Spring 5-7-2021

Understanding Hydrologic Risk Within the Context of Climate Variability and Atmospheric Moisture Pathways

Ali Aljoda
University of Maine, ali.aljoda@maine.edu

UNDERSTANDING HYDROLOGIC RISK WITHIN THE CONTEXT OF CLIMATE VARIABILITY AND ATMOSPHERIC MOISTURE PATHWAYS

By

Ali Aljoda

B.S. University of Baghdad, 2006

M.E. Universiti Putra Malaysia, 2009

A DISSERTATION

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

(in Civil Engineering)

The Graduate School

The University of Maine

May 2021

Advisory Committee:

Shaleen Jain, Professor of Civil and Environmental Engineering, Advisor

Willem Brutsaert, Professor Emeritus of Civil and Environmental Engineering

Huijie Xue, Professor of Marine Sciences

Kim Jong Suk, Professor of Hydrology and Water Resources

Sean Birkel, Research Assistant Professor and Maine State Climatologist

UNDERSTANDING HYDROLOGIC RISK WITHIN THE CONTEXT OF CLIMATE VARIABILITY AND ATMOSPHERIC MOISTURE PATHWAYS

By Ali Aljoda

Dissertation Advisor: Dr. Shaleen Jain

Nonstationarity in the means and extremes of water resources (e.g., streamflow) due to climate change presents challenges to water resources management and planning. Changes in hydrologic variables such as streamflow critically impact infrastructure, water use planning and adaptation strategies, and can lead to increased societal vulnerability. As such, existing water resources infrastructure designed based on limited records of historical flow may deteriorate under future hydro-climatic changes. Thus, important scientific tasks are: a) examining and redefining the risk, reliability, and return periods under nonstationary conditions, b) understanding the changes in extreme events statistics (i.e. frequency, and magnitude), and c) linking these changes in atmospheric moisture pathways and large-scale climatic processes. Doing so will help to make the decision of management and to mitigate the risk of extreme events, inform science-based decision tools for early-warning, and ascertaining the incidence of extreme events in a changing climate. This work seeks to inform decision management and risk mitigation of extreme events through two contributions.

First, nonstationarity in streamflow regime with linkages to climatic indices (i.e., ENSO and PDO) are analyzed using a case study from the Feather River in California, USA. This includes application of a simple storage-yield-reliability model to quantify the stationary-based risk in the system's design (i.e. reservoir's storage requirement) and

performance (i.e. reliability, resilience, vulnerability). Second, a comprehensive statistical framework for quantifying the nature of variability in the U.S. floods under the impacts of atmospheric rivers (AR) is presented. The approach includes utilizing a combined database of streamflow, atmospheric rivers, precipitation rate, surface air temperature, and water equivalent of accumulated snow depth data to: a) delineate the generating mechanism of a flood event, b) identify the track and source of a moisture-controlled flood event, and c) quantify the nature of variability in floods conditioned on the atmospheric moisture pathways and their oceanic sources.

The complementary research presented in this dissertation seeks to provide improved understanding of hydrologic risk in a variable and changing climate. As such, knowledge gained from this work has implications for engineering design, climate adaptation and decision-making at the local and national levels.

DEDICATION

For my number one, who is truly the only person who has shared every happiness and sadness with me: Nuha. To my littles one, Layanne and Adam. Your happy faces bring me joy every day and help me overcome life's challenges.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 Introduction

Changes in the earth's climate have introduced the nonstationarity - significant shifts in the statistical character of geophysical variables (e.g., precipitation, evapotranspiration, rivers discharge) over time - in the means and extremes of the water-relevant earth systems processes and variables (Jain and Lall, 2000; Khaliq et al., 2006; Milly et al., 2008). Since a few decades, climate change has played a significant role in the variability of water availability which implies that any existing water resources infrastructure designed based on historical flow statistics may be no longer able to be suitably functioning under present and future flows (Jain and Lall, 2001; Jain and Eischeid, 2008; Hossain et al., 2012). The overall aim of this work is to contribute to understanding nonstationarity in water resources and assessing its impacts on water management and decision making. Many hydro-climatic and socio-economic findings in the literature have motivated researchers to conduct work under the umbrella of temporal variation and nonstationarity. For example, studies have shown that climate change can cause significant impacts to freshwater arising from variations in precipitation, air temperature, and wind regimes. Paleohydrologic studies have also suggested that small changes in mean climate may produce large changes in extreme events. Such changes in extreme events poses challenges to infrastructure and water management. It is estimated that the annual global investment in water infrastructure exceeds 500 billion US dollars (Milly et al., 2008).

Consideration of flood events is an important aspect of water management and planning. Floods can cause numerous losses in human lives and properties, considerable damages to the infrastructures, and high stresses on water resources. One real example to this is what happened to Oroville Dam in winter 2016 - 2017. A heavy rain in Northern

California, posterior to five years of drought, caused severe damages in the primary and emergency spillways and rose the lake levels quickly. Due to the possibility of the dam collapsing, authorities evacuated 188,000 people from the Feather River Basin (White et al., 2019). This high-impact incidence highlights the need to address nonstationarity in water management and inspired this dissertation. This work makes two important contributions to the topic of nonstationarity in water resources as follows:

- Diagnosing of the variations in streamflow regimes to discover the temporal variation and nonstationarity in the regional hydrologic change. This includes investigating the variability of streamflow to climate variations, and examining the role of these linkages in a water system operating (e,g,. water supply reservoir).

- Understanding the recent changes in extreme events statistics (frequency, and magnitude) and linking these changes to the large-scale atmospheric phenomena (i.e. atmospheric rivers (AR)).

The remainder of this chapter is organized as follows. First, a brief background on changes in streamflow hydrological regime, hydroclimatic linkages, and the extreme events variability and ARs is provided. This is followed by a problem statement which briefly reviews the purpose of conducting this research. Finally, the section of aim, research questions, and scope that summarizes the details about the objectives of this work and highlights the data and methods used here.

1.2 Background

1.2.1 Changes in Hydrologic Regime

In the 20th century, the alternating of wet and dry spells in the western North American (WNA) region pushed the wheel towards understanding the consequences of climate variations and changes in hydrologic regimes. A diagnosis of the variations in streamflow regimes (most often characterized by the annual mean, interannual variability, and

persistence) is used to discover the temporal variation and nonstationarity in regional hydrologic change. These changes in flow regimes may imply that any existing water resources infrastructure designed based on historical flow statistics may be no longer able to be suitably functioning under the present and future flows, thus raising a question the reliability and the level of performance such systems may offer for critical objectives, such as water supply and flood control (Jain and Lall, 2001; Jain and Eischeid, 2008). As a result, changes in the runoff mean and variance both affect the reliability of streamflow and engender dramatic changes in storage requirement and system performance to meet water demands for multiple objectives (Jain et al., 2002, 2005; Jain and Eischeid, 2008). Furthermore, Jain et al. (2005) inferred a decrease in streamflow reliability in the late of twentieth – century with an increase in the incidence of synchronous flows across the major river basins in WNA region which resulted in expansive water resources stress. Also, they insisted that the detection, attribution, and projection of regional hydrologic change are induced by climate. Finally, the use of relatively short length of historical hydrologic records in design and planning limits the understanding of surprise changes in river flow regime given the issues related to water resources planning and management (Jain et al., 2002).

1.2.2 Hydro-climatic Linkages

The notion that the global climate has varied in the past and will vary in the future is not in doubt. Climate change, whether caused by natural or due to human action, has a precise impact on water resource systems. The nature of streamflow changes in the mean and variability will depend on the magnitude and direction of the climate change. Understanding the possible consequences of climate change on water resources systems is necessary to ensure adequate future water supplies. As such, climate variations can cause dramatic changes in geophysical variables: temperature, precipitation, streamflow, etc. that inevitably influence the system reliability. Therefore, it is critical to acknowledge the

range of such impacts in order to adopt appropriate solutions for risks mitigation in water resource systems.

Numerous studies on the hydrologic impacts of climate change have been conducted in the recent years. As such, the El Nino Southern Oscillation (ENSO), the Pacific Decadal Oscillation (PDO), the Pacific North American Index (PNA), the Atlantic Multidecadal Oscillation28 (AMO), the North Atlantic Oscillation (NAO), and the Arctic Oscillation (AO) are well-known indicators of climate variation to modulate the temperature, precipitation, and streamflow patterns across the United States (e.g., Ropelewski and Halpert 1987; Trenberth and Hurrell 1994; Mantua et al. 1997; Cayan et al. 1999; Dettinger et al. 2000; Jain and Lall 2001; Mallakpour and Villarini 2016; Dickinson et al. 2019). A focus on the correlation of streamflow with ENSO and PDO will be considered here to examine the temporal and spatial impacts of climate variations on the criteria of system operation. ENSO is a natural ocean-atmospheric variation phenomenon that involves fluctuating ocean temperatures in the tropical Pacific. On the other hand, PDO is the dominant year-round pattern of monthly North Pacific sea surface temperature (SST) variability. Although number of studies have investigated the linkages of these phenomena with streamflow, the relations are not well understood and their impacts on water resources systems have not addressed.

1.2.3 Extremes and Atmospheric Rivers

From a scientific standpoint, efforts to develop new approaches to integrative water and climate assessment are aided by the availability of high-resolution data from space-borne platforms. A striking example is the identification of atmospheric moisture pathways, known as atmospheric rivers (AR). ARs represent well-organized filamentary structures in the atmosphere that transport significant amounts of moisture over very long distances (often exceeding a thousand miles) (Ralph and Dettinger, 2011). Their typical horizontal dimensions can be several thousand kilometers long with width of approximately 500 km

(Guan and Waliser, 2015). At any given time there may be three to five ARs in each hemisphere (Zhu and Newell, 1998).

Atmospheric rivers are a part of the large extratropical cyclones system that transport heat and moisture from the tropics toward the poles. They are typically located within the low-level jet (LLJ), an area of strong winds in the lower levels of the atmosphere, ahead of the cold front in an extratropical cyclone. The moisture transportation through ARs occurs under particular combinations of wind, temperature, and pressure conditions. Thesw usually include high humidity levels, strong low level winds, and a moist neutral atmospheric profile, which allows for extensive precipitation production when air is lifted orographically over geographic features, such as mountain ranges (Ralph et al., 2004; Gimeno et al., 2014). ARs are monitored by satellites and predictable up to a week ahead of landfall and subsequent heavy rain events. While the existence of ARs has been known for 25 years (Newell et al., 1992), it is only recently that ARs have been recognized as an important hydroclimatic phenomenon (Guan and Waliser, 2015; Gimeno et al., 2016). Furthermore, significant societal impacts, especially due to flooding along the US west coast have been attributed to ARs (for example, Ralph and Dettinger 2011). Through careful integration of satellite data, in situ rainfall measurements, and model-based data assimilation, researchers at NASA Jet Propulsion Laboratory (NASA JPL), have developed new tools and suite of data products to enable assessment of AR events (Guan and Waliser, 2015) which are used to address the dissertation objectives in Chapter 3.

The frequency, magnitude, and timing of extreme precipitation have manifold consequences. ARs represent an important hydroclimatic phenomenon, not only from the standpoint of their predictability but also the potential to understand the cause of heavy rainfall and floods in a particular region. A review of the published literature reveals that the western, midwestern, central, and northeast US (Lavers and Villarini, 2013; Nakamura et al., 2013; Lu and Lall, 2016) are regions where the role and impact of ARs have been studied and confirmed. A large proportion of extreme precipitation events in the United

States and western Europe are now thought to be linked to the occurrence of ARs, including some famous historical floods (Lu and Lall, 2016; Konrad and Dettinger, 2017; Paltan et al., 2017).

1.3 Problem Statement

Although the concept of nonstationarity was born two decades ago (Jain and Lall, 2000), the literature is still lacking a concrete understanding of the impacts of changing climate on the planning and management of water resources systems (e.g., water supply reservoir). Changes in streamflow hydrologic regime and flooding records due to linkages to climate drivers variability have been explained over different scales of time and space (e.g., Jain and Lall 2000, 2001; Jain et al. 2005; Mallakpour and Villarini 2016; Dickinson et al. 2019). In particular, variations in reservoir storage requirements and in the reliability of meeting the water demand over decadal time-scale are examined. Jain et al. 2002 showed that the reliability of water supply in interior western US is contingent upon the nature of hydrologic regime. In other words, the ability of a reservoir storage to provide reliable water supply depends both on the scales of variability embedded in the limited length flow record used for design, and the nature of flow probabilities. Furthermore, Jain and Eischeid 2008 showed that reservoir storage requirements are very sensitive to the variability in mean, standard deviation, lag-1 serial correlation of the upper Colorado River basin streamflow - which is key linkages to climate variability. To this end, reliable results to understand the reflections of the hydrologic regime variations and climate variability on water resources management are vital to anticipate and mitigate the deterioration in the operation of the water resources structures.

Giving the impacts of changing hydrologic regime due to climate variation, a reliable measure of flooding nonstationarity enables the prediction of future extremes to insure life security and property protection. The statistical framework B17-B [Interagency Advisory Committee on Water Data (1982). "Guidelines for determining flood flow frequency:

Hydrology Subcommittee Bulletin 17B." Reston, VA: USGS] and its updated version the B17-C "Guidelines for determining flood flow frequency—Bulletin 17C" (England et al., 2018) were published to modulate the variability of floods in frequency and magnitude. The framework assumed that the flood records in a timeseries are homogeneous and independent and identically distributed (iid). Although the assumption is eventually violated when the mixed-population floods (heterogeneous) was introduced Waylen and Woo (1982); Hirschboeck (1991); Webb and Betancourt (1990) and in later studies the extreme floods are linked with the large atmospheric circulations (i.e. ARs), the risk of floods variation over time have not well understood. To this end, a robust methodology that deals with mixed-population floods and a unique approach to identify the source of floods variability are both needed to develop new modalities in flood frequency estimation in a changing climate.

1.4   Aim, Research Questions and Scope

The overall aim of this dissertation is to focus on understanding the nonstationarity in water resources and discovering its impacts on the water management and decision making. The research questions and the scope of this dissertation are splitted into two chapters (Chapter 2 and Chapter 3).

In Chapter 2, the Feather River inflow into Lake Oroville (FRI) is used to understand the changes in reservoir storage requirement and performance, stemming from the changing hydrological regime of the river basin and, to investigate the potential linkages between the runoff variability and climate variations. A simplified reservoir with specified storage and demand is used to quantify the risk of a stationary-based storage estimate, and to examine the reservoir performance under the uncertainties and risks of hydroclimatic variation. This framework leads the research to ask the following questions:

1. How is the hydrologic regime of the FRB changing historically?

2. What are the relative impacts of embedded temporal variations and nonstationarity to N-year blocks from the historical FRI and climate-driver records on the reservoir storage requirements and performance?

3. What are the long-term linkages between the historical FRI variability and the climate variations which may affect the system design and performance?

The observed and resolved multi-century long-term reconstructed records of the Feather River inflow into Lake Oroville (FRI) and climate indices (ENSO and PDO) are used to achieve the research objectives. The analyses are performed by using a simplified reservoir storage-yield-reliability (S-Y-R) model for reservoir storage estimations, and using the Reliability, Resilience, Vulnerability (RRV) metrics to evaluate the system performance. Further, application of wavelet and coherence analysis to identify the temporal and spatial periodic patterns as well as to introduce the possible linkages among the hydroclimatic variables.

Chapter 3 aims to determine satisfactory flood frequency estimates for mixed population floods across the US. This study examines the role of ARs and their moisture trajectories and sources on the heterogeneous flooding events of annual maximal flows in the US, and their impact on the magnitude and frequency estimates used for the design of flood structures. Therefore, three research foci related to the nature and variability of annual maximum flood events across U.S. are:

1. To develop a methodology for robust curvilinear estimation of AR events based on a Principal Curve-based approach, thus integrating the axes of maximal IVT variation and time progression over the event lifecycle.

2. To characterize the place-based mixed population of annual maxima flow (AMF) by delineating the constituent remote atmospheric moisture sources, thus enabling climate-informed partitioning of the flood record for risk analyses and estimation of hydroclimatic nonstationarity.

3. To systematically apply the new methodology towards a US-scale assessment of annual maximum floods to understand their regionality, and with a goal to develop new modalities in flood frequency estimation in a changing climate.

The analyses presented in this chapter used atmospheric reanalysis data sets to extract the timeseries for the point(s) within the watershed, as well as 623 USGS streamflow gauges to determine the AMF over 60-year period (1956-2015). To achieve the objectives, the current study introduce a comprehensive framework that aims to: a) identify major causes of annual floods using a hydroclimatic-based classification, b) determine the major tracks and sources of flood-induced moisture by employing the principal curve analysis to develop a new approach, and c) adopt new modality for risk analyses and estimation of hydroclimatic nonstationarity by climate-informed partitioning of the flood record.

Taken together, this piece of research adds two important contributions to the topic of nonstationarity in water resources. Tow original studies establish complimentary statistical frameworks that deal with managing water resources systems under the hydrologic impacts of climate change.

1.5   Thesis Structure

The rest of this thesis is structured as follows:

- Chapter 2 is formulated as a stand-alone research article. This chapter focuses on the impacts of changing hydrologic regime - climate change linkages on reservoir storage requirements and performance. It has three parts of analysis results which answered the research question in the article.

- Chapter 3 is formulated as a stand-alone research article. The main objective of this chapter is to determine satisfactory and reliable estimates of flood frequency in a US-scale analysis. The methodology section contains three phases of analysis to answer the research questions. Consequently, a comprehensive statistical framework

9

that starts by attributing the annual maxima flow records to their generating mechanisms and shows which are the major processes, next, investigate the impacts of ARs on flooding frequency and magnitude, and lastly, quantify the nature of flood variability based on moisture sources variation.

- Chapter 4 closes this dissertation by discussing the findings of this research; the perspectives for future research, and summarizing the major contributions and conclusions of the work.

CHAPTER 2

UNCERTAINTIES AND RISKS IN RESERVOIR OPERATIONS UNDER CHANGING HYDROCLIMATIC CONDITIONS

2.1 Chapter Abstract

Uncertainties and risks associated with hydroclimatic variations pose a challenge to the management and planning of water resources systems. This study demonstrates the importance of understanding the changing hydrologic regime of the Feather River Basin (FRB) and its impacts on water resources decision variables (i.e., storage requirement and performance of a water supply reservoir). A simple storage-yield-reliability model (S-Y-R) is used to quantify the risk of the stationary-based-designed reservoir under the temporal variation and nonstationarity in N-year blocks of the Feather River Inflow into Lake Oroville (FRI). Furthermore, the potential linkages of the long-term variability in the FRI to climate variations are investigated by applying wavelet spectrum and coherence analysis to the FRI and atmospheric-oceanic indices (e.g., ENSO and PDO). The results show substantial variations in the FRB hydrologic regime over different timescales with episodes of abrupt shifts toward significantly higher storage requirements, and decrease in the reservoir performance during historical periods of high FRI variance and lag-1 serial correlation. Although the mean inflows are high, the storage capacity is increased by (a) 38% and 48% due to the 5% and 20% increase in the FRI variance during the periods 1904-1953 and 1960-2009, respectively and, (b) 34% due to the increase in the serial correlation coefficient in the period of 1750-1799. Likewise, reservoir performance significantly decreased for the same reasons in the same critical periods. The reliability and resilience dropped to 74% and 29% (1904-1953) and to 76% and 50% (1960-2009 period) due to the increased variance of FRI, while vulnerability reached 70% during the high lag-1 correlations in 1532-1581 and 1564-1613, and 40% in 1904-1953 due to the high FRI

variance. Furthermore, the wavelet coherence analysis observes strong associations between the streamflow and climate teleconnection patterns in specific periodic cycles during the same critical periods which link the variability in FRI and decision variables to the hydroclimatic variations. These linkages give a primary indication for the reservoir storage requirement characterization.

2.2   Introduction

Dams and reservoirs are the most pervasive infrastructure elements that help achieve multiple societal objectives: reliable water supplies, flood control, recreation, and hydroelectricity. Large dams, such as, Oroville, Hoover, Grand Coulee, and Glen Canyon, have been catalysts for regional socioeconomic development. At the same time, the deleterious effects of dam building include harm to ecosystems, dislocation of people and cultures, and inundation of lands. Given this precarious balance, sound approaches to dam operations and management have the potential to ensure water security, as well as support efforts to restore and improve the ecosystem services (Poff and Olden, 2017).

Changing climate has the potential to cause nonstationarity (significant shifts in the statistical characteristics of geophysical variables over time) in the mean levels and extremes of temperature, precipitation, evapotranspiration, and streamflow (Jain and Lall, 2000; Khaliq et al., 2006; Milly et al., 2008). When such changes occur, accompanying increases in uncertainty and risk in the management and planning of water resources systems is also expected. As such, changes in the hydrological regime of a river basin under hydroclimatic variations implies that any existing water resources infrastructure that was designed based on limited historical flow records can fail under present or future flows (Jain and Lall, 2001; Jain and Eischeid, 2008; Hossain et al., 2012), thus raising the question of the reliability and the level of performance such systems may offer for critical objectives, such as water supply and flood control.

This work is motivated by a critical high flow event at the Oroville Dam, located on the Feather River, in February 2017. This incident forced 188,000 people downstream to evacuate as the authorities had concerns regarding potential dam failure, and eventually caused a loss of more than 1 billion US Dollars. Lake Oroville (a manmade multipurpose reservoir of 4.3 km^3 storage capacity) supplies fresh water to approximately 25 million Californians and about 3,035 km^2 of irrigated farmland (CADWR, 2009). In water year 2016 (1 Oct - 30 Sep), the Feather River Basin (FRB) received normal precipitation that filled Lake Oroville above historical levels to near its storage capacity. In the first ten days of February 2017, an extreme rainfall event (650 mm) in the FRB, damaging the primary spillway and raising the lake level quickly, which caused the water to flow over the emergency spillway for the first time in the dam's history.

Alongside episodic variations, climate teleconnection patterns such as the El Niño-Southern Oscillation (ENSO) modulate the regional precipitation extremes and contribute to flooding in the region of FRB. For instance, two of the wettest seasons in the FRB occurred during the 1983 and 1998 strong El Niño (warm) years, which were associated with large annual increase in water volume in Lake Oriville. Although water year 2017 is characterized by weak La Niña (cold) condition, the negative Multivariate Niño Index indicates a wet year in the northwest United States and the wettest season in the FRB since 1920 (White et al., 2019). In addition, Cayan et al. (1999) indicated that extreme precipitation and flooding in the FRB region may be more common with La Niña. However, the absence of such extreme events in the period of records of the structure's design put further limits upon the diagnosis of multidecadal and centennial hydroclimatic variations. As such, stationary-based designed infrastructures and policies which relied on relatively short hydrological records add more challenges to water resources management. This motivated the authors to use the long-term annually-resolved reconstructed data sets in the analysis of this study to investigate the surprising changes in the decision variables

that occurred in the past and may occur again in the future due to hydroclimatic variations.

In this paper, a simplified reservoir with specified storage and demand is used to quantify the risk of a stationary-based storage estimate, and to examine the reservoir performance under the uncertainties and risks of hydroclimatic variation. The analysis presented here uses the Feather River inflow into Lake Oroville (FRI) as a case study to understand the changes in storage requirement and reservoir performance, stemming from the changing hydrological regime of the FRB, for a given level of demand and reliability and, to investigate the potential linkages between the runoff variability and climate variations. By using annual data sets of observed and resolved multi-century long-term reconstructed records of the FRI and climate indices, we answer the following questions: a) How has the hydrologic regime of the FRB changed historically? b) What are the relative impacts of embedded temporal variations and nonstationarity to N-year blocks from the historical FRI and climate-driver records on the reservoir storage requirements and performance? c) Finally, what are the long-term linkages between the historical FRI variability and the climate variations which may affect the system design and performance? We explore these issues using a simplified reservoir storage-yield-reliability (S-Y-R) model for reservoir storage estimations, and using the Reliability, Resilience, Vulnerability (RRV) metrics to evaluate the system performance. Further, we apply wavelet and coherence analysis to identify the temporal and spatial periodic patterns as well as to introduce the possible linkages among the hydroclimatic variables.

2.3   Background

2.3.1   Reservoir Characteristics

The reservoir's function is to regulate the irregular natural flow to provide a regular rate of outflow to serve reservoir objectives. Several quantities are necessary for the reservoir's design and modeling (e.g., storage, yield, demand, etc.). Storage requirement is the volume

14

of storage that is needed to supply a given demand in a specified period under a selected level of reliability. The considered storage in this study is the active storage capacity, the difference between the maximum reservoir capacity at full supply level and the dead storage, the volume of water held below the lowest off-take valve. Active storage S_t ranges between zero and a maximum value C imposed by the reservoir size. The target draft or water demand D_t is the volume of water withdrawn from a reservoir to meet demand over a selected period. While the reservoir yield, draft or release R_t is the abstracted water during the same period of demand, and both have units of volume per specified time. Although the desirable yield is equal to the demand, it may fall below the target draft during the drought period and may exceed it in times of plenty. Yield can be decreased less than the target draft when the level of storage in a reservoir is low. Conversely, yield may increase over the target draft when the reservoir is full. The system starts to spill if it is filled to its maximum capacity and operating at its maximum level, and is fed by an inflow that is higher than its ultimate operating level. The base yield is the only yield component considered in this study. It is the lowest yield recorded when a reservoir is fed by an inflow while attempting to supply water to meet demand under a particular operating policy. However, the maximum abstracted base yield from a reservoir equals the target draft. The reservoir spill W_t is the excess volume of water that cannot be stored in the reservoir due to its maximum capacity C which usually occurs during periods of flood. These magnitudes can be determined in the following steady state equations (Vogel and Stedinger, 1987; **?**):

$$S_t = max[0, min(S_{t-1} + Q_t - D_t, C)] \,. \tag{2.1}$$

$$R_t = min[S_{t-1} + Q_t, D_t] \,. \tag{2.2}$$

$$
\begin{aligned}
W_t &= S_{t-1} - S_t + Q_t - R_t \\
&= max[0, S_{t-1} + Q_t - D_t - C] \,.
\end{aligned}
\tag{2.3}
$$

where Q_t is the net inflow per selected time. In the storage-yield relationship, the storage capacity is expressed by a ratio or a percentage of the mean annual flow or as a standardized capacity. The target draft is also denoted as a fraction of the mean annual flow. **?** used a large data set of monthly and annual streamflow for 729 rivers around the world to assess and compare the performance of certain storage requirement methods. In that study, statistics of reservoirs in three large countries - Australia, South Africa, and the United States - illustrated the range of variation in reservoir characteristics. This source indicated that the storage capacity could be times smaller or larger than annual mean flow. Nevertheless, the ratios of the storage and the draft as a fraction of the mean annual flow can be varied regionally and spatially. For instance, in Australia and South Africa, the capacity ratios are (>0.25 - 6) and (>0.7 - 3.3), respectively, while the target draft ratio in both countries is (>0.1 - 0.9). In the United States, both proportions are varied between the western and eastern regions. Reservoir storage capacities are less than the mean annual flow in the eastern regions but range from nearly zero to nearly 500% in some western regions. However, the demand ratio in some eastern regions is varied around (0.4 - 0.95) of the mean annual flow and is nearly uniformly distributed in some western regions.

Throughout this study, the 50-year period (1911–1960) prior to the construction of the Oroville Dam is used to compute the reference hydrology (baseline) period. An annual demand of 80% and maximum reservoir capacities of 25%, 50%, and 75% of the baseline mean inflow are used in the reservoir S-Y-R computations. The impact of evaporation on the S-Y-R analysis is not considered due to the lack of long-term and reconstructed estimates. However, the analysis approach presented here can be readily incorporate evaporative losses, when available, as a component of water demand. Furthermore, the terms standard deviation and variance are used interchangeably to describe interannual variability. Similarly, we use persistence in the runoff, lag-1 correlation and serial correlation synonymously.

2.3.2 Streamflow Variability in a Changing Climate

Earth's climate has undergone substantial variations in the past and will vary in the future. Climate change, whether caused by natural phenomena or by human action, will have a certain impact on water resource systems. The nature of streamflow changes in the mean and variability will depend on the magnitude and direction of the climate change. Understanding the possible consequences of climate change on water supply systems is necessary to ensure adequate future supplies. However, climate variations can cause dramatic changes in geophysical variables: temperature, precipitation, streamflow, etc. which inevitably influence the system performance. It is important to acknowledge the range of such impacts in order to adopt appropriate planning and mitigation measures for water resource systems.

Numerous studies in the literature have been conducted in the last three decades on the hydrologic impacts of climate change. The El Niño-Southern Oscillation (ENSO) and the Pacific Decadal Oscillation (PDO) are both well-known indicators of climate variation that modulate the temperature, precipitation, and streamflow patterns across the United States (e.g., Ropelewski and Halpert 1987; Trenberth and Hurrell 1994; Mantua et al. 1997; Cayan et al. 1999; Dettinger et al. 2000; Jain and Lall 2001). ENSO is a natural ocean-atmospheric variation phenomenon that involves fluctuating ocean temperatures in the tropical Pacific. On the other hand, PDO is the dominant year-round pattern of monthly North Pacific sea surface temperature (SST) variability. Of particular interest in this study is the time-varying relationship between climatic indices (such as, ENSO and PDO), and the extent to which reservoir performance may mirror climatic variability.

2.4 Data

Annual observations (1906–2012) and tree-ring based reconstruction (900AD–2012) of the FRI in California with multiple sets of climate (ocean–atmospheric) data are analyzed here. Both of the observed inflows of the Feather River into Oroville Dam and the

17

reconstructed streamflow data, which was updated by David Meko and Ramzi Touchan (University of Arizona Laboratory of Tree-Ring Research) in 2013-2014, are provided by the California Department of Water Resources (https://www.treeflow.info/content/feather-river-inflow-oroville-reservoir-ca-update).

The observed ocean-atmospheric data sets are the monthly Southern Oscillation Index (SOI) for the period (1876–2017) and the monthly Niño 3.4 SST Index (1870–2017) are provided by the National Weather Service – Climate Prediction Center (https://www.cpc.ncep.noaa.gov/data/indices/). The Pacific Decadal Oscillation Index (PDO) monthly timeseries are provided and updated by Mantua (http://research.jisao.washington.edu/pdo/PDO.latest.txt). On the other hand, the reconstructed datasets of climatic indices used in this analysis are the 700-year El Niño/Southern Oscillation (ENSO) Niño3.4 index reconstruction (1301–2005) (Li et al., 2013), the 700-year tree-ring ENSO index reconstructions (1300–2006) (Cook et al., 2008), and the Pacific Decadal Oscillation reconstruction data for the past millennium (993–1996) (MacDonald and Case, 2005). It is worthy noting that the current study is treated the uncertainties in the long-term reconstructed data by using the observed records to generate robust conclusions.

2.5   Methods

2.5.1   Storage-Yield-Reliability Model

The Gould-Dincer (G-D) formulation is a simple S-Y-R model for a single reservoir. It uses annual inflow statistics to compute the over-year (carry over) capacity. Mean annual inflow and standard deviation are used to assess the water storage based on a variable and changing climate. The model has three sets of formula based on annual inflows distribution; Normal, Lognormal, and Gamma. The reader is encouraged to refer to McMahon et al. (2007) and Jain and Eischeid (2008) for details about the theories and assumptions of the three sets. The current study is limited to the G-D Normal suite in its analysis. This

18

model assumes that annual inflows are normally distributed and independent. The model accounts for the inflow persistence by using the lag-1 serial correlation as follows;

$$C = \frac{Z_p^2}{4(1-\alpha)} C_v^2 \mu \frac{1+\rho}{1-\rho} \ . \tag{2.4}$$

where C is the required storage, Z_p is the reliability, α is the target draft fraction, $C_v = \frac{\sigma}{\mu}$ is the coefficient of variation, μ is the mean annual inflow, σ is the standard deviation, and ρ is the lag-1 serial correlation.

Two checks are adopted here to ensure that the storage estimates are consistent with the over-year storage assumption based on standardized net inflow or drift, $\mu = \frac{1-\alpha}{C_v} < 1$ and the critical time is greater than one year as follows;

$$n_{crtl.} = \frac{Z_p^2}{4(1-\alpha)} C_v^2; \qquad n_{crtl.} > 1 \ . \tag{2.5}$$

where, $n_{crtl.}$ is the time taken by the reservoir to empty from a fully-filled state.

2.5.2 Criteria of Reservoir Performance Evaluation

The evaluation of reservoir performance in this work is carried out by applying three metrics: reliability, resiliency, and vulnerability (RRV) introduced by Hashimoto et al. (1982). Although these criteria were defined based on the assumption of stationarity, the distribution is time-invariant, the present analysis uses them to evaluate the dynamic risk of a reservoir's performance in a changing climate.

Reliability is the number of satisfactory events, when the targeted demand is met during the simulation time and it can be determined as follows:

$$R_s = \frac{N_s}{N}; \qquad 0 < R_s \leq 1 \ . \tag{2.6}$$

where R_s is the time-based reliability, N_s is the number of satisfied years or events, N is the total number of events or the whole period of simulation.

Resilience measures how quickly the reservoir will recover when it has already failed to meet the target draft. The expression used to find the reservoir resilience in the current study is defined as follows:

$$r = \frac{f}{N_f}; \qquad N_f \neq 0 \ . \tag{2.7}$$

where r is the resilience, f is the number of individual continuous sequences of failure, and N_f is the total duration of all the failures.

The dimensionless vulnerability variable measures the severity of reservoir shortfall during the period of failure. It is defined here as follows;

$$\upsilon = \frac{\sum_{i=1}^{N_t} max(S_i)}{D_t N_t}; \qquad N_t \neq 0 \ . \tag{2.8}$$

where υ is the dimensionless vulnerability, S_i is the volumetric shortfall during the ith continuous failure sequence, D_t is the target draft, and N_t is the number of failure sequences. Both metrics, resilience and dimensionless vulnerability, are on the interval of [0,1] and undefined for non-failure system. The reader can refer to Hashimoto et al. (1982) for more details on the theories and mathematical expressions and to McMahon et al. (2006) for applications and analytic example.

2.5.3   Wavelet and Coherence Analysis

Wavelet transform (WT) is a well-known analysis tool to study the multi-scale, nonstationary processes occurring over finite spatial and temporal domains (Lau and Weng, 1995). Since its introduction by Morlet in 1983, WT has found wide application in signal and image processing, medicine, geophysics, astronomy, and economics. Numerous studies in geophysics have used WT in different fields of research. It has been used for hydroclimatic and oceanic variables (e. g., ENSO, PDO, temperature, precipitation, etc.) (Lau and Weng 1995; Minobe 2000; Grinsted et al. 2004; Ho et al. 2017). A completely detailed description of WT applications in geophysical research can be found in

Foufoula-Georgiou et al. (1995), while a theoretical explanation of WT analysis is given in Torrence and Compo (1998).

Wavelet analysis is a common tool for analyzing localized variations of power within a time series. By decomposing a time series into time-frequency space, one can determine both the dominant modes of variability and how those modes vary in time (Torrence and Compo, 1998). There are two classes of wavelet transforms: Continuous Wavelet Transform (CWT) and its discrete counterpart (see Grinsted et al. (2004) for details). CWT is commonly used for analyzing localized intermittent oscillations in a time-series and examining two time-series together that may be expected to be linked in some way. In this paper, we used the CWT to expand the time series into time-frequency space to find localized intermittent periodicities. Furthermore, we applied Wavelet Coherence (WTC) between two time-series to find a significant association, although the common power is low. However, the wavelet analysis presented here is carried out by using the WaveletComp package (Rösch and Schmidbauer, 2016) in the R environment programming software.

2.6   Results and Discussions

2.6.1   FRB: Historical Changes in The Hydrologic Regime

Hydroclimatic variation in the FRB occurred over different ranges from interannual, to multidecadal and centennial time scales. The streamflow statistics (mean, variance, lag-1 correlation) became particularly interesting in the context of the inflow variability and the reliability of water supplies in FRB (see Figure 2.1). Both the mean inflow and standard deviation estimates interestingly described the high and low runoff regime over time as some characterized by high year-to-year variability, and others with relatively low interannual variability. Multidecadal and centennial time-scale variations were also introduced in the historical FRI records. The long-term flow regime variability can be clearly seen throughout the sequence of high- and low-flow regime periods throughout eleven centuries (see Figure 2.1). An example of the multidecadal time-scale variability in

21

FRI is the high-flow period in 1100-1150, the highest mean state in the entire record, followed by a low mean-state period in the second half of the twelfth century. An example of centennial time-scale variability in the FRI is the relatively high mean inflow estimates of the fourteenth century that are followed by a century of relatively low mean runoff.

Figure 2.1 shows the underlying temporal variations and nonstationarity in the FRI characteristics as they vary in time over a variety of time scales. These changes in the runoff characteristics dramatically affected the storage requirement and the metrics of the system performance (discussed in detail in the next section). Finally, it is noteworthy that the 50-year moving window average for the mean inflow and variance interestingly showed an upward trend of high interannual variability with a relatively high mean state in the last 30 years of the record, which makes it the period of the highest variability over the entire 11 centuries. This result is reinforced by other studies which have also noted a late 20th-century trend towards higher variance of streamflow across the western United States (Jain et al., 2005; Jain and Eischeid, 2008). As a result, increasing year-to-year variability implied a higher incidence of elevated aridity and wetness relative to the mean state; however, both factors may negatively alter reservoir storage and performance.

2.6.2   Impacts of FRI Nonstationarity on a Hypothetical Reservoir Applications

The embedded nonstationarity in the FRI is examined by using reservoir applications of storage requirements and performance indices (RRV). Individually, the 50-year moving window average estimates of the three key statistics in Figure 2.1 (mean, variance, and persistence) provide a sampling distribution of these metrics to compute reservoir storage requirements and performance metrics with time.

2.6.2.1   Reservoir Storage Requirements

In this analysis, the Gould–Dincer Normal-model (McMahon et al., 2007; Jain and Eischeid, 2008) is used to estimate the storage requirements of a hypothetical reservoir
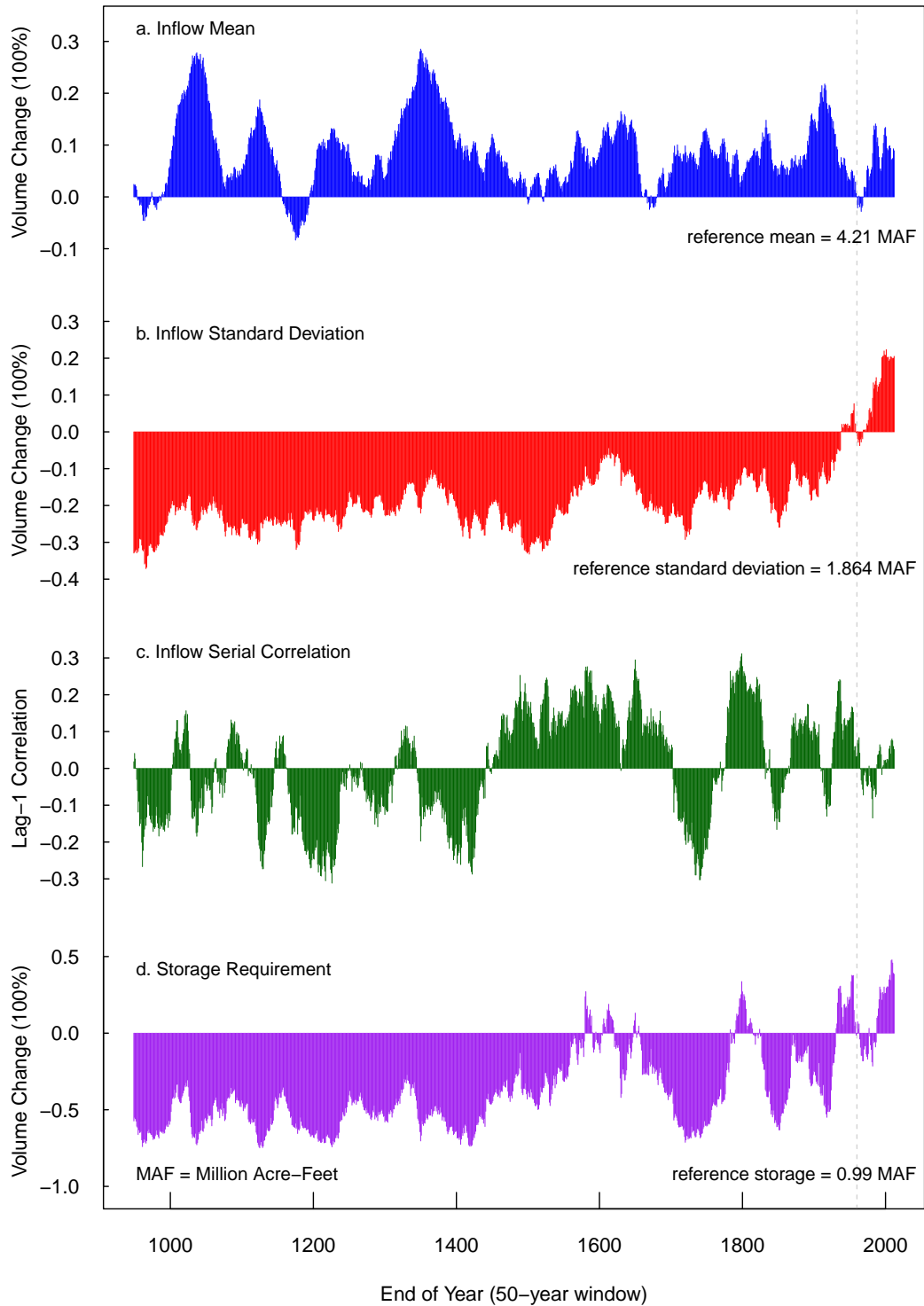
Figure 2.1: Long-term variations in the FRI and storage requirement (900AD-2012). Results in (a, b, & d) are shown as percentages for the departure from the reference: mean, standard deviation, and storage requirement, respectively, of the hydrologic baseline period (1911-1960). A gray vertical dashed line highlights the mean, standard deviation, serial correlation, and storage computed for the hydrologic reference period. 1 MAF = 1.234 KM³.

serving a water demand of 80% of the baseline mean annual inflow with a 95% reliability of supply. Figure 2.1(d) demonstrates the nonstationarity in the storage requirements due to the impacts of changing runoff regime. The mean and standard deviation of the FRI (Figure 2.1(a b)) showed that the reservoir's capacity is indirectly related to the mean inflow and is inversely associated with the variance. Changes in the inflow mean and interannual variability can cause significant variations in the reservoir's storage requirements. For instance, rapid and abrupt changes in the storage requirements during the 16th and 20th centuries (see Figure 2.1(d)). At the same time, the shift in both statistics can offset their effects on storage requirements, e.g., the decrease in the inflow standard deviation during the twelfth century nullified the effect of the decreasing runoff mean on the storage requirements (Figure 2.1).

It is critical to state that the storage requirements in the 50-year periods of (1904-1953) and (1960-2009) abruptly increased by 38% and 48% of the baseline storage because of the significant changes in the FRI variance during the same periods (see Table 2.1). However, the impact of the lag-1 serial correlation coefficient on the storage requirements appears through $((1 + \rho)/(1 - \rho))$. Changes in the inflow persistence can individually alter the storage estimation, i.e., although the inflow mean was relatively high and the interannual variability was low in the second half of the 15th century (1750-1799), the reservoir storage was 34% higher than the baseline storage due to the effect of high serial correlation values (see Figure 2.1(c & d) and Table 2.1). Further, periods of the 16th and early of 17th centuries display an upward trend toward high storage requirements due to the increased values of the lag-1 correlation coefficient during the same periods (Table 2.1). Thus, the results of the S-Y-R model show that the sensitivity of the reservoir storage requirements are related to the changes in the FRI variance and persistence.

Table 2.1: The relationship of the storage requirements rate of change to the changes in FRI statistical characteristics with respect to the hydrologic references.

Period	$\frac{\mu - \mu_0}{\mu_0}$ (%)	$\frac{\sigma - \sigma_0}{\sigma_0}$ (%)	ρ	$\frac{C - C_0}{C_0}$ (%)
1532−1581	6.7	-9.3	0.28	27.1
1564−1613	13.4	-4.4	0.22	18.9
1601−1650	12.6	-14.0	0.29	13.1
1750−1799	4.0	-11.7	0.31	33.7
1904−1953	5.2	5.0	0.17	37.7
1960−2009	7.3	20.1	0.08	47.9

Reference period $= 1911 - 1960$), Mean $= 4.21$ MAF, Standard Deviation $= 1.86$ MAF, Lag-1 Correlation $= 0.032$, Storage Requirement $= 0.99$ MAF.
MAF $=$ Million Acre-feet
1 MAF $= 1.234$ KM^3

2.6.2.2   Reservoir Performance Evaluation

The performance of a hypothetical reservoir with three different capacities is evaluated by using the reliability, resilience, and vulnerability indices (Figure 2.2). Fluctuations in the reliability curves throughout the entire record captures the changes in the hydrological regime of the FRI as shown in Figure 2.1. It is clear that the reliability of the water supply improved under a condition of low interannual variability with a relatively high mean state. It is noteworthy that the reliability of the water supply during the periods of 1904-1953 and 1960-2009 reached higher rates of failure to fulfill the demand in a 50-year window throughout the 11 centuries (see Table 2.2). These changes in the system reliability are explained by the elevated level of the streamflow variance during these periods (Figure2.1(b)). Furthermore, the influence of the streamflow persistence on the reservoir reliability is observed during the periods of high lag-1 serial correlation coefficient values such as in the 16th and 18th centuries when the system became less reliable for water supply (Table 2.2).

However, the results presented above are limited in the following manner; a) potential events of shortfalls are counted regardless of the persistence and severity of deficit in water supply, b) the potential need for alternative water sources and strategies such as conservation were not well informed by these results. Thus, this prompted us to look for

alternate indices, such as the distribution of failure events (resilience and vulnerability) that more reliably illustrate the persistence and severity of events, and thus clarify the target strategies to achieve sustainable solutions.

The resilience and vulnerability metrics of a water supply reservoir are also functions of the runoff regime variability. Both indices improved under a condition of low inflow variance and/or a high mean inflow state. Figure 2.2 shows that the reservoir recovered more quickly from failure and became less sensitive to failure in the 14th and 15th centuries due to the low inflow variability. Interestingly, in the same two time-periods of the 20th century the system displayed low resilience and high vulnerability in its performance due to the high year-to-year variability in the FRI during these periods, even though they were periods of relatively high mean states, which demonstrates the key role of the runoff variance in reservoir performance (see Figure 2.1(b) and Table 2.2). On the other hand, the positive high coefficients of FRI lag-1 serial correlation affected both metrics in the opposite fashion. As such, the system spent more time in failure and became more prone to fail, such as during periods in the 1500s and 1800s (Table 2.2).

These curves in Figure 2.2 address the critical contribution of storage requirements in the sustainability of a water supply system. Changing the reservoir capacity and holding the demand constant significantly affected the reliability, resilience, and vulnerability of the system. Increasing storage requirements always improved the system reliability and resilience and made the reservoir less vulnerable to failure. However, the influences of an incremental increase in the storage requirements of the reservoir performance in some instances could be nonlinear. For example, the incremental 25% increase in the reservoir capacity in the 1500s and 1900s (Table 2.2) resulted in nonlinear improvement in the RRV metrics. In other words, the high variance and serial correlation of the FRI in these instances caused the non-linear system responses to the storage improvement. Contrary to what was expected from the results, the resilience and vulnerability in Figure 2.2 did not always increase and decrease with an increase in the capacity of the reservoir for a given

26

target draft, e.g., both metrics of the system in the latter periods of the twentieth century did not improve relative to the increased storage requirement (Table 2.2). This was the result of the effect of averaging in the definitions of both indices as described previously (Section 2.5.2). Two aspects can explain the effect of averaging on both metrics: (a) the resilience could actually decrease if the effect of an increase in the reservoir capacity resulted in decrease in the number of continuous failure sequences without a significant decrease in the total duration of failure, and (b) the vulnerability could effectively increase if the effect of the reservoir capacity increase was to lower the number of continuous failure sequences without a dramatic decrease in the maximum volume of water shortage in each failure sequence.

In short, the inflow interannual variability and lag-1 serial correlation played important roles in the reservoir design and operation. They also represented the thresholds for determining a reliable inflow for sustaining a water resources system. To this end, the abovementioned results showed that the RRV indices were very sensitive to the temporal variation and nonstationarity that were embedded in the inflow characteristics (i.e., mean, variance, persistence) which increased the risk of failure in the system due to the hydroclimatic variations and added more challenges to the planning and management of the water resources systems.

2.6.3   Long-Term Variations of FRI and Potential Climate Linkages

In the context of understanding the long-term variations in the FRI, it is useful to explore its embedded nonstationarity over different frequencies. The wavelet power spectrum analysis provided by Torrence and Compo (1998) decompose a time series into time-frequency domain to determine the dominant modes of variability and how those modes vary in time without a prior specification of the span of the window average. The FRI time-frequency structure is examined over the instrumental records in Figure 2.3(a). It shows that the dominant frequency associated with inflow has undergone changes over the

Table 2.2: Reservoir performance metrics during the critical periods of FRI, based on storage requirements of 25%−, 50%−, and 75%−reference storage.

Period	Reliability (%)			Resilience (%)			Vulnerability (%)		
	25%	50%	75%	25%	50%	75%	25%	50%	75%
1532−1581	84	98	98	62	100	100	27	70	39
1564−1613	90	98	98	60	100	100	38	70	39
1601−1650	96	98	100	100	100	−	22	4	−
1750−1799	80	90	96	50	60	100	36	24	5
1904−1953	74	86	96	54	29	50	20	25	40
1960−2009	76	88	92	50	67	50	34	27	33

period of record, that is, two to three years during the first decade, weak interannual activity during 1915-1937, two to four years during 1938-1965, two to seven years during 1970s-2000s, and a long lower frequency signal with 10-18 year cycle in the post-1940 period. Interestingly, signals with longer periodicities are discovered in the long-term resolved FRI (Figure 2.4(a)) such as the 30-50 year cycle in the 16th and 17th centuries as well as the 50-70 year cycle during the last two centuries of the record. As such, the existence of the periodicity in the FRI records over a variety of time-scales increases the risks and uncertainties in reservoir design and operating policy, as the likelihood of causing high and low water volumes under streamflow nonstationarity is higher.

These dominant frequencies in the FRI exhibited the same behavior as the periodic large-scale climate variations (e.g., ENSO and PDO) (Figure 2.3). Therefore, it is interesting to examine the time-varying frequency-range of the climate indices such as SOI, NINO3.4, PDO to determine the potential association of the long-term variations in the FRI with the teleconnection patterns. Both the SOI and NINO3.4 historical time series (Figure 2.3(b & c)) showed cycles of two to eight years in the pre-1920 and 1940-1960 periods, and an increase in the frequency band to include longer-term variations in the post-1970 period. While the PDO time series (Figure 2.3(d)) showed two to eight year frequency during 1930-1960 period, and two to 15 years in the post-1980 period. Furthermore, the long-term signals with low frequencies are also observed in the wavelet
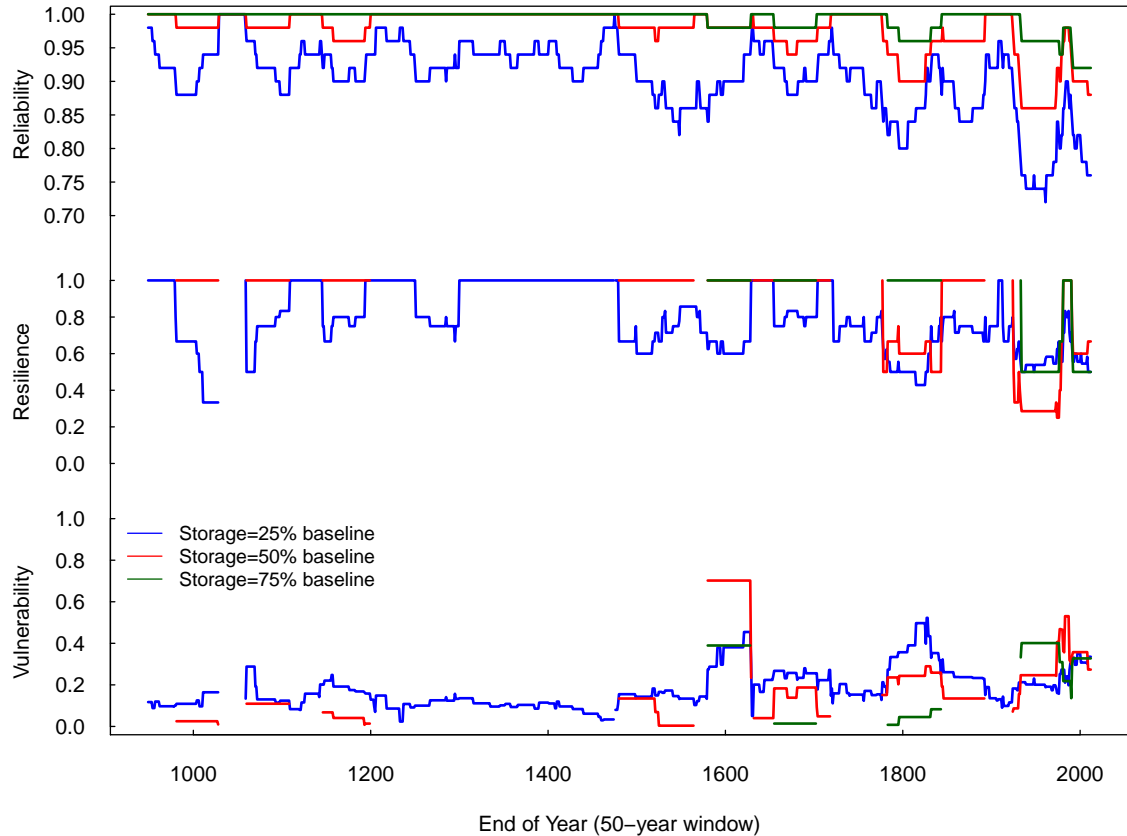
Figure 2.2: Long-term variations in the reservoir performance (900AD-2012). Reliability, resilience, and vulnerability indices are computed by 50-year moving window inflow segments for three hypothetical reservoirs with maximum capacities 0.25, 0.5, 0.75 of inflow mean baseline, and serving a water demand of 80% of inflow mean baseline. Resilience and Vulnerability are not defined when the reservoir is 100% reliable.

power spectrum for the resolved long-term reconstructed time series of the climate indices (Figure 2.4). However, these results suggest potential historical relationships between the FRI variability and the large-scale climate variations (which will be explored next).

Previous studies (e.g., Kahya and Dracup 1992; Dettinger et al. 2000; Sagarika et al. 2015) found weak correlations between northern California streamflows and the atmospheric-oceanic indices. However, using the wavelet coherence analysis between those variables introduced statistically significant associations over different time-varying frequency-ranges. The wavelet coherence spectrum may reveal interesting relationships in time-frequency space between the two time series as its values can be considered to be the

local correlation coefficient in the time-frequency domain (Grinsted et al., 2004). The coherence analysis for the observations of inflow with SOI and NINO3.4 (Figure 2.5(a & b)), respectively, reveal significant association in the two to eight year frequency band during the pre-1925 and 1940s-1960 periods of the records, and eight to 16 year cycle in the post-1970 period. However the inflow-PDO coherency (Figure 2.5(c)) did not display significant variation over the period of records. It is worth noting that coherence analysis for the long-term reconstructed data sets show significant coherence for the inflow records with the climate indices in long-term low frequency signals (Figure 2.6). For example, it displays cycles of 20-30 year and 50-70 year periodic signals in the inflow-NINO3.4 coherency while lower frequency cycle of 132 years can be seen in the coherence of inflow-PDO. Hence, having longer records overcomes the limitations of introducing low frequency multidecadal and centennial signals which can markedly improve the system design and robustness. To this end, although the large-scale climate drivers (i.e., ENSO, PDO) show linkages with the FRI during the critical periods (Table 2.1) and over different time-scales, the signal-to-noise ratio is relatively modest; nevertheless, that does not diminish the fact that the observed streamflow variability and changes are linked to proximate meteorological variables and climatic variations. In particular, it is noteworthy that historical variability in streamflow results in surprising changes in reservoir storage requirements and performance (discussed in detail in the next section).

2.6.4   The Impacts of Climate Teleconnections on The Reservoir Applications

Visual inspection of the storage requirements and the FRI variance and persistence (Figure 2.7(a)) reveals the impacts of the variation in the empirical distributions of the runoff variance and serial correlation (Figure 2.1(b & c)) on the reservoir capacity. The largest storage requirements (green triangles), which are larger than the baseline storage, are observed in the periods of positive high lag-1 serial correlation values and relatively high states of the runoff standard deviation. On the other hand, the lower storage events
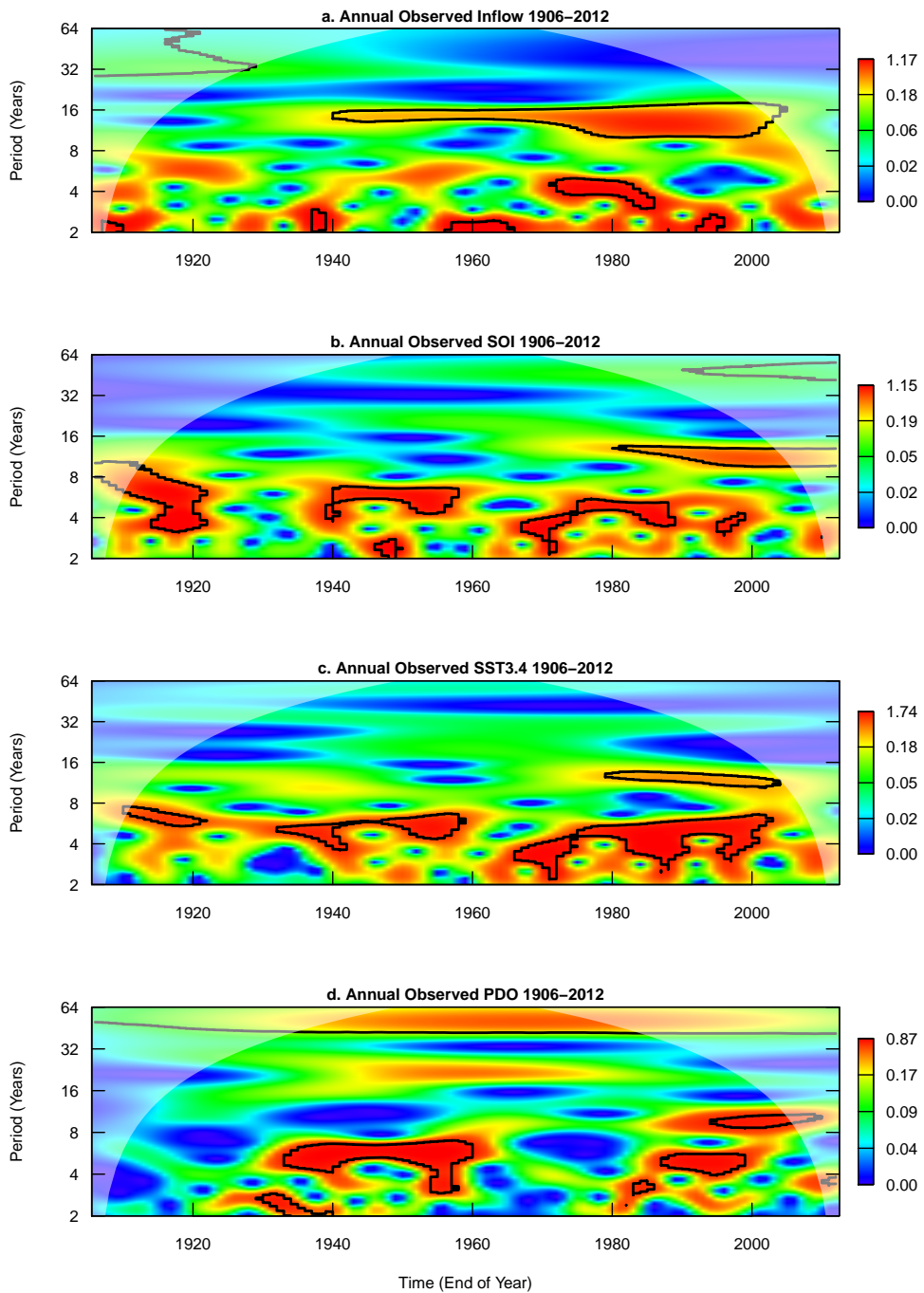
Figure 2.3: The wavelet power spectrum of the annual observed data (1906-2012). (a) the standardized time series of Feather River inflow into the Lake Oroville, (b) Southern Oscillation Index, (c) Nino 3.4 index, and (d) Pacific Decadal Oscillation. In all of abovementioned figures, black lines are the 10% significance level using the auto-regressive lag-1 model. Ghosted regions show where edge effects may become important.
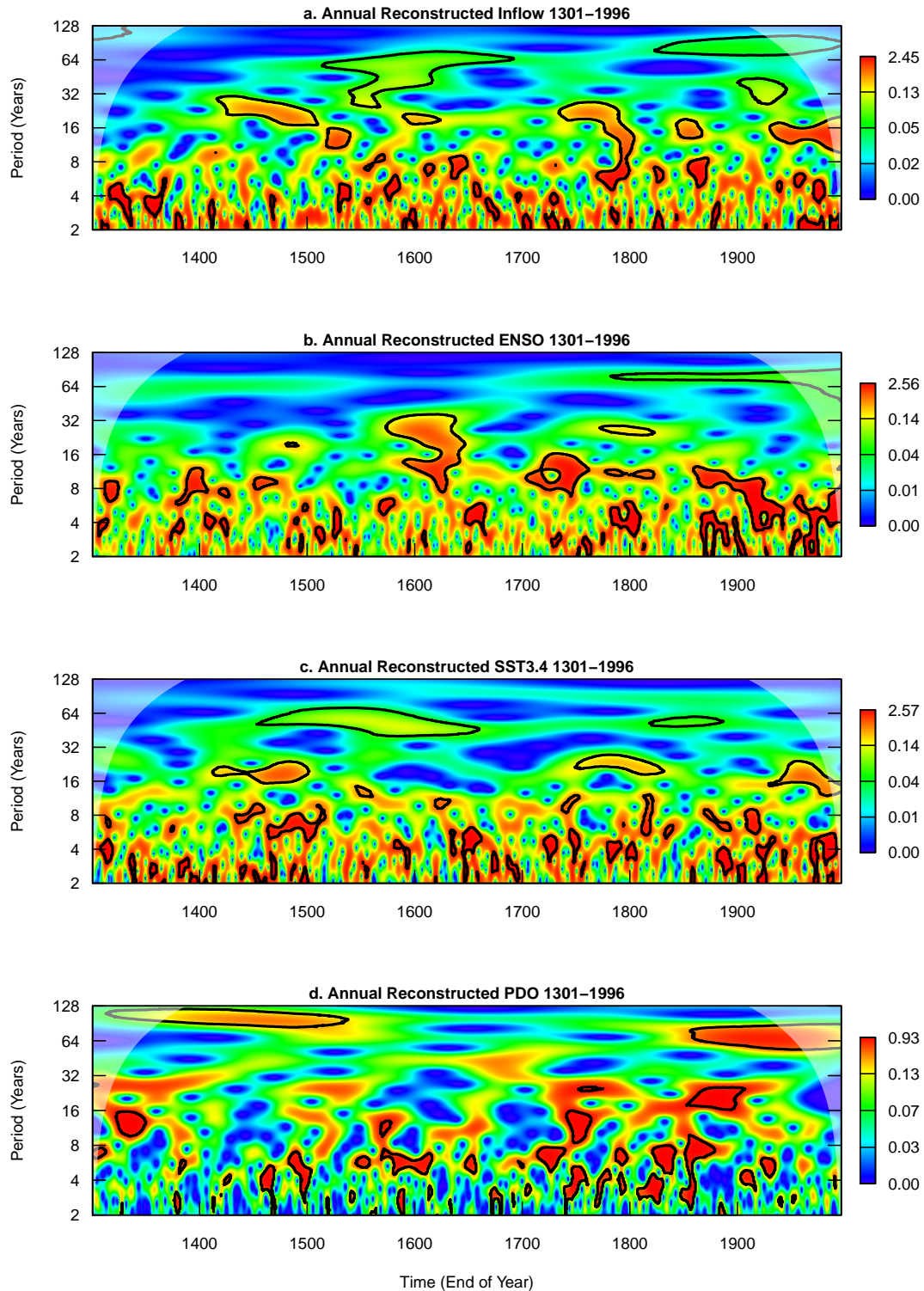
Figure 2.4: The wavelet power spectrum of the long-term reconstructed annual data 1301AD-1996. (a) the standardized time series of Feather River inflow into the Lake Oroville, (b) El Nino/Southern Oscillation Index (ENSO), (c) Nino 3.4 Index, and (d) Pacific Decadal Oscillation Index (PDO). In all of abovementioned figures, black lines are the 10% significance level using the auto-regressive lag-1 model. Ghosted regions show where edge effects may become important.
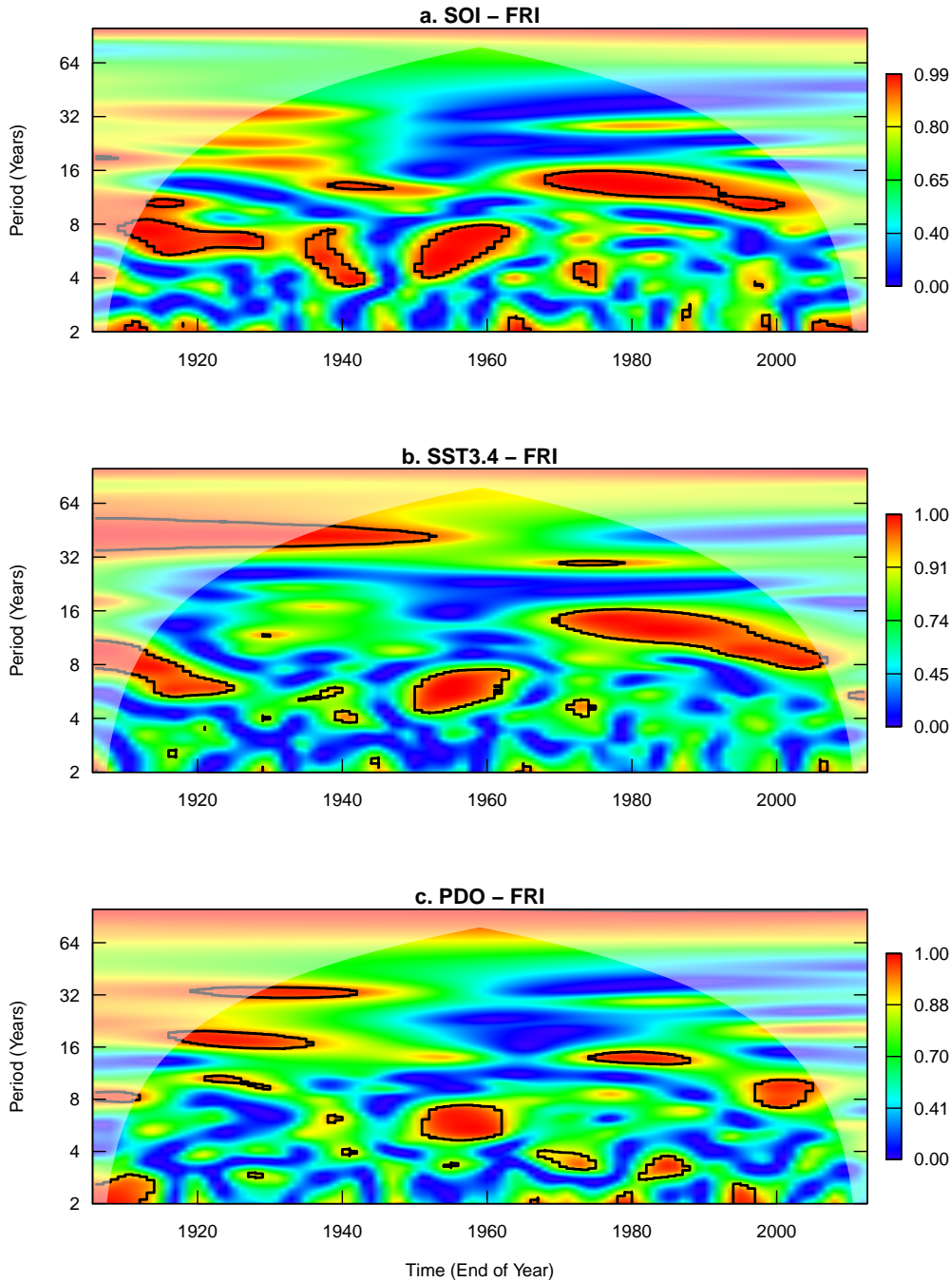
Figure 2.5: The wavelet coherence spectrum of the standardized FRI with the climate indices. (a) Southern Oscillation Index, (b) Nino 3.4 index, and (c) Pacific Decadal Oscillation (1906-2012). Contours are wavelet squared coherency. Black lines are the 10% significance level using the auto-regressive lag-1 model. Ghosted regions show where edge effects may become important.
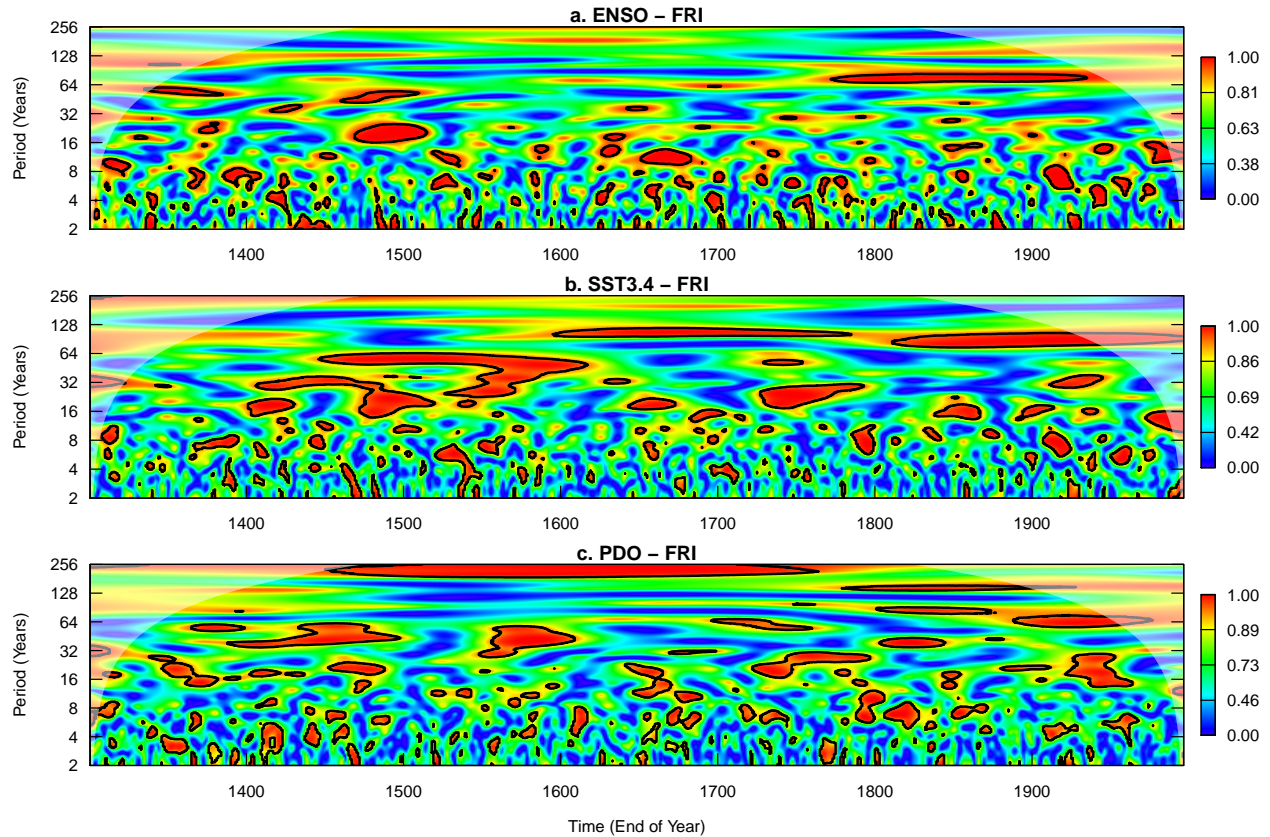
Figure 2.6: The wavelet coherence spectrum of the long-term reconstructed annual time series 1301AD-1996 between the standardized FRI and the climate indices. (a) ENSO, (b) Nino 3.4, (c) PDO. Contours are wavelet squared coherency. Black lines are the 10% significance level using the auto-regressive lag-1 model. Ghosted regions show where edge effects may become important.

(brown dots) occur during the periods of low inflow variance. As a result, having a high year-to-year variability due to a significant increase in the runoff variance/serial correlation results in abrupt changes in the storage requirements, which leads the system to failure.

The nonparametric kernel density estimation method is used to develop the relationship between storage requirements and the variability of the FRI and the characteristics of climate drivers. The red contours in Figure 2.7 show the joint probability density estimates of the mean and standard deviation of; a) FRI, b) ENSO, c) NINO3.4, and d) PDO with the storage requirements conditioned in the higher correlation values. The gray contours in Figure 2.7 represent the unconditioned relationships. It can be observed in Figure 2.7(a)

that more than 75% of the high storage requirements with positive high lag-1 serial correlation occur when the streamflow variance is more than 1.48 MAF. However, the unconditioned joint probability density estimates in Figure 2.7(a-d) reveal a weak dependency between the mean and the variance for the four variables.

To what extent did the hydroclimatic variations encoded in the mean and variance of the large-scale climate drivers conspire to produce the rich variety of fluctuations in the decision variables? The high coherence within the low frequency signals between FRI and ENSO components in 1906-1955 and 1970-2010 periods (Figure 2.5(a & b)) occurred alongside a period of high inflow variance; higher variance is consistent with abrupt increases in the storage requirements and decrease in the reservoir performance in the periods of 1904-1953 and 1960-2009. As a result, deterioration in the system performance, low reliability and resilience and high vulnerability, are observed during these periods. On the other hand, the results in Figure 2.7(b & c) show that the higher storage requirements (green dots) are associated with the high standard deviation of the both ENSO and NINO3.4 indices over the whole period of record. Furthermore, the higher correlations of streamflow with ENSO and NINO3.4 (triangles) are linked to periods of lower storage requirements. In terms of the longer-term variability, the results of the PDO (Figure 2.7(d)) show that, while the lowest storage requirements (brown dots) over the entire record are associated with the low mean of the PDO, the higher storage requirements (green dots) are clustered in periods with high mean and standard deviation of PDO. Also, the periods of higher correlations of the streamflow with PDO (triangles) are represented by moderate storage requirements. To this end, the appear to be consistent variations in the observed hydroclimatic variations and fluctuation in FRB storage requirement over the past eleven centuries. As such, for a reservoir with a fixed storage, the temporal variations in storage requirement signify changes in reliability of such systems, and likely increases in the risk of extreme events (for example, for spills).

35
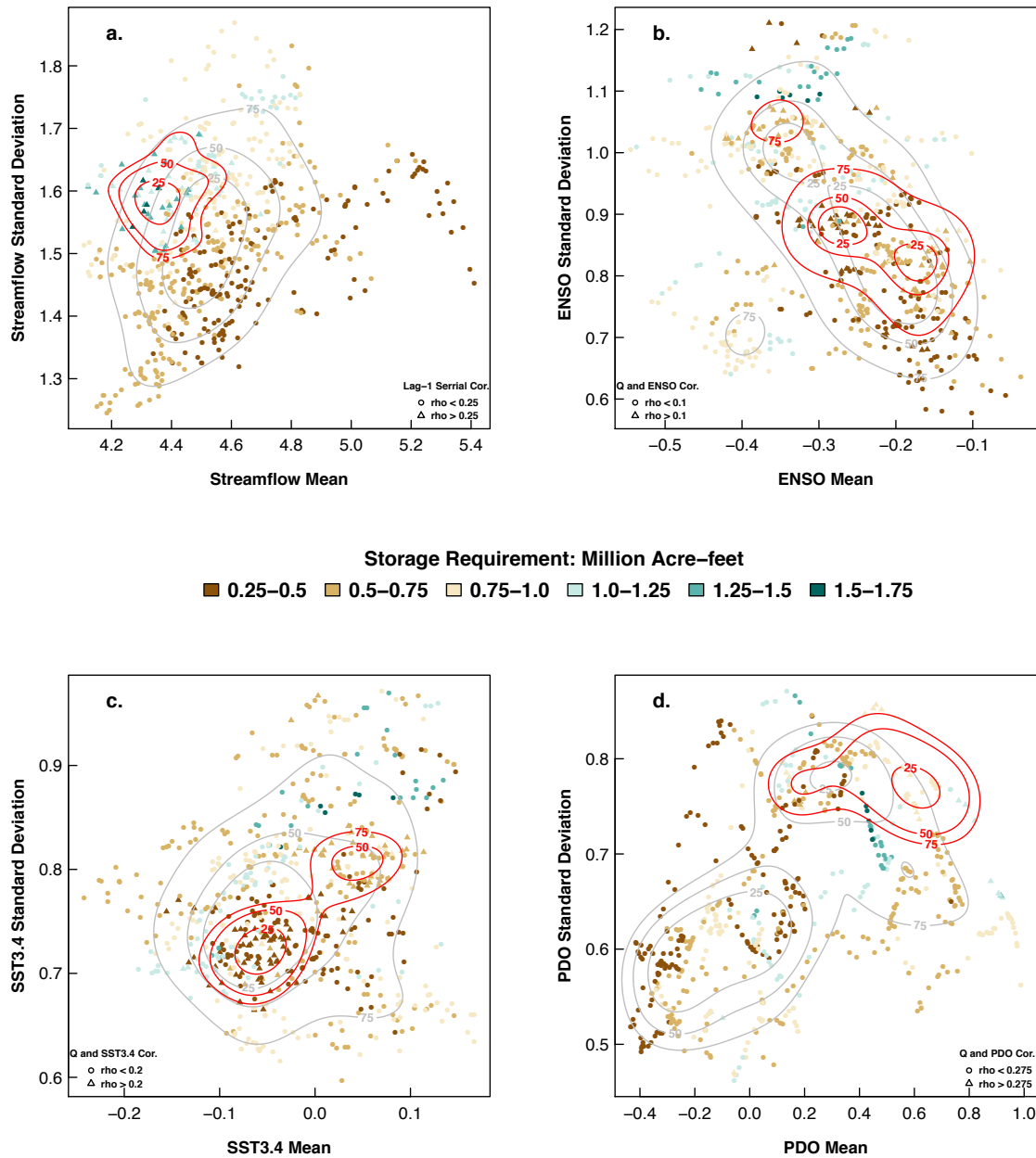
Figure 2.7: The long-term impacts of the hydroclimatic variability on the storage requirement. Color gradient refers to the storage volume. Symbols show the data points with high and low correlation coefficient as stated in the legend of figures (a, b, c, d). Contour lines: (gray) joint probability density estimation for the entire record, (red) joint probability density estimation for the data points with correlation coefficient assigned as triangles.

2.7   Summary and Conclusions

The results presented here are obtained by using the annually resolved records and the long-term reconstructed data sets of the FRI and climate indices (e.g., ENSO and PDO). The FRB hydrologic regime shows substantial variations in interannual to multidecadal and centennial time-scales over the entire record. These variations in the 50-year window average of the FRI statistics (mean, variance, persistence) indicate that the temporal variation and nonstationarity are embedded in the streamflow records. The FRI time-series demonstrate that the last three decades are a relatively wet period and have a highest variance than other periods over the past eleven centuries.

The results of the reservoir applications show that the FRI interannual variability and persistence play a key role in the system decision variables. This leads to significant changes in the system decision variables on short- to long-time scales. As such, the storage requirement during the last three decades is abruptly increased by 50% from the baseline storage. On the other hand, the results show that the storage requirement can be affected by the streamflow persistence, e.g., the 40% higher storage requirements in the late 18th century were caused by the high serial correlation values during that period. In terms of the system performance, reliability and resilience of the reservoir decrease and the system becomes more vulnerable to shortfalls during the periods of high inflow variability and persistence. Furthermore, the reservoir performance metrics (RRV) responded nonlinearly to the incremental increase in storage requirements during the periods in which variance and lag-1 serial correlation of FRI are high.

By using the wavelet power spectrum and coherence analysis, the authors demonstrate that, in many instances and periods, changes in reservoir storage requirements and performance (reliability, resilience, and vulnerability) could be readily linked to the changing co-variability between the FRI and climate teleconnection patterns. The above-mentioned results allow a qualitative assessment of the relationship between streamflow and climate indices (e.g., ENSO, PDO) which is resolved at low-frequency

bands from interannual to multidecadal and centennial. The coherence estimates shown above provide a clearer interpretation of swings in the decision variables that occur during eras of high coherence between the streamflow and climate drivers at the select timescale. While the results provide an interesting perspective regarding FRI and climatic phenomenon, it is worth noting the diagnostic studies do not imply causal relations. The moderate co-variability between climate indices and streamflow merits attention, in particular related to high frequency atmospheric phenomenon, such as the atmospheric rivers—a key moisture delivery mechanism for the US west coast.

The correlations of the FRI-climate indices and the statistical characteristics (mean, standard deviation) of the ENSO and PDO also show some indication of the systematically varying storage requirement characteristics. These results lead to the conclusion that the large storage requirements are associated with high variance of the climate drivers. On the other hand, smaller storage requirements occur in periods of high correlation between the FRI and the ENSO components. To this end, the expectations of an increase in the extreme events of ENSO flavors, which cause the extreme weather events, in the 21st century (Cai et al., 2014) may lead to changes in the storage requirements to maintain requisite reliability levels.

All in all, it is reasonable to draw the conclusion that the reconstructed hydroclimatic records lend useful insights regarding the underlying streamflow variability over different timescales, thus any historical record of shorter length will only contain or represent a fraction of the variability seen here. Therefore, the timescales that are not represented in a record of limited length are likely to be a source of system's deterioration if they occur in the future. Our ability to anticipate future hydrology and integrate that knowledge into design and planning is thus well informed by analysis of the type presented here. It is hoped that, alongside other emerging work on the topic of nonstationarity and its applications to water resources planning and management (for example, Ho et al. 2017), this work will aid

in providing a fresh perspective. Much remains to be done to clarify and adopt systematic approaches to decision-making under uncertain and changing climate conditions.

CHAPTER 3

IDENTIFYING THE SOURCES OF NATURE VARIABILITY IN THE U.S. FLOODS: THE ROLE OF MOISTURE SOURCES

3.1 Chapter Abstract

Floods often can be catastrophic and costly. Therefore, in order to assess and anticipate flood potential at a given location, reliable estimate of flood frequency and magnitude are vital to engineering design and planning. Placed-based estimation of flood requires knowledge of the hydrologic process and carefully-designed statistical estimation methodologies, such as the Bulletin 17-C (England et al., 2018). The assumption of stationarity is central to statistical methodologies for flood estimation. However, hydroclimatic variability and change may significantly impact flood statistics, and induce trends and long-term variations. To better understand if the observed trend stem from climatic variability and or are artifacts of random weather-climate fluctuations, an important first step is the characterization of floods within context of hydroclimate. The current study introduces a comprehensive framework that aims to: a) identify major causes of annual floods using a hydroclimate-based classification, b) determine the major tracks and sources of flood-caused moisture by employing the principal curve analysis to develop a new approach, and c) adopts a new modality for risk analyses and estimation of hydroclimatic nonstationarity by climate-informed partitioning of the flood record. The results show that, in 623 reference USGS stream gauges across the conterminous United States over the 1956-2015 period, a total of 25,725 ($\sim$ 73%) out of 37,380 annual floods are caused by atmospheric rivers (AR). As such, most basins in the Northwest, West, Northeast, and Southeast climate regions observed 70-100% of their annual flood records caused by ARs while it is widely varying below 70% in the Central, South, East North Central and 50% in the Southwest. However, ARs cause less 15% of annual floods in the

West North Central. In terms of magnitude, station with heterogeneous floods and highly affected by ARs have the higher ratios of 100-year flood of AR-generated floods to Non-AR floods. Finally, characterization of place-based mixed population floods by ARs moisture sources provide satisfactory fit of flood frequency curves and determine sources associated with high risk floods. These results explain the nature of variability in flooding risks to provide reliable estimates of flood frequency.

3.2 Introduction

Nonstationarity in flood events magnitude and timing due to the changing hydrologic regime under climate variation is a key role in water resources management. A reliable estimation of flood frequency and magnitude to anticipate future extremes is vital for life security and property protection. Thus, precise knowledge about the nature of floods variability will contribute to improved strategies for future water resources management and flood risk mitigation. After 25 year of publishing the B17-B "Guidelines for determining flood flow frequency: Hydrology Subcommittee Bulletin 17B" framework by the Interagency Advisory Committee on Water Data, in March 1982, Stedinger and Griffis 2008 recommended the use of the Expected Moment Algorithm (EMA) to address the limitations of the adapted statistical framework in B17-B. Later, B17-C "Guidelines for determining flood flow frequency—Bulletin 17C" England et al. (2018), the updated version of B17-B, improved the statistical framework and enabled using longer data by using the EMA to deal with uncertainties in historical information, zero and low floods, interval data, and confidence limits. To this end, B17-C is still limited in addressing the floods variability as the committee group listed this objective in the framework's future studies.

Numerous studies in the literature of the dynamic risk of floods have attempted to modulate the changes in floods risks. For example, modification of the mathematical model of the Log-Pearson Type III distribution (LP3) to account for co-variate (e.g., time, SST, etc.) has been used to improve flood risk estimation (Jain and Lall, 2000; Kashelikar and

Griffis, 2008; Stedinger and Griffis, 2011). Conversely, Aljoda and Jain 2020 showed moderate co-variability between climate indices and streamflow merits attention, in particular related to high frequency atmospheric phenomenon, such as atmospheric rivers — a key moisture delivery mechanism for the US west coast.

On the other hand, flood frequency analysis in different regions of the conterminous United States (US) often contains annual floods generated by distinctive different hydrologic and hydroclimatic processes (Waylen and Woo, 1982; Hirschboeck, 1991; Webb and Betancourt, 1990; Berghuijs et al., 2016). As such, a fundamental assumption, i.e. iid, which considers floods timeseries as time sample of random homogeneous events (as in B-17B and B-17C) does not apply in regions of mixed population flood events. Among the different flood-generating mechanisms, atmospheric rivers (AR) are responsible for large, regional-scale floods (Ralph et al., 2006; Dettinger, 2011; Neiman et al., 2011; Lavers and Villarini, 2013; Barth et al., 2017).

Recent changes in the seasonality and magnitude of extreme precipitation in North America critically impact societal vulnerability, infrastructure design, as well as planning and adaptation. An important scientific task is understanding the recent changes in extreme events statistics (frequency, and magnitude) and linking these changes to the large-scale atmospheric phenomena (i.e. ARs). Doing so will inform science-based decision tools for early-warning, as well as develop new modalities in flood frequency estimation in a changing climate. This study examined the role of ARs and their moisture trajectories and sources on the heterogeneous flooding events of annual maximal flows (AMF) in the US, and their impact on the magnitude and frequency estimates used for the design of flood structures. Three research foci related to the nature and variability of annual maximum flood events across US are:

1. To develop a methodology for robust curvilinear estimation of AR events based on a Principal Curve-based approach, thus integrating the axes of maximal IVT variation and time progression over the event lifecycle.

2. To characterize the place-based mixed population of AMF by delineating the constituent remote atmospheric moisture sources, thus enabling climate-informed partitioning of the flood record for risk analyses and estimation of hydroclimatic nonstationarity.

3. To systematically apply the new methodology towards a US-scale assessment of annual maximum floods to understand their regionality, and with a goal to develop new modalities in flood frequency estimation in a changing climate.

A comprehensive statistical framework with three phases of analysis is introduced to achieve these goals. Phase I, attributes the AMF records to their generating mechanisms and shows the major processes. Phase II, identify the moisture tracks and sources and investigate their impacts on flooding frequency and magnitude. Phase III, quantify the nature of flood variability based on moisture sources variation.

The remainder of this chapter is organized as follows. First, a brief background shows the original contribution of this work to the literature of flood frequency analysis. Second, the data section. Third, explain the three phases of analysis in more details in the methods section. This is followed by a complete review for the results of this research. Finally, the section of discussion and conclusions.

3.3   Background

During the last decade of the 20th century, developments in the understanding of the atmospheric moisture pathways led to the recognition of the large-scale moisture delivery as a precursor to floods (Hirschboeck, 1991). After ARs were introduced in the literature by Zhu and Newell 1998, scientists studied the links between ARs, precipitation and floods Ralph et al., 2006; Neiman et al., 2008. Since 2010, studies have started to look at the variability in floods from a hydroclimatic and metrologic perspectives (e.g., Dettinger, 2011; Lavers et al., 2011, 2012; Lavers and Villarini, 2013; Berghuijs et al., 2016; Lu and Lall, 2016; Barth et al., 2017; Konrad and Dettinger, 2017; Dickinson et al., 2019; Brunner

et al., 2020) to address the variability of flooding risks. However, the nature of flood variability in these studies and others in the literature has not well explained due to limitations in approaches, and incomplete assumptions. To this end, this work introduces a unique methodology to explain the nature of flood variability compared to recent studies on flood risk in the literature as summarized in Table **??**.

3.4   Data

3.4.1   Streamflow Records

Streamflow data are obtained from the U.S. Geological Survey (USGS) National Water Inventory System (NWIS, waterdata.usgs.gov) database for a 60-year period for water years (October 1 through September 30) 1956 to 2015. This work selected reference sites from the GAGES II (Falcone, 2011), which is a subset of USGS streamflow gauges that have no anthropogenic interference (e.g. effects of dams, diversions, water withdrawals etc.). To examine the role of ARs and their trajectories and sources on the flood frequency and magnitude in the conterminous United States (US), 623 gauges from the GAGES II dataset are selected using criteria for completeness of the average daily flow record at each stream gauge. These gauges are well spatially distributed across the study region (Figure 3.1). The annual maximum flow (AMF), the largest daily streamflow during a water year, for all stations over the 60-year records is determined within 20% or less of total number of days with missing data. The stream gauge with more than 13 missing AMF is dropped from the subset. The percentage of the missing AMF in the selected 623 gauges over the whole record is 5.5%. The study period is limited to 1956-2015 based on the length of the climate data. The start date of the streamflow record is optimized to get enough number of stations that well spatially distributed across the study region with the most possible length of record. The majority of watersheds' area are between 100-2500 km^2 and catchments along the US coasts are relatively small compared to those located in the interior part of the country (Figure 3.1).

Table 3.1: Comparison between the abilities of the current study and the solid studies in the literature of flooding risks to explain the nature of floods variability.

Study	Study region & time period	Hydrologic variable	AR-event related characterization	AR episode	Moisture sources	Extremes analysis modality
Ralph et al. (2006)	• Russian river, northern CA, USA. • 10/01/1997 -02/28/2006.	Daily mean streamflow exceeded the monitor-stage flood threshold.	Single point-based index.	None	None	Investigate the ARs impact on flood by monitoring both of water vapor and rain magnitudes in coastal mountain near the Russian River.
Dettinger (2011)	• California, USA. • Periods of 20 years in the 20 & 21 centuries.	Daily water vapor, winds, and temperature available in IPCC model.	None	None	None	Study the ARs duration, intensity, frequency, and seasonality under changing climate projections.
Lavers and Villarini (2013)	• Central US. • 1979-2011.	USGS NWIS: Flood Data (AMF; $n = 1105$).	Single point-based index.	Max IVT per time-step to form the major axis of a partial life-cycle AR.	Limited to identify moisture sources.	Investigate ARs role in floods by calculating IVT and using metrological variables.

Table 3.1: Continued.

Study	Study region & time period	Hydrologic variable	AR-event related characterization	AR episode	Moisture sources	Extremes analysis modality
Lu and Lall (2016)	• Northeastern U.S. • 1989-2010.	Flood events by the Dartmouth Flood Observatory.	Single point-based index	Time-step point-based moisture trajectories	Pre-defined sources: • Pineapple Express. • Great Plains. • Gulf Stream. • West Pacific.	Examine the relationship between floods and the tropical moisture exports based on the moisture sources.
Barth et al. (2017)	• Western US. • 1900s-2010.	USGS NWIS: Flood Data (Instantaneous peak flow; $n = 1375$).	Single point-based index.	Max IVT per time-step to form the major axis of a partial life-cycle AR.	Limited to identify moisture sources.	Performed the EMA-MGBT algorithm on the mixed and homogenous populations to show the role of ARs in the floods frequency and magnitude.

Table 3.1: Continued.

Study	Study region & time period	Hydrologic variable	AR-event related characterization	AR episode	Moisture sources	Extremes analysis modality
Barth et al. (2019)	• Western US. • 1900s–2010	USGS NWIS: Flood Data (Instantaneous peak flow; $n = 43$).	Single point-based index.	Max IVT per time-step to form the major axis of a partial life-cycle AR.	Limited to identify moisture sources.	Applied the weighted mixed population approach to quantify the floods frequency of a AR/Non-AR annual peak flows.
Schlef et al. (2019)	• US, AK, HI, PR. • 1874–2014 (Water year: 1 Oct–30 Sep).	USGS NWIS: Flood Data (AMF, POT; $n = 681$)	Use self-organizing maps to identify atmospheric circulation patterns associated with floods.	Limited to track the atmospheric phenomena in temporal and spatial scales.	• South and north pineapple express in the west. • Great Plain and Gulf of Mexico in the central. • Gulf of Mexico and Atlantic moisture for east.	Assess the flood characteristics (e.g., frequency, spatial domain, event size, and seasonality) specific to each circulation pattern.

Table 3.1: Continued.

Study	Study region & time period	Hydrologic variable	AR-event related characterization	AR episode	Moisture sources	Extremes analysis modality
Dougherty and Rasmussen (2019)	• Conterminous US. • 2002-2013.	Comprehensive US floods database.	Merge flood reports grouped by causative meteorological event with stream gauge-indicated floods database.	Moisture pathways not utilized.	Moisture sources not identified.	Summarized the seasonal and spatial distribution of the flash−, slow−rising−, and hybrid−flood.
Current study	• Conterminous U.S. • 1956−2015 (Water year: 1 Oct−30 Sep).	USGS NWIS: Flood Data (AMF; $n = 623$)	Watershed-based index.	Lagrangien integrated episodic AR trajectory.	• Pacific* (TCPO, TEPO, sub-TPO, extra-TPO). • Atlantic* (TAO, sub-TAO, extra-TAO). • Caribbean Sea and Gulf of Mexico in the central. • Local.	Develop a climate-based flood frequency estimation by delineating the moisture trajectories and sources for the AMF-AR event.

* = Ocean

3.4.2 Atmospheric Data

The shape index of AR in the NCEP/NCAR reanalysis product provided by the NASA Jet Propulsion Laboratory (NASA JPL) (Guan and Waliser, 2015) was used to determine which ARs are associated with floods events (AMF records). The shape index is a unique number that is given to each observed AR at each time step (6 hours) to distinguish between several ARs available over the globe. This data set is available in a 6-hourly global scale with $2.5° \times 2.5°$ resolution for the period 1948-2015. The NCEP/NCAR $2.5° \times 2.5°$ 6-hourly integrated water vapor transport (IVT) time series (1948-2017) by Rutz et al. (2014) is used to eliminate the grid points of AR with IVT<250 kg/m/s from the determined polygon of AR shape. Finally. to determine the flood generating mechanism of each event we obtain daily timeseries of the surface air temperature (SAT), precipitation rate (PR), and water equivalent of accumulated snow depth (WEASD) from the NCEP/NCAR reanalysis data provided by the National Oceanic and Atmospheric Administration (NOAA) Physical Sciences Laboratory (PSL) (https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.html). These variables are available in global scale with different resolutions for the period 1984-2020.

Table 3.2: The US hydrologic regions (HUC).

No.	Name	HUC	No.	Name	HUC
1	New England	01	10	Missouri	10
2	Mid−Atlantic	02	11	Arkansas−White-Red	11
3	South Atlantic-Gulf	03	12	Texas−Gulf	12
4	Great Lakes	04	13	Rio Grande	13
5	Ohio	05	14	Upper Colorado	14
6	Tennessee	06	15	Lower Colorado	15
7	Upper Mississippi	07	16	Great Basin	16
8	Lower Mississippi	08	17	Pacific Northwest	17
9	Souris−Red-Rainy	09	18	California	18

Figure 3.1: Map of the study region. The colors represent the US climate regions, the red dots are the locations of 623 USGS reference streamflow gauge stations distributed across the conterminous US, gray line is the interior borders of the states, the black line represents the borders of the hydrologic unit regions (see Table 3.2 for regions' names). N_s is the number of stations in the climate region. Bar charts are for the drainage areas within each climate region.

3.5 Methods

3.5.1 Climate Regions Representatives

As the scope of this research work is to study the role of ARs in flood frequency and magnitude across the US by performing at-site hydroclimatic analysis, it is impractical to show numerous plots that demonstrate the same results for different locations within the study area. Therefore, the study area is divided into regions to generalize the facts to regional perspective. However, we select regional representative basins in the regional analysis to avoid the visual complexity of showing all the regional results in a single plot. Two of the common methods to divide the US into sub-regions are the hydrologic regions and climate regions. The USGS divided the conterminous US into 18 geographical regions based on the drainage area of a major river such as Missouri region or series of rivers such as Texas Gulf region (see Figure 3.1 and Table 3.2 for regions locations and names). The second method of division is the US climate regions. Karl and Koss (1984) performed temporal and spatial climatological analyses on statewide temperature and precipitation data to subdivide the coterminous US into nine climate regions (Figure 3.1). Consequently, scientists in NOAA - National Centers for Environmental Information have identified nine climatically consistent regions within the contiguous US (https://www.ncdc.noaa.gov/monitoring-references/maps/us-climate-regions.php) and used them for climate monitoring and providing data for the public services and private sectors. Since the main focus of this study is the impact of the large−scale atmospheric patterns on floods, we select the climate regions division to perform the analysis.

The study criteria to select the representative station of each climate region are that stream gauge should has no missing AMF over the 60-year record and, it is most correlated with the other stations within the region. The methodology may not choose the most correlated station in some cases to fulfill the record completeness condition. As shown in Figure 3.1 and Figure 3.2, each climate region contains a group of contiguous whole states. Each of the 623 stations is assigned to a climate region based on the geographical location.

The drainage basin is considered to belong to the climate region based on the geographical location of the stream gauge, which may lead to have a basin that most of its area in a specific region to be in different region where its outlet is located. Then, the 60-year records of AMF of each station is correlated with all others in the same region to find the station with the possible higher median of correlation coefficient values and complete record. As a result, the missing AMF records in each climate region is less than 10% of the total numbers of annual record. It is noted that all of the representative stations have complete records. Figure 3.2 shows the location of the representative station of each climate region and Table 3.3 tabulates the details of each representative station. Histograms in Figure 3.2

Table 3.3: Streamflow Network Information.

No.	Station ID	Name	Climate Region
1	05501000	North River at Palmyra, MO	Central
2	04056500	Manistique River near Manistique, MI	East North Central
3	01350000	Schoharie Creek at Prattsville, NY	Northeast
4	14222500	East Fork Lewis River near Heisson, WA	Northwest
5	08164000	Lavaca River near Edna, TX	South
6	02059500	Goose Creek near Huddleston, VA	Southeast
7	09081600	Crystal River ABV Avalanche C, near Redstone, CO	Southwest
8	10343500	Sagehen C. near Truckee, CA	West
9	06360500	Moreau River near Whitehorse, SD	West North Central

HUC	Area (km^2)	Longitude	Latitude	Elevation (m)	Correlations Median
07	922.8	9132'45.7"W	3949'01.2"N	144.384	0.21
04	2945.9	8609'40"W	4601'50"N	185.459	0.30
02	612.5	7426'12"W	4219'10"N	344.768	0.39
17	323.9	12227'54"W	4550'13"N	108.753	0.46
12	2124.0	9641'10"W	2857'35"N	4.298	0.32
03	485.4	7931'14"W	3710'23"N	180.719	0.41
14	432.9	10713'39"W	3913'57.5"N	2104.644	0.55
16	27.6	12014'13"W	3925'54"N	1926.336	0.62
10	12655.0	10050'33"W	4515'21"N	506.419	0.44

show that the majority of stations in each climate region are moderately correlated and have positive median of correlation coefficient values less than 0.5. The probability density function (PDF) curves (red curves in Figure 3.2) show that the representative station is
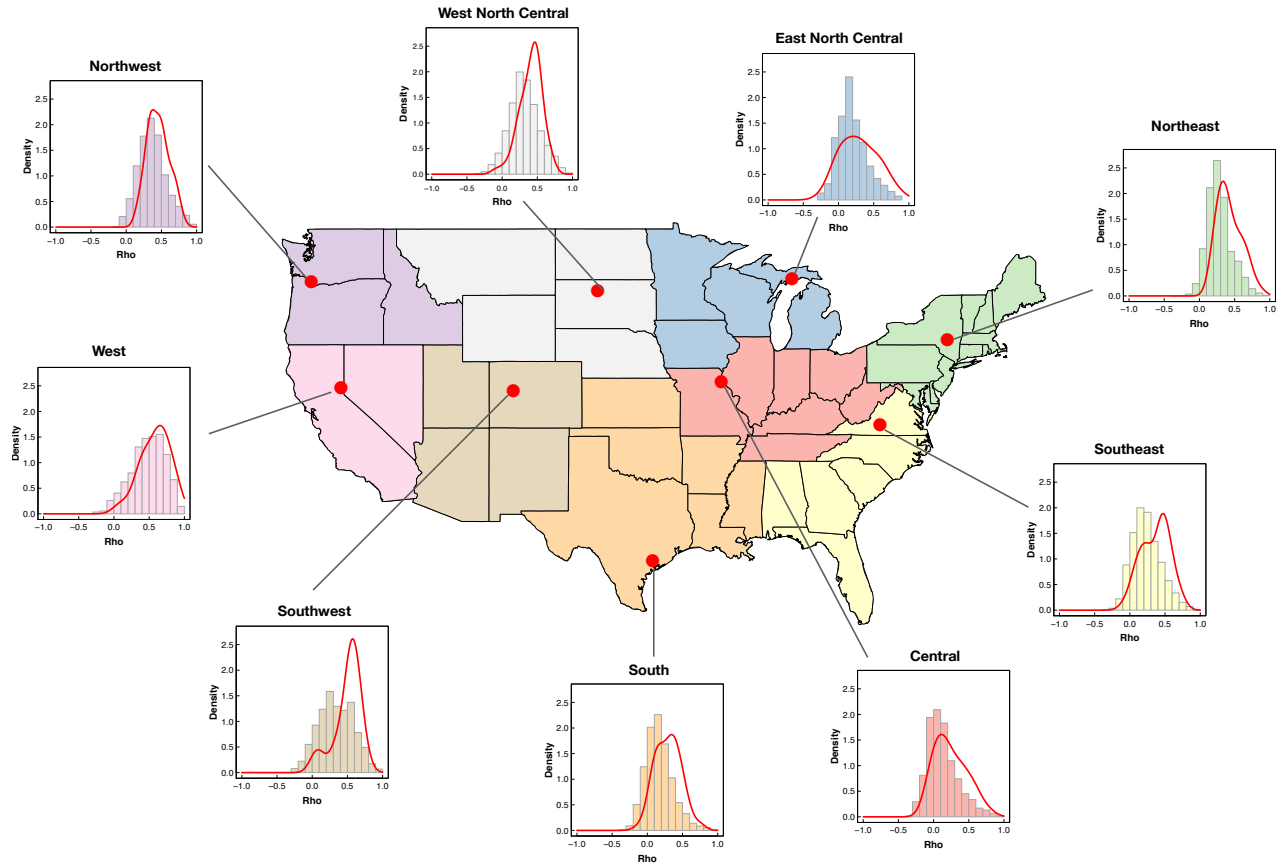
Figure 3.2: The locations of the representative streamflow gauge stations for the US climate regions. Colors represent the US climate regions. Red dots refer to the locations of the representative stations (see Table 3.3 for details). Red lines represent the PDFs for the coefficients of correlations between the AMFs of the representative and the other stations in the same region. Histograms are for the coefficients of correlations of each station with others within the region.

most correlated with the other stations in the region as the higher density peaks of PDFs are shifted toward higher values of the correlation coefficient.

3.5.2   The Flood Events Separation

The analysis in this chapter is divided into three separated phases. Phase I of the analysis is to determine the flood mechanism of each AMF event. The next phase of the analysis is to determine the moisture major axis (trajectories) and sources for a group of events identified in Phase I and, apply the flood frequency analysis on these events in the third phase analysis (see section 3.5.3 for details on Phase II and Phase III analyses).

The Phase I analysis, as shown in the flowchart of Figure 3.3, starts by selecting the watersheds and determining the AMF for the 623 stream gauges (section 3.4.1) and, assigning the grid point(s) from each atmospheric variable (section 3.4.2) that is/are included within the watershed's boundaries. An intersection between the latitudes and longitudes of a watershed polygon and a layer of the grid data results in which grid point(s) is/are located within and/or on the drainage basin boundaries. As a result, daily timeseries for each atmospheric variables in each watershed are constructed by averaging or taking the maximum of the daily values for the assigned grid points. The analysis considers a specific number of days before and after the AMF date to account for the effects of the atmospheric variables in the flood generating mechanism. The AR shape and IVT timeseries are examined one day before and one day after the AMF date to detect whether is there an AR event with IVT$\geq$250 kg/m/s stalling over the watershed or not. Other variables such as the IVT alone and PR are examined at three before and one day after the AMF date, while the SAT and WEASD are examined at five days before and one day after the AMF date. Then, the cause of each AMF event is assessed based on the stated conditions in the last step of the Phase I analysis in the flowchart (Figure 3.3), and assigned to a specific category. As such, ARs with more than 1 mm rainfall occur within 5 days of the AMF (1-day after, to consider the time difference, and 3-days before) over a watershed are considered the cause of that AMF event. For instance, if all the examined variables for any AMF in a selected station met their specified thresholds (here are: AR Shape$\neq 0$, $IVT \geq 250$ kg/m/s, PR>1 mm, SAT≥ 0 °C, WEASD>1), then, the generating process of that AMF event will be the category of AR-rain over the snow. To this end, five categories of the flood-generating mechanisms are resulted with one additional category to contain all the events that not belonged to the five main categories. Figure 3.4 shows that majority of the AMF events across the conterminous US are controlled by the flood-generating mechanisms: AR-rain on snow and AR-rain only. Therefore, this study will focus on the role of ARs in the flood frequency and magnitude.

3.5.3   The Detection of Moisture Pathways and Sources

Two categories from the previous analysis are merged into one group to perform the Phase II analysis; the AMFs caused by AR-rain on snow and AR-rain only. The AR shape and IVT timeseries are examined at a 6-hour time-scale three days before and one day after the AMF event as shown in the Phase II analysis in Figure 3.5. All intersected ARs with the watershed within those five days are stored to be tracked back to 10 days. Intersection between two successive stored ARs is applied to identify if they are continuous AR or separated ARs. A lack of a 6-hour period between the stored ARs indicates they are different events. To determine the AR trajectories (moisture major axis) we store all the connected 6-hour stages of a complete AR event (from its birth until it disappears). Then, they are combined into one element and the repeated grid points with the smaller IVT values and those with IVT<250 kg/m/s are eliminated. The weighted principal curve (WPC) approach is applied on the element to produce the AR trajectory. The principal curve (Hastie and Stuetzle, 1989) is a nonparametric nonlinear one-dimensional smooth curve that is defined as each point of the curve is the average of the observations projecting there, i.e., for which that point is the closest point on the curve. Therefore, this study adds weight to the principal curve by replicating each point in the element proportional to its IVT value. The weight or the number of times to replicate the point is calculated with the following equation:

$$W = \frac{IVT}{IVT_{thr}} \times 10 \ . \tag{3.1}$$

where W is the weight, IVT_{thr} is the threshold of IVT for identifying AR. The IVT_{thr} is 250 kg/m/s in this study. The WPC approach has the ability to draw a one integrated pathway for the moisture movement of a complete AR event. It is worth noting that up to the time of writing this chapter there have been no studies in the existing literature of ARs and their hydrological impacts were able to determine the trajectories and sources of the ARs causing the AMFs. Few studies (e.g., Lavers et al. 2011, 2012; Lavers and Villarini
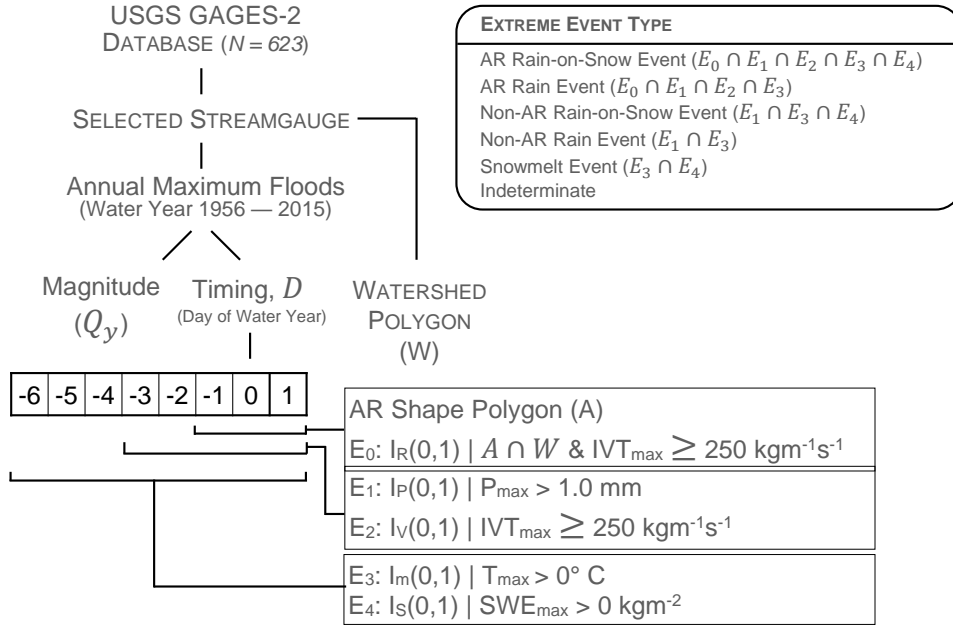
Figure 3.3: Phase I Analysis: Flood-generating Mechanism.

2013; Barth et al. 2017 have identified the major axis of ARs causing floods as the grid point with the higher IVT value within each time step, but they are limited to determine the AR trajectory and the source of origin.

The other part of the Phase II analysis is about determining the AR source of origin (Figure 3.5). Only the last 6-hour stage of a complete AR (initial AR stage) is stored for all events contributing to a single AMF event. To determine the source of origin for each AR, we determine the trajectory of the AR initial stage by applying the WPC approach. In this case, the IVT threshold will be the minimum value in that 6-hour time-step. The weight (W) or the number of replicated points is determined as:

$$W = \frac{IVT}{IVT_{min}} \times 10 \; . \tag{3.2}$$

Then, we intersect the resulting track with the nine sources of moisture. These sources are: (20S-20N) (210W-280W) Tropical Eastern Pacific Ocean (TEPO), (20S-20N) (160E-210W) Tropical Central Pacific Ocean (TCPO), (20N-40N) Sub-tropical Pacific

a) AR with rain on snow **b) AR with rain** **c) Rain only**

d) Snowmelt **e) Rain on snow** **f) Unknown process**

Percentage of AMF Events (WY: 1956–2015)

☐ [0,1) ☐ [1,20) ☐ [20,40) ■ [40,60) ■ [60,80) ■ [80,100]
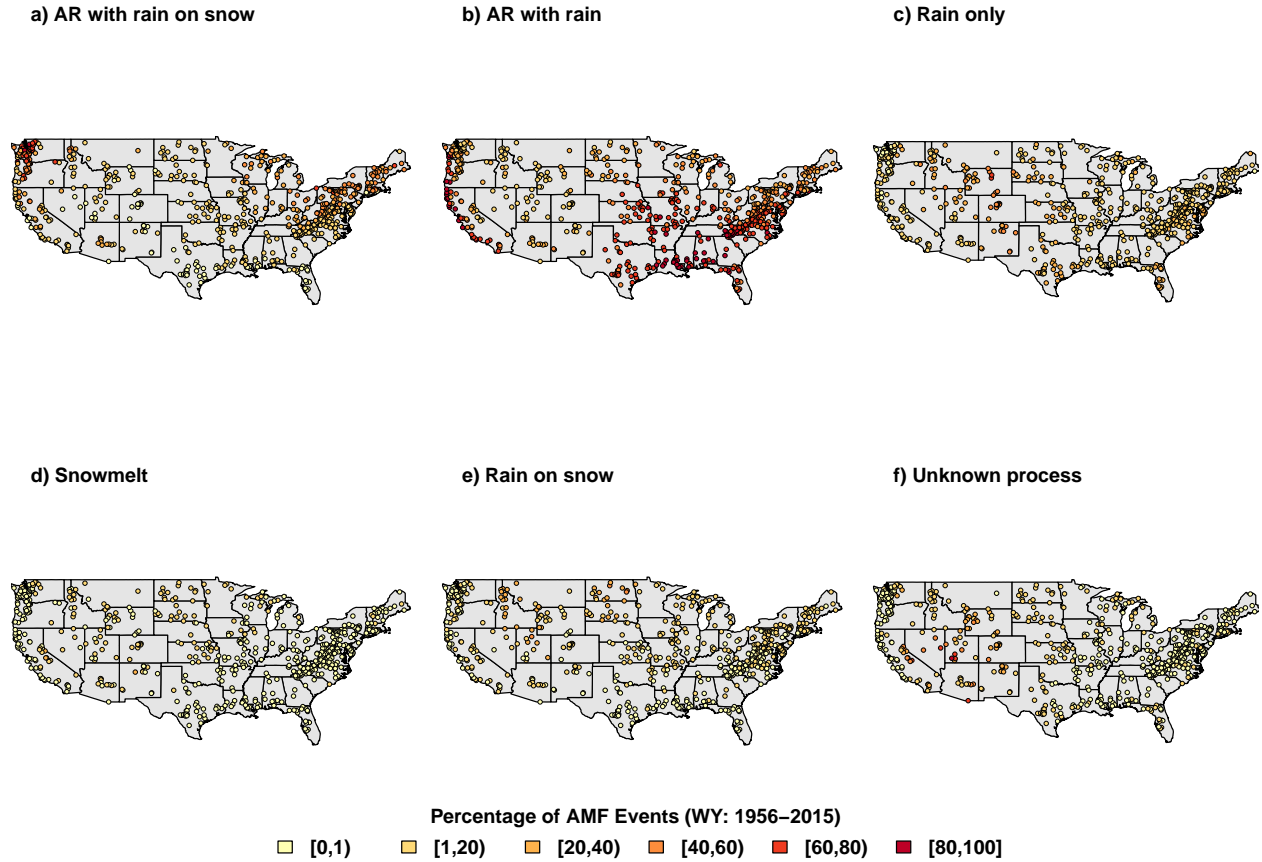
Figure 3.4: The percentage of events for each AMF-generating mechanism (water year 1956-2015).

Ocean (Sub-TPO), (>40N) Extra-tropical Pacific Ocean (Extra-TPO), (20S-20N) Tropical Atlantic Ocean (TAO), (20N-40N) Sub-tropical Atlantic Ocean (Sub-TAO), (>40N) Extra-tropical Atlantic Ocean (Extra-TAO), Caribbean Sea (CS), and Gulf of Mexico (GM). In the case that the AR initial stage trajectory is intersected with more than one source, the source with the lower latitude is considered to be the event origin. If there is no intersected source with the track, the AR is considered to have originated over the interior water bodies (e.g., lakes) which considers as a local source in this study. It is worth noting that a number of ARs with different trajectories and sources of moisture may contribute to produce a single AMF event within a few days. Figure 3.6(a) shows that the AMF-AR event at the North River in Missouri in May 11, 2003 (see Figure 3.2 & Table 3.3 for more information about the station) was caused by five different AR events. The sum of their

```
┌──────────────────────┐
│  AMF-AR Date, D      │──┐
└──────────────────────┘  │
                          │   ┌───────────────┐   ┌───────────────┐
┌──────────────────────┐  │   │ Intersected   │   │ Track back IP │
│ Watershed Polygon, W │──┼──▶│ Polygon       │──▶│ each 6-hour   │
└──────────────────────┘  │   │ IP = A ∩ W    │   │ for up to     │
                          │   └───────────────┘   │ 10 days       │
┌──────────────────────┐  │                       └───────────────┘
│ 6–hour AR Shape      │──┘
│ Polygon, A           │
│ Day=-3:D:+1          │
│ A≠ 0 & P > 1 mm &    │
│ IVT ≥ 250 kg/m/s     │
└──────────────────────┘
```

Combine all 6-hour polygons of the same AR and chose only points with IVT ≥ 250 kg/m/s → **Perform the weighted principle curve method on the location points of the combine data set.**
$$W = \frac{IVT}{250} \times 10$$

Store the polygon of the last 6-hour stage of AR → **Perform the weighted principle curve method on the location points of the combine data set.**
$$W = \frac{IVT}{MinIVT} \times 10$$
→ **Intersect the curve with the moisture sources polygons and chose the source with lower latitude**

Figure 3.5: Phase II Analysis: The detection of moisture trajectories and sources of AMF-induced ARs.

effects were reflected on the AMF day as the streamflow discharge had a significant spike (Figure 3.6(b)). Furthermore, the algorithm of m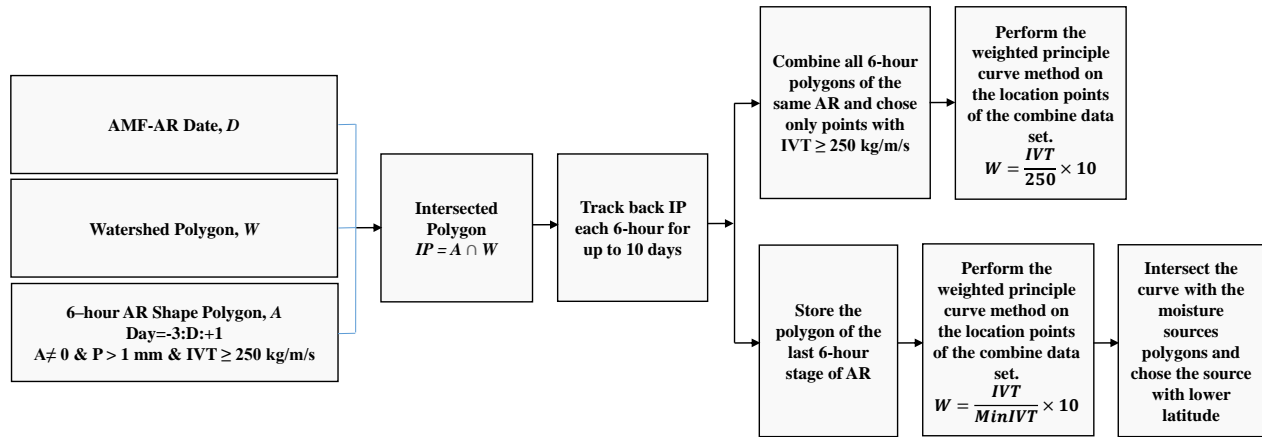ethodology has the ability to track the impact of each AR on the magnitude of the river discharge within selected period before and after the day of AMF event as shown in Figure 3.6(c-g). However, one challenge in the algorithm is to plot a continuous curve for an AR trajectory with a gap between two successive points. As the methodology of determining ARs trajectories in this study is conditioned by eliminating all the grid point of the AR shape with IVT<250 kg/m/s, a threshold of 500 km distance between two successive points is considered as a solution. Nevertheless, sometimes this distance is larger than the specified threshold which results in a shorter trajectory curve. On the other hand, selecting a high threshold may result in merge two different tracks. Furthermore, a merge of separate AR events in the AR shape reanalysis data due to the cell large size affects the algorithm's accuracy to determine the separate tracks. As such, the algorithm will consider these separate ARs as a single event that results in a misleading track which may go over different sources such as the Pacific and Atlantic oceans. Finally, the Phase III analysis is performed by applying the
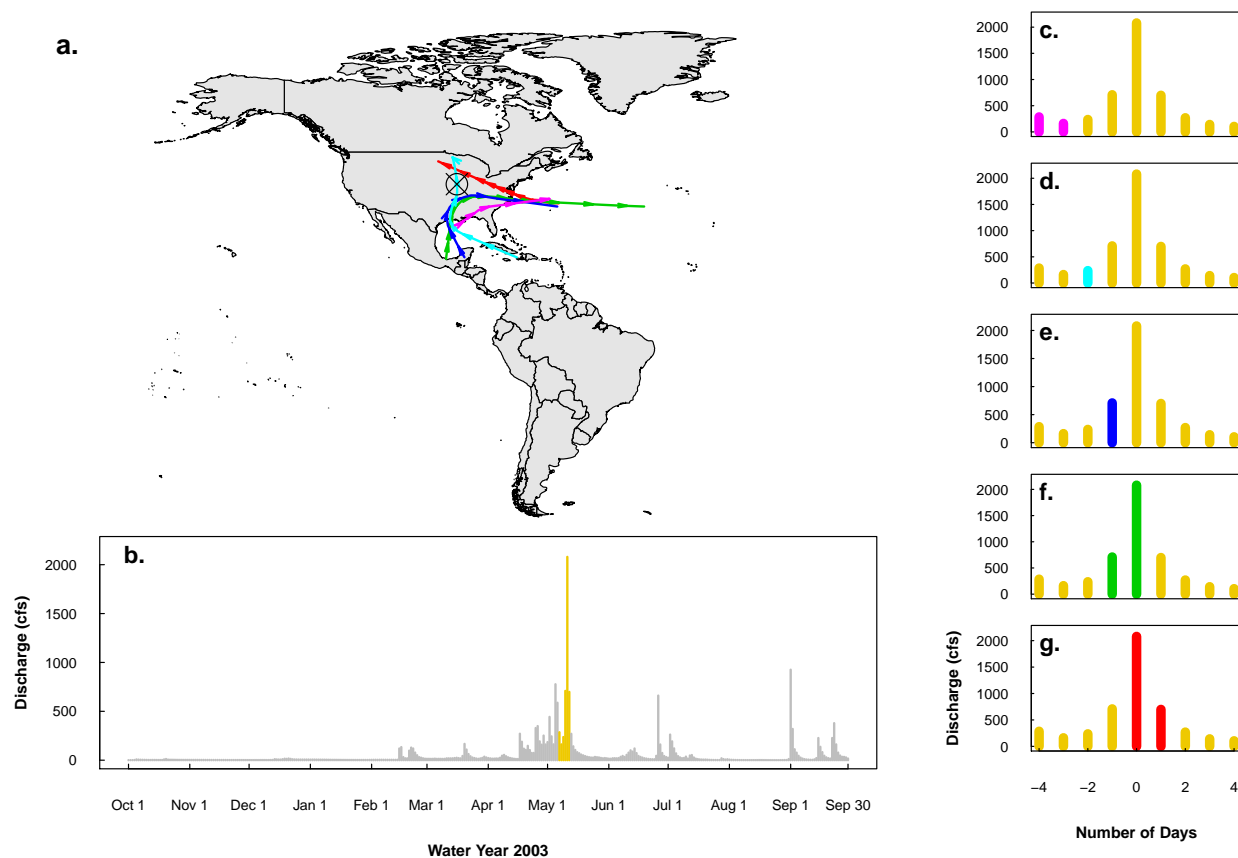
Figure 3.6: The effects of AR on streamflow. (a) The moisture tracks and sources of five different AR events which caused the AMF−AR event at the North River, MO (USGS station 05501000) in May 11, 2003. (b) The daily average discharge (cfs) during the water year of 2003. (c−g) The effect of each AR event on the magnitude of streamflow within four days before and after the day of AMF.

Log-Pearson type III (LP3) distribution to the results of Phase II to investigate the role of ARs and their sources in the flood frequency and magnitude.

3.6   Analysis Results

3.6.1   The availability of moisture over the conterminous US

Rutz et al. (2020) shows that ARs are more frequent in mid-latitude ocean basins than over land and other latitudes and their maxima are in the extratropical North Pacific/Atlantic, southeastern Pacific, and South Atlantic in 1979-2015. They calculated

the AR frequency as the percentage of reanalysis time-steps when the grid cell is within the boundary of an AR at each grid cell. Here the work counts at each grid cell (25N-50N, 65W-125W) the annual number of days that observed AR (AR Shape$\neq$0 & IVT$\geq$250 kg/m/s) and rainfall (PR$>$1 mm) over the period of 1956-2015. Figure 3.7(a & b) show the long-term mean and standard deviation of the annual number of days with an AR effect. The 60-year long-term mean of the annual AR-Days in the eastern part of the US is the highest in the US and ranges between 60-70 days. The long-term mean ranges between 20-55 days in the central and western parts of the country. However, the lower mean of annual AR-Days is below 20 day in the southwest US. On the other hand, the long-term variance (Figure 3.7(b)) in the annual AR-Days over the period 1956-2015 follows the AR-Days mean. The AR-Days variance is the highest 10-12 days in the eastern rigion, 6-10 in the western and central regions, and below 6 in the southwest region of the US.

Since there are more physical processes that generate rain storms other than ARs, the rain days are separated throughout the year between ARs and Non-ARs rain day categories. To determine which rain process is more frequent across the study area, the method counts the Non-AR-Days during the year which have rain (PR$>$1 mm) as well as the AR-Days at each grid cell. Figure 3.7(c) shows the long-term median for the annual number of days with and without ARs. Most of the rainy days in the western parts of the US are caused by ARs, and the number of AR-Days are much higher than Non-AR-Days Figure 3.7(c). Furthermore, the coastal areas of the western US have average or high number of AR-Days through the year compared to the mid-west and southwest regions which observe fewer wet days. Most of the wet days in the eastern and central parts of US are due to the Non-AR processes. Generally speaking, the 60-year median of the annual number of wet days in the eastern half of the US are significantly high comparing to the western half (Figure 3.7(c)).
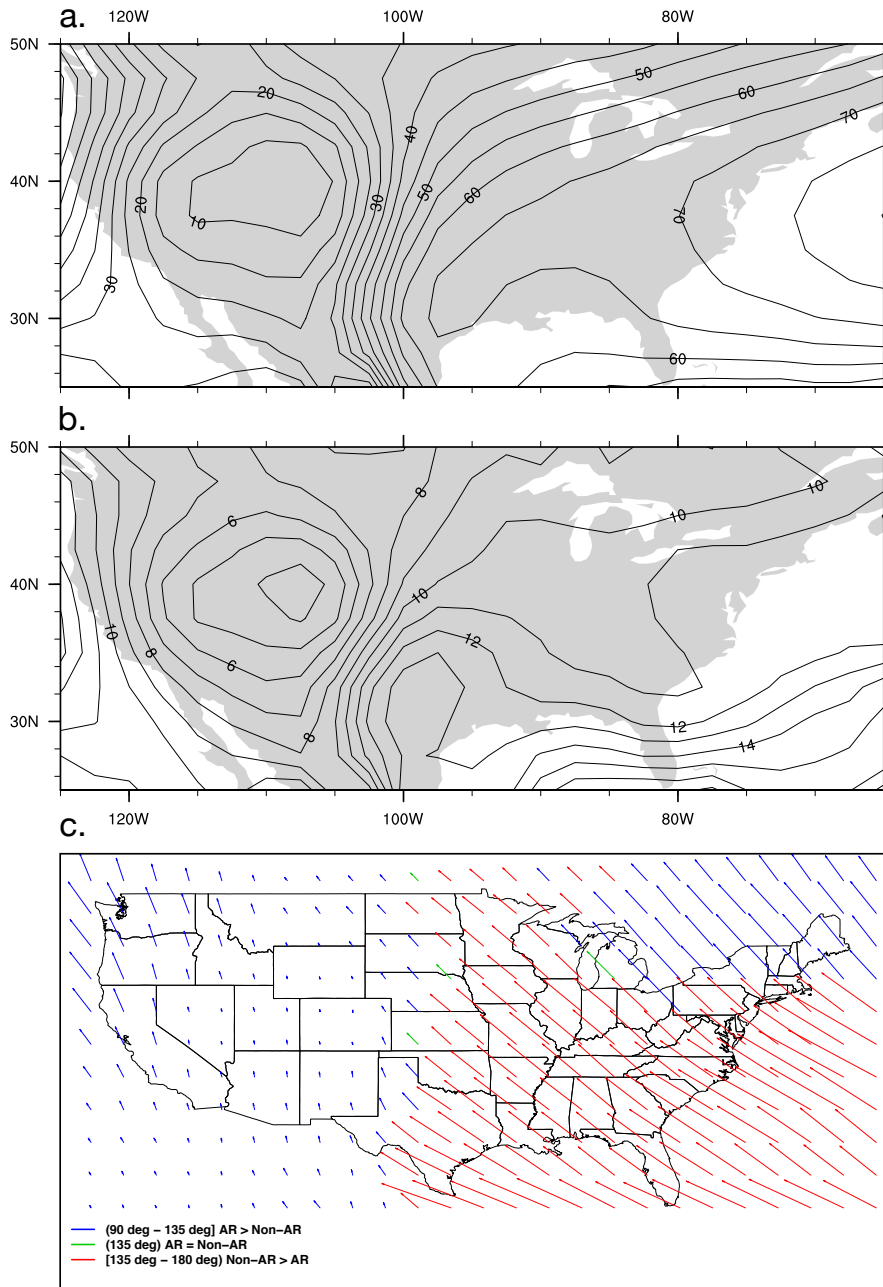
Figure 3.7: The annual number of wet days caused by AR vs. Non-AR. (a, b) Contour maps of the long-term mean and standard deviation for the annual number of days with AR event. Contour intervals are 10 to 80 by 5 and 4 to 17 by 1, respectively. (c) Grid points map of the long−term median for the number of AR and Non−AR days . Arrow's length and angle are scaled to the number of days. Arrow's length refers to the magnitude of the long−term median. Arrow's angle refers to the difference in the number of days. Arrow's color refers to which is greater number of days the AR (blue) or Non−AR (red). In (a, b, & c), only days with rainfall $> 1mm$ and IVT $\geq 250kg/m/s$ are counted.

3.6.2 The role of detected at-site ARs on the AMF frequency and magnitude

3.6.2.1 The Spatial and Fractional Contributions of ARs in the AMFs

Across vast swathes of the conterminous US, this work can identify that flooding is impacted by ARs at various degrees (Figure 3.8). The results show that a total of 25,725 out of 37,380 ($\sim 73\%$) AMF events occur due to the contribution of ARs across the study region in 1956-2015. The majority of the basins which are located in the Northwest and West climate regions have 40-60 (70-100%) AMF events over 60-year caused by ARs, and fewer stations have only 20-40 (30-70%) events. Furthermore, most basins in the Northeast and Southeast and many in the South, Central and East North Central climate regions have 40-60 (70-100%) AMF-AR events out of 60. This shows the important role of ARs in floods is spatially expanded across the US. The range of AMF-AR events is 10-40 (15-70%) in the rest basins in the South, Central and East North Central climate regions. Lastly, ARs appear to have the least effect in the Southwest and West North Central climate regions as the number of AMF-AR events ranges 10-30 (15-50%) in some basins of the first region and fewer in the latter, and the rest basins observed less than 10 (15%) AMF-AR events over the 60-year record.

These findings diverge from some studies (e.g., Lavers and Villarini 2013; Barth et al. 2017). However, they have a biased estimation of AR impacts on floods due to: inaccurate methodology of detecting ARs, and assumed a basin is affected by AR if its major axis is within a specific distance at any side to the stream gauge location. As such, a flood can be contributed to an AR although the AR is out of the watershed boundaries; or a peak of discharge can be attributed to Non-AR generated process due to lack of AR detection. Conversely, the methodology to examine the ARs impacts on floods in the current study intersects the AR and basin boundaries to relate both events. As a result, accurate percentages of AMF-AR related events are identified across the contiguous US.

62

3.6.2.2   The impacts of ARs on the AMFs magnitude

The investigation of ARs impacts on flood magnitude is performed in an at-site scale. To examine the ARs effects on the river discharge we analyze the AMFs for the nine climate regions separately. Figure 3.9 shows the AMFs quantiles for the AR and Non-AR generated events. The boxplots demonstrate the impacts of ARs on the AMF across different quantiles. It is clear that the peak discharge flows in the nine regions gauges are influenced by ARs as there is an increase in the medians and interquartiles of the AMF-AR events. The upper quantiles of AMF-AR events in the selected stream gauges of the Northeast, Northwest, and West climate regions are increased compared to the AMF-Non-AR events. Further, the AMFs interquartiles in arid and semi-arid regions such as the South, Southwest, and West are significantly increased due to the ARs impacts. However, there are no or very slight impacts of ARs on the lower quantiles of AMFs across the nine regions. As a result, although studies in the literature (e.g., Ralph et al. 2006; Lavers and Villarini 2013; Barth et al. 2017) showed that ARs are a major cause of the largest floods in different regions, the type of analysis presented here concluded that ARs have effects on small and large floods.

The risk in flood frequency and magnitude under the impacts of ARs is also measured by the ratio of the 100-year flood of AMF-AR to AMF-Non-AR by using LP3 estimates. The distribution is performed when there are at least 10 events in each group to minimize the level of uncertainty in the 100-year flood estimation. Ratio greater than 1.1 means ARs are responsible for causing the large floods while ratio less than 0.9 indicates that other flood-generating mechanisms control the large AMFs. However, the interval 0.9-1.1 is considered as even effects, and ratios are not determined for catchments observing homogeneous or weak mixed population floods.

Figure 3.10, shows that stations with distinct mixed population floods are found across the nine climate regions (see Figure 3.4). As such, 49% of the total stations are heterogenous floods stations while the homogenous catchments contain 45% stations of
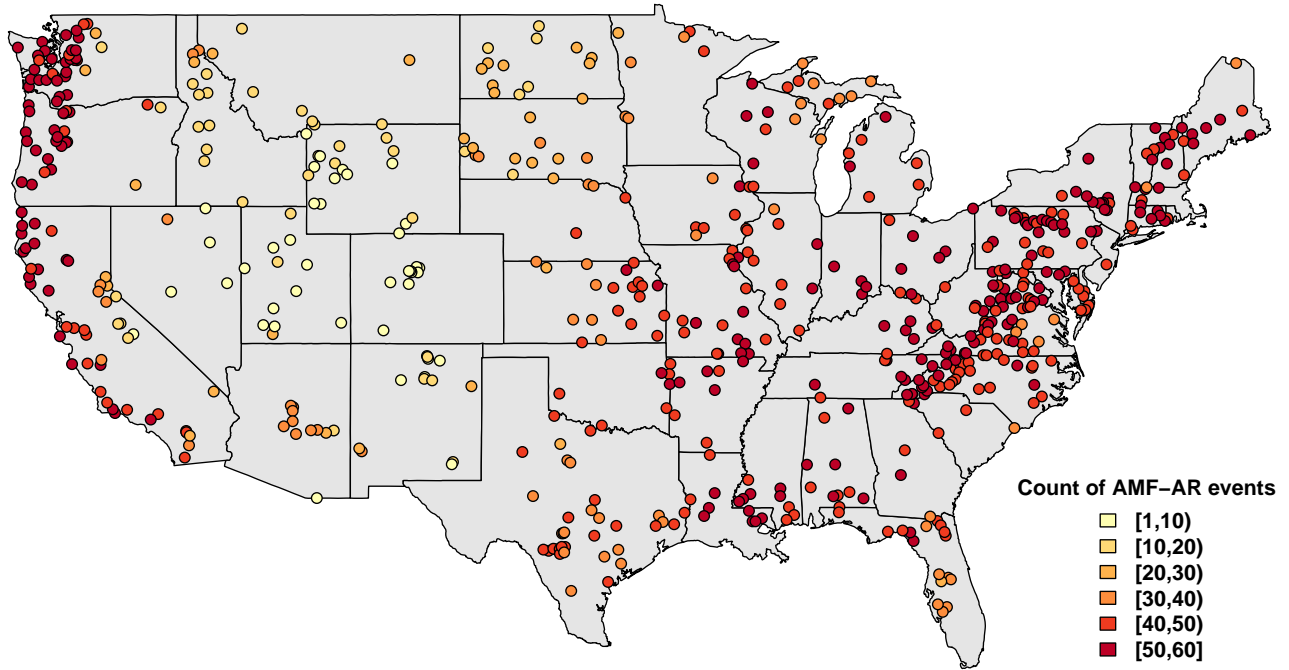
Figure 3.8: The total number of AMF-AR events over 60-year (WY: 1956-2015) in the selected streamflow gauges.

AR-generated floods and 6% stations of Non-AR generated floods. Most of the stations with mixed population floods in the Northwest and West show high ratios (1.4-2.3) and (3-25), respectively. The other high AR-impacted regions such as the South and Central indicate lower number of stations with high ratios and more with 1.4-2.2 ratios. In the Northeast and Southeast regions, although the AMF-AR events are highly frequent in the regions, the 100-year flood ratios show that most of heterogeneous flood records have moderate low or high ratios as there are other major flood generating mechanisms in the regions that can cause large floods such as the tropical cyclones and large snow pack volume. However, a mix of moderate ratios are found in the East North Central and West North Central regions, while some stations in the Southwest indicate that annual floods magnitude are highly induced by ARs. Therefore, it is clear that the key moisture delivery patterns significantly affect the US flood frequency and magnitude, but the risk of flood variability up to this point has not been accurately quantified.

Figure 3.9: Boxplots of the AMF−AR and AMF−Non−AR events (WY: 1956−2015) for the US climate regions. N_s (N_e, N_n) is the total number of stations (AMF−AR events, AMF−non−AR events) within the climate region.
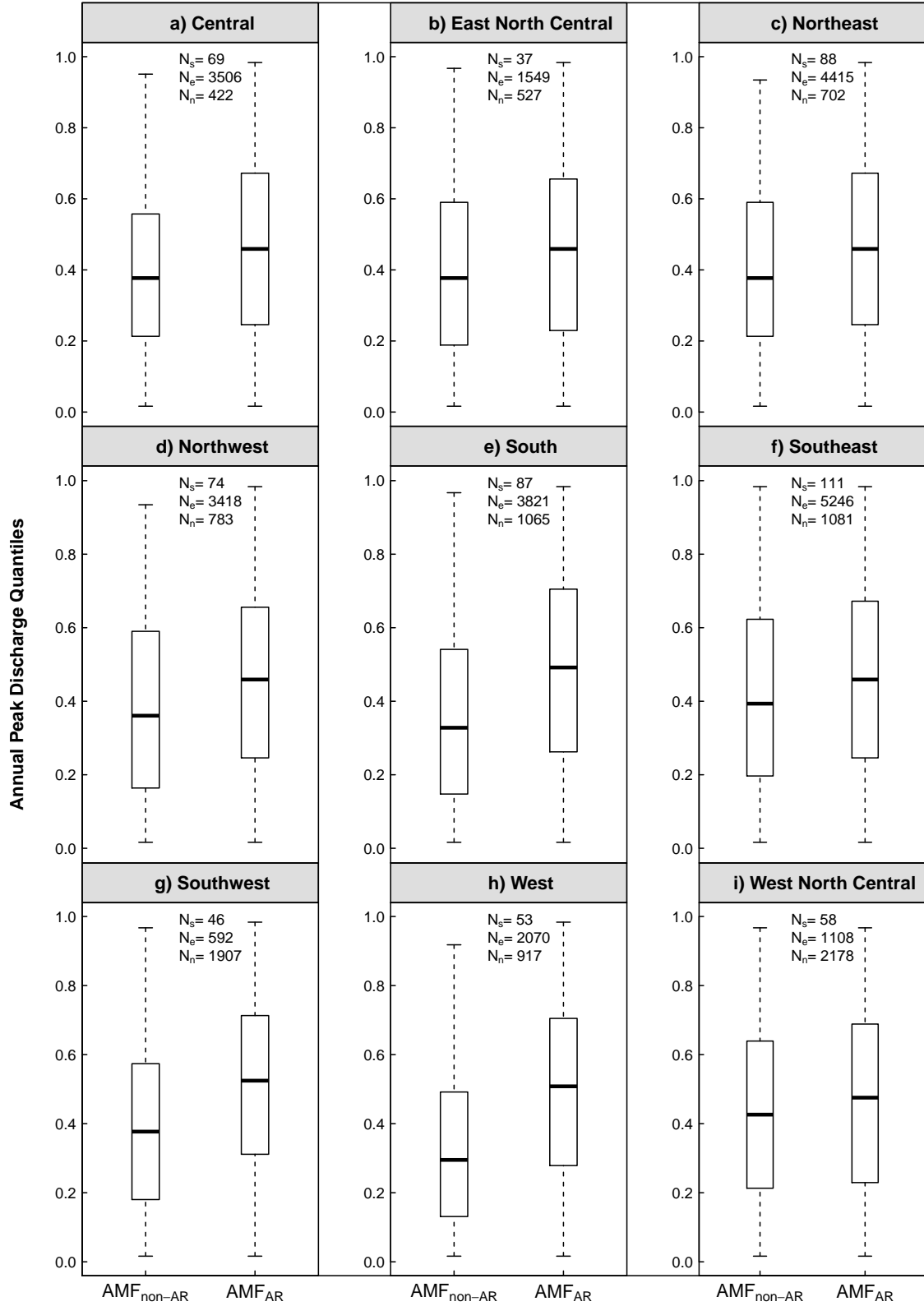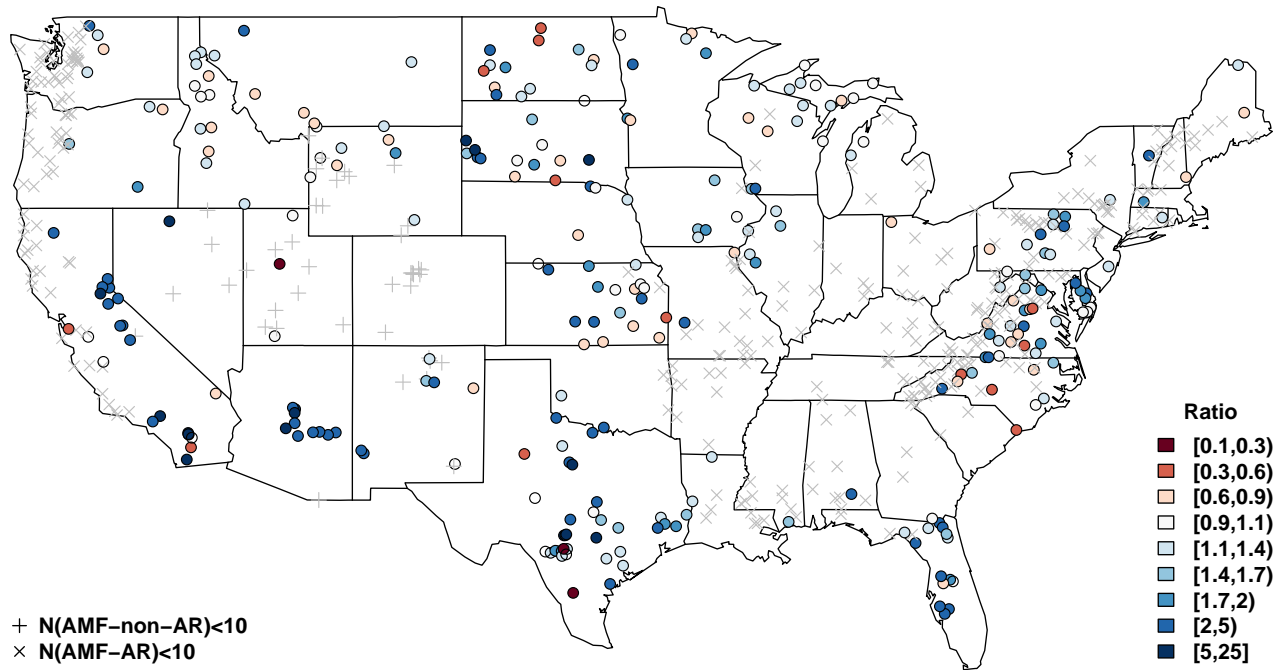
Figure 3.10: The 100-year flood (cfs) ratios for the AMF-AR to the AMF-Non-AR events of the selected stream gauges across the US climate regions.

3.6.3 The trajectories and sources of moisture associated with the AMF events

3.6.3.1 At-site identification of the AMF-causded ARs tracks and sources

Based on the methodology in section 3.5.3 and as a result of the Phase II analysis (Figure 3.5), we draw the major axis of the moisture trajectories (pathways) and identify the moisture sources of the ARs that driven AMF events across the conterminous US. The major axis is the integrated smoothed weighted curve of all grid points with IVT$\geq$250 kg/m/s of a complete AR episode. The ARs trajectories and sources of the AMF-AR events for the nine representative basins are shown in Figure 3.11. For all the nine representative basins, ARs tend to converge over the conterminous US and many of them are passed over the related study basin. Table 3.4 sorts the AR events that contributed to the AMFs for the nine representatives based on the sources of moisture. The representative basin of the Central climate region observed 55 AMF-AR events that were induced by 73 different AR episodes during the period 1956-2015 (Table 3.4). The Pacific Ocean and the

Gulf of Mexico are the major sources of AMF-cause moisture which are responsible about 27% and 21%, respectively, of the total AMF-contributed ARs over 60-year record. The Atlantic Ocean and the Caribbean Sea have less contribution to the AMF-AR events. However, about 38% of the contributed ARs to the flood events in 1956-2015 originated over the land mass. In the East North Central representative, 42 ARs contributed to cause 31 of AMF-AR events. The Gulf of Mexico is the major source which contributed 26% of the total events while local sources of moisture have the higher impacts on floods with a contribution of 40%. The West North Central representative observed 27 AMF-AR events that were caused by 39 AMF-caused ARs. The Pacific Ocean and the land mass are the major sources of moisture with contributions of 31% and 51%, respectively. The Northeast representative tend to have the highest number of contributed ARs which caused 50 AMF-AR events. Both of the Atlantic Ocean and the Gulf of Mexico are evenly participated to produce 46% of the total contributed ARs while the local sources produced 27%. The representative basin of the Northwest climate region has the highest frequency (59 out of 60) of AMF-AR events across the nine basins. The floods are dominated by the Pacific Ocean moisture as it contributed to 93% of the total AMF-caused ARs. On the other side, the West representative is a unimodal source dominant as the Pacific Ocean contributed to all 25 AMF-AR events in the basin. However, both of the South and Southeast region representatives have their 39 and 46 AMF-AR events highly impacted by the moisture of the Caribbean Sea and the Gulf of Mexico. Both sources contributed 38% and 49% of the total AMF-caused ARs in both basins, respectively. Nevertheless, the local-originated moisture has a contribution of 31% to the AMF-AR events in the South representative basin. Finally, the Southwest region representative basin is the less impacted basin by ARs when it is observed only five AMF-AR events that are dominated by the moisture of the Pacific Ocean over 60 year.

Table 3.4: The total number of AMF-induced AR events in the climate region representative station. ARs are sorted based on the source of moisture.

No.	Total AMF-AR	TCP	TEP	sTP	eTP	TA	sTA	eTA	CS	GM	Local	Total AR
1	55	0	6	14	0	0	2	1	7	15	28	73
2	31	0	0	7	0	0	2	3	2	11	17	42
3	50	1	1	4	0	0	11	6	4	17	20	64
4	59	2	0	54	6	0	0	1	0	1	3	67
5	39	0	6	9	1	2	1	1	8	16	20	64
6	46	0	6	6	0	1	4	3	11	17	9	57
7	5	0	0	3	2	0	0	0	0	1	2	8
8	25	1	0	25	0	0	0	0	0	0	0	26
9	27	0	3	7	2	0	0	0	3	4	20	39

Climate Regions: 1=Central, 2=East North Central, 3=Northeast, 4=Northwest, 5=South, 6=Southeast, 7=Southwest, 8=West, and 9=West North Centrl.

3.6.3.2 The moisture sources controlled the regional AMF-AR events

After quantifying the regionally impacts of ARs on the AMF events (Figure 3.8) and sorting them based on the major sources of moisture (Figure 3.11), the regional contribution of ARs to floods and the spatial extent of these storms can be determined. The boxplots in Figure 3.12 show the percentages of contribution of a specific source of moisture to the AMF-AR events of each basin within the region. The percentage of a single basin in the region is determined by counting the frequency of ARs related to a specific source of moisture that induced the AMF events over 60 year in that watershed. The small red circles represent the contribution percentage of a selected source of moisture to the AMF-AR events of the climate region representative basin. Table 3.5 lists the total number of influenced stations by a specific source of moisture within the climate region. Both results in Figure 3.11 and Table 3.5 illustrate the regional influence of the major sources of moisture on the flooding frequencies across the conterminous US. In other words, the source of moisture that generated AMF-AR events in a selected climate region is considered to be a major source of AMF-caused ARs for that region when causes as many as stations in that region to have high number of AMF-AR events. As such, majority of the basins in the Central climate region are influenced by all the sources (Table 3.5), but only
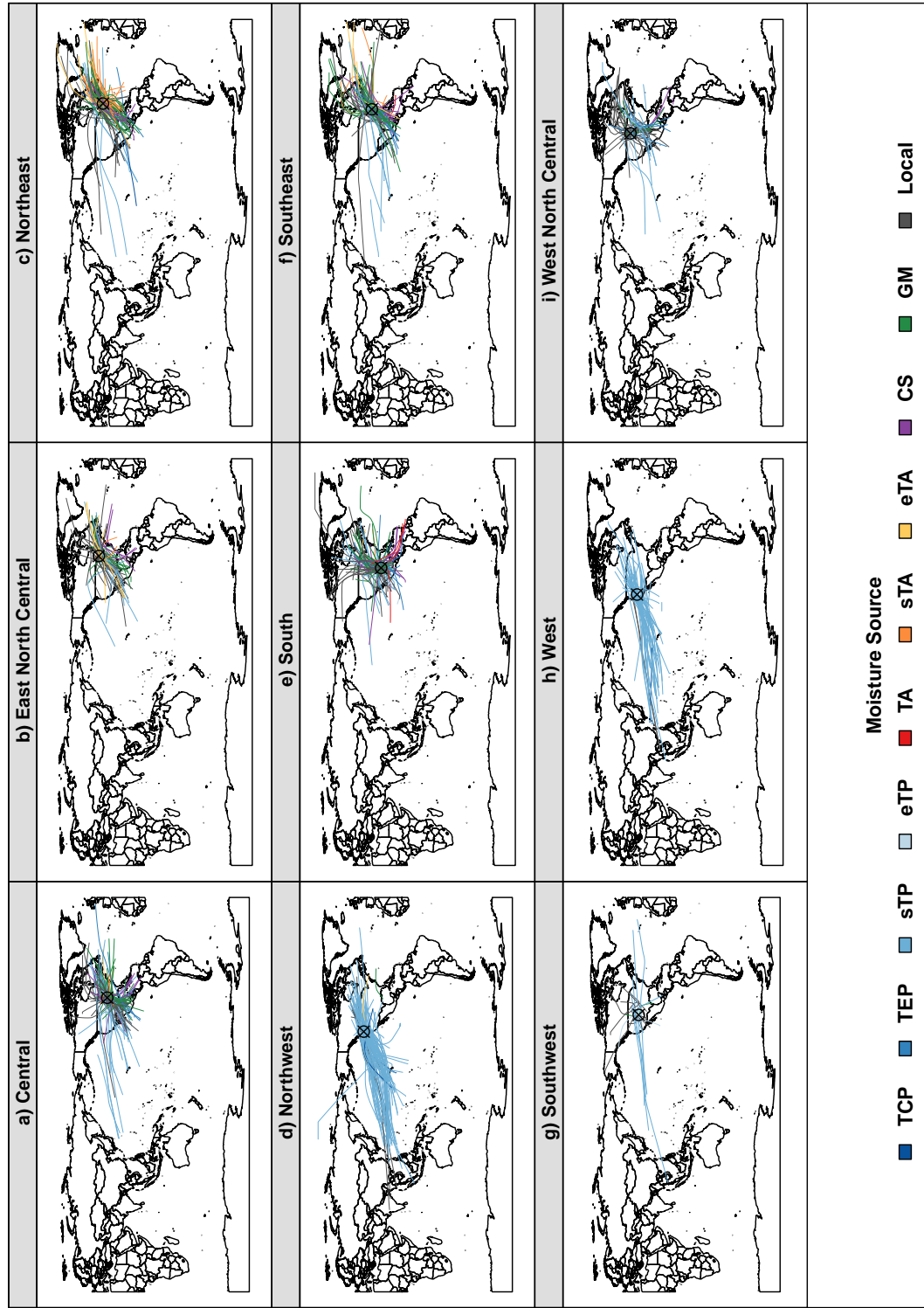
Figure 3.11: The major axis and sources of moisture that caused the AMF−AR events (WY: 19562015). The sources of moisture in the representative stations of the US climate regions are: (TCP) Tropical Central Pacific Ocean, (TEP) Tropical Eastern Pacific Ocean,(sTP) Sub−tropical Pacific Ocean, (eTP) Extra−tropical Pacific Ocean, (TA) Tropical Atlantic Ocean, (sTA) Sub−tropical Atlantic Ocean, (eTA) Extra−tropical Atlantic Ocean, (CS) Caribbean Sea, (GM) Gulf of Mexico, (Local) In −land moisture.

the Gulf of Mexico and the local moisture are caused high percentages of medians 38% and 32%, respectively, of generating AMF-AR events over the entire record (Figure 3.12). In the same manner, the AMF-AR events in the climate regions of Northeast, South, and Southeast are dominated by the Gulf of Mexico AMF-caused ARs and the local moisture except the latter region which is only affected by the Gulf of Mexico moisture. As a result, the medians of AMF-AR events in the Northeast, South, and Southeast climate regions attributed to ARs originating in the Gulf of Mexico are 38%, 40%, 48%, and those generated by the landmass moisture are 34%, 45%, 22%, respectively. On the other hand, the ARs that originated in the sub-tropical Pacific Ocean and those created over the land influenced the all basins in in the East North Central and West North Central regions (Table 3.5), and are attributed to AMF-AR events with medians 32%, 48% for the ocean and 52%, 62% for the land, respectively. In the Northwest and West regions, the moisture of the sub-tropical Pacific Ocean influenced all basins within the two regions (Table 3.5). Further, the generation of AMF-AR events are dominated by the s-TP with medians 95% and 98%, respectively. Additionally, the sub-tropical Pacific Ocean is also considered a main source of AMF-caused ARs in the Southwest climate region in addition to the local moisture. Both sources affected all basins in the region (Table 3.5), and cause high number of AMF-AR events with medians of 70% and 50% (Figure 3.12).

In comparison, estimation of moisture source contribution to AMF-AR events in the representative basins is closely match those for the regions in many cases (Figure 3.12). For instance, both estimations in Figure 3.12(d and h) show that the sub-tropical Pacific Ocean is the main source of moisture in the region. However, the single-basin based estimation of source contribution is occasionally under or over estimated the regional contribution of the moisture source. As such, the single-basin based estimation is overestimated the contribution of the extra-tropical Pacific Ocean moisture in the Southwest and is underestimated the contribution of sub-tropical Pacific Ocean moisture in the West North

Central region (Figure 3.12(g and i). In conclusion, although the estimation of the source contribution may not reflect the actual response of the majority of basins in the region due to insufficient number of detected AMF-AR events in the record of the representative basin, the results of estimations confirm the major sources for each climate region.

Table 3.5: The total number of influenced stations by a specific source of moisture within the climate regions.

No.	Climate Region	Total Stations	TCP	TEP	sTP	eTP	TA	sTA	eTA	CS	GM	Local
1	Central	69	33	69	69	7	53	69	57	69	69	69
2	East North Central	37	4	31	37	16	9	32	31	37	37	37
3	Northeast	88	22	82	87	44	55	87	87	88	88	88
4	Northwest	74	53	1	74	71	0	3	14	19	15	69
5	South	87	7	85	87	19	73	70	63	87	87	87
6	Southeast	111	21	111	107	5	74	110	106	111	111	110
7	Southwest	46	12	19	45	17	0	2	6	20	30	44
8	West	53	40	17	53	24	0	1	0	1	23	48
9	West North Central	58	3	27	58	29	11	16	11	30	40	57

3.6.4 The impact of moisture sources variation on the floods frequency and magnitude

3.6.4.1 Floods frequency

Figures 3.13-3.21 show the Log-Pearson Type III (LP3) flood frequency curves for the AMF events in populations of: a) mixed events, b) Non-AR generated events, c) AR generated events, and sort the latter based on the main sources of moisture, d) Pacific AR generated events, e) Atlantic AR generated events, f) Caribbean Sea and Gulf of Mexico AR generated events, and g) local AR generated events for the nine representative basins. A threshold of five AMF events assigns to apply the LP3 on the data points. In the Central region representative basin, the flood frequency curve of the mixed population events (Figure 3.13(a)) shows that floods are dominated by ARs at varying degrees. Further, the fitted LP3 curve shows a poor fit in the right-hand tail of the distribution. However, the separation of events based on Non-AR and AR generating floods (Figure 3.13(b and c))
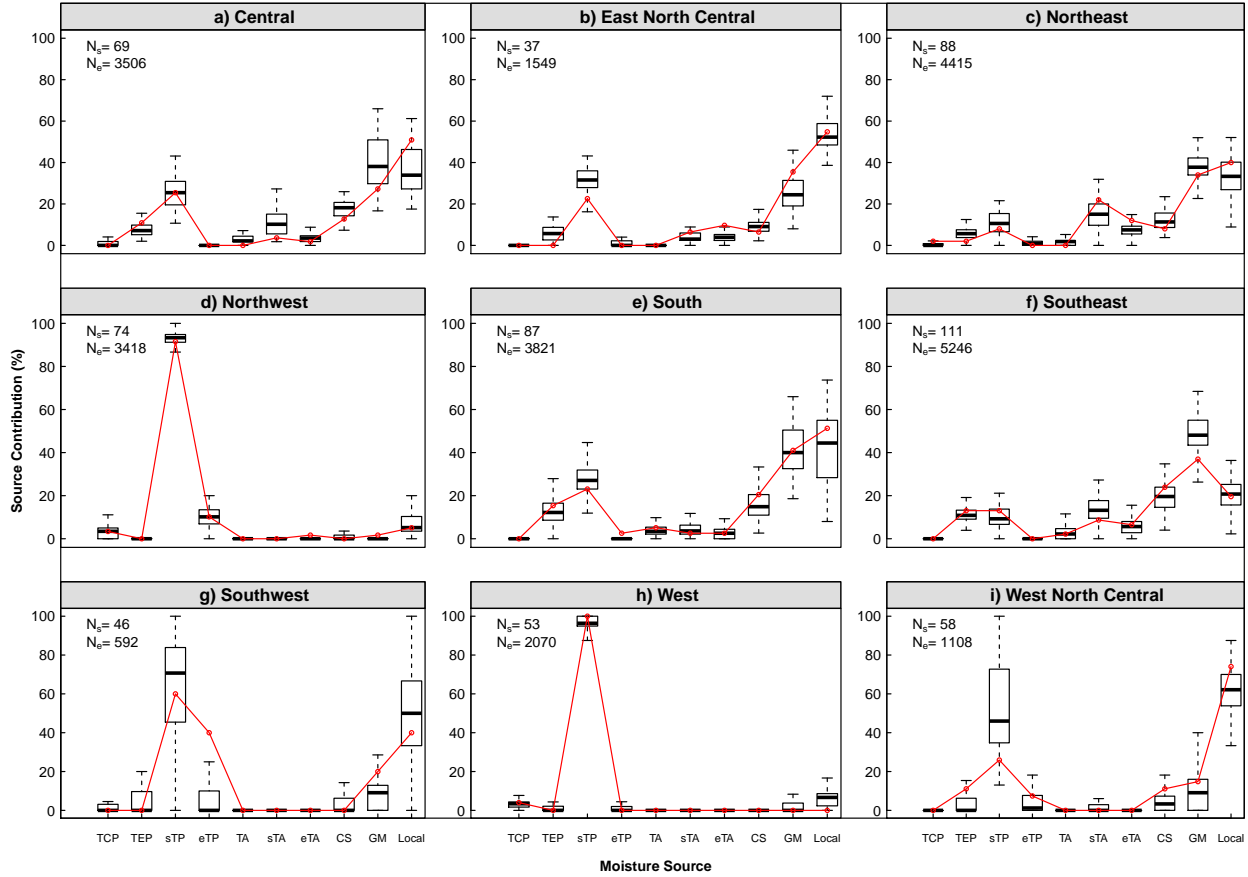
Figure 3.12: Moisture source contribution. boxplots (red lines) are the percentages of total number of AMF−AR events which caused by a specific source of moisture for the all stations (representative station) in each climate region (see Table 3.4 and Table 3.5). N_s (N_e) is the total number of stations (AMF−AR events) within the climate region. Where the sources of moisture are the same as in Figure 3.12

helps to have a better fit to the AMF-Non-AR events although this raises the level of uncertainty due to the lack of data points. Therefore, could not improve the distribution fit line for the AMF-AR events. Yet, the AR-source based separation for the AR generating floods population (Figure 3.13(d, f, g)) show the improved flood frequency curves among the AR generated population of the Pacific Ocean, Caribbean Sea and Gulf of Mexico, and local moisture. The LP3 fit line for the mixed population in the East North Central representative (Figure 3.14) underestimated the large AR-floods. As such, AR-source based separation of AMF-AR events improves the fit line in the upper tail. However, the flood frequency curve overestimates the Non-AR events. Although, the LP3 fit line for the mixed

population events in the Northeast representative (Figure 3.15) is maintained the data distribution, the events separation fits the data points more precisely especially for those having sufficient number of events such as AMF-AR and AMF events generating by ARs from the Caribbean Sea and Gulf of Mexico. As almost all AMF events in the Northwest representative generated by ARs and most of them from the Pacific Ocean, the three flood frequency curves in Figure 3.16 are properly fitted the data points. In the South region representative, the AMF-AR events dominate both tails in the distribution (Figure 3.17). While the LP3 distribution of the mixed population events observe poor fit, the events separation improve the distribution fit line for the AMF-Non-AR and AMF-AR events such as the events generated by the effects of the Caribbean Sea and Gulf of Mexico. In the Southeast region representative, the flood frequency curve has a poor fit to the mixed population event (Figure 3.18). Although events separation improves the AMF-Non-AR observations fit line, the distribution shows a poor fit in the right-hand tail. On the other hand, the source-based separation of the AMF-AR events improves the distribution fit for the large floods. Contrary to all other regions representatives, the Southwest region representative basin is dominated by the AMF-Non-AR events at varying degrees (Figure 3.19). The LP3 distribution shows a poor fit to the mixed population events. However, events separation does not improve neither the AMF-Non-AR events as they may come from different flooding mechanisms (Figure 3.4) nor the AMF-AR events due to insufficient number of events. In the West region representative basin, although the AMF events are evenly generated in the mixed population (Figure 3.20), the right-hand tail of the LP3 distribution is dominated by the AMF-AR generated events. The distribution of the mixed population shows a good fit for the data points as most of the flood events are very closed in their magnitudes. Furthermore, events separation improve the flood frequency curve of the AMF-AR events as most events are caused by the Pacific Ocean moisture. However, the LP3 distribution fit overestimates the AMF-Non-AR events. Last, the representative basin of the West North Central Climate region shows a poor fit in the

73

upper tail of the heterogeneous events distribution (Figure 3.21). Although the events separation indicates underestimated LP3 fit to the Non-AR generated floods, the upper tail of the distribution is improved to fit the flood events caused by ARs that affected the basin and their sources.

In conclusion, eight out of nine basins have their largest floods events over 60 year are generated by ARs no matter what their sources are. But, the Southwest representative basin which is the least impacted by ARs has its third largest flood over the entire record caused by ARs. Further, although the events separation is biased the fit line of the AMF-Non-AR events distribution in some cases as they may be caused by different flood generating processes, the flood frequency curves of the floods generated by ARs and their sources are improved to fit the data.

3.6.4.2 Floods magnitude

To examine the impacts of ARs and their sources on the floods magnitude, we perform the Log-Pearson Type III distribution on the AMF events to calculate the 100-year flood which typically used for the design of flood structures. The impact of ARs on flood magnitude is determined by calculating the ratios of the 100-flood of the AMF-AR and AMF-AR-Sources based (Pacific Ocean, Atlantic Ocean, Caribbean Sea and Gulf of Mexico, Local) events to the 100-year flood estimated from the AMF-Non-AR events. For statistical confidence of the results, we resample the actual AMF timeseries 1000 times and split them into subgroups based on their flood-generating processes. Then, the 100-year floods for each population is calculated to find the required ratios to construct their probability density function (PDF) estimates. As a result, the more actual ratio of a selected population deviates from ratio equal to one in the PDF, the more confidence the actual ratio is.
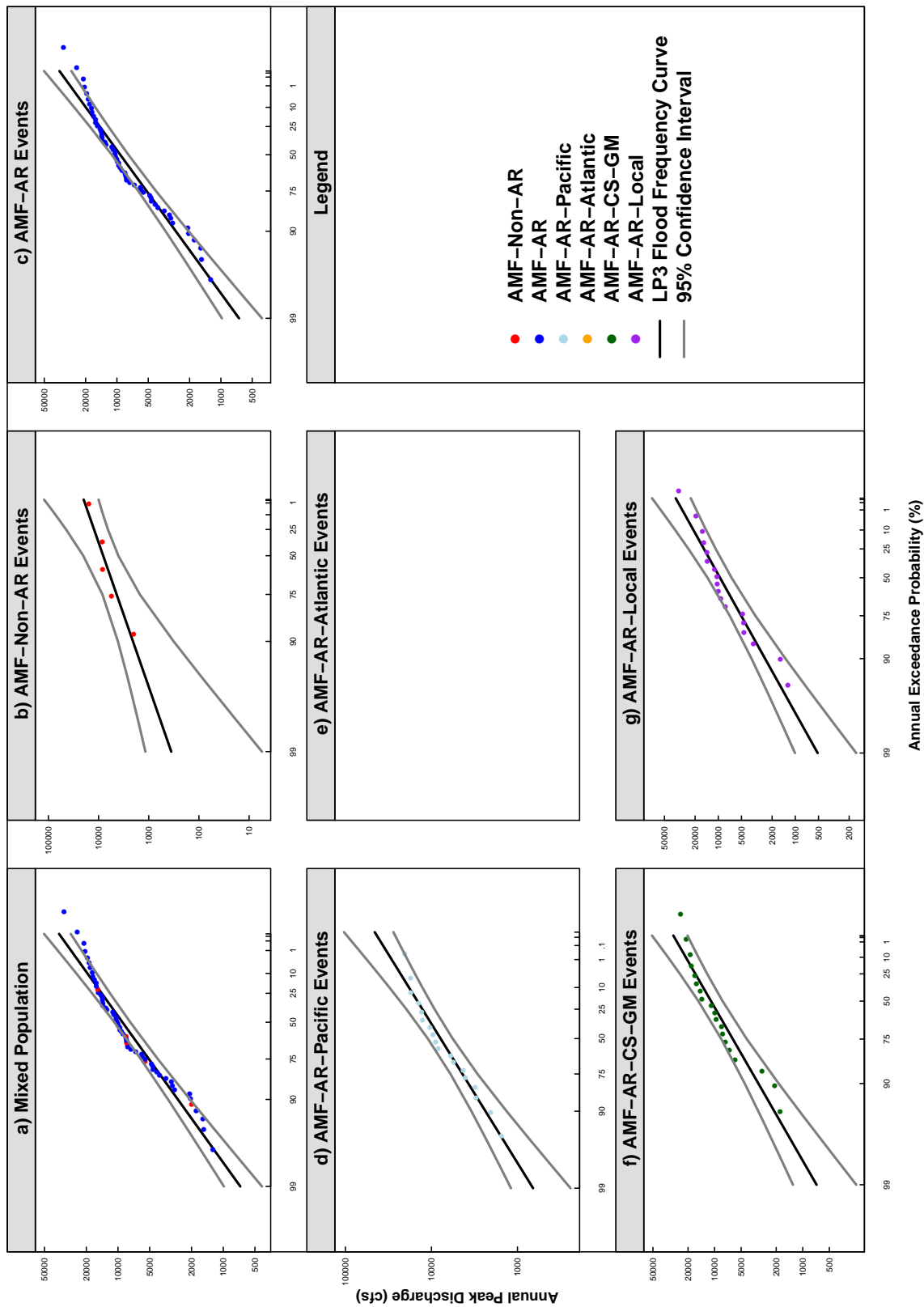
Figure 3.13: The Log-Pearson Type III flood frequency curve for the representative station of Central climate region. (a) AMF mixed population, (b) AMF-Non-AR, (c) AMF-AR, and (d, e, f, g) AMF-AR separated by the four major sources of moisture.

Figure 3.14: The Log-Pearson Type III flood frequency curve for the representative station of East North Central climate region. (a) AMF mixed population, (b) AMF-Non-AR, (c) AMF-AR, and (d, e, f, g) AMF-AR separated by the four major sources of moisture.

Figure 3.15: The Log-Pearson Type III flood frequency curve for the representative station of Northeast climate region. (a) AMF mixed population, (b) AMF-Non-AR, (c) AMF-AR, and (d, e, f, g) AMF-AR separated by the four major sources of moisture.

Figure 3.16: The Log-Pearson Type III flood frequency curve for the representative station of Northwest climate region. (a) AMF mixed population, (b) AMF-Non-AR, (c) AMF-AR, and (d, e, f, g) AMF-AR separated by the four major sources of moisture.
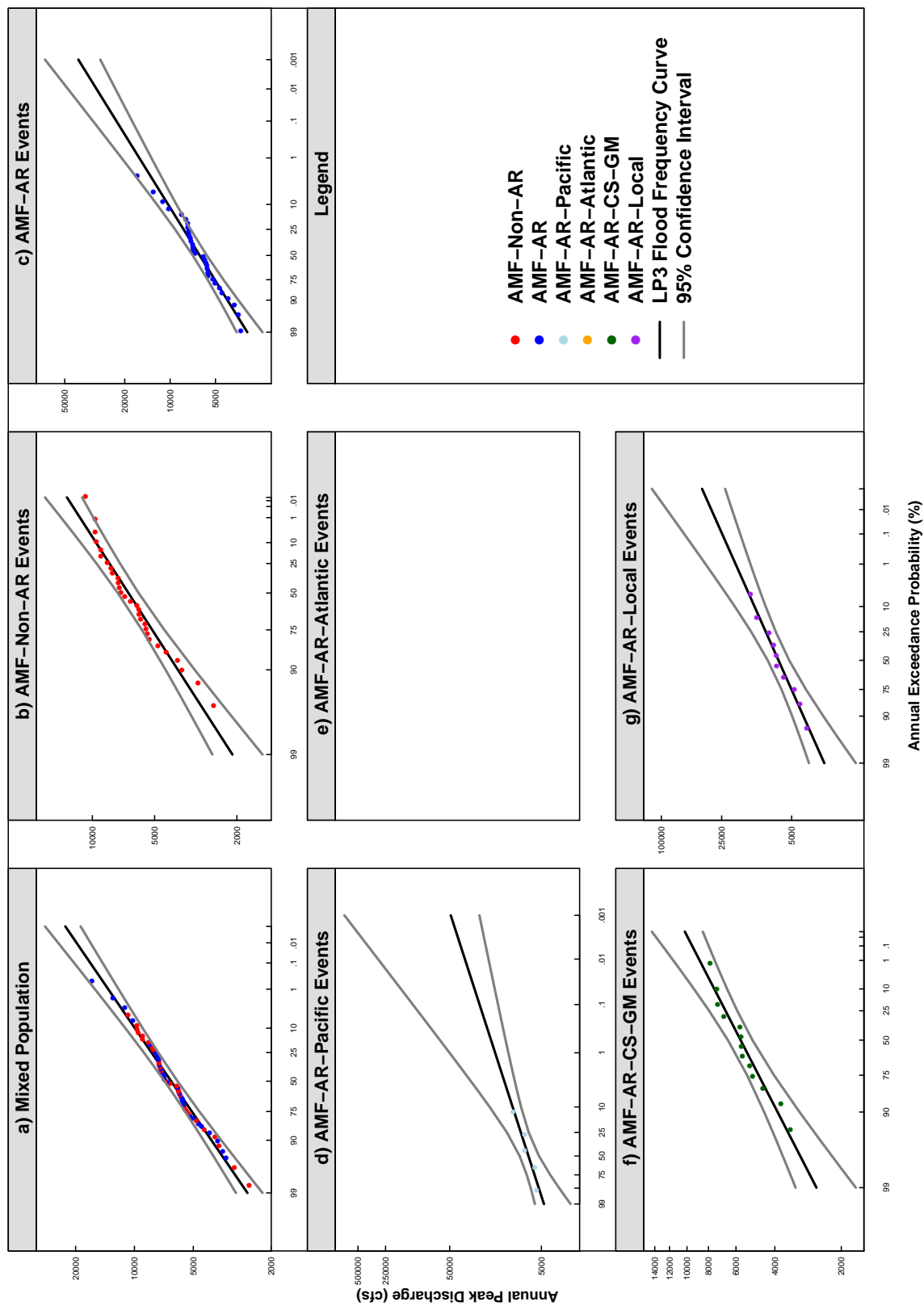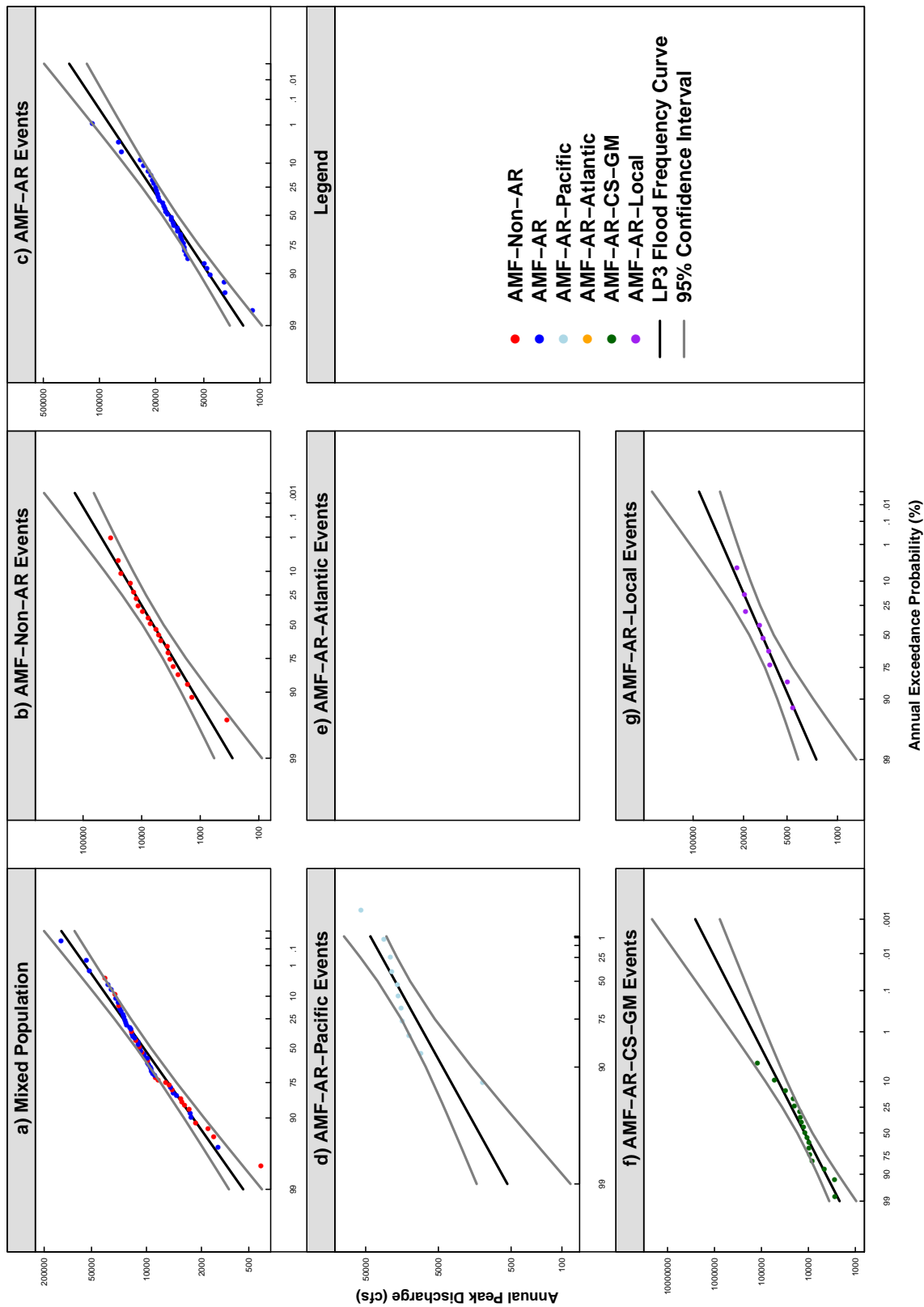
Figure 3.17: The Log-Pearson Type III flood frequency curve for the representative station of South climate region. (a) AMF mixed population, (b) AMF-Non-AR, (c) AMF-AR, and (d, e, f, g) AMF-AR separated by the four major sources of moisture.
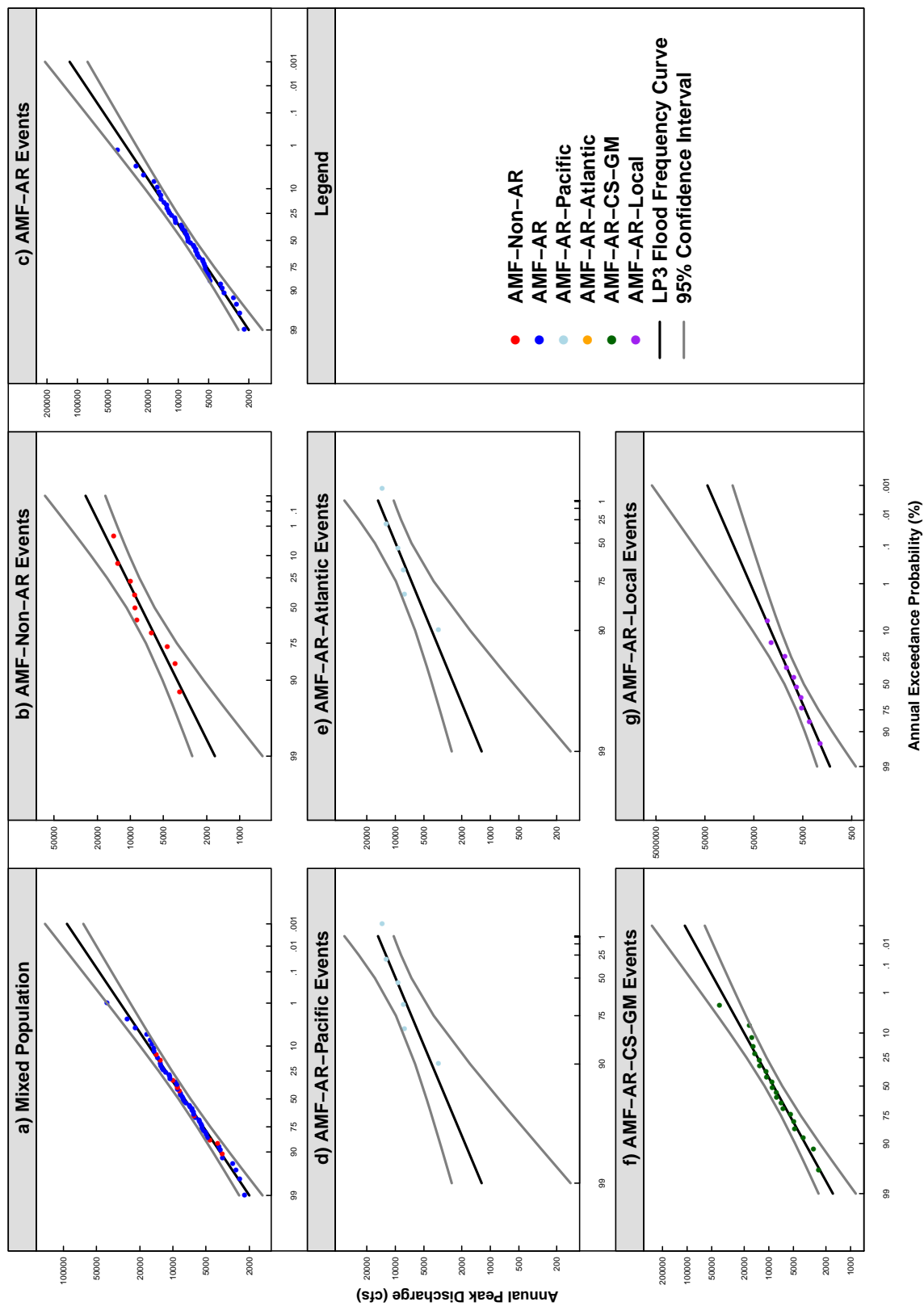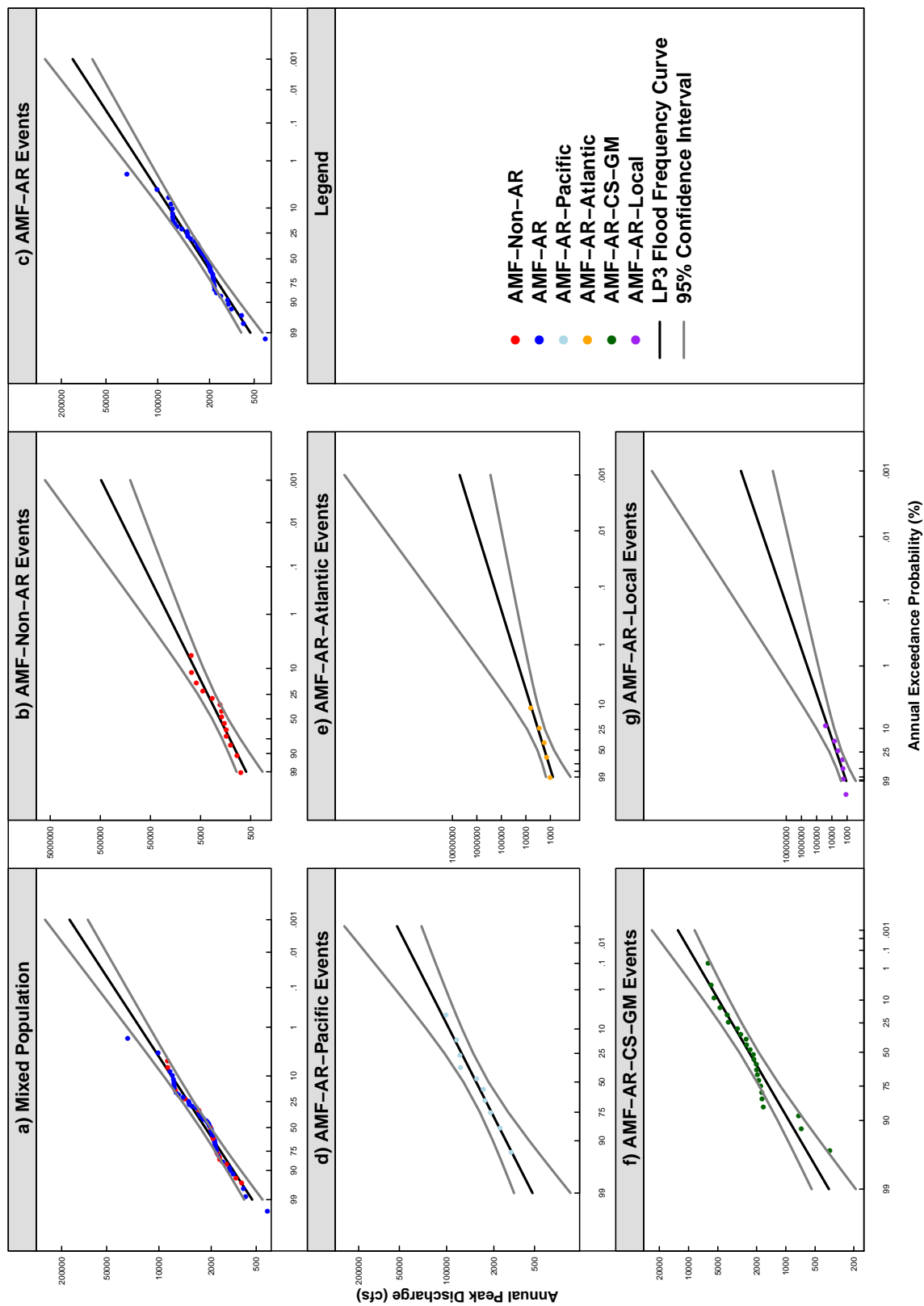
Figure 3.18: The Log-Pearson Type III flood frequency curve for the representative station of Southeast climate region. (a) AMF mixed population, (b) AMF-Non-AR, (c) AMF-AR, and (d, e, f, g) AMF-AR separated by the four major sources of moisture.
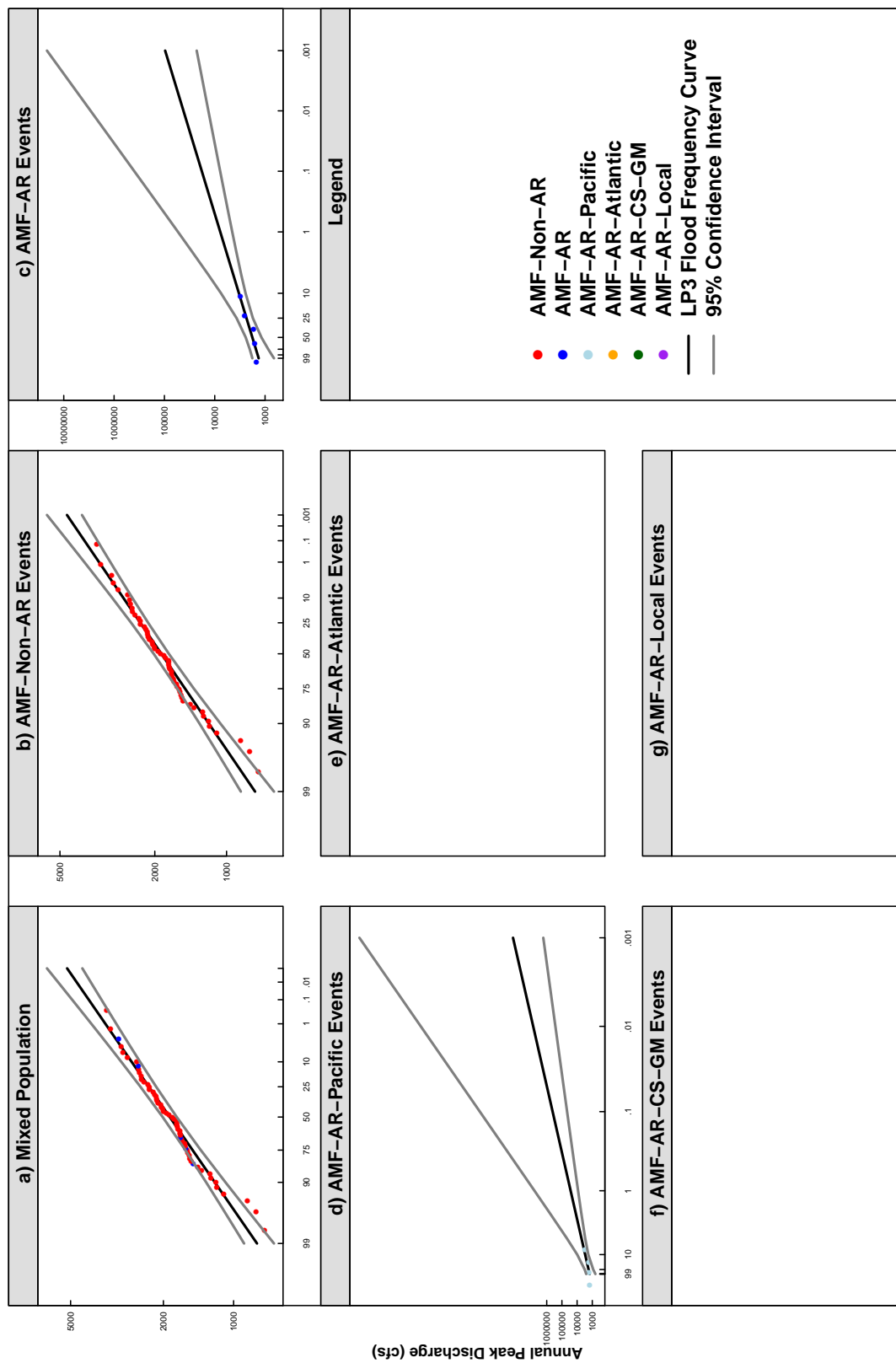
Figure 3.19: The Log-Pearson Type III flood frequency curve for the representative station of Southwest climate region. (a) AMF mixed population, (b) AMF-Non-AR, (c) AMF-AR, and (d, e, f, g) AMF-AR separated by the four major sources of moisture.
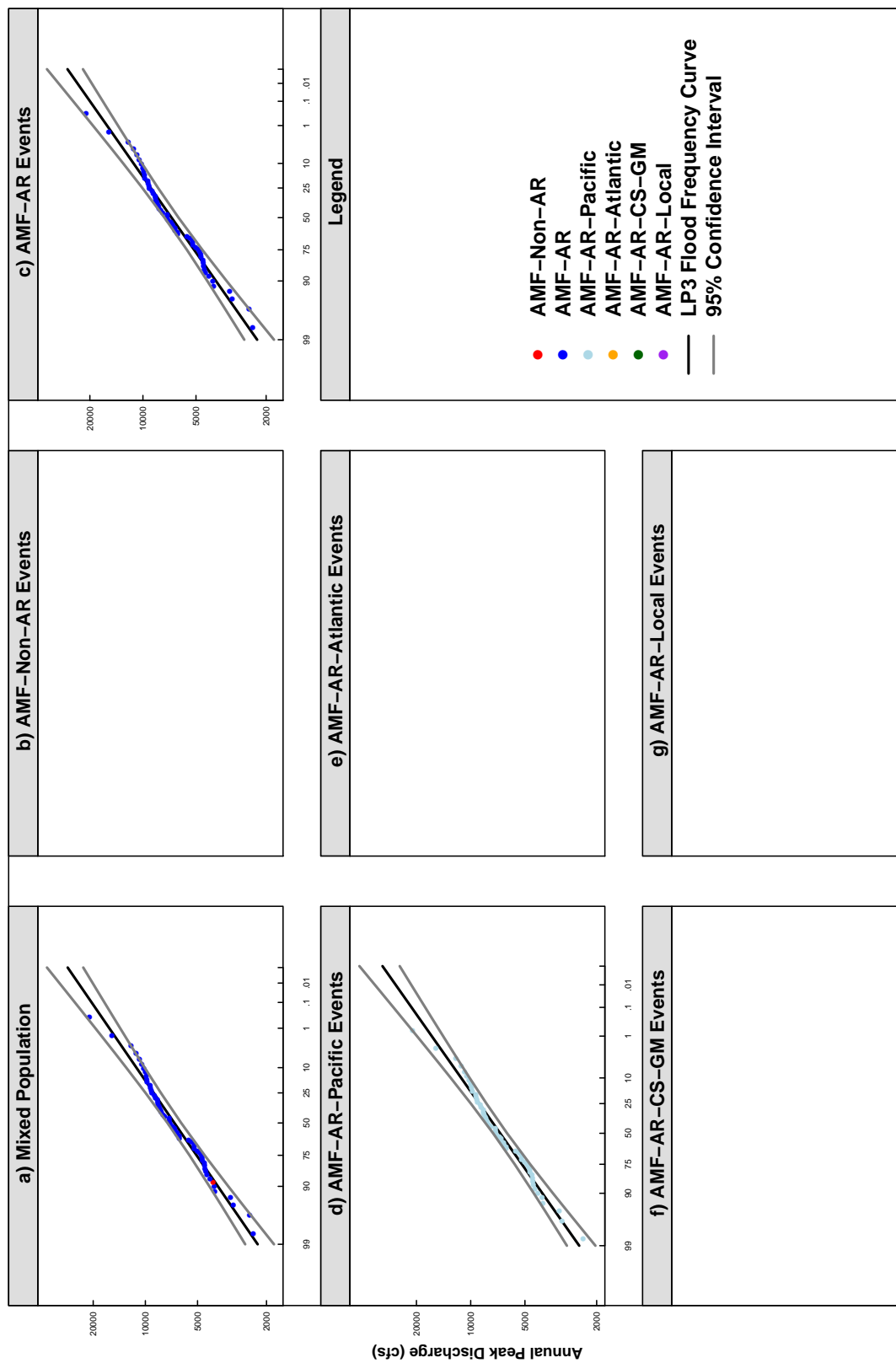
Figure 3.20: The Log-Pearson Type III flood frequency curve for the representative station of West climate region. (a) AMF mixed population, (b) AMF-Non-AR, (c) AMF-AR, and (d, e, f, g) AMF-AR separated by the four major sources of moisture.

Figure 3.21: The Log-Pearson Type III flood frequency curve for the representative station of West North Central climate region. (a) AMF mixed population, (b) AMF-Non-AR, (c) AMF-AR, and (d, e, f, g) AMF-AR separated by the four major sources of moisture.

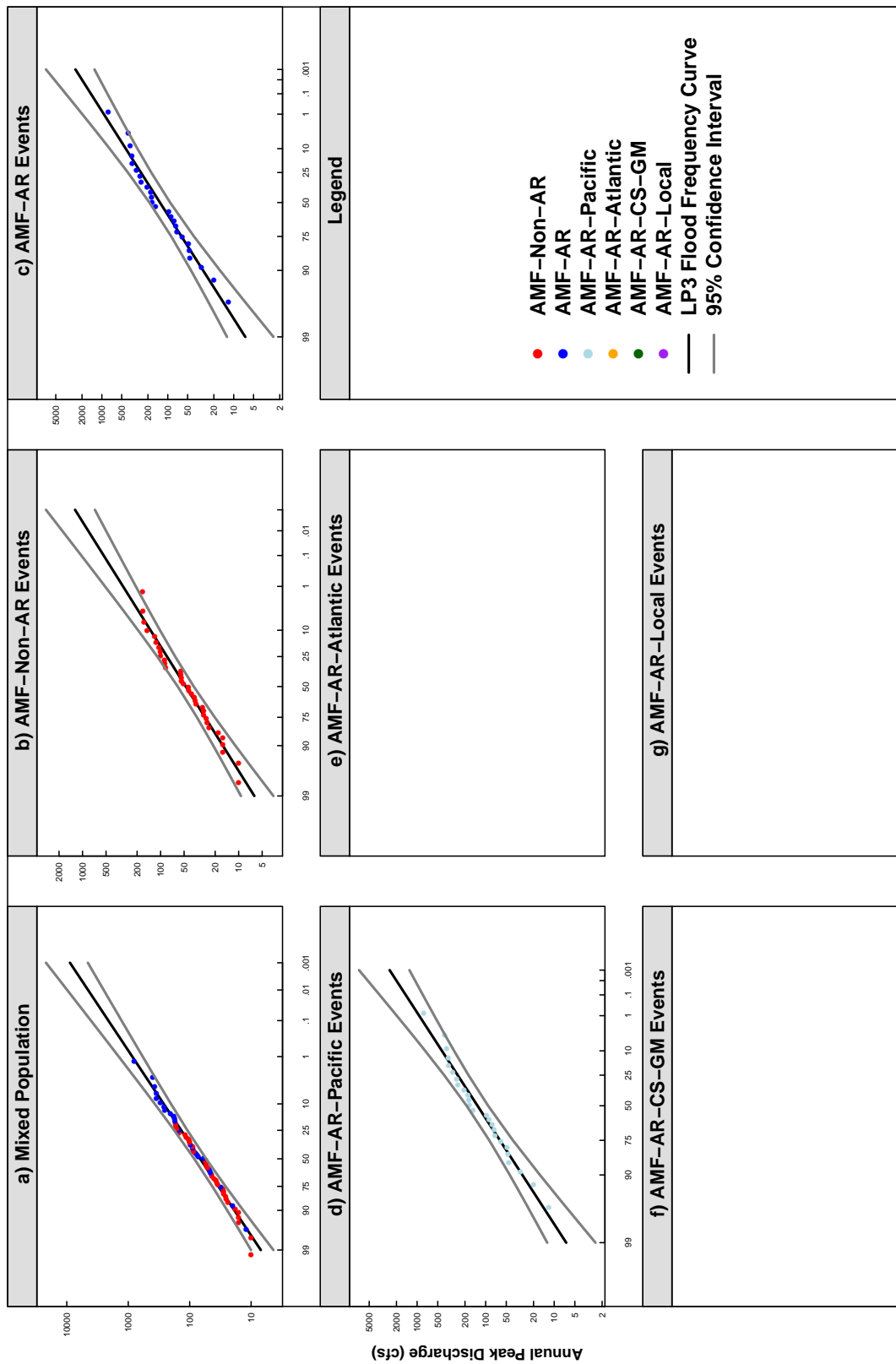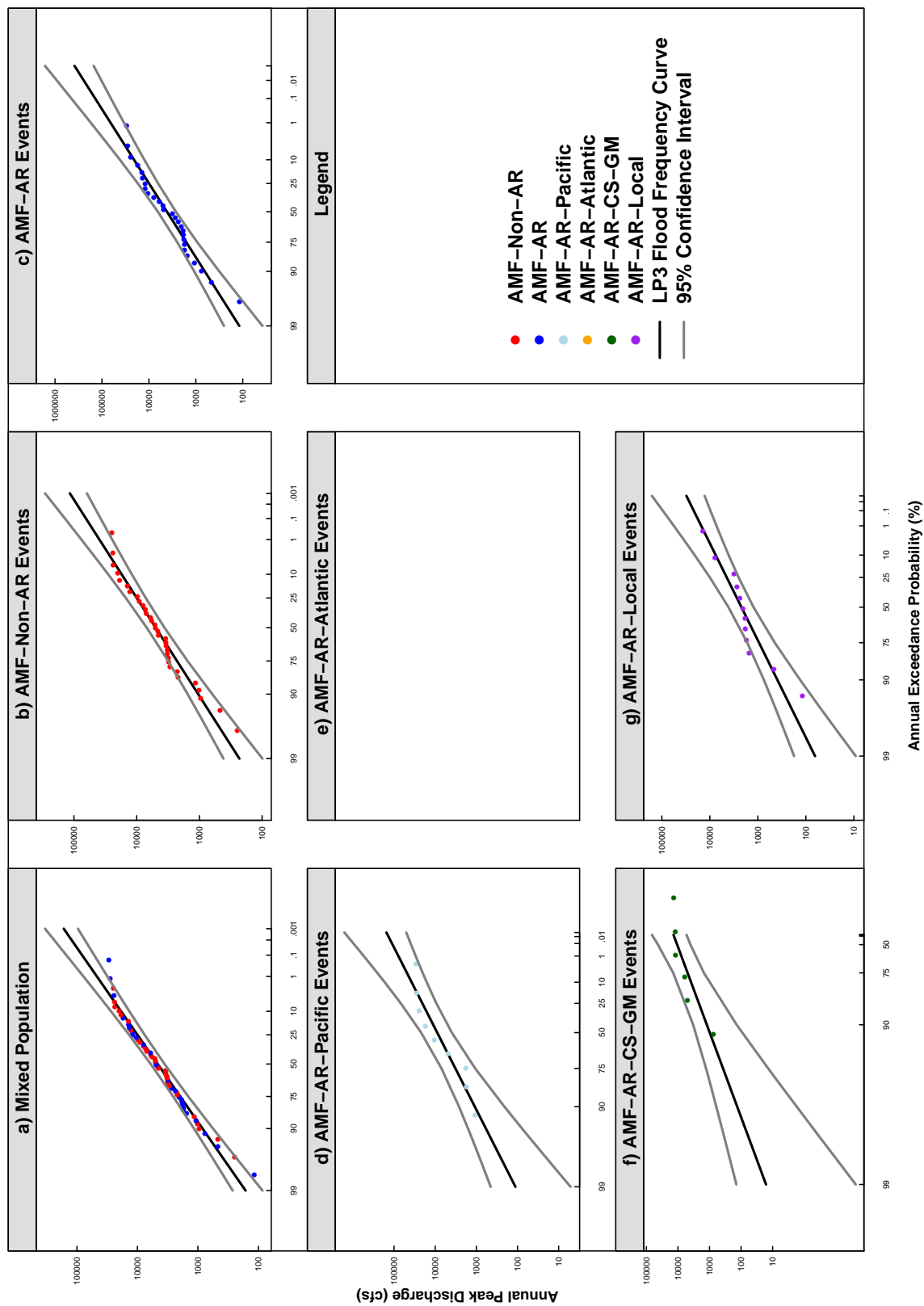Figure 3.22 shows the ratio of the 100-year flood from AMF-AR and their sources to the 100-year flood of the AMF-Non-AR events, and the probability density function of the ratios calculated from the resampled actual timeseries of the AMF events for the nine climate region representative basins. Table 3.6 lists the ratios of the 100-year flood estimated from the real floods generated by ARs and their sources to those which generated by Non-AR, as well as in which quantile is the actual ratio located with regard to the ratios from resampled events in the nine representative basins. As a result, all basins have at least one actual ratio greater than one which ensures the impact of ARs and their sources on the floods magnitude (Figure 3.22 and Table 3.6). As such, the magnitude of a 100-year flood estimated from the AMF-AR population is 50% larger than the 100-year flood of the AMF-Non-AR events in the representatives basin of the Central region. While the impact of ARs originated in the Atlantic Ocean on the 100-year flood of the East North Central region representative is four times larger than the magnitude of the 100-year flood estimated from the AMF-Non-AR population. In the same manner, the ARs of the Caribbean Sea and Gulf of Mexico and the local moisture double the magnitude of the 100-year flood in the Northeast representative basin. Furthermore, the ARs of the Caribbean Sea and Gulf of Mexico increases the 100-year flood magnitude by 350% in the South representative basin. The most impacted ARs on the 100-year flood in the Southeast representative basin is the local moisture which increases the magnitude to 600%, while the moisture of the Pacific Ocean rises the 100-year flood by 183% in the Southwest representative basin. In addition, the Pacific also levels the magnitude of 100-year flood up to 345% and 298% in the West and West North Central representative basins, respectively. However, ratios for the Northwest representative cannot be calculated since all the annual floods in the records are generated by ARs.

Figure 3.23, emphasizes on the role of ARs and their sources in the floods magnitude. It is obvious that ARs increase the flood magnitude as the real ratio, in many cases is placed in or above the upper quartile of the ratios calculated from the resampled flood

Figure 3.22: The PDFs of the 100-year flood ratios for the AMF-AR to the AMF-Non-AR events of the US climate region representative station based on multi and single populations of moisture source. The PDF (vertical line) refers to the ratios of resampled AMFs (real AMFs). Colors refer to the moisture source controlled the annual floods: gray refers to all ARs, blue (Pacific Ocean), red (Atlantic Ocean), green (Caribbean Sea and Gulf of Mexico), and purple (local moisture).

records. Further, some regions are dominated by a one distinct source of moisture, i.e. the West climate region representative. In conclusion, there are distinct differences in the magnitude of the 100-year flood estimated based on event's generated mechanism. The separation of the AMF-AR events based on the source of moisture shows the key difference in the AMF-AR magnitude. Furthermore, PDFs in Figure 3.22 proof that the actual ratios are happened not accidentally which is ensuring the impacts of ARs and their originating sources on the floods magnitude and frequency.
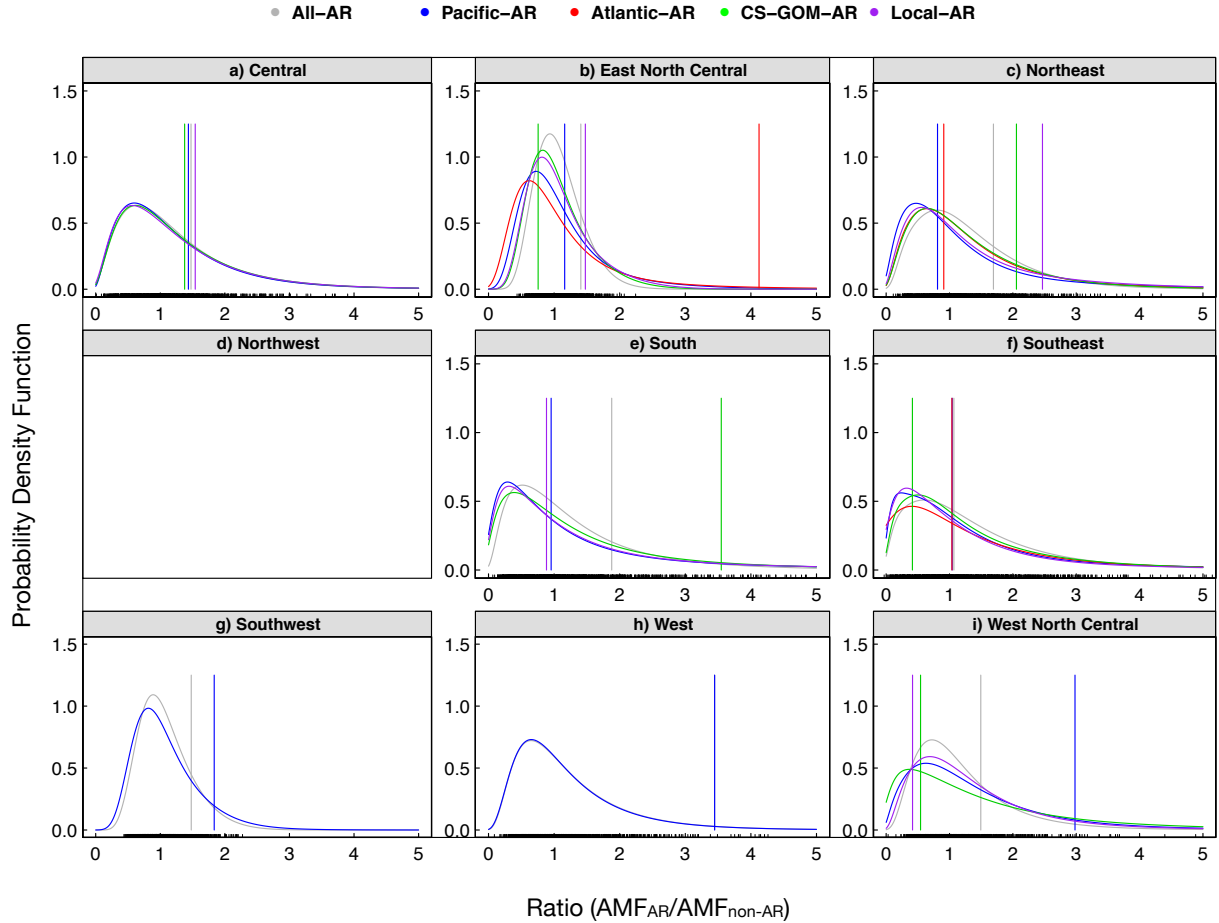
Figure 3.23: The boxplots of the 100-year flood ratios for the AMF-AR to the AMF-Non-AR events of the US climate region representative station based on multi and single populations of moisture source. Boxplot (star) refers to the ratios of resampled AMFs (real AMFs). Colors refer to the moisture source controlled the annual floods: gray refers to all ARs, blue (Pacific Ocean), red (Atlantic Ocean), green (Caribbean Sea and Gulf of Mexico), and purple (local moisture).

3.7 Discussion and Conclusions

This study examined the role of ARs and their moisture trajectories and sources on the heterogeneous flooding events of annual maximal flows in the conterminous United States, and their impact on the magnitude and frequency estimates used for the design of flood structures. Flood frequency analysis in different regions of the country often contains annual floods generated by distinctive different hydrologic and hydroclimatic processes (Waylen and Woo, 1982; Hirschboeck, 1991; Webb and Betancourt, 1990; Berghuijs et al.,
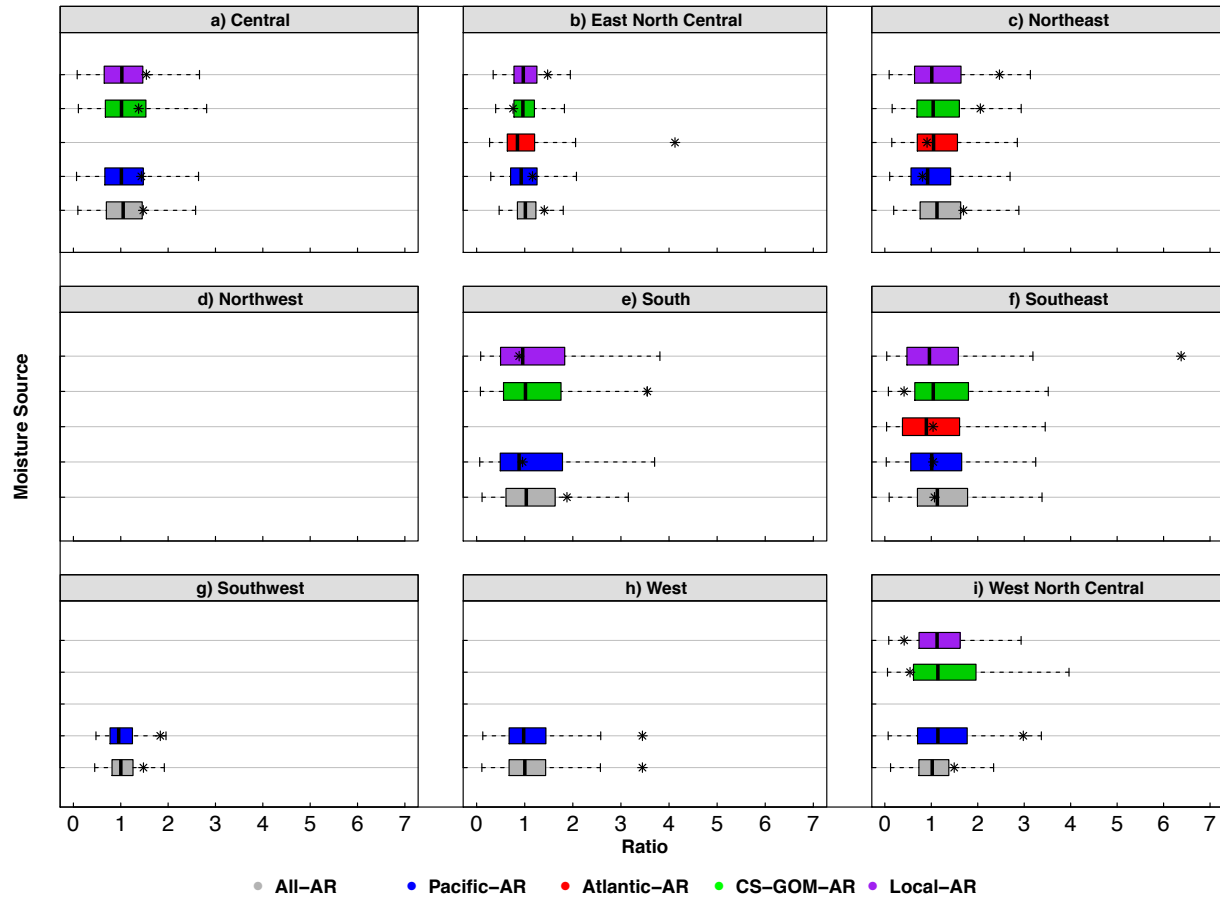
Table 3.6: The ratio of 100-year flood (cfs) for the AMF−AR to the AMF−Non−AR events of the US climate region representative station based on multi and single populations of moisture source. Number in parenthesis is the quantile of actual ratio with regarding to the quantiles of sampling ratios. Bold numbers refers to quantile 0.75 or higher.

No.	All-AR/ Non-AR	PO-AR/ Non-AR	AO-AR/ Non-AR	CS-GM-AR/ Non-AR	Local-AR/ Non-AR
1	1.48 (**0.76**)	1.44 (0.73)	–	1.38 (0.69)	1.54 (**0.78**)
2	1.41 (**0.87**)	1.16 (0.69)	4.12 (**0.98**)	0.76 (0.22)	1.48 (**0.85**)
3	1.69 (**0.77**)	0.81 (0.41)	0.91 (0.40)	2.06 (**0.86**)	2.47 (**0.86**)
4	–	–	–	–	–
5	1.88 (**0.81**)	0.95 (0.53)	–	3.55 (**0.92**)	0.88 (0.47)
6	1.07 (0.46)	1.04 (0.51)	1.04 (0.56)	0.41 (0.15)	6.38 (**0.93**)
7	1.48 (**0.88**)	1.83 (**0.94**)	–	–	–
8	3.45 (**0.99**)	3.45 (**0.99**)	–	–	–
9	1.49 (**0.80**)	2.98 (**0.94**)	–	0.54 (0.20)	0.42 (0.7)

Climate Regions: 1=Central, 2=East North Central, 3=Northeast, 4=Northwest, 5=South, 6=Southeast, 7=Southwest, 8=West, and 9=West North Centrl.

2016). Among the different flood-generating mechanisms, ARs are responsible for large, regional-scale floods.

The analysis started by classifying the annual maximal flows of 623 USGS stream gauges, well distributed across the study area over the period 1956-2015, into five main flooding agents: a) Snowmelt, b) Local rain, c) Local rain on snow, d) AR-rain, and e) AR-rain on snow. Each flood event assigned to one category based on the thresholds of the AR Shape index, IVT, PR, SAT, and WEASD within a specific time of the AMF in the catchment (Figure 3.3). The results showed that AR is a major source of floods in large swathes of the US (Figure 3.4). Therefore, we examined the annual number of wet days across the study region and we concluded that ARs are flying over the country and considered as main source of moisture in the western half of the US (Figure 3.7).

The study calculated the fraction of AMF-AR events in the 623 USGS stream gauge stations. Within a window of five days (3 days before the AMF day and 1 day after), all ARs intersected the catchment boundaries were retained and tracked back to their time of birth. Then, their trajectories and sources were determined by using the weighted principal

curve approach. As a result, a single flood event can be induced by different AR episodes (Figure 3.6). As such, we showed that ARs contributed to a total of 25,725 out of 37,380 ($\sim$ 73%) AMF events. Further, the high impacts of ARs on AMF events (70-100%) found within the Northwest, West, Northeast, Southeast, South, Central, and East North Central climate regions, while lower effect of ARs were found in the West North Central and Southwest climate regions (Figure 3.8). In contrast, few studies (e.g., Lavers et al. 2011, 2012; Lavers and Villarini 2013; Barth et al. 2017) examined the AR-generated floods based on a single-point index over smaller spacial-scale. They considered an event as AR-generated flood where the peak discharge is within a specific distance at any side of the AR axis within 10 days which cause overestimated role of ARs. Furthermore, the current analysis showed the ARs affect the floods magnitude across different quantiles (Figure 3.9).

The new statistical framework based on the weighted principal curve approach was used to determine, for the first time, the integrated trajectories of the flood-induced ARs within their full episodes, to track these ARs back to their origins (time and place of birth). The type of analyses presented here enabled to plot the major axis or pathway of the flood-caused ARs as they count for the maximum IVT grid points in each time-step of the AR life cycle to draw the curvilinear pathway. Furthermore, fractional contribution of AMF-AR events were able to be calculated based on the source of moisture. As a result, the major sources of moisture that control floods in the US climate regions have been identified. The percentages of the sources contribution to AMF-AR events and their trajectories were shown in Section 3.6.3 (see Figure 3.11, Figure 3.12, Table 3.4, and Table 3.5). Few studies in the literature tried to determine the sources and tracks of the flood-induced moisture. Some studies (e.g., Lavers et al. 2011, 2012; Lavers and Villarini 2013; Barth et al. 2017) constructed the AR axis by retaining the point with the maximum value of the IVT and others found the moisture sources by using digital metrological maps and images (Schlef et al., 2019). However, these studies may give incomplete or inaccurate results about the important roles of ARs in flood frequency analysis.

To employ the new framework in the flood frequency analysis, the current study showed that separating the mixed population events to AR and Non-AR generated floods agents did not improve the LP3 fit unless most events were caused by only one source of moisture such as the case in the Northwest region (Figure 3.16(d)). Conversely, the results indicated satisfactory fit for the LP3 distribution of the AR-source based floods (Figure 3.16). Furthermore, the results showed the impact of source-based ARs analysis on the designated magnitude of the 100-year flood. As a result, the AR-source based floods showed that the 100-flood magnitude could be up to 650% larger than the 100-year flood estimated based on the AMF-Non-AR events (see (Figure 3.22) and Table 3.6). However, what other studies (e.g., Barth et al. 2017, 2019) have done was to separate the mixed population peak flows to Non-AR and AR homogeneous floods which improved the LP3 fit for only specific cases as mentioned early in the section.

A reliable estimation of flood frequency and magnitude to anticipate future extremes is vital for life security and property protection. Furthermore, having precise knowledge about the nature of floods variability will improve the present strategies for future water resources management and flood risks mitigation. A fundamental assumption, i.e. iid, which considers floods timeseries as time sample of random homogeneous events (as in Bulletin 17-B and 17-C) does not apply in regions of mixed population flood events. Therefore, the type of results presented here gave a comprehensive understanding about the flood variability by accounting for distinctly different hydrologic and hydroclimatic flooding processes in at-site scale based analyses.

CHAPTER 4
SUMMARY AND CONCLUSIONS

This chapter closes the dissertation with an overview of the research findings (Section 4.1). The future research directions are discussed in Section 4.2. Section 4.3 summaries the original contributions of this research, while Section 4.4 provides a concluding statement for this thesis.

4.1   Research Overview

As described in Chapter 1, the main objective of this research is understanding nonstationarity in water resources and assessing its impacts on water management and decision making. As Earth's climate is increasingly undergoing anthropogenic and natural changes, water resources management and flood risk control become more challenging. As such, changes in hydrologic regime negatively impact water resources reliability. Consequently, the operation of the water resources structures (e.g., water supply reservoir) deteriorate under the uncertainties and risks associated with streamflow variability. This work has sought to address these issues in two separated chapters.

The results presented in Chapter 2 demonstrate a comprehensive diagnosis study of streamflow temporal variation and nonstationarity impacts on a reservoir's operation by underscoring three points. First, the eleven centuries record length of Feather River inflow into Lake Oroville (FRI) shows substantial variations in interannual to multidecadal and centennial time-scales in the Feather River Basin (FRB) hydrologic regime. As a result, these variations in the 50-year rolling window of the FRI statistics (mean, variance, persistence) indicate that the temporal variation and nonstationarity are embedded in the streamflow records. Second, changes in the FRI variance and persistence of 50-year blocks are the key of variability and source of surprise in the system decision variables. As such,

the reservoir's storage requirement and performance (RRV) significantly change due to slight differences in the FRI variance and/or persistence.

Lastly, the wavelet coherence analysis demonstrates that changes in reservoir storage requirements and performance (RRV) in many instances and periods are due to the changing co-variability between the FRI and climate teleconnection patterns (e.g., ENSO and PDO). The correlations of the FRI-climate indices and the statistical characteristics (mean, standard deviation) of ENSO and PDO also show some indication of the systematically varying storage requirement characteristics. These results lead to the conclusion that large storage requirements are associated with high variance of climate drivers. On the other hand, smaller storage requirements occur in periods of high correlation between the FRI and the ENSO components. To this end, the expectations of an increase in the extreme ENSO, which in turn cause extreme weather events, in the 21$^{\text{st}}$ century (Cai et al., 2014) may lead to changes in the storage requirements to maintain requisite reliability levels.

Chapter 3 underscores the issues of nonstationarity by developing a comprehensive statistical framework to provide satisfactory estimates of flood frequency under climate change conditions. The analysis started by classifying the annual maxima flows (AMF) of 623 USGS stream gauges, well distributed across the study area over the period 1956-2015, based on hydro-meterological databases into five main flooding agents: a) snowmelt, b) local rain, c) local rain on snow, d) AR-rain, and e) AR-rain on snow. Each flood event assigned to one category based on the thresholds of the AR Shape index, IVT, PR, SAT, and WEASD in a specific time of the AMF within the catchment boundaries. The results show that AR is a major source of floods in large swathes of the US. As a result, a total of 25,725 out of 37,380 ($\sim 73\%$) AMF events in the 623 stations were generated by ARs over a 60 year period. Results show ARs have a higher impact on AMF events (70-100%) found within the Northwest, West, Northeast, Southeast, South, Central, and East North Central climate regions, while ARs have less impact on the West North Central and Southwest

climate regions. Furthermore, the results show that ARs affect flood magnitude across different quantiles with higher influences in the arid and semi-arid regions (e.g., West and Southwest).

In terms of quantifying the risk of flood variability, the new statistical framework enables tracking of the flood-induced ARs back to origin. As a results, the major sources of moisture that control the floods in the US climate regions are identified and the percentages of source contribution to AMF-AR events and their trajectories are determined. Distinct source(s) of moisture are identified as the major cause of floods in a selected region. Thus, determining the nature of variability in floods.

Finally, to employ the new framework in the flood frequency analysis, this work showed that separating the mixed population events to AR and Non-AR generated floods agents did not improve the LP3 fit unless most events were caused by only one source of moisture such as the case in the Northwest region . The results indicated satisfactory fit for the LP3 distribution of the AR-source based floods in most cases. Furthermore, the results showed the impact of source-based ARs analysis on the designated magnitude of the 100-year flood indicating that the 100-flood magnitude could be up to 650% larger than the 100-year flood estimated based on the AMF-Non-AR events.

Taken together, the results presented in this dissertation introduce new insights and augments the knowledge base for water resources management and flood risk mitigation in the era of nonstationarity and climate change.

4.2   Further Research Directions

The hope is to continue the efforts on the topic of nonstationarity and its applications to water resources planning and management. As a fresh prospective on uncertainties and risks in water resources systems operation under climate change conditions has been introduced in Chapter 2 in a regional scale, a US-scale analysis is needed to plot a national

map that shows the changing climate impacts on water resources reliability. Doing so will provide climate-based inform about anticipated deterioration in systems operation.

Given the consideration of mixed population floods in the flood frequency analysis, the use of long-term records is needed for performing the LP3 on flood categories to have reliable estimate of floods frequency based on hydrological processes classification. Finally, studying the seasonality of flood-induced ARs and investigation of the shift and change in time and the IVT magnitude is a natural expansion to Chapter 3 that aid in the understanding of nonstationarity.

4.3   Original Research Contributions

Taking the work in chapters 2 and 3 together, this dissertation has provided contributions to the areas of water resources management and floods risk mitigation. As many studies in the topic of nonstationarity have examined the impacts of climate change on the hydrologic regime, insights on the historical impacts of climate variability on water resources systems are needed to anticipate future hydrological changes and adopt management strategies. The results in Chapter 2 allow for a quantitative assessment of the relationship between streamflow and climate indices (e.g., ENSO, PDO) which is resolved at low-frequency bands from interannual to multidecadal and centennial time-scales. The coherence estimates provide a clearer interpretation of swings in the decision variables that occur during eras of high coherence between the streamflow and climate drivers at the select timescale. As a result. a comprehensive framework that evaluate a water system's operation under changing climate condition is established to inform reliable and robust management of water resources systems.

In the field of flood frequency, the Work Group that updated the guidelines for performing flood frequency in B-17C did not evaluate methods to account for climate variability. As such, in the 'Future Studies' section, the Work Group of B-17C identified the need for the identification and treatment of mixed distributions, including those based

on hydrometeorological or hydrological conditions (England et al., 2018). The new methodological developments described in Chapter 3 specifically addresses this limitation. The improvements associated with this statistical framework not only leads to a better characterization of the observational records and identification of the process-based flood frequency analysis, but it also quantifies flood variability by determining the nature of this variability. By addressing this issue listed in B-17C 'Futures Studies' section, the studies conducted in this work add to a growing body of literature continuing to build a more solid framework based on physical processes to be used in the revisions of B-17C.

4.4   Concluding Statement

Before conducting the current research, the literature of water resources managements lacked complementary frameworks that addressed the issues of nonstationarity. The original contributions of this work to the other studies in the topic of nonstationarity added two concrete frameworks: a) the knowledge of streamflow-climate linkages into water resources management to avoid system's deterioration under the future variability, and b) large atmospheric circulation to inform the nature of variability in floods. Taken together, a whole picture of nonstationarity impacts on water resources management is drawn and an integrated framework to solve these issues is built.

All in all, our ability to anticipate future hydrology and integrate that knowledge into design and planning is thus well informed by analysis of the type presented here. It is hoped that, alongside other emerging work on the topic of nonstationarity and its applications to water resources planning and management, this work will aid in providing a fresh perspective.

REFERENCES

Aljoda, A., Jain, S., 2020. Uncertainties and risks in reservoir operations under changing hydroclimatic conditions. Journal of Water and Climate Change doi:`10.2166/wcc.2020.133`.

Barth, N.A., Villarini, G., Nayak, M.A., White, K., 2017. Mixed populations and annual flood frequency estimates in the western united states: The role of atmospheric rivers. Water Resources Research 53, 257–269.

Barth, N.A., Villarini, G., White, K., 2019. Accounting for mixed populations in flood frequency analysis: bulletin 17c perspective. Journal of Hydrologic Engineering 24, 04019002.

Berghuijs, W.R., Woods, R.A., Hutton, C.J., Sivapalan, M., 2016. Dominant flood generating mechanisms across the united states. Geophysical Research Letters 43, 4382–4390.

Brunner, M.I., Gilleland, E., Wood, A., Swain, D.L., Clark, M., 2020. Spatial dependence of floods shaped by spatiotemporal variations in meteorological and land-surface processes. Geophysical Research Letters 47, e2020GL088000.

CADWR, 2009. The State Water Project Delivery Reliability Report. Technical Report. California State Department of Water Resources. Sacramento, CA. URL: `http://baydeltaoffice.water.ca.gov/swpreliability/index.cfm`.

Cai, W., Borlace, S., Lengaigne, M., Van Rensch, P., Collins, M., Vecchi, G., Timmermann, A., Santoso, A., McPhaden, M.J., Wu, L., et al., 2014. Increasing frequency of extreme El Niño events due to greenhouse warming. Nature Climate Change 4, 111.

Cayan, D.R., Redmond, K.T., Riddle, L.G., 1999. Enso and hydrologic extremes in the western united states. Journal of Climate 12, 2881–2893.

Cook, E., D'Arrigo, R., Anchukaitis, K., 2008. Enso reconstructions from long tree-ring chronologies: Unifying the differences, in: Talk presented at a special workshop on Reconciling ENSO Chronologies for the Past.

Dettinger, M., 2011. Climate change, atmospheric rivers, and floods in california–a multimodel analysis of storm frequency and magnitude changes 1. JAWRA Journal of the American Water Resources Association 47, 514–523.

Dettinger, M.D., Cayan, D.R., McCabe, G.J., Marengo, J.A., 2000. Multiscale streamflow variability associated with El Nino/Southern oscillation. Cambridge University Press.

Dickinson, J.E., Harden, T.M., McCabe, G.J., 2019. Seasonality of climatic drivers of flood variability in the conterminous united states. Scientific reports 9, 1–10.

Dougherty, E., Rasmussen, K.L., 2019. Climatology of flood-producing storms and their associated rainfall characteristics in the united states. Monthly Weather Review 147, 3861–3877.

England, J., Cohn, T., Faber, B., Stedinger, J., Thomas, W., Veilleux, A., Kiang, J., Mason, R., 2018. Guidelines for determining flood flow frequency—Bulletin 17C. Technical Manual - Survey Techniques and Methods 4 - B5. U.S. Geological Survey. URL: https://doi.org/10.3133/tm4B5.

Falcone, J.A., 2011. GAGES-II: Geospatial attributes of gages for evaluating streamflow. Technical Report. US Geological Survey.

Foufoula-Georgiou, E., Kumar, P., Mukerji, T., Mavko, G., 1995. Wavelets in geophysics. Pure and Applied Geophysics 145, 374–375.

Gimeno, L., Dominguez, F., Nieto, R., Trigo, R., Drumond, A., Reason, C.J., Taschetto, A.S., Ramos, A.M., Kumar, R., Marengo, J., 2016. Major mechanisms of atmospheric moisture transport and their role in extreme precipitation events. Annual Review of Environment and Resources 41, 117–141.

Gimeno, L., Nieto, R., Vázquez, M., Lavers, D.A., 2014. Atmospheric rivers: A mini-review. Frontiers in Earth Science 2, 2.

Grinsted, A., Moore, J.C., Jevrejeva, S., 2004. Application of the cross wavelet transform and wavelet coherence to geophysical time series. Nonlinear Processes in Geophysics 11, 561–566.

Guan, B., Waliser, D.E., 2015. Detection of atmospheric rivers: Evaluation and application of an algorithm for global studies. Journal of Geophysical Research: Atmospheres 120, 12514–12535.

Hashimoto, T., Stedinger, J.R., Loucks, D.P., 1982. Reliability, resiliency, and vulnerability criteria for water resource system performance evaluation. Water Resources Research 18, 14–20.

Hastie, T., Stuetzle, W., 1989. Principal curves. Journal of the American Statistical Association 84, 502–516.

Hirschboeck, K.K., 1991. Climate and floods. US Geol. Surv. Water-Supply Pap 2375, 67–88.

Ho, M., Lall, U., Sun, X., Cook, E.R., 2017. Multiscale temporal variability and regional patterns in 555 years of conterminous us streamflow. Water Resources Research 53, 3047–3066.

Hossain, F., Degu, A.M., Yigzaw, W., Burian, S., Niyogi, D., Shepherd, J.M., Pielke Sr, R., 2012. Climate feedback–based provisions for dam design, operations, and water management in the 21st century. Journal of Hydrologic Engineering 17, 837–850.

Jain, S., Eischeid, J.K., 2008. What a difference a century makes: Understanding the changing hydrologic regime and storage requirements in the Upper Colorado River basin. Geophysical Research Letters 35.

Jain, S., Hoerling, M., Eischeid, J., 2005. Decreasing reliability and increasing synchroneity of western north american streamflow. Journal of Climate 18, 613–618.

Jain, S., Lall, U., 2000. Magnitude and timing of annual maximum floods: Trends and large-scale climatic associations for the blacksmith fork river, utah. Water Resources Research 36, 3641–3651.

Jain, S., Lall, U., 2001. Floods in a changing climate: Does the past represent the future? Water Resources Research 37, 3193–3205.

Jain, S., Woodhouse, C.A., Hoerling, M.P., 2002. Multidecadal streamflow regimes in the interior western united states: Implications for the vulnerability of water resources. Geophysical Research Letters 29.

Kahya, E., Dracup, J.A., 1992. The relationship between enso events and california streamflows, in: AIP Conference Proceedings, American Institute of Physics. pp. 86–95.

Karl, T., Koss, W.J., 1984. Regional and national monthly, seasonal, and annual temperature weighted by area, 1895-1983 .

Kashelikar, A., Griffis, V., 2008. Forecasting flood risk with bulletin 17b lp3 model and climate variability, in: World Water and Environmental Resources Congress, American Society of Civl Engineers, edited by Babcock RW and Walton R, Honolulu, Hawaii.

Khaliq, M., Ouarda, T., 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, 534–552.

Konrad, C.P., Dettinger, M.D., 2017. Flood runoff in relation to water vapor transport by atmospheric rivers over the western united states, 1949–2015. Geophysical Research Letters 44, 11–456.

Lau, K., Weng, H., 1995. Climate signal detection using wavelet transform: How to make a time series sing. Bulletin of the American Meteorological Society 76, 2391–2402.

Lavers, D.A., Allan, R.P., Wood, E.F., Villarini, G., Brayshaw, D.J., Wade, A.J., 2011. Winter floods in britain are connected to atmospheric rivers. Geophysical Research Letters 38.

Lavers, D.A., Villarini, G., 2013. Atmospheric rivers and flooding over the central united states. Journal of Climate 26, 7829–7836.

Lavers, D.A., Villarini, G., Allan, R.P., Wood, E.F., Wade, A.J., 2012. The detection of atmospheric rivers in atmospheric reanalyses and their links to british winter floods and the large-scale climatic circulation. Journal of Geophysical Research: Atmospheres 117.

Li, J., Xie, S.P., Cook, E.R., Morales, M.S., Christie, D.A., Johnson, N.C., Chen, F., D'Arrigo, R., Fowler, A.M., Gou, X., et al., 2013. El Niño modulations over the past seven centuries. Nature Climate Change 3, 822.

Lu, M., Lall, U., 2016. Tropical moisture exports, extreme precipitation and floods in northeast us. Hydrology and Earth System Sciences Discussions , 1–40.

MacDonald, G.M., Case, R.A., 2005. Variations in the Pacific Decadal Oscillation over the past millennium. Geophysical Research Letters 32.

Mallakpour, I., Villarini, G., 2016. Investigating the relationship between the frequency of flooding over the central united states and large-scale climate. Advances in Water Resources 92, 159–171.

Mantua, N.J., Hare, S.R., Zhang, Y., Wallace, J.M., Francis, R.C., 1997. A Pacific Interdecadal Climate Oscillation with Impacts on Salmon Production. Bulletin of the American Meteorological Society 78, 1069–1079.

McMahon, T.A., Adeloye, A.J., Zhou, S.L., 2006. Understanding performance measures of reservoirs. Journal of Hydrology 324, 359–382.

McMahon, T.A., Pegram, G.G., Vogel, R.M., Peel, M.C., 2007. Review of gould–dincer reservoir storage–yield–reliability estimates. Advances in Water Resources 30, 1873–1882.

Milly, P.C., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., Stouffer, R.J., 2008. Stationarity is dead: Whither water management? Science 319, 573–574.

Minobe, S., 2000. Spatio-temporal structure of the pentadecadal variability over the north pacific. Progress in Oceanography 47, 381–408.

Nakamura, J., Lall, U., Kushnir, Y., Robertson, A.W., Seager, R., 2013. Dynamical structure of extreme floods in the us midwest and the united kingdom. Journal of Hydrometeorology 14, 485–504.

Neiman, P.J., Ralph, F.M., Wick, G.A., Kuo, Y.H., Wee, T.K., Ma, Z., Taylor, G.H., Dettinger, M.D., 2008. Diagnosis of an intense atmospheric river impacting the pacific northwest: Storm summary and offshore vertical structure observed with cosmic satellite retrievals. Monthly Weather Review 136, 4398–4420.

Neiman, P.J., Schick, L.J., Ralph, F.M., Hughes, M., Wick, G.A., 2011. Flooding in western washington: The connection to atmospheric rivers. Journal of Hydrometeorology 12, 1337–1358.

Newell, R.E., Newell, N.E., Zhu, Y., Scott, C., 1992. Tropospheric rivers?–a pilot study. Geophysical research letters 19, 2401–2404.

Paltan, H., Waliser, D., Lim, W.H., Guan, B., Yamazaki, D., Pant, R., Dadson, S., 2017. Global floods and water availability driven by atmospheric rivers. Geophysical Research Letters 44, 10–387.

Poff, N.L., Olden, J.D., 2017. Can dams be designed for sustainability? Science 358, 1252–1253.

Ralph, F., Dettinger, M., 2011. Storms, floods, and the science of atmospheric rivers. Eos, Transactions American Geophysical Union 92, 265–266.

Ralph, F.M., Neiman, P.J., Wick, G.A., 2004. Satellite and caljet aircraft observations of atmospheric rivers over the eastern north pacific ocean during the winter of 1997/98. Monthly Weather Review 132, 1721–1745.

Ralph, F.M., Neiman, P.J., Wick, G.A., Gutman, S.I., Dettinger, M.D., Cayan, D.R., White, A.B., 2006. Flooding on california's russian river: Role of atmospheric rivers. Geophysical Research Letters 33.

Ropelewski, C.F., Halpert, M.S., 1987. Global and regional scale precipitation patterns associated with the el niño/southern oscillation. Monthly Weather Review 115, 1606–1626.

Rösch, A., Schmidbauer, H., 2016. Waveletcomp: A guided tour through the r-package, 2014.

Rutz, J.J., Guan, B., Bozkurt, D., Gorodetskaya, I.V., Gershunov, A., Lavers, D.A., Mahoney, K.M., Moore, B.J., Neff, W., Neiman, P.J., et al., 2020. Global and regional perspectives, in: Atmospheric Rivers. Springer, pp. 89–140.

Rutz, J.J., Steenburgh, W.J., Ralph, F.M., 2014. Climatological characteristics of atmospheric rivers and their inland penetration over the western united states. Monthly Weather Review 142, 905–921.

Sagarika, S., Kalra, A., Ahmad, S., 2015. Interconnections between oceanic–atmospheric indices and variability in the us streamflow. Journal of Hydrology 525, 724–736.

Schlef, K.E., Moradkhani, H., Lall, U., 2019. Atmospheric circulation patterns associated with extreme united states floods identified via machine learning. Scientific reports 9, 1–12.

Stedinger, J.R., Griffis, V.W., 2008. Flood frequency analysis in the united states: Time to update.

Stedinger, J.R., Griffis, V.W., 2011. Getting from here to where? flood frequency analysis and climate 1. JAWRA Journal of the American Water Resources Association 47, 506–513.

Torrence, C., Compo, G.P., 1998. A practical guide to wavelet analysis. Bulletin of the American Meteorological society 79, 61–78.

Trenberth, K.E., Hurrell, J.W., 1994. Decadal atmosphere-ocean variations in the pacific. Climate Dynamics 9, 303–319.

Vogel, R.M., Stedinger, J.R., 1987. Generalized storage-reliability-yield relationships. Journal of Hydrology 89, 303–327.

Waylen, P., Woo, M.k., 1982. Prediction of annual floods generated by mixed processes. Water Resources Research 18, 1283–1286.

Webb, R.H., Betancourt, J.L., 1990. Climatic effects on flood frequency: An example from southern arizona .

White, A.B., Moore, B.J., Gottas, D.J., Neiman, P.J., 2019. Winter storm conditions leading to excessive runoff above california's oroville dam during january and february 2017. Bulletin of the American Meteorological Society 100, 55–70.

Zhu, Y., Newell, R.E., 1998. A proposed algorithm for moisture fluxes from atmospheric rivers. Monthly Weather Review 126, 725–735.

BIOGRAPHY OF THE AUTHOR

Ali Aljoda was born and raised in Almansoor City, Baghdad, Iraq, on August 21, 1983. He graduated from Almansoor Elementary – Mid - High Schools in 1995, 1998, and 2001. Ali attended the University of Baghdad, Baghdad, Iraq, and graduated in 2006 with a Bachelor's of Science degree (B.Sc.) in Water Resources Engineering. He attended the Universiti Putra Malaysia, Serdang, Selangor Darul Ehsan, Malaysia, and graduated in 2009 with a Master of Engineering degree (M.Eng.) in Structural Engineering and Constructions. Ali joined the center of restoration of Iraq marshlands and wetlands – the Iraqi Ministry of Water Resources, Baghdad, Iraq in 2010 as a senior engineer to conduct research and perform projects on restoring the Iraqi marshes. During his employment, Ali traveled for three months from September – December 2012 to the kingdom of Netherlands to attend a post diploma program in project management at the Maastricht School of Management, Maastricht, The Netherlands. In October 2014, he traveled to the United States to attend an academic English language program for the graduate study preparations at the Intensive English Institute (IEI) - the University of Maine, Orono, Maine, USA. He then joined the University of Maine in Summer 2015 as a Doctor of Philosophy student in the department of Civil and Environmental Engineering as a sponsored student from the High Committee for Education Development in Iraq (HCED-Iraq). Ali is a candidate for the Doctor of Philosophy degree in Civil Engineering from the University of Maine in May of 2021. Ali Aljoda is a candidate for the Doctor of Philosophy degree in Civil Engineering from The University of Maine in May 2021.