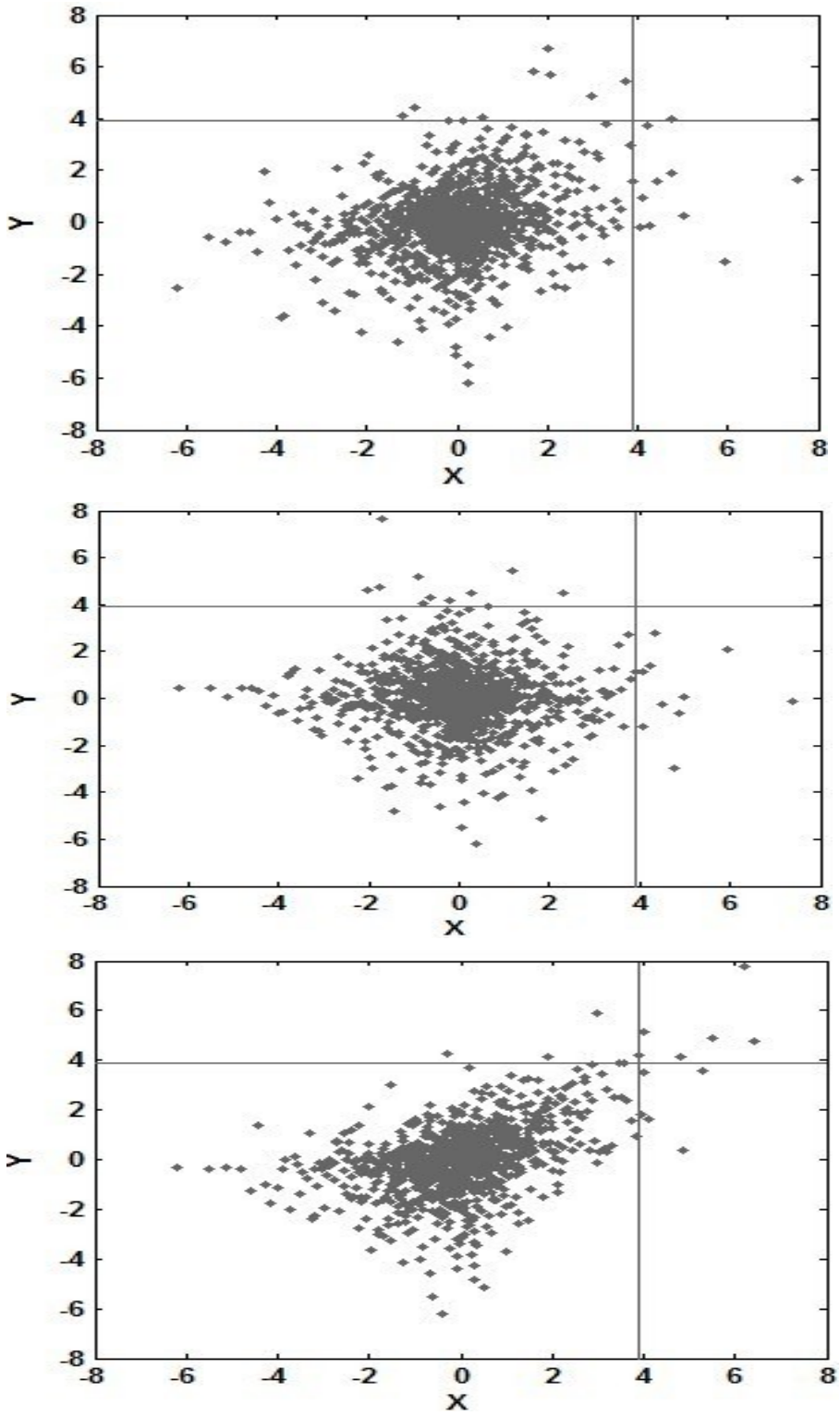

Figure 5.2 Scatter plots of randomly generated data of sample size 1000. The marginal threshold level, corresponding to the 0.99 quantile is shown in grey lines. Data are shown on Laplace transformed scales. Asymptotic Independence (upper), Independence (middle), Asymptotic dependence (lower).

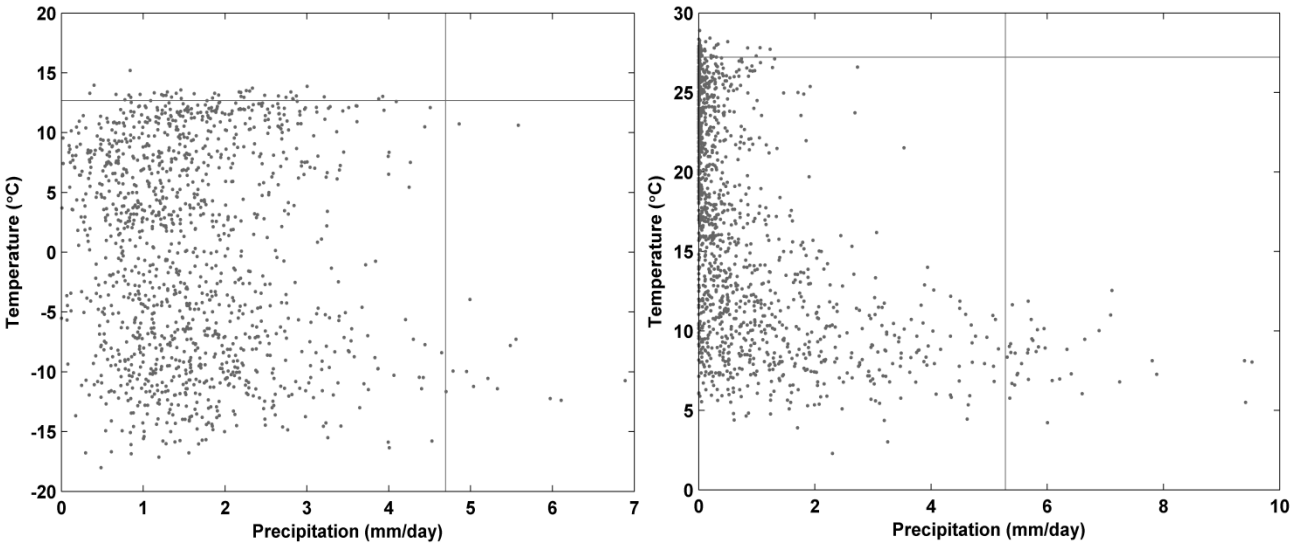

Figure 5.3 Scatter plots of Precipitation and Temperature data at Boulder, Colorado, U.S.A. (left), and Los Angeles, California, U.S.A. (right). The marginal threshold level for temperature, corresponding to the 0.99 quantile at both locations; quantile of 0.99 at Boulder and quantile of 0.97 at Los Angeles for the precipitation, respectively, are shown in grey lines.

Table 5.4 Comparison results from fitting the conditional extreme value model to real data using the empirical Bayes estimation approach and *texmex*

Los Angeles	**Exceedance**	α	β	σ_1	ξ_1	σ_2	ξ_2
Empirical Bayes	**40**	**-0.065**	**-0.559**	**0.733**	**-0.367**	**0.820**	**0.193**
	Std. Dev	**0.070**	**0.214**	**0.028**	**0.043**	**0.044**	**0.046**
texmex	**40**	**-0.071**	**-0.568**	**0.688**	**-0.338**	**0.849**	**0.140**
Boulder	**Exceedance**	α	β	σ_1	ξ_1	σ_2	ξ_2
Empirical Bayes	**39**	**-0.230**	**-0.270**	**0.683**	**-0.143**	**0.890**	**-0.251**
	Std. Dev	**0.074**	**0.334**	**0.033**	**0.027**	**0.842**	**0.060**
texmex	**39**	**-0.249**	**-0.284**	**0.648**	**-0.149**	**0.802**	**-0.143**

CHAPTER 6: An Empirical Bayes Conditional Extreme Value Model for Detecting Changes in the Hydrological Cycle

6.1 Introduction

Atmospheric greenhouse gases have been increasing since the industrial revolution, leading to the Earth warming through an increase in downwelling infrared radiation (Trenberth 1999). The warmer atmosphere's greater water holding capacity will likely intensify the hydrological cycle (Huntington 2006; ACIA 2004; Ding et al., 2001; NAST 2001). The Clausius–Clapeyron relation, which characterizes transition between two phases of single-constituent matter, indicates that specific humidity increases approximately exponentially with temperature (Joshi et al., 2008; Willett et al., 2007; Trenberth et al., 2003). Therefore, climate warming is theoretically expected to cause increases in evaporation and precipitation (Allen and Ingram, 2002) and ultimately acceleration in water cycle processes (McCarthy et al., 2001; Held and Soden, 2000; Trenberth, 1999; Loaciga et al., 1996; Del Genfo et al., 1991). This is already shown in some studies; for example, one modeling study suggests that precipitation could increase by about 3.4% per degree Kelvin (Allen and Ingram, 2002).

However, other important research questions remain regarding whether increases in global temperatures have actually changed the hydrological cycle significantly in the past century and whether the cycle will intensify in the future (Ohmura and Wild, 2002). Empirical trends in observed precipitation, temperature, snow-water equivalent, and soil moisture at regional to continental scales confirm climate changes over time (Huntington 2006; Robock et al., 2000; Brown 2000). Observations show increased precipitation in northern Australia, central Africa, Central America and parts of southwest Asia, and drying trends in western United States and the Mediterranean region (Damberg and AghaKouchak 2013). Due to rising temperatures and

subsequent increases in both atmospheric moisture content and moisture transport into storms, the probable maximum precipitation (PMP) is expected to increase globally (Kunkel et al., 2013). Even concurrent extremes (i.e., warm-dry and warm-wet conditions) have increased significantly in the second half of the 20th century (Hao et al., 2013). Such changes in the hydrological cycle will affect water availability, frequency and intensity of tropical storms, floods, droughts, and potentially amplify warming through water vapor feedback (Wu et al., 2013; Wehner, 2013; Schubert et al., 2013; Brekke and Barsugli, 2013; Huntington 2006; Deardorff 1978). For this reason, methods for assessing and predicting potential changes in the hydrological cycle have received a great deal of attention recently (e.g., Wu et al., 2013; Trenberth et al., 2007; Zhang et al., 2007; Huntington 2006; Trenberth 1999).

Several methods for evaluating changes in climatic variables have been developed, including the forest machine-learning approach (Loosvelt et al., 2012; Liaw and Wiener, 2002), fingerprint, multivariate and multi-fingerprint techniques (Marvel and Bonfils, 2013; Santer et al., 2013), the generalized Bayesian approach based on subjective probabilities (Schnur and Hasselmann, 2005; Hasselmann 1998) and others (see also, Lee et al., 2005; Barnett et al., 1999; Hegerl et al., 1996). Numerous indices for monitoring changes in climatic extremes have also been developed (e.g., Zhang and Zwiers, 2013; Zhang et al., 2011). Most of these methods, indices, and trend studies focus on changes in one variable at a time. However, hydrologic variables are dependent and a change in one variable can alter extreme and non-extreme values of other variables (Leonard et al., 2014). Temperature, for example, drives the hydrologic cycle and has a profound effect on rainfall (Gyasi-Agyei 2013). Concurrent climatic extremes have been evaluated using both empirical methods (e.g., Hao et al., 2013; Beniston, 2009) and multivariate extreme value analysis (Salvadori and De Michele, 2013). The latter can analyze two or more concurrent

extremes (e.g., concurrent extreme precipitation and temperature). However, it cannot assess changes in one full distribution conditioned on extremes of another variable (e.g., changes in the full distribution of precipitation including non-extreme values conditioned on extreme temperature). To address this major research gap in current available methods, we present a generalized statistical framework for assessing conditional extremes.

Heffernan and Tawn (2004) introduce a new approach for modeling multivariate extreme values through a conditional distribution framework, which has been used in economics and finance literature. A drawback of this methodology however, is that without assuming a specific dependence structure, there is no simple closed-form conditional distribution. Therefore, we offer an empirical Bayes estimation strategy for modeling conditional multivariate extremes based on the previous model. Our proposed methodology circumvents the prior method's drawback, quantifies uncertainty of all parameters involved simultaneously, computes statistics of extremes efficiently, and is applicable across different spatial scales.

The proposed methodology uses the empirical likelihood technique developed and proved by Owen (1988) as a robust alternative to classical likelihood approaches or the bootstrap (Mengersen et al., 2013; Owen 1990). This concept is different from how empirical Bayes has been used in previous studies — to define an appropriate prior distribution (e.g., Smith, Marshall, and Sharma, 2014; Goodman 2010). When the likelihood function is numerically unavailable or not entirely known, empirical likelihood method appears to be a valuable approach even for limited observations (Li, 1995; Jing, 1995; Chen, 1994; Qin and Lawless, 1994; Hall and La Scala, 1990). Therefore, empirical likelihood provides a unique avenue to tackle the difficulty in the conditional likelihood estimation. This chapter highlights the value of empirical Bayes conditional extreme value analysis as a tool for simulating and assessing

conditional extremes (e.g., changes in the distribution of precipitation conditioned on extreme temperature).

6.2 Methodology

Refer to the Chapter 5.2.

6.3 Data and Study Areas

In this study, we propose the empirical Bayes estimation approach to explore changes in the precipitation given extreme temperatures. Monthly observations of precipitation and temperature from the Climatic Research Unit (CRU; Mitchell and Jones, 2005; New et al., 2000) regridded to a common 2 × 2-degree spatial resolution are used. CRU observations are validated and used in numerous climate studies (e.g., Hao et al., 2013; Tanarhte et al., 2012). To investigate changes in precipitation conditioned on extreme temperature, the proposed framework is first applied to five selected locations in the central and western U.S., northern China, Europe and Australia. The methodology is also applied to the continental United States and Australia to show example applications of the presented method at large spatial scales. The five selected locations are close to urban areas in Austin, Texas (Latitude 30.25° N, Longitude 97.75° W), Los Angeles, California (34.05° N, 118.25° W), Beijing (39.91° N, 116.39° E), Paris (48.86° N, 2.35° E) and Canberra (35.31° S, 149.12° E), near long-term observation stations. In both local scale and continental scale applications, the data from 1910 to 2009 is used and separated equally into two periods: 1910-1959 (first 50 years) and 1960-2009 (second 50 years).

6.4 Results

Using our approach, the observed precipitation and temperature during 1960-2009 are evaluated against the baseline period (1910-1959). The conditional distributions (i.e., precip. | temp. $>$ u,

with u being 0.95 quantile) simulated by the proposed model, for the five selected points are presented in Figure 6.1. The 0.95 quantile (i.e., u) of monthly temperature in the five selected locations are approximately 29.53°C (Austin), 26.32°C (Los Angeles), 24.85°C (Beijing), 19.30°C (Paris), and 20.85°C (Canberra). Figure 6.1 shows the univariate distributions of precipitation (1st column), temperature (2nd column), and the distribution of precip. | temp. > u (3rd column) for the two fifty-year periods: 1910-1959 (blue) and 1960-2009 (red). The figure shows how distributions have changed in the target period relative to the baseline.

The t-test is applied to the distributions in Figure 6.1 to examine whether the distributions have changed significantly (at the $\alpha = 0.05$ significance level) in the two periods. The test examines the null hypothesis of no change between the two distributions (i.e., no statistically significant changes in the means and standard deviations). While there are differences between the univariate precipitation and temperature distributions, the t-test results do not reject the null hypothesis of no change in the univariate distributions (1st and 2nd columns in Figure 6.1) at the 95% confidence level ($\alpha = 0.05$). In other words, the test indicates no statistically significant change in the univariate distributions of recitation and temperature in the selected locations. However, the t-test rejects the null hypothesis of no change for the distributions of the precipitation conditioned on high temperature (i.e., precip. | temp. > u) indicating a statistically significant change at the 95% confidence level – 3rd column in Figure 6.1. The five examples show that while the entire distributions of precipitation has not changed significantly in the selected areas, the distribution of precipitation conditioned on high temperature has changed significantly. This is consistent with the theoretical expected change in precipitation in a warming climate (Trenberth, 1999; Allen and Ingram, 2002). The results show that the presented

methodology reveals information that otherwise cannot be obtained from univariate statistics of individual variables.

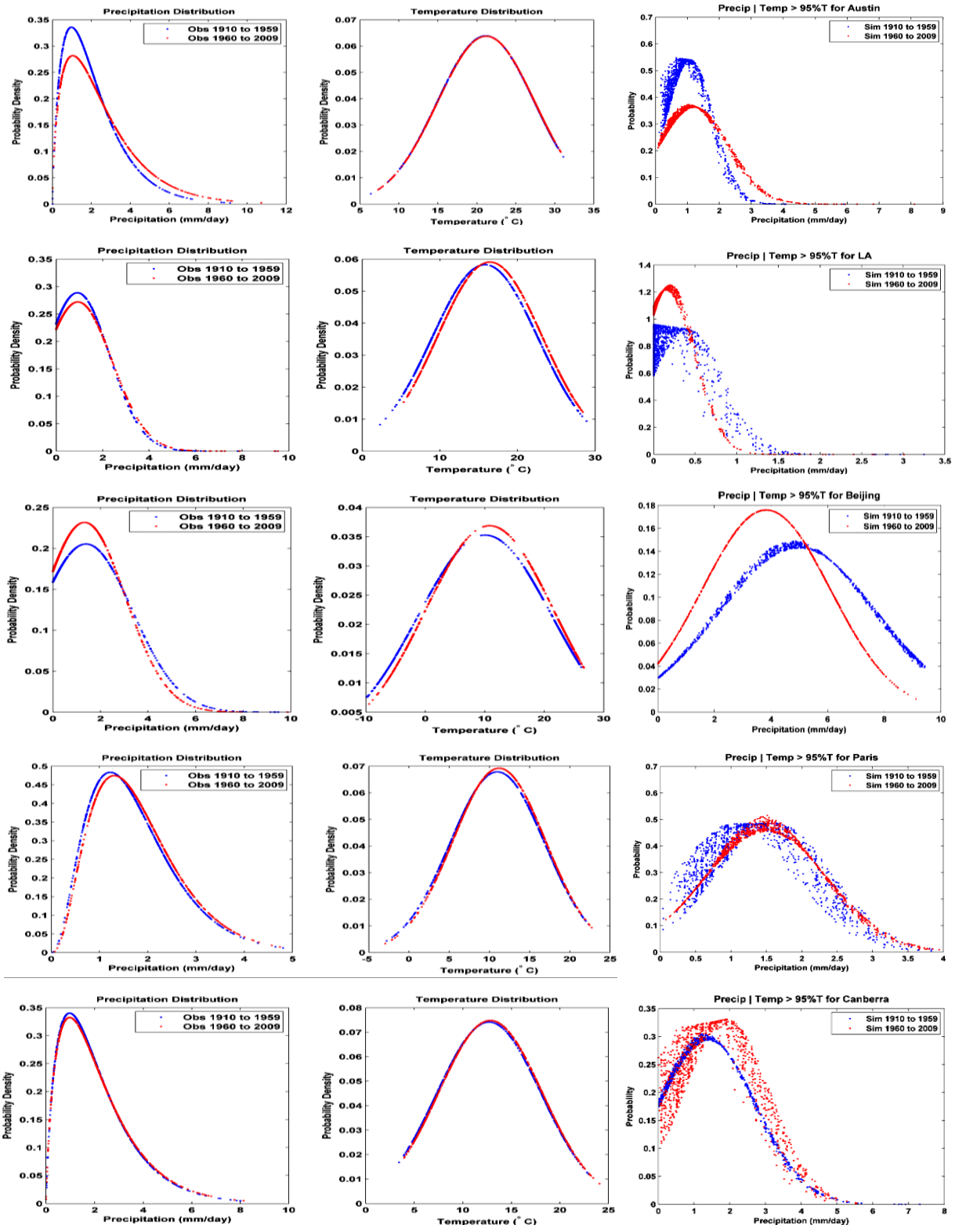

Figure 6.1 Distributions of precipitation (left), temperature (middle) and precipitation conditioned on high temperature (right) for 1910 to 1959 (blue) and 1960 to 2009 (red) over the selected five locations: Austin, TX, Los Angeles, CA, Beijing, China, Paris, France, and Canberra, Australia.

To demonstrate example applications over large spatial scales, the presented methodology is applied over the United States and Australia. The results are cross-checked with the commonly used Mann-Kendall (Kendall 1976; Mann 1945) trend analysis, which is widely used in hydrology (Zhang et al. 2004). For the two periods, mentioned in Figure 6.1, changes in precipitation, temperature and precipitation conditioned on high temperature are evaluated over summer and winter separately. In the Mann-Kendal test, the null hypothesis of no trend is rejected if the test statistic is significantly different from zero. Figure 6.2 and 6.3 present the Mann-Kendal trend test results for summer and winter, respectively. In these figures, a positive trend (blue) means an increasing pattern in precipitation, while a negative trend (red) indicates a decreasing precipitation pattern. The white areas in both figures do not show a statistically significant trend at the 95% confidence level. Overall, in summer, more areas exhibit a statistically significant trend in precipitation (Figure 6.2) compared to winter (Figure 6.3). In Figure 6.2, the western U.S. exhibits a decreasing trend (red), whereas part of southern U.S. and western Mexico show an upward trend. More importantly, comparing the trends in the two periods (left and right panels in Figure 6.2) indicates that a similar trend appears in both periods. On the other hand, winter data (Figure 6.3) shows an increasing trend in precipitation in the second 50-year period (compare the two panels in Figure 6.3).

Considering the precipitation conditioned on high temperature reveals substantially different patterns of change in summer and winter (Figure 6.4). Using the t-test, the significance of changes in precipitation conditioned on high temperature is evaluated for summer and winter in Figures 6.4 (left) and 6.4 (right), respectively. In these figures, the blue pixels indicate that the mean of the conditional rainfall distribution has changed significantly, and that the rainfall has increased in the second period compared to the baseline. On the other hand, red pixels

correspond to areas where precipitation conditioned on high temperature has decreased significantly at the 95% confidence level. Unlike Figure 6.2 and 6.3 where trends are evaluated for the entire distribution of precipitation, Figure 6.4 only focuses on precipitation during warm months. One can see that there are substantial differences between spatial patterns of univariate statistics (Figure 6.2 and 6.3) and the presented conditional extremes (Figure 6.4). For example, while the entire distribution of precipitation shows a decrease in the overall precipitation in the western U.S., precipitation during warm months has increased significantly in the target period relative to the baseline. This can be explained by potentially more moisture transport in the warm months from the Pacific Ocean toward the western United States. On the contrary, the areas around the Rocky Mountains exhibit and increase in the overall precipitation. However, precipitation conditioned on high temperature has decreased substantially relative to the baseline. Figure 6.4 also highlights opposite signs of change in summer and winter around Midwest United States. The patterns indicate more precipitation during summer and less during winter in the latter 50 years of the observations. It is worth pointing out the opposite sign of change in Midwest winter precipitation in Figure 6.3 that shows the entire distribution of data. Finally, Figure 6.4 indicates opposite signs of change in conditional precipitation over the southeastern United States that is dominated by hurricanes and tropical storms.

Both the Mann-Kendall trend test and the presented conditional extreme value model have been applied to Australia to evaluate changes in precipitation. The Mann-Kendall trend analyses results are presented in Figure 6.5 and 6.6, for summer and winter, respectively. One can see that the trends in the two periods are similar over Australia. In both winter and summer, most of Australia exhibits a drying trend with respect to the full distribution of precipitation (Figure 6.5 and 6.6), except the southern part of Australia where an upward trend can be observed. The

proposed empirical Bayes conditional framework, however, reveals significantly different spatial patterns of change for precipitation conditioned on high temperature (Figure 6.7). The results show that precipitation conditioned on high temperature increases in eastern Australia while it decreases in the western part. Separated by the Macdonnell Ranges and the Great Victoria Desert, most areas on the eastern part including Queensland, New South Wales and Victoria exhibit a wetting pattern during warm months, while the areas around the Western Plateau exhibit opposite patterns of change in the conditional distribution of precipitation. Unlike the United States, in Australia the precipitation distributions conditioned on high temperature are relatively similar in both summer and winter.

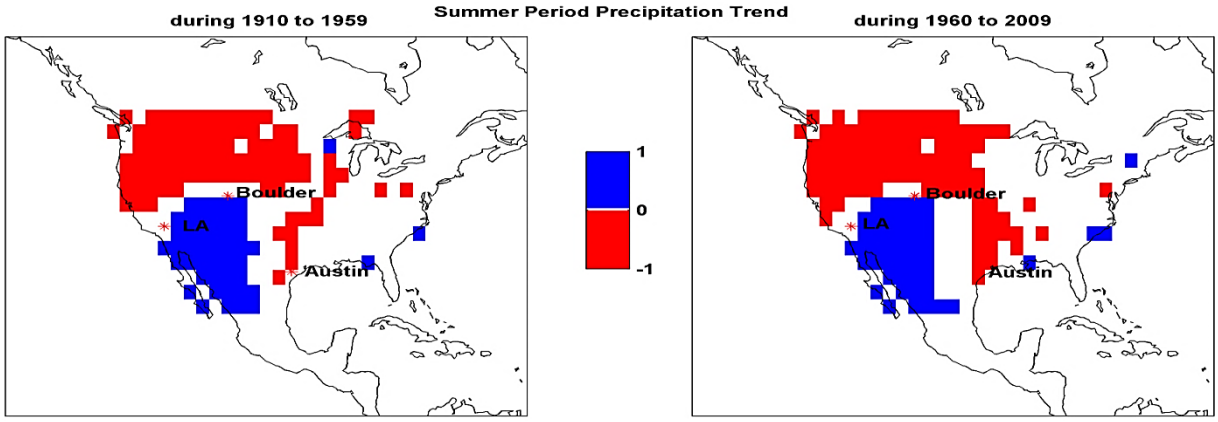

Figure 6.2 Summer (June, July and August) precipitation trends in 1910-1959 (left) and 1960-2009 (right) over United States.

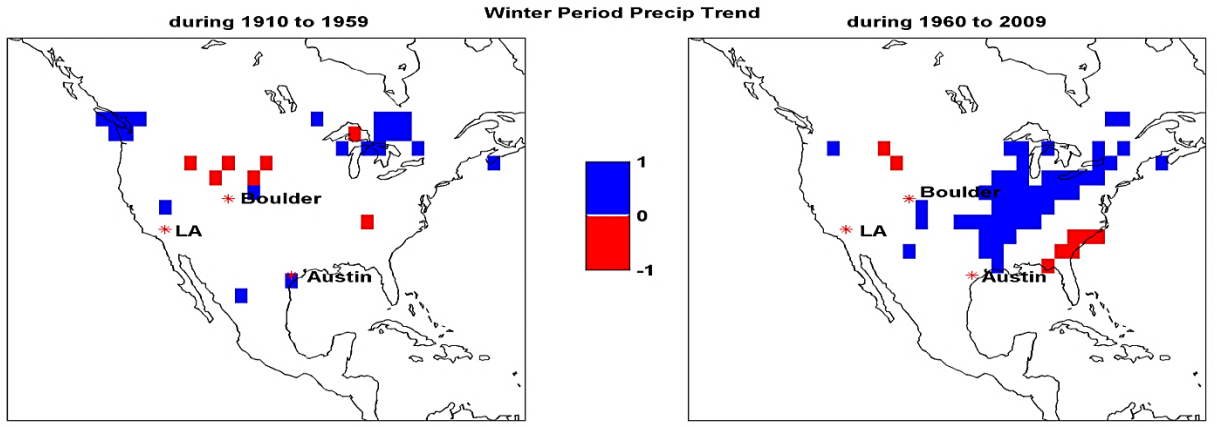

Figure 6.3 Winter (June, July and August) precipitation trends in 1910-1959 (left) and 1960-2009 (right) over United States.

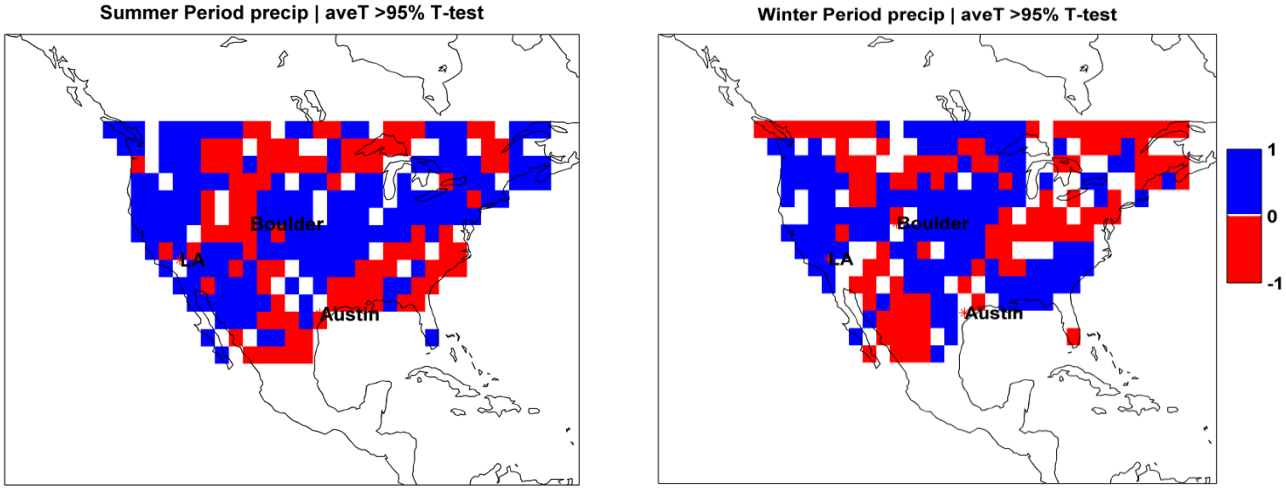

Figure 6.4 Changes in precipitation conditioned on temperature higher than its 95th quantile in summer (June, July and August) (left) and winter (December, January and February) over United States. In blue pixels, during 1960-2009, the mean of the conditional rainfall distribution has increased, whereas in red pixels it has decreased relative to 1910~1959.

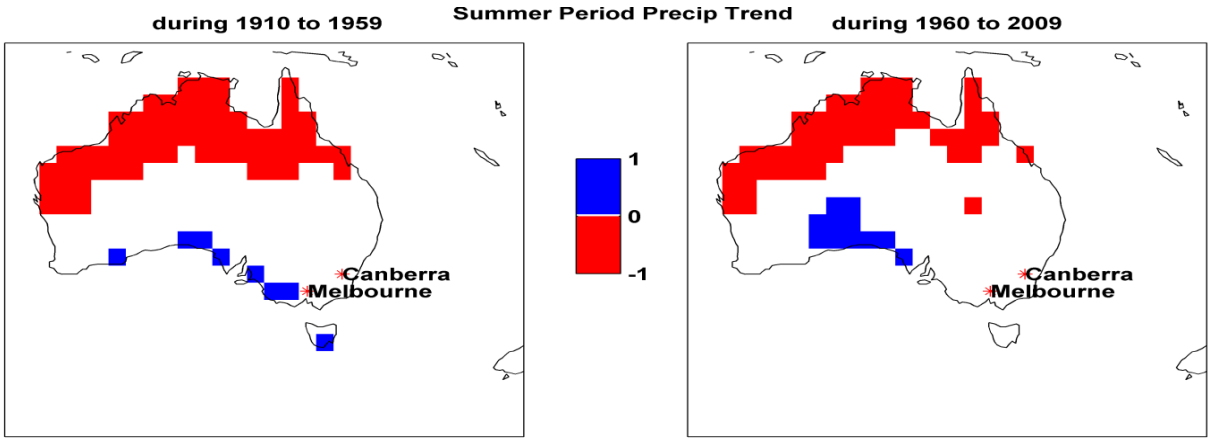

Figure 6.5 Summer (June, July and August) precipitation trends in 1910-1959 (left) and 1960-2009 (right) over Australia.

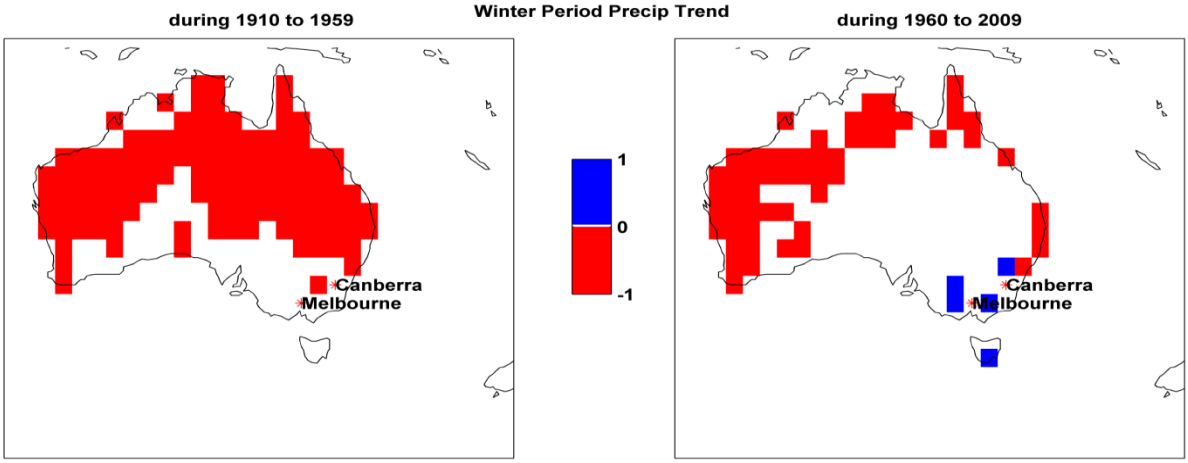

Figure 6.6 Summer (June, July and August) precipitation trends in 1910-1959 (left) and 1960-2009 (right) over Australia.

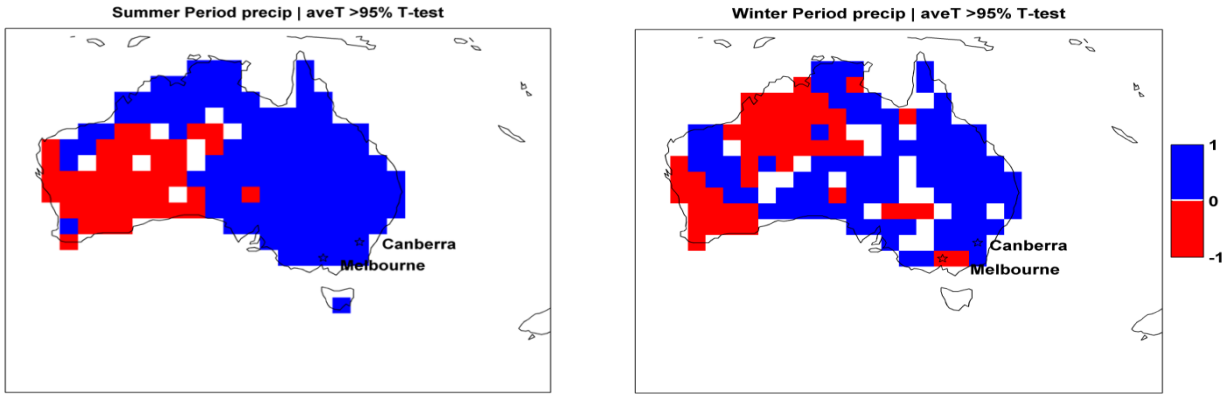

Figure 6.7 Changes in precipitation conditioned on temperature higher than its 95th quantile in summer (June, July and August) (left) and winter (December, January and February) over Australia. In blue pixels, during 1960-2009, the mean of the conditional rainfall distribution has increased, whereas in red pixels it has decreased relative to 1910~1959.

CHAPTER 7: A Methodology for Deriving Ensemble Response from Multi-model Simulations

7.1 Introduction

Several national and international efforts, such as the Intergovernmental Panel on Climate Change (IPCC; IPCC 2007), provide data sets of historical and future climate. However, climate simulations are subject to uncertainties arising from uncertainties in boundary, and initial conditions, parameters and model structure (Reichler and Kim, 2008; Feddema et al., 2005; Brekke and Barsugli, 2013; Foley 2010; Deque et al., 2007; Liepert and Previdi, 2012; Wehner 2012; AghaKouchak et al., 2013; John and Soden, 2007). Multi-model ensembles have been widely employed to quantify uncertainties of climate simulations (Meehl et al., 2007; Yun et al., 2003; Tebaldi and Knutti, 2007). Model simulations are also used to force hydrologic and land-surface models to derive hydrology projections. Previous studies have confirmed that a multi-model ensemble approach increases the skill of model simulations (Doblas-Reyes et al., 2003; Cantelaube and Terres, 2005). Regardless of the method of estimation, an ensemble consists of a number of realizations (individual climate simulations), each of which representing a probable climate condition that can occur. While a multi-model ensemble approach increases the skill of model simulations, one may need to know which ensemble member is more likely to be true, particularly when the ensemble is spread out over a wide area. It is customary to derive the ensemble response or prediction quantity (hereafter, climate response) of multi-model ensembles by taking the arithmetic mean of simulated ensemble members (Min et al., 2007) where an equal weight is given to each ensemble member. Masson and Knutti (2011) stressed that strong similarities exist between several models (members of an ensemble) which may cause biased climate response toward models with strong similarities. One way to combine simulations of

climate models is to weight ensemble members based on their performance in simulating past and present climate (e.g., Krishnamurti et al., 2000). Knutti et al. (2010) argues that while the ensemble mean provides useful information, there exist the need for more quantitative approaches to assess model simulations in order to maximize the value of multi-model ensemble climate simulations. In recent years, Bayesian model averaging has also been used to derive the climate response of multi-model ensembles (e.g., Smith et al., 2009; Robertson et al., 2004; Tebaldi et al., 2004; Min et al., 2007). Limitations of the Bayesian methodology, when applied to climate projections, are addressed in Tebaldi and Knutti (2007). For a weighted average approach, quantifying the weights requires an index of model skill in order to estimate the weights accordingly. Several studies have tackled this issue and contradicting results are presented on the best method to combine climate model projections (see Tebaldi and Knutti (2007) and references therein). Among many reasons, the choice of model skill, and strong dependencies and similarities of ensemble members are the main challenges in deriving a meaningful climate response. In order to resolve this limitation, a model is proposed for deriving the climate response of climate model simulations. In the proposed method, ensemble members are weighted based upon their performance in simulating observations using the so-called Expert Advice algorithm (Cesa-Bianchi and Lugosi, 2006). The goal of this methodology is to derive the weights (predicting models) such that at every time step the climate response is equal or better (less error) than the best model. In most studies that rely on climate model simulations, simulated anomalies are used instead of the absolute values to remove biases in individual model simulations (e.g., Collins et al., 2011). However, in hydrology and water resources studies, often the absolute values of model simulations are necessary. For example, to run a hydrologic model with climate simulations as forcing (e.g., Ficklin et al., 2009), the original model simulations are

used and not the anomalies. Similarly, when temperature and/or precipitation simulations are used for multivariate analysis (Hao and AghaKouchak, 2013), climate impact assessment (Madani and Lund, 2010), drought analysis (Madadgar and Moradkhani, 2011), etc. The suggested algorithm can be applied to both original ensemble simulations and their anomalies.

7.2 Data

Climate Model Simulations: In this study, 41 Coupled Model Intercomparison Project Phase 5 (CMIP5) historical annual temperature simulations from 1951 to 2005 are used for analysis. These data sets represent the most extensive and ambitious multi-model simulations that contribute to the World Climate Research Programme's CMIP multi-model dataset (Meehl and Bony, 2011; Taylor et al., 2012). For an overview of the climate models and the experiment, the interested reader is referred to Taylor et al. (2012). Ground-Based Observations: Annual observations of temperature provided by the Climatic Research Unit (CRU, Mitchell and Jones, 2005; New et al., 2000), available in a 0.5^{o} spatial resolution, are used as reference data. The CRU gridded temperature data are based on an archive of monthly mean temperatures provided by more than 4000 weather stations distributed across the globe. CRU observations have been validated and used in numerous studies of historical climate variability (e.g., Tanarhte et al., 2012). For consistency, the CMIP5 model simulations and CRU observations all are gridded to a common 2 × 2-degree resolution. This study focuses on global land areas (excluding Antarctica) for which the CRU observations are available.

7.3 Methodology and Results

The concept of estimation using Expert Advice (EA) algorithm (Cesa-Bianchi and Lugosi, 2006) has been successfully applied in the financial sector and game theory (e.g., DeSantis et al., 1988). The goal of the methodology is to weight the predictors (ensemble members) such that at any

given time, the composite climate response is superior to the best model. Let's assume that Θ is a finite set of climate observations, and Ψ is a set of climate simulations over the period for which observations are available. In other words, Ψ is the set of all probability measures on Θ:

$$\Psi := \Pr(\Theta) \tag{7.1}$$

After Vovk and Zhdanov (2009), an error (loss) function $\lambda(\omega, \gamma)$ is defined as the vector product of observations and climate simulations:

$$\lambda(\omega, \gamma) = \sum_{o \in \Theta} (\gamma\{o\} - \delta_\omega\{o\})^2 \tag{7.2}$$

where ω = individual variables in observations space Θ, γ = individual variables in climate simulations space Ψ, $\delta_\omega \in Pr(\Phi)$ = probability measure concentrated at ω, $\gamma\{o\}$ = difference $(\Psi - \Theta)$, $\delta_\omega\{\omega\} = 1$ for $o = \omega$, meaning $\gamma\{o\} = 0$, $\delta_\omega\{o\} = 0$ for $o \neq \omega$, meaning $\gamma\{o\} \neq 0$.

Having a finite number (n time steps) of observations ($\omega_n \in \Theta$), the objective of EA algorithm is to derive the best predictor (γ_n, climate response) given $k = 1, 2, \ldots, K$ climate simulations ($\gamma_n^k \in \Psi$). Throughout this chapter, a common statistical convention is used in which uppercase and lowercase characters denote random variables and their specified variables, respectively. Figure 7.1 displays the flowchart of the proposed algorithm. As shown, first the loss function is computed (Equation 7.2). Then, the initial values of weights at the beginning are set to 1: $\omega_0^1, \ldots \omega_n^K = 1$, where $\omega_0^1, \ldots \omega_n^K$ are weights corresponding to $k = 1, 2, \ldots, K$ climate simulations (ensemble members). In other words, at the beginning of the analysis, the model assumes all climate simulations are as equally representative, and thus a similar weight will be assigned to each ensemble member. Then, the EA algorithm decreases the weights (ω_n^k) of ensemble members ($k = 1, 2, \ldots, K$) exponentially with the increase of error (loss) function ($\lambda(\omega_n, \gamma_n^K)$). The weight function

$$\Phi_n(\omega) = -\ln(\sum_{k=1}^{K} \omega_{n-1}^{k} \times e^{-\lambda(\omega_n, \gamma_n^{\kappa})}) \quad (7.3)$$

where ω_n^k and ω_{n-1}^k refer to weights of ensemble member k at time steps n and $n-1$, respectively. Vovk (2001) mathematically proves that there is a unique s (Figure 7.1) that can be derived through optimizing $\gamma_n(\omega)$. Then, the weighting factors (ω_n^k) at time step n can be obtained for each ensemble member based on the performance of climate simulations with respect to observations up to time step $n-1$:

$$\omega_n^k = \omega_{n-1}^k \times e^{-\lambda(\omega_n, \gamma_n^{\kappa})} \quad (7.4)$$

This indicates that EA algorithm learns from the past and adjusts itself to derive the best ensemble response. In this approach, each ensemble member (e.g., K^{th} member of the ensemble) would have its own cumulative error function (E_n^K). Having K expert advice (climate simulations or ensemble members), the objective of the algorithm is to obtain the best prediction at time step n with the least cumulative error over the past $n-1$ time steps (E_{n-1}) where observations are available.

$$E_n^K = E_{n-1}^K + \lambda(\omega_n, \gamma_n^{\kappa}) \quad (7.5)$$

As shown in Figure 7.1 (right flowchart), the initial values of error (loss) functions are set to zero (i.e., $\mathrm{E}_0^1, \ldots, \mathrm{E}_0^{\mathrm{K}} = 0$). The cumulative loss (error) for each ensemble member at time step n can then be obtained by accumulating the error (loss) function in the past $\mathrm{n}-1$ time steps (see Figure 7.1 (right flowchart)). The algorithm guarantees that for all $\mathrm{n} = 1, 2, \ldots$, the cumulative error function ($\mathrm{E_n}$) will be less or equal to the best model plus a constant - depending on the number of climate simulations (Vovk (2001)):

$$E_n \leq \min_{k=1,\ldots,K} E_n^K + \ln\mathrm{K} \quad (7.6)$$

The proposed methodology is used to derive the climate response of the multi-model CMIP5 temperature simulations. Figure 7.2 displays the global annual mean temperature (1951-2005)

based on (a) EA algorithm; and (b) the multi-model ensemble mean. Both Figure 7.2a and 7.2b are derived using 41 CMIP5 historical temperature simulations. One can see the spatial patterns of both are very similar. However, the EA algorithms leads to smaller mean absolute error (MAE) compared to the ensemble mean (compare Figure 7.2c and 7.2d). As shown, the MAE of the ensemble mean exceeds 2 °C over certain regions, while the MAE of the EA algorithm remains primarily below 1 °C. In most climate change and variability studies, anomalies are used instead of the absolute values of, here, temperature, to account for biases in climate model simulations (e.g., Collins et al. (2011)). Figure 7.2e and Figure 7.2f display the MAE of the EA algorithm and ensemble mean, respectively. In these figures, CMIP5 temperature anomalies are derived based on CRU observations (1951-2005). As shown, even considering the temperature anomalies, the EA algorithm leads to a smaller error than the ensemble mean. Figure 7.3 shows the climate response of the global annual temperature based on the CMIP5 multi-model ensemble for three years: 1960, 1980, and 2000. The first and second rows in the figure display the results using the EA algorithm and the corresponding error, respectively. In Figure 7.3, the third and fourth rows show the same result for the ensemble mean. Similar to the results presented in Figure 7.2, at the three time steps, the EA algorithm leads to a smaller error compared to the ensemble mean.

To further investigate the performance of the proposed climate response algorithm, the time series of the CMIP5 ensemble members, and the ensemble response based on the arithmetic mean and the EA algorithm are provided for the western United States, Europe, eastern China and eastern Australia (see the highlighted regions in Figure 7.4). In Figure 7.5, the solid black line represents the CRU annual mean temperature, whereas the gray lines show the individual CMIP5 ensemble members (41 models). The dashed blue and solid red lines respectively show

the ensemble mean and the EA algorithm. As shown, the EA algorithm is in much better agreement with the observed historical data compared to the ensemble mean, especially in the western United States and eastern China. In the EA algorithm, the ensemble members that are in better agreement with observations and lead to the smaller cumulative loss function receive higher weights in estimating the climate response. Figure 7.6 plots the mean absolute error (temperature ^{o}C) values for the ensemble arithmetic mean and EA algorithm. The figure confirms that the EA algorithm leads to less error with respect to observed historical data. Technically, the proposed algorithm can be used with different data sets. Application of the algorithm to CMIP5 precipitation data is also presented (Figure 7.7 to 7.9). As shown the behavior of the EA algorithm relative to the ensemble median is similar to temperature data (compare Figure 7.8 and 7.9 with Figure 7.5 and 7.6). It should be noted that CMIP5 simulations are not forced with the observed sea surface temperature, and hence their monthly or annual values (especially extremes) are not expected to match with the observations. We do not claim that this method leads to a climate response that can represent the observed monthly or interannual variability. Neither do we claim that the proposed algorithm would remove the underlying biases. The suggested algorithm provides an ensemble response consistent with the average statistics of the observations. The final product should be used and interpreted the way climate model simulations are used in the literature. That is, understanding the long-terms means, statistics, trends, responses to changes in forcing, etc. Finally, the application of this algorithm is not limited to climate model simulations and is not designed for a specific data set or variable. It can potentially be applied to other applications including deriving ensemble response of multi-model streamflow simulations, hurricane tracks, etc.

7.4 Conclusions and Discussion

In this dissertation, a methodology is proposed for deriving the climate response of multi-model climate simulations. The suggested approach is an alternative to the arithmetic mean of ensemble members. The methodology is based on the concept of Expert Advice (EA) algorithm that has been widely used in the financial sectors. The objective of the EA algorithm is to derive weights of predictors (here, individual ensemble members) such that at every time step the ensemble response (here, climate response) is equal or better than the best model. The model was tested using the CMIP5 historical temperature simulations (1951-2005), and the results showed that the EA algorithm led to smaller mean absolute error (MAE) values compared to the ensemble mean. The MAE values were smaller for both the original simulations and the temperature anomalies derived based on CRU observations. The suggested climate response model could also be used with climate projections, assuming that the performance of the models in future will be the same as in the past. That is, the final set of weights obtained based on historical data would be used for deriving ensemble response of projections. The authors acknowledge that modeling observed historical data accurately does not guarantee that the model can produce reliable climate response. Nonetheless, the importance of representing historical observations cannot be ignored. It is worth mentioning that the proposed methodology is more suitable when absolute values of climate model simulations are needed. Using anomalies one can avoid biases and look into relative changes simulated by individual models. However, for practical applications such as climate change impact assessment on the water cycle and ecosystem, one needs the absolute values of climate variables. It is well-known that the multi-model ensembles are not necessarily symmetrical around observations. The proposed algorithm can capture the asymmetries in the ensemble, leading to a response that matches the observations best rather than a response in the

center of the ensemble. In most studies, uncertainties of climate projections are described/quantified by a measure of spread across the ensemble mean (Furrer et al. (2007); Tebaldi and Knutti (2007); Masson and Knutti (2011); Lopez et al. (2006)). For example, in a review study, Knutti et al. (2008) describes the uncertainty of the global temperature projections as one standard deviation of the multi-model response ensemble around the ensemble mean. In other words, most uncertainty models, assume a symmetrical uncertainty space around the climate response. However, there is no reason to believe that uncertainty space of future projections is symmetrical around a given ensemble mean (climate response). While the Gaussian assumption of uncertainty is widely being used mainly due to its simplicity, the distribution of uncertainty space is completely arbitrary. Current efforts are underway by the authors to use a non-Gaussian uncertainty model based on AghaKouchak et al. (2010) around the suggested climate response model (EA algorithm). This would allow deriving the probability of exceedance of a certain condition above/below the climate response given an asymmetrical spread of the uncertainty (ensemble).

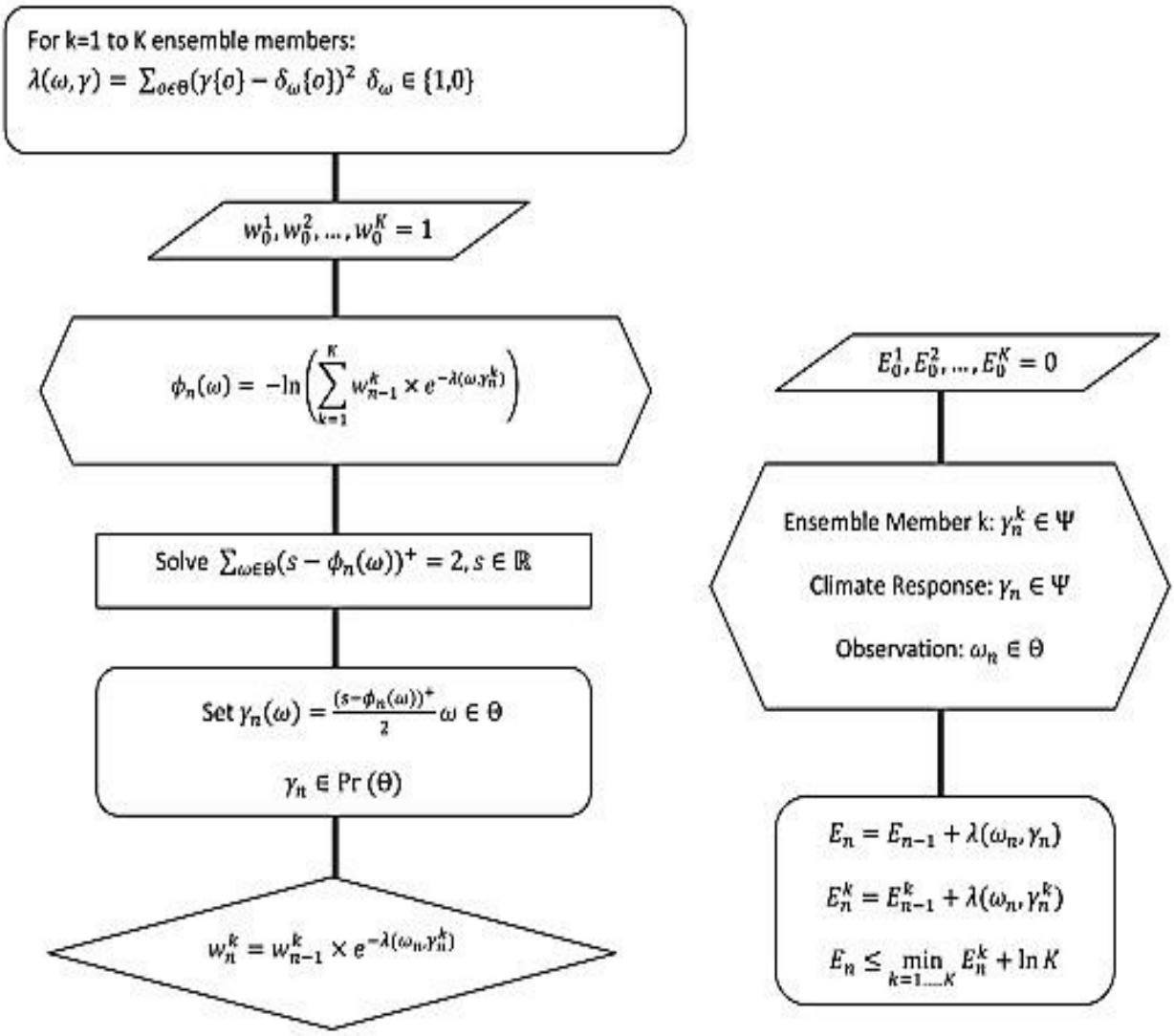

Figure 7.1 The proposed algorithm for estimation of climate response weights (left), and cumulative error (right).

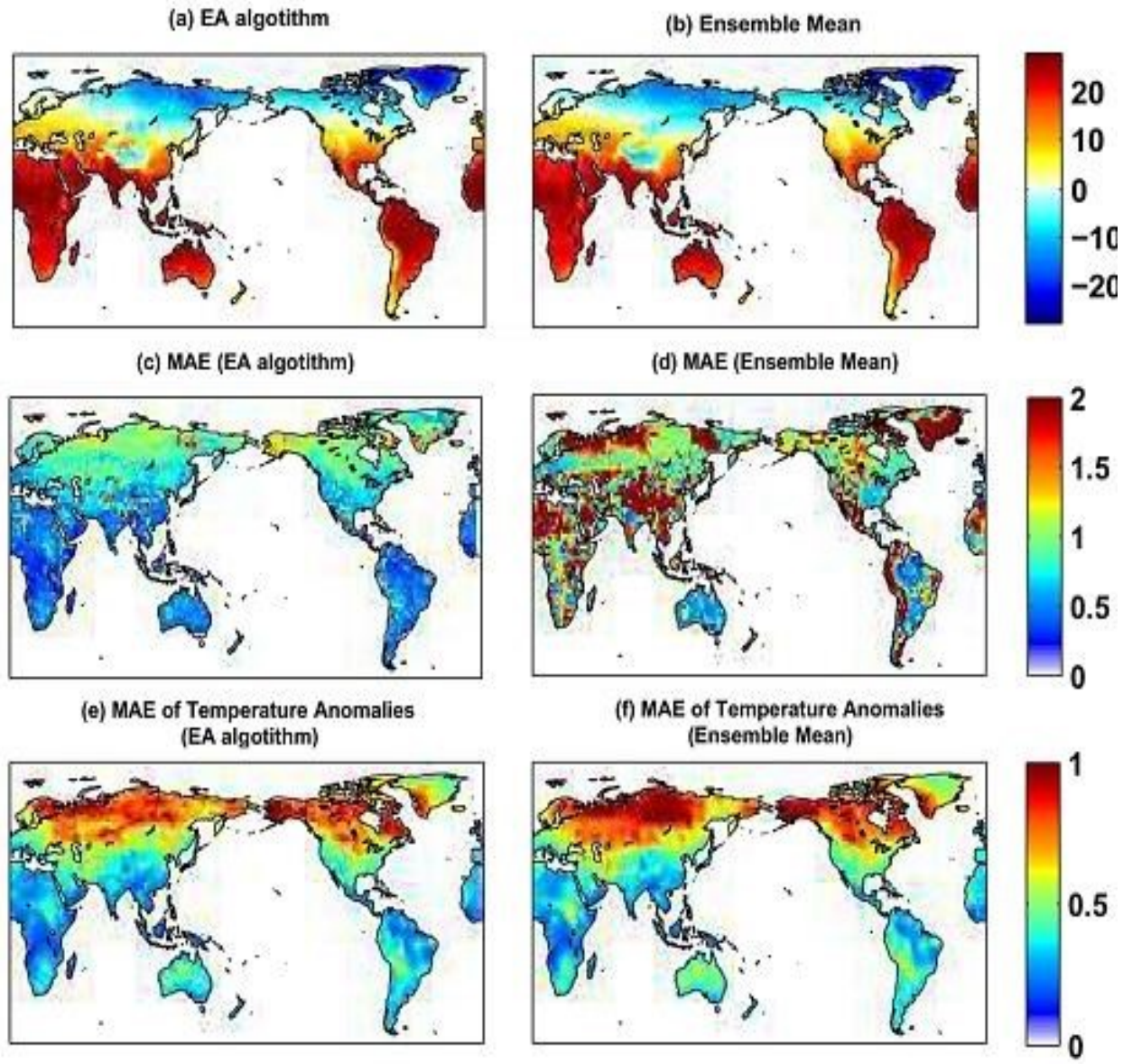

Figure 7.2 The global annual mean temperature (1951-2005) based on the EA algorithm (a) and the multi-model ensemble mean(b), and their corresponding mean absolute error (MAE) maps relative to the CRU observations (MAE for absolute temperature values (c) and (d) and temperature anomalies (e) and (f)).

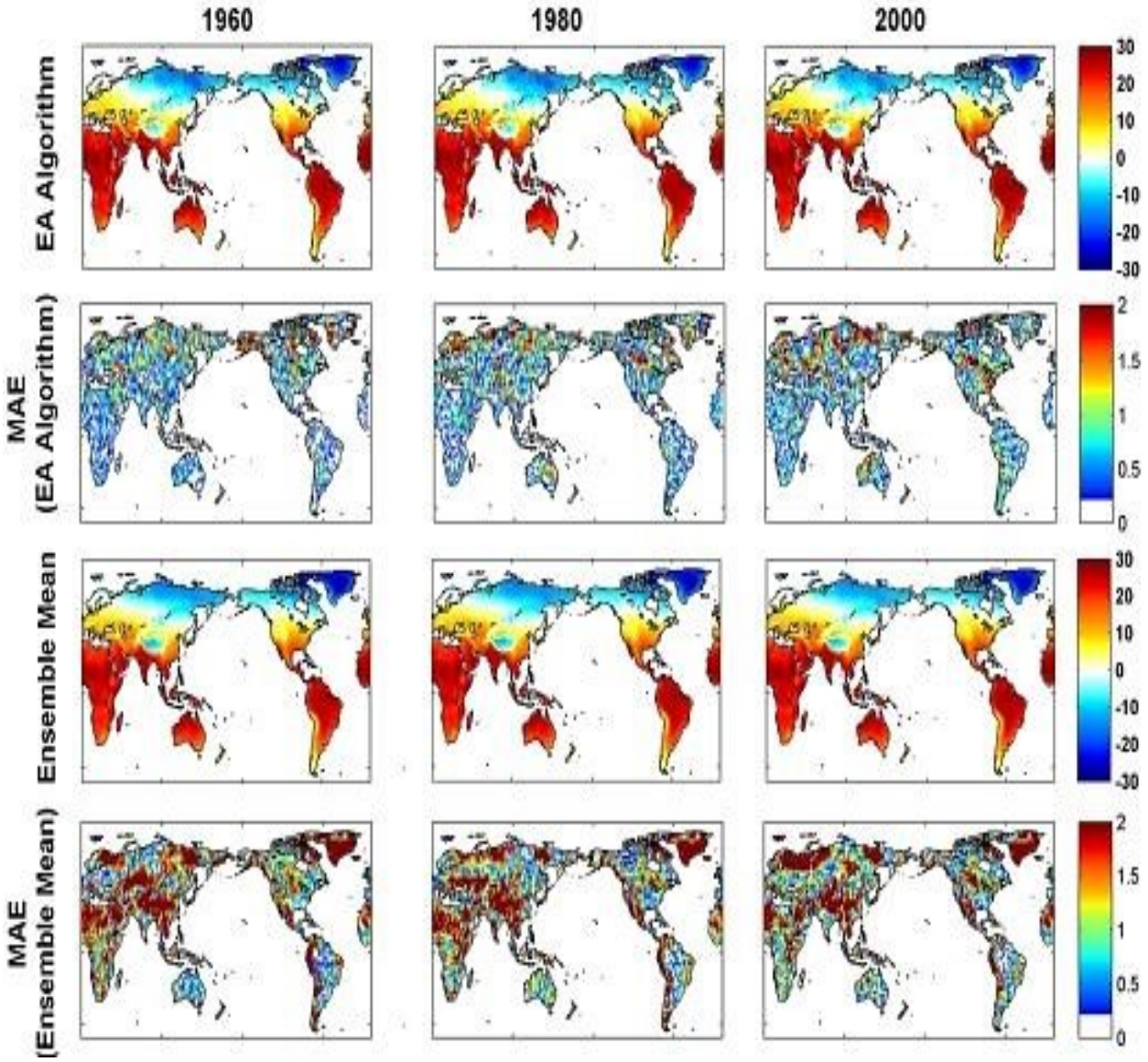

Figure 7.3 The climate response of the global annual temperature based on the CMIP5 multi-model ensemble for three years: 1960, 1980, and 2000 (the 1st and 2nd row are based on the EA algorithm, and the 3rd and 4th rows are based on the ensemble mean).

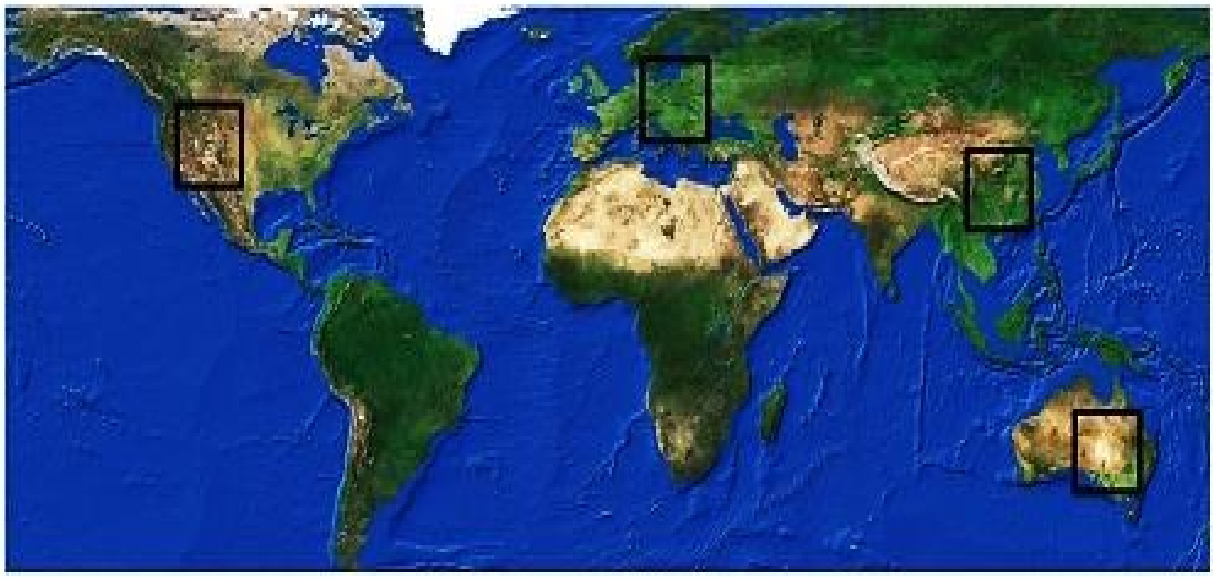

Figure 7.4 Selected regions for time series analysis.

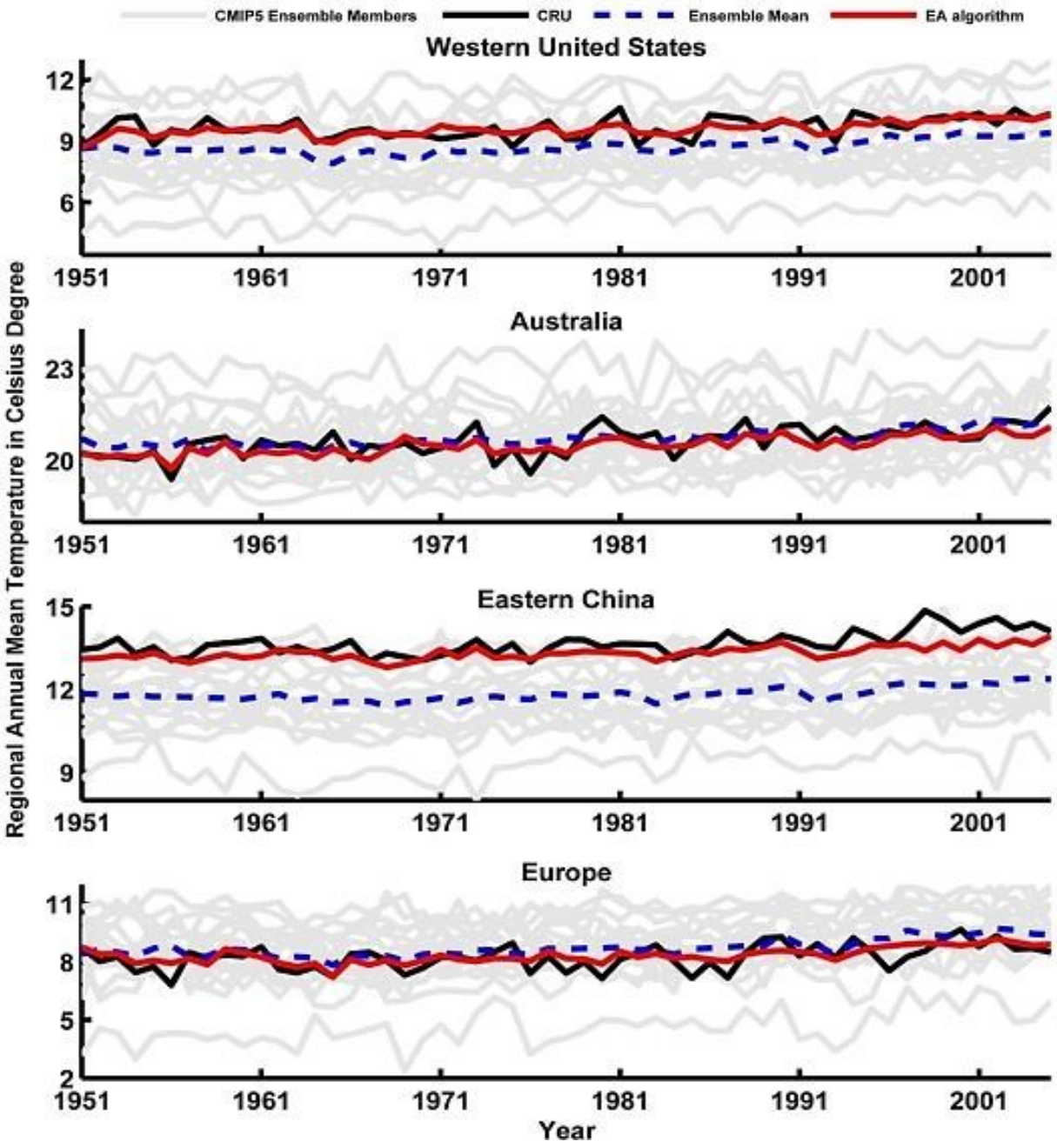

Figure 7.5 Time series of the CMIP5 annual mean temperature, and the ensemble response based on the arithmetic mean and the EA algorithm for the western United States, Europe, eastern China and eastern Australia. The solid black line represents the CRU annual mean temperature, whereas the gray lines show the individual CMIP5 ensemble members (41 models). The dashed blue and solid red lines respectively show the ensemble mean and the EA algorithm.

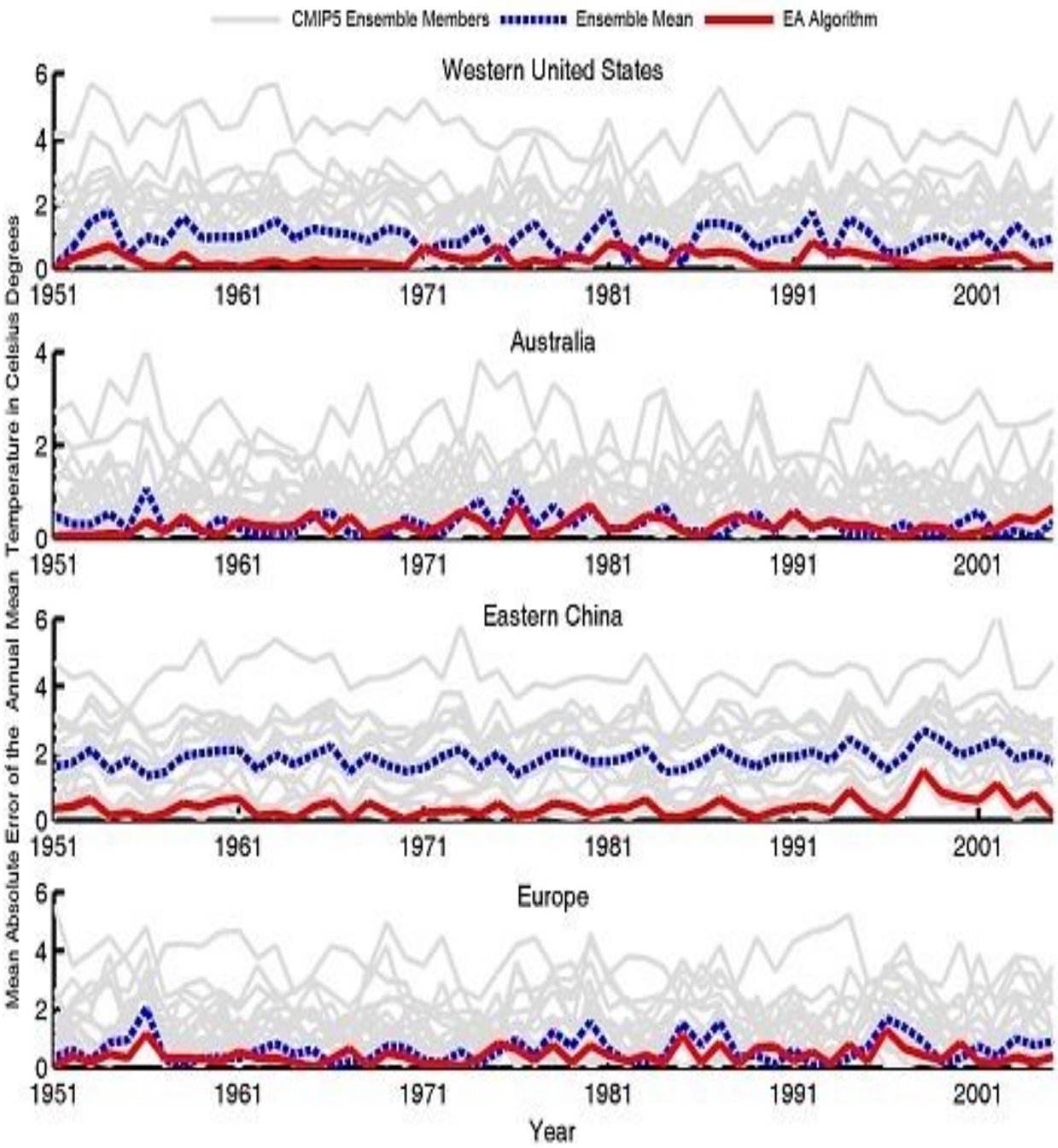

Figure 7.6 Mean absolute error (temperature ^{o}C) values for the ensemble arithmetic mean and EA algorithm shown in Figure 7.5.

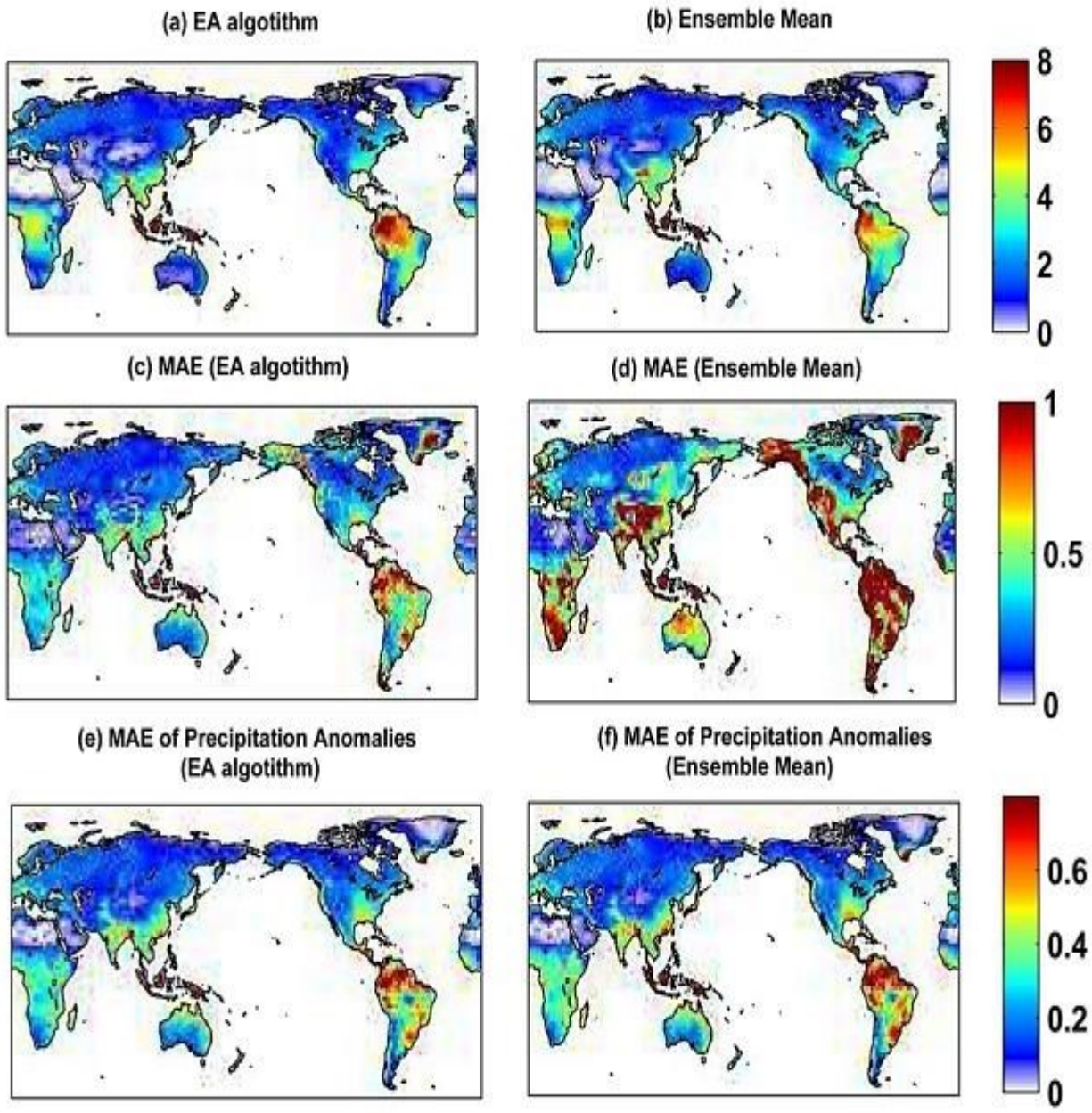

Figure 7.7 The global annual mean precipitation (1951-2005) based on the EA algorithm (a) and the multi-model ensemble mean(b) in mm/day, and their corresponding mean absolute error (MAE) maps relative to the CRU observations (MAE for absolute temperature values (c) and (d) and temperature anomalies (e) and (f)).

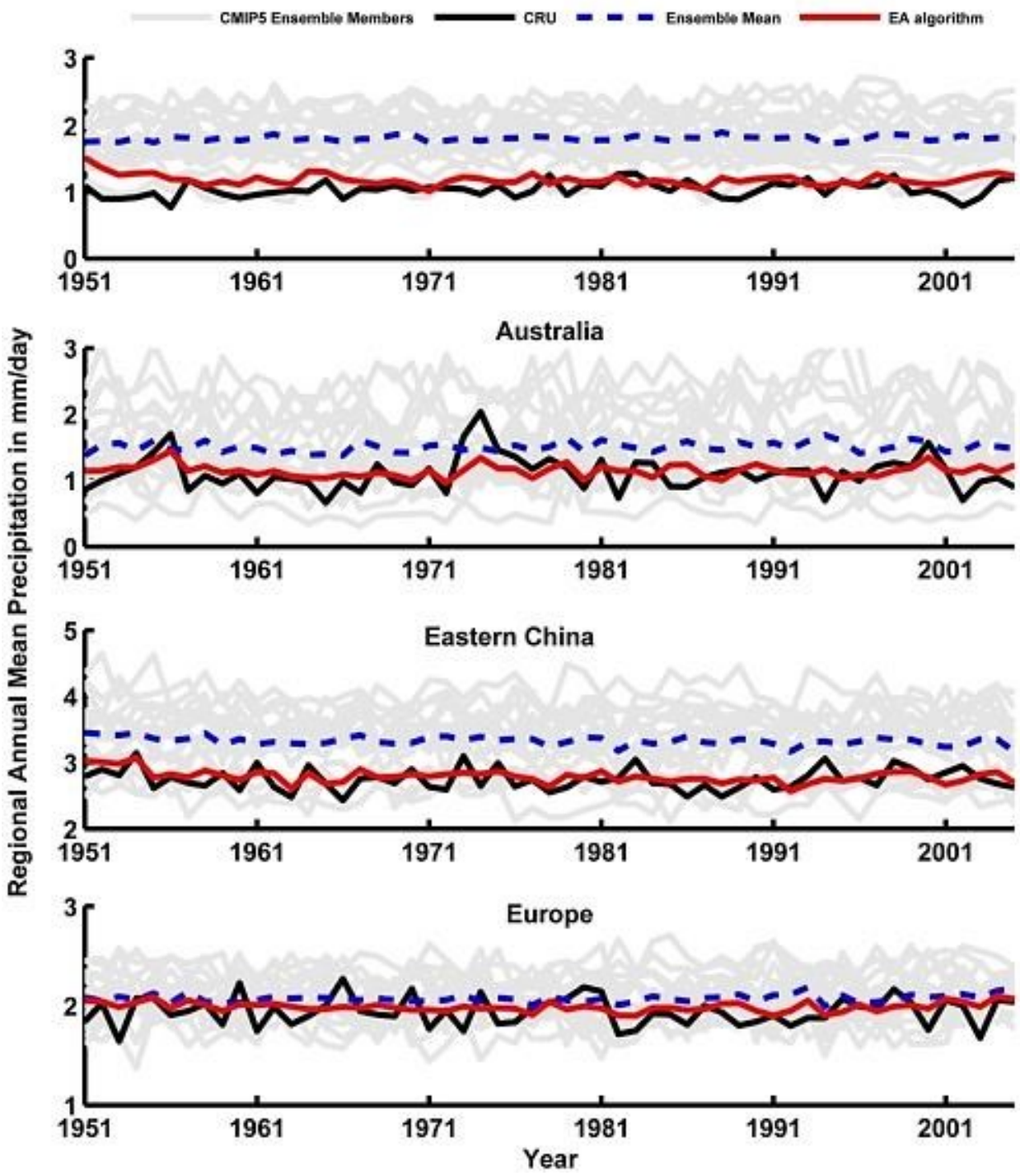

Figure 7.8 Time series of the CMIP5 annual mean precipitation, and the ensemble response based on the arithmetic mean and the EA algorithm for the western United States, Europe, eastern China and eastern Australia. The solid black line represents the CRU annual mean precipitation, whereas the gray lines show the individual CMIP5 ensemble members (41 models). The dashed blue and solid red lines respectively show the ensemble mean and the EA algorithm (similar to Figure 7.5 in the main dissertation, but for precipitation).

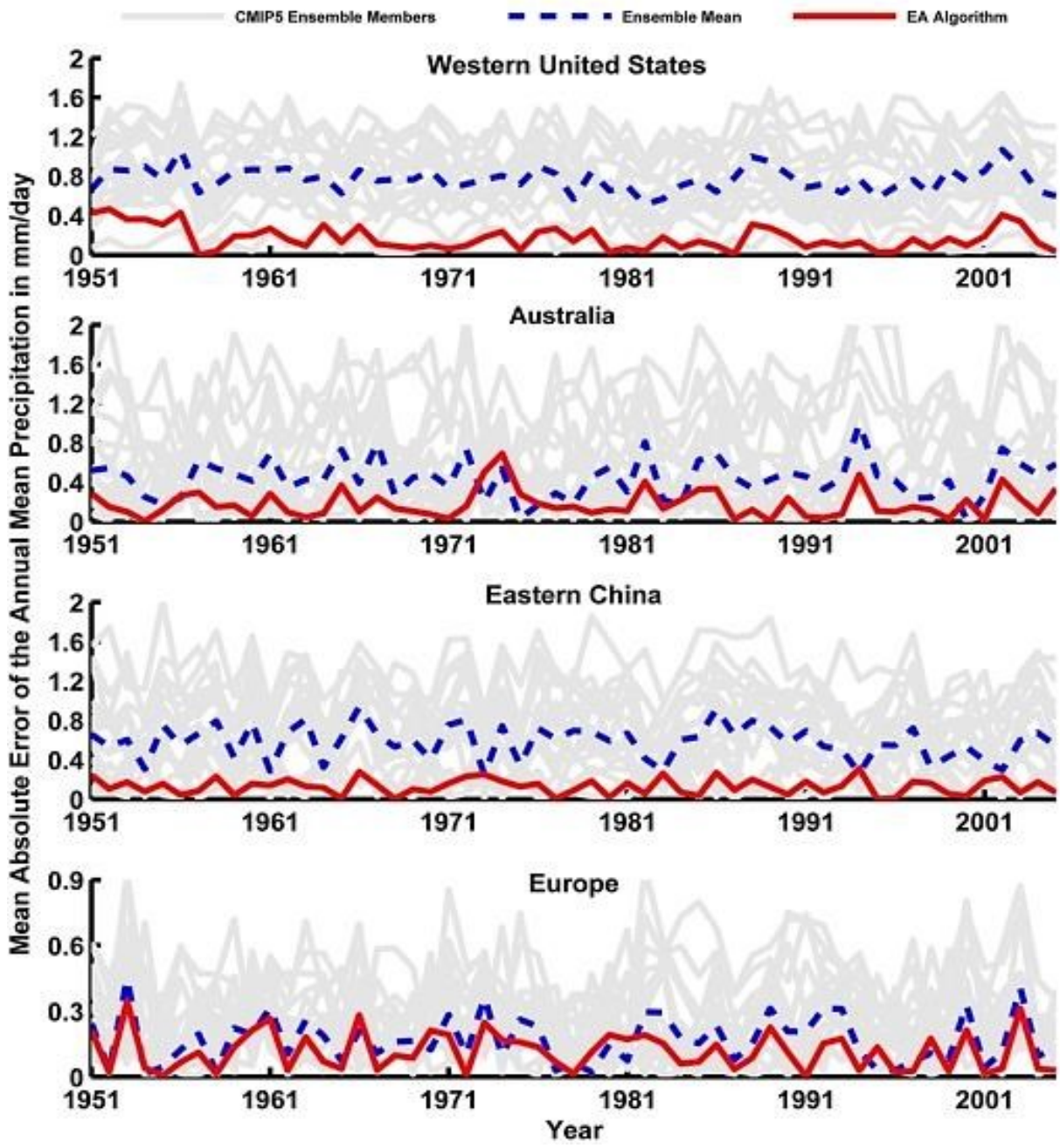

Figure 7.9 Mean absolute error (precipitation in mm/day) values for the ensemble arithmetic mean and EA algorithm shown in Figure 7.8 (similar to Figure 7.6 in the main dissertation, but for precipitation).

CHAPTER 8: Summary and Main Conclusions

- In Chapter 2, substantial evidence shows that the climate is non-stationary, possibly due to anthropogenic changes in the environment. The assumption of stationarity in extreme value analysis is therefore questionable and statistical models that explicitly allow for non-stationarity are much needed. Specifically, statistical models that can provide estimates of return levels under non-stationary conditions are essential for design and risk assessment purposes. In this study, a practical package named the Non-stationary Extreme Value Analysis (NEVA) package is introduced for assessing extremes in a changing climate. Both stationary and non-stationary components of the package are evaluated using Climatic Research Unit (CRU) observations. The results indicate that NEVA simulates GEV-based return levels consistent with empirical observations. While the focus of this chapter is on climate extreme value analysis, the methodology can potentially be used in different areas (hydrology, ecology, and economics) and with different data sets.

- In Chapter 3, NEVA is applied to evaluate to what extent the CMIP5 climate model simulations can represent observed warm monthly temperature extremes under a changing climate. The biases of simulated annual temperature maxima are quantified for the selected CMIP5 models. Furthermore, the 2-, 10-, 25-, 50, and 100-year return levels of the annual temperature maxima from CMIP5 simulations are compared with those derived from CRU observations. Overall, the results of this study indicate that the models capture the spatial patterns of temperature extremes well, but that individual models are biased relative to the CRU observations. Considering the global averages, most models overestimate the simulated return levels of the annual temperature maxima, while fewer models (e.g.,

FGOALS-g2, INMCM4_esm, NorESM1-ME) underestimate the temperature extremes. Among the models, FGOALS-s2, CanESM2 and MIROC5 exhibit the highest global averages of the mean error of the annual temperature maxima. Most models either systematically overestimate or underestimate the extreme return levels, except the BCC model experiments in which the shorter return levels (2- and 10-year) are underestimated and the longer ones are overestimated. The differences between CMIP5 simulations and CRU observations indicate that while local differences can be large, most differences (between 25^{th} and 75^{th} percentiles) fall within ±2 Celsius degrees.

- In Chapter 4, we show that ignoring the stationary assumption could lead to substantial underestimation of extremes, especially at sub-daily durations (e.g., 1-hr, 2-hr). We also outline a novel framework using NEVA to create the next generation of IDF curves to be incorporated into infrastructure design. However, infrastructure design and construction requires substantial investment over a long period of time and effective integration of this methodology as well as development of adaptive design frameworks will require collaborative and interdisciplinary research with engineers, policy makers, economists, climate scientists and decision makers.

- In Chapter 5 and Chapter 6, a methodology is presented for analyzing conditional extremes (i.e., one climatic event conditioned on another extreme event). The presented model can be potentially applied in a wide variety of science fields including finance, earth science, environmental science, and biology. Particularly, this model can be used for assessing spatial climatic extremes.

- In Chapter 7, a methodology is proposed for deriving the climate response of multi-model climate simulations. The suggested approach is an alternative to the arithmetic mean of ensemble members. The methodology is based on the concept of the Expert Advice (EA) algorithm that has been widely used in the financial sector. The objective of the EA algorithm is to derive weights of predictors (here, individual ensemble members) such that at every time step the ensemble response (here, climate response) is equal or better than the best model. The model was tested using the CMIP5 historical temperature simulations (1951-2005), and the results showed that the EA algorithm led to smaller mean absolute error (MAE) values compared to the ensemble mean. The MAE values were smaller for both the original simulations and the temperature anomalies derived based on CRU observations. The suggested climate response model could also be used with climate projections, assuming that the performance of the models in the future will be the same as in the past. That is, the final set of weights obtained based on historical data would be used for deriving the ensemble response of projections.

REFERENCES

AghaKouchak, A., D. Easterling, K. Hsu, S. Schubert, and S. Sorooshian (2013)“Extremes in a Changing Climate”, Springer, Springer Netherlands.

AghaKouchak, A., N. Nasrollahi (2010)"Semi-parametric and parametric inference of extreme value models for rainfall data." Water Resources management 24: 1229-1249.

AghaKouchak, A., B´ardossy, A., Habib, E., 2010. Conditional simulation of remotely sensed rainfall data using a non-gaussian v-transformed copula. Advances in Water Resources 33 (6), 624–634.

Alexander, L.V., et al. (2006)"Global observed changes in daily climate extremes of temperature and precipitation." Journal of Geophysical Research 111.D5.

Allen, M. R., & Ingram, W. J. (2002). Constraints on future changes in climate and the hydrologic cycle. Nature, 419(6903), 224-232.

Alley, R. B., Clark, P. U., Huybrechts, P., Joughin, I. (2005). Ice-sheet and sea-level changes. Science 310 (5747), 456–460.

Assessment, A. C. I. (2004). Impacts of a Warming Arctic-Arctic Climate Impact Assessment. Impacts of a Warming Arctic-Arctic Climate Impact Assessment, by Arctic Climate Impact Assessment, pp. 144. ISBN 0521617782. Cambridge, UK: Cambridge University Press, December 2004., 1.

Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. Nature,438(7066), 303-309.

Barnett, T. P., Hasselmann, K., Chelliah, M., Delworth, T., Hegerl, G., Jones, P., ... & Tett, S. (1999). Detection and attribution of recent climate change: a status report. Bulletin of the American Meteorological Society, 80(12), 2631-2659.

Beguería, S., et al. (2011)"Assessing trends in extreme precipitation events intensity and magnitude using nonstationary peaks-over-threshold analysis: a case study in northeast Spain from 1930 to 2006.” International Journal of Climatology 31: 2102-2114.

Beniston M 2009 Trends in joint quantiles of temperature and precipitation in Europe since 1901 and projected for 2100 Geophys. Res. Lett. 36 L07707.

Beniston, M., et al. (2007)"Future extreme events in European climate: an exploration of regional climate model projections." Climatic Change 81:71-95.

Berg, P., Moseley, C., & Haerter, J. O. (2013). “Strong increase in convective precipitation in response to higher temperatures”. Nature Geoscience, 6, 181–185.

Bonnin, G., Martin, D., Lin, B., Parzybok, T., Yekta, M., Riley, D. (2004) "Precipitation-frequency atlas of the United States." NOAA atlas 14 (2).

Brekke, L., Barsugli, J., 2013. Uncertainties in projections of future changes in extremes. In: Extremes in a Changing Climate. Springer, doi: 10.1007/978-94-007-4479-0 11.

Brown, R. D. (2000). Northern Hemisphere snow cover variability and change, 1915-97. Journal of climate, 13(13).

Chahine, M. T. (1992). The hydrological cycle and its influence on climate.Nature, 359(6394), 373-380.Chen, S. X., & Cui, H. (2003). An extended empirical likelihood for generalized linear models. Statistica Sinica, 13(1), 69-82.

Cantelaube, P., Terres, J.-M., 2005. Seasonal weather forecasts for crop yield modelling in Europe. Tellus A 57, 476487.

Cannon, A.J. (2011) "GEVcdn: An R package for nonstationary extreme value analysis by generalized extreme value conditional density estimation network.”Computers & Geosciences 37:1532-1533.

Cesa-Bianchi, N., Lugosi, G., 2006. Prediction, Learning, and Games. Cambridge University Press, Cambridge, England.

Chen, S. X., (1994)“Comparing empirical likelihood and bootstrap hypothesis tests,” Journal of Multivariate Analysis 51(2), 277–293.

Chen, S. X. and Cui, H., (2003)“An extended empirical likelihood for generalized linear models,” Statistica Sinica 13(1), 69–82.

Coles, S. G. and Tawn, J. A., (1991)“Modelling extreme multivariate events.” J. R. Statist. Soc. B 53, 377–392.

Coles, S., (2001)“An introduction to statistical modeling of extreme values” Springer, London.

Coles, S., E.A. Powell (1996)"Bayesian methods in extreme value modelling: a review and new developments." International Statistical Review: 119-136.

Collins, M., Booth, B. B. B., Bhaskaran, B., Harris, G. R., Murphy, J. M., Sexton, D. M. H., Webb, M. J., MAY 2011. Climate model errors, feedbacks and forcings: a comparison of perturbed physics and multi-model ensembles. Climate Dynamics 36 (9-10), 1737–1766.

Clarke, R. T., 2002. Estimating trends in data from the Weibull and a generalized extreme value distribution. Water Resources Research 38 (6), 1089.

Cooley, D. (2013)“Return periods and return levels under climate change”, Extremes in a Changing Climate. Springer Netherlands.

Cooley, D. (2009) "Extreme value analysis and the study of climate change."Climatic Change 97: 77-83.

Cooley, D., Nychka, D., Naveau, P. (2007)“Bayesian spatial modeling of extreme precipitation return levels”, Journal of the American Statistical Association, 102, 824-840.

Cunderlik, J. M., & Ouarda, T. B. (2006). “Regional flood-duration–frequency modeling in the changing environment”. Journal of Hydrology, 318(1), 276-291.

Damberg L., AghaKouchak A., 2013, Global Trends and Patterns of Droughts from Space, Theoretical and Applied Climatology, doi: 10.1007/s00704-013-1019-5.

Das, Tapash, et al. (2011)"Potential increase in floods in California’s Sierra Nevada under future climate projections." Climatic Change 109.1: 71-94.

Davison, A.C., R.L. Smith. (1990)"Models for exceedances over high thresholds." Journal of the Royal Statistical Society. 393-442.

de Elia, R. and H. Cote, 2010: Climate and climate change sensitivity to model configuration in the Canadian RCM over North America. Meteorologische Zeitschrift, 19(4), 325–339.

Deardorff, J. W. (1978). Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation. Journal of Geophysical Research: Oceans (1978–2012), 83(C4), 1889-1903.

De Michele, C., & Salvadori, G. (2003). A Generalized Pareto intensity-duration model of storm rainfall exploiting 2-Copulas. Journal of Geophysical Research: Atmospheres (1984–2012), 108(D2).

Del Genfo, A. D., Lacis, A. A., & Ruedy, R. A. (1991). Simulations of the effect of a warmer climate on atmospheric humidity. Nature, 351(6325), 382-385.

Deque, M., Rowell, D. P., Luethi, D., Giorgi, F., Christensen, J. H., Rockel, B., Jacob, D., Kjellstrom, E., de Castro, M., van den Hurk, B., 2007. An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections. Climatic Change 81 (1), 53–70.

DeSantis, A., Markowsky, G., Wegman, M., 1988. Learning probabilistic prediction functions. In: 29th Annual Symposium on Foundations of Computer Science (IEEE Cat. No.88CH2652-6). IEEE, pp. 110–19.

Diffenbaugh, N. S., & Giorgi, F. (2012)"Climate change hotspots in the CMIP5 global climate model ensemble". Climatic change, 114(3-4), 813-822.

Ding, Y. D. J. G., Griggs, D. J., Noguer, M., van der LINDEN, P. J., Dai, X., Maskell, K., & Johnson, C. A. (2001). Climate change 2001: the scientific basis(Vol. 881). Cambridge: Cambridge university press.

Doblas-Reyes, F., Pavan, V., Stephenson, D., 2003. The skill of multi-model seasonal forecasts of the wintertime North Atlantic Oscillation. Climate Dynamics 21 (5-6), 501–514.

Easterling, D.R., et al. (2000) "Climate extremes: observations, modeling, and impacts." Science 289: 2068-2074.

El Adlouni, S., et al.(2007)"Generalized maximum likelihood estimators for the nonstationary generalized extreme value model". Water Resources Research, 43.

Emori, S., and Brown, S. J. (2005). "Dynamic and thermodynamic changes in mean and extreme precipitation under changed climate". Geophysical Research Letters, 32(17).

Endreny, Theodore A., and Nana Imbeah (2009)"Generating robust rainfall intensity–duration–frequency estimates with short-record satellite data." Journal of hydrology 371.1: 182-191.

Feddema, J., Oleson, K., Bonan, G., Mearns, L., Washington, W., Meehl, G., Nychka, D., OCT 2005. A comparison of a GCM response to historical anthropogenic land cover change and model sensitivity to uncertainty in present-day land cover representations. Climate Dynamics 25 (6), 581–609.

Ficklin, D. L., Luo, Y., Luedeling, E., Zhang, M., 2009. Climate change sensitivity assessment of a highly agricultural watershed using swat. Journal of Hydrology 374 (1), 16–29.

Field, C.B., et al., (2012)"Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change". Cambridge University Press.

Fisher, R., Tippett, L. (1928)"Limiting forms of the frequency distribution of the largest or smallest member of a sample". Proceedings of the Cambridge Philosophical Society 24, 180–190.

Foley, A. M., OCT 2010. Uncertainty in regional climate modelling: A review. Progress in Physical Geography 34 (5), 647–670.

Frei, C., et al. (2006)"Future change of precipitation extremes in Europe: Intercomparison of scenarios from regional climate models." Journal of Geophysical Research, 111.D6.

Furrer, R., Knutti, R., Sain, S. R., Nychka, D. W., Meehl, G. A., 2007. Spatial patterns of probabilistic temperature change projections from a multivariate Bayesian analysis. Geophysical Research Letters 34 (6), L06711.

Gelman, A., K. Shirley. (2011)"Inference from simulations and monitoring convergence" Handbook of Markov Chain Monte Carlo: 163-174.

Gelman, A., et al. (2003) "Bayesian data analysis" CRC Press.

Gilks, W. R., Richardson, S. and Spiegelhalter, D. J., (1996)"Introducing Markov chain Monte Carlo," Markov chain Monte Carlo in practice , 1–19.

Gilleland, E., Brown, B. G. and Ammann, C. M., (2013)"Spatial extreme value analysis to project extremes of large-scale indicators for severe weather," Environmetrics 24(6), 418–432.

Gilleland, E., Katz, R. W. (2006)"Analyzing seasonal to interannual extreme weather and climate variability with the extremes toolkit". In: 18th Conference on Climate Variability and Change, 86th American Meteorological Society (AMS) Annual Meeting. Vol. 29.

Gilleland, E., Katz, R.W. (2011)"New software to analyze how extremes change over time" Eos, 92(2), 13—14.

Gnedenko, B., 1943. Sur la distribution limite du terme maximum d'une s´erie al´eatoire. Ann. Math. 44, 423–453.

Goodman, D. (2010). 1 2 Taking the Prior Seriously: Bayesian Analysis without Subjective Probability. The Nature of Scientific Evidence: Statistical, Philosophical, and Empirical Considerations, 379.

Grimaldi, S., & Serinaldi, F. (2006). Asymmetric copula in multivariate flood frequency analysis. Advances in Water Resources, 29(8), 1155-1167.

Gräler, B., et al., (2013). Multivariate return periods in hydrology: a critical and practical review focusing on synthetic design hydrograph estimation. Hydrology & Earth System Sciences, 17(4).

Groisman, Pavel Ya, Richard W. Knight, and Thomas R. Karl (2012)"Changes in Intense Precipitation over the Central United States." Journal of Hydrometeorology 13.1 .

Groisman, P. Y., Knight, R. W., Karl, T. R., Easterling, D. R., Sun, B., & Lawrimore, J. H. (2004). Contemporary changes of the hydrological cycle over the contiguous United States: Trends derived from in situ observations. Journal of hydrometeorology, 5(1).

Groisman, Pavel Ya, et al. (2005) "Trends in intense precipitation in the climate record." Journal of climate 18.9.

Gumbel, E.J. (1958)"Statistics of Extremes", Mineola, NY: Dover.—, (1958)"Statistics of extremes". Columbia University Press, New York.

Gumbel, E., (1942)"On the frequency distribution of extreme values in meteorological data". B. Am. Meteorol. Soc., 23.

Gyasi-Agyei, Y. (2013). Evaluation of the effects of temperature changes on fine timescale rainfall. Water Resources Research, 49(7), 4379-4398.

Hall, P. and La Scala, B., (1990)"Methodology and algorithms of empirical likelihood," International Statistical Review/Revue Internationale de Statistique, 109–127.

Hanel, M., Buishand, T. A., & Ferro, C. A. (2009). "A nonstationary index flood model for precipitation extremes in transient regional climate model simulations". Journal of Geophysical Research: Atmospheres (1984–2012), 114(D15).

Hansen, J., Ruedy, R., Sato, M., Lo, K., 2010. Global surface temperature change. Reviews of Geophysics 48 (4), RG4004.

Hao, Z., AghaKouchak, A., Phillips, T. (2013)"Changes in Concurrent Monthly Precipitation and Temperature Extremes, Environmental Research Letters, 8(4), 034014.

Hassanzadeh, Elmira, Alireza Nazemi, and Amin Elshorbagy. (2013) "Quantile-Based Downscaling of Precipitation using Genetic Programming: Application to IDF Curves in the City of Saskatoon." Journal of Hydrologic Engineering, doi: 10.1061/(ASCE)HE.1943-5584.0000854.

Hasselmann, K. (1998). Conventional and Bayesian approach to climate-change detection and attribution. Quarterly Journal of the Royal Meteorological Society,124(552), 2541-2565.

Heffernan, J. E. and Resnick, S. I., "Limit laws for random vectors with an extreme component," The Annals of Applied Probability 17(2), 537–571 (2007).

Heffernan, J. E. and Tawn, J. A., "A conditional approach for multivariate extreme values (with discussion)," Journal of the Royal Statistical Society: Series B (Statistical Methodology) 66(3), 497–546, ISSN 1467-9868 (2004).

Hosking, J., and J. Wallis, 1997: Regional frequency analysis: an approach based on L-moments. Cambridge University Press, Cambridge.

Hegerl, G. C., voN SToRcH, H., Hasselmann, K., Santer, B. D., Cubasch, U., & Jones, P. D. (1996). Detecting greenhouse-gas-induced climate change with an optimal fingerprint method. Journal of Climate, 9(10), 2281-2306.

Held, I. M., & Soden, B. J. (2000). Water vapor feedback and global warming 1.Annual review of energy and the environment, 25(1), 441-475.

Huntington, T. G. (2006). Evidence for intensification of the global water cycle: review and synthesis. Journal of Hydrology, 319(1), 83-95.

IPCC (2007)"Climate Change 2007: The Physical Science Basis", Working Group 1, IPCC Fourth Assessment Report, Cambridge University Press.

Jakob, D. (2013). Nonstationarity in Extremes and Engineering Design. Extremes in a Changing Climate, Springer Netherlands.

Jing, B.-Y., "Two-sample empirical likelihood method," Statistics & probability letters 24(4), 315–319 (1995).

John, V. O., Soden, B. J., SEP 25 2007. Temperature and humidity biases in global climate models and their impact on climate feedbacks. Geophysical Research Letters 34 (18), L18704.

Jonathan, P., Ewans, K. and Flynn, J., "Joint modelling of vertical profiles of large ocean currents," Ocean Engineering 42, 195–204 (2012).

Jonathan, P., Ewans, K. and Randell, D., "Joint modelling of extreme ocean environments incorporating covariate effects," Coastal Engineering 79, 22–31 (2013).

Jongman, B., et al., (2014). Increasing stress on disaster risk finance due to large floods. Nature Climate Change, doi: 10.1038/NCLIMATE2124.

Joshi, M. M., Gregory, J. M., Webb, M. J., Sexton, D. M., & Johns, T. C. (2008). Mechanisms for the land/sea warming contrast exhibited by simulations of climate change. Climate dynamics, 30(5), 455-465.

Karl, Thomas R., Jerry M. Melillo, and Thomas C. Peterson, eds. (2009) "Global climate change impacts in the United States". Cambridge University Press.

Karl, T., Knight, R., 1998. Secular trends of precipitation amount, frequency, and intensity in the usa. Bulletin of the American Meteorological Society 79, 231–241.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. Journal of the american statistical association, 90(430), 773-795.

Katz, R.W. (2013)"Statistical methods for nonstationary extremes". Extremes in a Changing Climate, Springer Netherlands.

Katz, R. (2010), Statistics of extremes in climate change, Climatic Change, 100(1), 71-76.

Katz, R., et al., (2002)"Statistics of extremes in hydrology". Advances in Water Resources, 25, 12871304.

Keef, C., Papastathopoulos, I. and Tawn, J. A., "Estimation of the conditional distribution of a multivariate variable given that one of its components is large: Additional constraints for the heffernan and tawn model," Journal of Multivariate Analysis 115, 396–404 (2013a).

Keef, C., Svensson, C. and Tawn, J. A., "Spatial dependence in extreme river flows and precipitation for Great Britain," Journal of Hydrology 378(3–4), 240 – 252, ISSN 0022-1694 (2009a).

Keef, C., Tawn, J. and Svensson, C., "Spatial risk assessment for extreme river flows," Journal of the Royal Statistical Society: Series C (Applied Statistics) 58(5), 601–618, ISSN 1467-9876 (2009b).

Keef, C., Tawn, J. A. and Lamb, R., "Estimating the probability of widespread flood events," Environmetrics 24(1), 13–21 (2013b).

Kendall, M.G. (1976)"Rank Correlation Methods". 4thEd. Griffin.

Khaliq, M. N., Ouarda, T. B. M. J., Ondo, J.-C., Gachon, P., & Bobée, B. (2006). Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review. Journal of Hydrology, 329(3-4), 534–552.

Kharin, V., Zwiers, F., Zhang, X., Wehner, M., 2013. Changes in temperature and precipitation extremes in the CMIP5 ensemble. Climatic Change, 1–13.

Kharin, V., Zwiers, F., Zhang, X., Hegerl, G. (2007)"Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations." Journal of Climate 20: 1419-1444.

Kharin, V.V., F.W. Zwiers. (2005)"Estimating extremes in transient climate change simulations." Journal of Climate 18:1156-1173.

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.

Kilbourne, E., 1997. Heat waves and hot environments. The Public Health Consequences of Disasters, 245–269.

Kjær, K., et al., 2012. Aerial photographs reveal late–20th-century dynamic ice loss in northwestern Greenland. Science 337 (6094), 569–573.

Kjellstrom, E. and K. Ruosteenoja, 2007: Present-day and future precipitation in the Baltic Sea region as simulated in a suite of regional climate models. Climatic Change, 81(1), 281–291.

Klein T., et al. (2009) Guidelines on Analysis of extremes in a changing climate in support of informed decisions for adaptation. WMO-TD 1500, 56 pp.Leadbetter, M., et al. (1983)“Extremes and related properties of random sequences and processes”.

Krabill, W., et al., 2004. Greenland ice sheet: increased coastal thinning. Geophysical Research Letters 31 (24), L24402.

Krishnamurti, T., et al., 2000. Multimodel ensemble forecasts for weather and seasonal climate. Journal of Climate 13 (23), 4196–4216.

Kundzewicz, Z. W., & Robson, A. J. (2004). Change detection in hydrological records—a review of the methodology/revue méthodologique de la détection de changements dans les chroniques hydrologiques. Hydrological Sciences Journal, 49(1), 7-19.

Kunkel, K., 2013. Uncertainties in observed changes in climate extremes. In: Extremes in a Changing Climate. Springer, doi: 10.1007/978-94-007-4479-0 10.

Kunkel, Kenneth E., et al. (2013)"Probable maximum precipitation and climate change." Geophysical Research Letters 40.7: 1402-1408.

Kunkel, K. E., et al. (2013) Monitoring and understanding trends in extreme storms: State of knowledge. Bulletin of the American Meteorological Society, 94(4), 499-514.

Knutti, R., et al., 2008. A review of uncertainties in global temperature projections over the twenty-first century. Journal of Climate 21 (11), 2651–2663.

Knutti, R., Furrer, R., Tebaldi, C., Cermak, J., Meehl, G. A., 2010. Challenges in combining projections from multiple climate models. Journal of Climate 23 (10), 2739–2758.

Lamb, R., Keef, C., Tawn, J., Laeger, S., Meadowcroft, I., Surendran, S., Dunning, P. and Batstone, C., "A new method to assess the risk of local and widespread flooding on rivers and coasts," Journal of Flood Risk Management 3(4), 323–336, ISSN 1753-318X (2010).

Ledford, A. W. and Tawn, J. A., "Statistics for near independence in multivariate extreme values," Biometrika 83(1), 169–187 (1996).

Leadbetter, M., Lindgren, G., Rootz´en, H., 1983. Extremes and related properties of random sequences and processes.

Lee, T. C., Zwiers, F. W., Zhang, X., Hegerl, G. C., & Tsao, M. (2005). A Bayesian approach to climate change detection and attribution. J. Clim, 18, 2429-2440.

Legates, D. R., & Willmott, C. J. (1990). Mean seasonal and spatial variability in gauge-corrected, global precipitation. International Journal of Climatology,10(2), 111-127.

Leonard, M., Westra, S., Phatak, A., Lambert, M., van den Hurk, B., McInnes, K., ... & Stafford- Smith, M. (2014). A compound event framework for understanding extreme impacts. Wiley Interdisciplinary Reviews: Climate Change, 5(1), 113-128.

Li, G., "Nonparametric likelihood ratio estimation of probabilities for truncated data," Journal of the American Statistical Association 90(431), 997–1003 (1995).

Li, Y., Cai, W., & Campbell, E. P. (2005). Statistical modeling of extreme rainfall in southwest Western Australia. Journal of Climate, 18(6).

Liepert, B. G., Previdi, M., JAN-MAR 2012. Inter-model variability and biases of the global water cycle in CMIP3 coupled climate models. Environmental Research Letters 7 (1), 014006.

Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. R news, 2(3), 18-22.

Loaiciga, H. A., Valdes, J. B., Vogel, R., Garvey, J., & Schwarz, H. (1996). Global warming and the hydrologic cycle. Journal of Hydrology, 174(1), 83-127.

Loosvelt, L., Peters, J., Skriver, H., Lievens, H., Van Coillie, F., De Baets, B., & Verhoest, N. E. (2012). Random Forests as a tool for estimating uncertainty at pixel-level in SAR image classification. International Journal of Applied Earth Observation and Geoinformation, 19, 173-184.

Lopez, A., Tebaldi, C., New, M., Stainforth, D., Allen, M., Kettleborough, J., OCT 2006. Two approaches to quantifying uncertainty in global temperature changes. Journal Of Climate 19 (19), 4785–4796.

Madadgar, S., Moradkhani, H., 2011. Drought analysis under climate change using copula. Journal of Hydrologic Engineering.

Madani, K., Lund, J., 2010. Estimated impacts of climate warming on californias high-elevation hydropower. Climatic Change 102 (3), 521–538.

Makkonen, L., 2006. Plotting positions in extreme value analysis. Journal of applied meteorology and climatology 45 (2), 334–340.

Mann, H.B. (1945)"Nonparametric tests against trend", Econometrica, 13:245-259.

Marvel, K., & Bonfils, C. (2013). Identifying external influences on global precipitation. Proceedings of the National Academy of Sciences, 110(48), 19301-19306.

Mays, L., 2010. Water Resources Engineering. Wiley, pp. 920.

McCarthy, J. J. (Ed.). (2001). Climate change 2001: impacts, adaptation, and vulnerability: contribution of Working Group II to the third assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press.

McMichael, A. J., 2003. Climate change and human health: risks and responses. World Health Organization.

Mengersen, K. L., Pudlo, P., & Robert, C. P. (2013). Bayesian computation via empirical likelihood. Proceedings of the National Academy of Sciences, 110(4), 1321-1326.

Meehl, G., Bony, S., 2011. Introduction to CMIP5. Clivar Exchanges 16 (2), 4–5.

Meehl, G. A., et al., SEP 2007. The WCRP CMIP3 multimodel dataset - A new era in climate change research. Bulletin of the American Meteorological Society 88 (9), 1383+.

Meehl, G.A., et al. (2000)"An Introduction to Trends in Extreme Weather and Climate Events: Observations, Socioeconomic Impacts, Terrestrial Ecological Impacts, and Model Projections" B. Am. Meteorol. Soc. 81:413-416.

Mengersen, K. L., Pudlo, P. and Robert, C. P., "Bayesian computation via empirical likelihood," Proceedings of the National Academy of Sciences 110(4), 1321–1326 (2013).

Mitchell, T. and Jones, P., "An improved method of constructing a database of monthly climate observations and associated high-resolution grids," International Journal of Climatology 25(6), 693–712 (2005).

Milly, P.C.D., et al. (2008), Stationarity is Dead: Whither Water Management? Science 319:573-574.

Min, S. K., Zhang, X., Zwiers, F. W., & Hegerl, G. C. (2011). Human contribution to more-intense precipitation extremes. Nature, 470(7334), 378-381.

Min, S.-K., Simonis, D., Hense, A., 2007. Probabilistic climate change predictions applying Bayesian model averaging. Philosophical Transactions of the Royal Society A - Mathematical Physical and Engineering Sciences 365 (1857), 2103–2116.

Mitchell, T., Jones, P., 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. International Journal of Climatology 25 (6), 693–712.

Morak, S., Hegerl, G. C., Christidis, N., 2013. Detectable changes in the frequency of temperature extremes. Journal of Climate 26, 15611574.

NOAA, 1980. Impact assessment: U.S. social and economic effects of the great 1980 heat wave and drought. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, Environmental Data and Information Service, Center for Environmental Assesment Services, Washington, DC. September.

NOAA, 1995. Impact assessment: The July 1995 heat wave natural disaster survey report. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service, Silver Spring, MD, December.

New, M., et al. (2000)"Representing twentieth-century space-time climate variability. Part II: Development of 1901-96 monthly grids of terrestrial surface climate." Journal of Climate 13: 2217-2238.

Papalexiou, S.M., D. Koutsoyiannis (2013)"Battle of extreme value distributions: A global survey on extreme daily rainfall." Water Resources Research.

Parey, S., Hoang, T. T. H., Dacunha-Castelle, D., (2010)"Different ways to compute temperature return levels in the climate change context."Environmetrics 21: 698-718.

Ohmura, A., & Wild, M. (2002). Is the hydrological cycle accelerating?.Science, 298(5597), 1345-1346.

Owen, A., “Empirical likelihood ratio confidence regions,” The Annals of Statistics 18(1), 90–120 (1990).

Owen, A. B., “Empirical likelihood ratio confidence intervals for a single functional,” Biometrika 75(2), 237–249 (1988).

Qin, J. and Lawless, J., “Empirical likelihood and general estimating equations,” The Annals of Statistics , 300–325 (1994).

Ramanathan, V. C. P. J. K. J. T. R. D., Crutzen, P. J., Kiehl, J. T., & Rosenfeld, D. (2001). Aerosols, climate, and the hydrological cycle. science,294(5549), 2119-2124.

R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, URL http://www.R-project.org (2013).

Reeh, N., 1989. Parameterization of melt rate and surface temperature on the Greenland ice sheet. Polarforschung 59 (3), 113–128.

Reichler, T., Kim, J., MAR 5 2008. Uncertainties in the climate mean state of global observations, reanalyses, and the GFDL climate model. Journal of Geophysical Research-Atmospheres 113 (D5), D05106.

Reiss, R.-D., Thomas, M., 2007. Statistical analysis of extreme values: with applications to insurance, finance, hydrology and other fields. Springer.

Ren, D., Fu, R., Leslie, L., Chen, J., Wilson, C., Karoly, D., 2011. The Greenland ice sheet response to transient climate change. Journal of Climate 24 (13), 3469–3483.

Renard, B., et al. (2013)“Bayesian methods for non-stationary extreme value analysis”, Extremes in a Changing Climate, Springer.

Renard, B., et al. (2006)"An application of Bayesian analysis and Markov chain Monte Carlo methods to the estimation of a regional trend in annual maxima." Water resources research 42.

Ribatet, M. A., A User’ s Guide to the POT Package (Version 1.0), URL http://cran.r-project.org/ (2006).

Robertson, A. W., Lall, U., Zebiak, S. E., Goddard, L., 2004. Improved combination of multiple atmospheric GCM ensembles for seasonal perdition. Monthly Weather Review 132, 27322744.

Robock, A., Vinnikov, K. Y., Srinivasan, G., Entin, J. K., Hollinger, S. E., Speranskaya, N. A., ... & Namkhai, A. (2000). The global soil moisture data bank. Bulletin of the American Meteorological Society, 81(6), 1281-1299.

Rootzén, H., R.W. Katz. (2013)"Design Life Level: quantifying risk in a changing climate".Water Resources Research 49:5964-5972.

Rosbjerg, R. and Madsen, H. (1998)“Design with uncertain design values, Hydrology in a Changing Environment”, Wiley, 155-163.

Salas, J. D., and J. Obeysekera (2013), Revisiting the concepts of return period and risk for nonstationary hydrologic extreme events, J. Hydrol. Eng., doi:10.1061/(ASCE)HE.1943-5584.0000820.

Salvadori, G., & De Michele, C. (2010). Multivariate multiparameter extreme value models and return periods: A copula approach. Water resources research, 46(10).

Santer, B. D., Painter, J. F., Mears, C. A., Doutriaux, C., Caldwell, P., Arblaster, J. M., ... & Zou, C. Z. (2013). Identifying human influences on atmospheric temperature. Proceedings of the National Academy of Sciences, 110(1), 26-33.

Schlather, M., 2002. Models for stationary max-stable random fields. Extremes 5 (1), 33–44.

Schnur, R., & Hasselmann, K. (2005). Optimal filtering for Bayesian detection and attribution of climate change. Climate Dynamics, 24(1), 45-55.

Schubert, S.D., Lim, Y.K. (2013)“Climate Variability and Weather Extremes: Model-Simulated and Historical Data” Extremes in a Changing Climate, Springer Netherlands.

Seidou, O., Ouarda, T. B. M. J., Barbet, M., Bruneau, P., & Bobee, B. (2006). “A parametric Bayesian combination of local and regional information in flood frequency analysis”. Water resources research, 42(11).

Sillmann, J., Kharin, V., Zhang, X., Zwiers, F., Bronaugh, D., 2013. Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. model evaluation in the present climate. Journal of Geophysical Research: Atmospheres 118 (4), 1716–1733.

Simonovic, Slobodan P., and Angela Peck (2009) Updated rainfall intensity duration frequency curves for the City of London under the changing climate. Department of Civil and Environmental Engineering, The University of Western Ontario.

Smith, T., Marshall, L., & Sharma, A. (2014). Predicting hydrologic response through a hierarchical catchment knowledgebase: A Bayes empirical Bayes approach. Water Resources Research, 50(2), 1189-1204.

Smith, R. L., Tebaldi, C., Nychka, D., Mearns, L. O., 2009. Bayesian Modeling of Uncertainty in Ensembles of Climate Models. Journal of the American Statistical Association 104 (485), 97–116.

Smith, R., (2001)“Extreme value statistics in meteorology and environment. Environmental Statistics”, Chapter 8, 300–357.

Smith, R.L. (1989)"Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone." Statistical Science 4: 367-377.

Smith, R.L. (1987)"Estimating tails of probability distributions." Ann. Stat.: 1174-1207.

Solomon, S., et al. (2007) Projections of future climate change. Climate change 2007: the physical science basis. Contribution of working group I to the fourth assessment report of the Intergovermental Panel on Climate Change, Cambridge University Press, New York, NY.

Southworth, H. and Heffernan, J., “texmex: Threshold exceedences and multivariate extremes,” R package version 1 (2010).

Stephenson, A., J. Tawn. (2004)"Bayesian inference for extremes: accounting for the three extremal types." Extremes 7: 291-307.

Stocker, Thomas F., Q. Dahe, and Gian-Kasper Plattner (2013)"Climate Change 2013: The Physical Science Basis." Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Summary for Policymakers (IPCC, 2013)

Storn, R. and Price, K., “Differential evolution–a simple and effcient heuristic for global optimization over continuous spaces,” Journal of global optimization 11(4), 341–359 (1997).

Tanarhte, M., Hadjinicolaou, P. and Lelieveld, J., “Intercomparison of temperature and precipitation data sets based on observations in the Mediterranean and the Middle East,” Journal of Geophysical Research 117(D12), D12102 (2012).

Taylor, K. E., Stouffer, R. J., Meehl, G. A., 2012. An Overview of CMIP5 and the Experiment Design. Bulletin of the American Meteorological Society 93 (4), 485–498.

Tebaldi, C., Knutti, R., 2007. The use of the multi-model ensemble in probabilistic climate projections. Philosophical Transactions of the Royal Society A-Mathematical Physical and Engineering Sciences 365 (1857), 2053–2075.

Tebaldi, C., Hayhoe, K., Arblaster, J., Meehl, G., 2006. Going to the extremes. Climatic Change 79 (3), 185–211.

Tebaldi, C., Mearns, L., Nychka, D., Smith, R., 2004. Regional probabilities of precipitation change: A Bayesian analysis of multimodel simulations. Geophysical Research Letters 31 (24), L24213.

Ter Braak, C.JF, J.A. Vrugt. (2008)"Differential evolution Markov chain with snooker updater and fewer chains." Statistics and Computing 18:435-446.

Ter Braak, C.J.F. (2006)"A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: easy Bayesian computing for real parameter spaces." Statistics and Computing 16:239-249.

Ter Braak, C., 2004. Genetic algorithms and Markov Chain Monte Carlo: Differential Evolution Markov Chain makes Bayesian computing easy. Biometris, Wageningen UR.

Towler, E., B. Rajagopalan, et al., (2010)"Modeling hydrologic and water quality extremes in a changing climate: A statistical approach based on extreme value theory", Water Resour. Res., 46.

Trenberth, Kevin E (2011)"Changes in precipitation with climate change." Climate Research 47.1: 123.

Trenberth, K. E., & Guillemot, C. J. (1998). Evaluation of the atmospheric moisture and hydrological cycle in the NCEP/NCAR reanalyses. Climate Dynamics, 14(3), 213-231.

Trenberth, K. E. (1999). Conceptual framework for changes of extremes of the hydrological cycle with climate change. In Weather and Climate Extremes (pp. 327-339). Springer Netherlands.

Trenberth, K. E., Dai, A., Rasmussen, R. M., & Parsons, D. B. (2003). The changing character of precipitation. Bulletin of the American Meteorological Society, 84(9).

Trenberth, K. E., Smith, L., Qian, T., Dai, A., & Fasullo, J. (2007). Estimates of the global water budget and its annual cycle using observational and model data. Journal of Hydrometeorology, 8(4).

Villarini, G., et al. (2011)"Annual maximum and peaks-over-threshold analyses of daily rainfall accumulations for Austria." Journal of Geophysical Research 116.D5.

Vose, R.S., D.R. Easterling, B. Gleason. (2005)"Maximum and minimum temperature trends for the globe: An update through 2004."Geophysical Research Letters 32.

Vovk, V., 2001. Competitive on-line statistics. International Statistical Review 69 (2), 213–248.

Vovk, V., Zhdanov, F., 2009. Prediction With Expert Advice For The Brier Game. Journal of Machine Learning Research 10, 2445–2471.

Vrugt, J.A., et al. (2009)"Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling" International Journal of Nonlinear Sciences and Numerical Simulation 10:273-290.

Wehner, M. (2013)"Methods of Projecting Future Changes in Extremes." Extremes in a Changing Climate. Springer Netherlands, 223-237.

Westra, S., Alexander, L. V., & Zwiers, F. W. (2013). “Global Increasing Trends in Annual Maximum Daily Precipitation”. Journal of Climate, 26(11).

Winkler, R.L. (1973)“A Bayesian approach to nonstationary processes”. IIASA.

Willett, K. M., Gillett, N. P., Jones, P. D., & Thorne, P. W. (2007). Attribution of observed surface humidity changes to human influence. Nature, 449(7163), 710-712.

Wu, P., Christidis, N., & Stott, P. (2013). Anthropogenic impact on Earth/'s hydrological cycle. Nature Climate Change, 3(9), 807-810.

Wu, Z., et al. (2007)"On the trend, detrending, and variability of nonlinear and nonstationary time series." Proceedings of the National Academy of Sciences104.38: 14889-14894.

Yilmaz, A. G., and B. J. C. Perera. (2013)"Extreme Rainfall Non-Stationarity Investigation and Intensity-Frequency-Duration Relationship." Journal of Hydrologic Engineering.

Yun, W., Stefanova, L., Krishnamurti, T., 2003. Improvement of the multimodel superensemble technique for seasonal forecasts. Journal of Climate 16 (22), 3834–3840.

Zhang, X., Zwiers, F., 2013. Statistical indices for the diagnosing and detecting changes in extremes. In: Extremes in a Changing Climate. Springer, doi: 10.1007/978-94-007-4479-0 1.

Zhang, X., F.W. Zwiers, G. Li. (2004)"Monte Carlo experiments on the detection of trends in extreme values." Journal of Climate 17: 1945-1952.

Zhang, X., et al. (2001)“Spatial and temporal characteristics of heavy precipitation events over Canada” Journal of Climate, 14(9):1923-1936.

Zhang, X., J. Wang, F. W. Zwiers, and P. Y. Groisman, (2010) “The Influence of Large-Scale Climate Variability on Winter Maximum Daily Precipitation over North America”. Journal of Climate, 23, 2902–2915.

Zwiers, F. W., Kharin, V. V., 1998. Changes in the extremes of the climate simulated by CCC GCM2 under CO2 doubling. Journal of Climate 11 (9), 2200–2222.

Zhang, X., Zwiers, F. W., Hegerl, G. C., Lambert, F. H., Gillett, N. P., Solomon, S., ... & Nozawa, T. (2007). Detection of human influence on twentieth-century precipitation trends. Nature, 448(7152), 461-465.

Zhang, X., Alexander, L., Hegerl, G. C., Jones, P., Tank, A. K., Peterson, T. C., ... & Zwiers, F. W. (2011). Indices for monitoring changes in extremes based on daily temperature and precipitation data. Wiley Interdisciplinary Reviews: Climate Change, 2(6), 851-870.