THE RESPONSE OF DROUGHT AND PRECIPITATION VARIABILITY TO REGIONAL CLIMATE FORCING IN NORTHEAST PUERTO RICO

by

CRAIG ALLEN RAMSEYER

(Under the Direction of Thomas L. Mote)

ABSTRACT

Northeast Puerto Rico is home to a wide range of terrestrial and aquatic climate-sensitive ecosystems. Climate disturbances such as extreme events (e.g. hurricanes) and drought have cascading impacts on the biota. The biota responds to changes in precipitation variability on daily and sub-daily time scales. As a result, high temporal and spatial resolution climate data are needed to adequately assess climate impacts on the ecological process occurring in the region. This dissertation analyzes past, present, and future precipitation variability at a highly localized scale in northeast Puerto Rico. Additionally, a more comprehensive understanding of the regional climate forcing on precipitation variability is achieved. Artificial neural networks are used to downscale synoptic scale atmospheric variables to precipitation. These tools allow for modeling precipitation and determining the atmospheric processes driving precipitation variability

This dissertation finds that precipitation throughout Puerto Rico is driven primarily by variability in specific humidity and wind shear in the low-troposphere. The driest daily precipitation in northeast Puerto Rico is observed in synoptic environments with high wind shear and low moisture at 700 hPa. Both of these atmospheric variables

are driven by changes in the north Atlantic sea-surface temperature, the Saharan Air Layer, and the North Atlantic Subtropical High. The historical record shows little linear trend in total precipitation, however, precipitation variability is shown to be changing especially during the early rainfall season. It is posited that this increase in variability could be in part due to changes in the mechanisms driving the Caribbean Mid-Summer Drought. Future precipitation in northeast Puerto Rico is likely to be more variable with an overall drying trend. The highest magnitude changes are expected to occur in the early rainfall season as the trade wind inversion strengthens and wind shear across the region increases. These changes will cause disruptions to precipitation processes across several scales of motion, from tropical storm development to deep, moist convection. These trends in precipitation will likely cause significant impacts to the ecosystems of northeast Puerto Rico.

INDEX WORDS: Tropical Climatology, Climate Modeling, Precipitation Variability, Puerto Rico Climate, Climate Change, Artificial Neural Networks

THE RESPONSE OF DROUGHT AND PRECIPITATION VARIABILITY TO REGIONAL CLIMATE FORCING IN NORTHEAST PUERTO RICO

by

CRAIG ALLEN RAMSEYER

B.S., James Madison University, 2009

M.S., University of Georgia, 2011

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2016

THE RESPONSE OF DROUGHT AND PRECIPITATION VARIABILITY TO REGIONAL CLIMATE FORCING IN NORTHEAST PUERTO RICO

by

CRAIG ALLEN RAMSEYER

Major Professor:	Thomas L. Mote
Committee:	J. Marshall Shepherd
	Andrew J. Grundstein
	Alan P. Covich
	Douglas W. Gamble

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2016

DEDICATION

This dissertation is dedicated to my family. They have provided me with a lifetime full of support, love, and appreciation for my education. Also, this journey would not have been possible without the extensive and unconditional support of my partner and best friend. This dissertation is also dedicated to my son whom gave me newfound perspective on what is important in life.

ACKNOWLEDGEMENTS

I would like to thank Dr. Tom Mote an outstanding mentor and teacher. I consider it a privilege and honor to have been mentored by him. I thank him for his extensive support, encouragement, and advice. Additionally, I would like to thank my committee, Drs. Marshall Shepherd, Andy Grundstein, Alan Covich, and Doug Gamble for their insight and suggestions. I also extend a special thank you to Drs. Mote, Shepherd, and Grundstein for writing valuable letters of recommendation on my behalf. This dissertation was made possible due to all of the wonderful and encouraging educators I have had dating back to Kindergarten. They all played a role in making this possible. Lastly, I would like to thank all of my colleagues in the Climatological Research Laboratory, whom provided support, inspiration, and lasting friendship.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1. Introduction

Numerous studies have examined ecosystem responses in the Luquillo Experimental Forest (LEF) to disturbances and climate change. These responses indicate that ecosystems are sensitive to perturbations in climatic variability, particularly precipitation. For example, prolonged drought can alter rainfall and throughfall nutrient fluxes (e.g., Heartsill-Scalley et al. 2007) and reduce streamflow, creating barriers for migratory species (e.g., Pringle 1997; Benstead et al. 1999; Covich et al. 2006). These effects may become more frequent in tropical rainforests if climate change increases drought frequency (Larsen 2000; Giannini et al. 2001; Covich et al. 2003, Crook et al. 2009). Moreover, balancing water needs, between ecological integrity and extraction by expanding human development outside of the LEF, is an increasing problem due to low water storage capacity, the flashy nature of storm flows, and a high rate of loss from pipes (Pringle and Scatena, 1998). A water budget for the El Yunque National Forest developed in 2004, showed that on a typical day, 70 percent of water generated within the forest is diverted into municipal water supplies before reaching the ocean – up from an estimated 54% in 1994 (Crook et al. 2007).

Precipitation changes are likely to be significant and are likely more important than temperature change for wet tropical ecosystems (Mulligan and Burke 2005). The response of the ecosystem to these precipitation changes will be important for the future

of the LEF. This dissertation will define changes in the global and regional climate that produce the observed climate variability over the past several decades, as well as generate climate scenarios for sites within the Luquillo Experimental Forest (LEF). These high resolution climate analyses will provide valuable context for future changes to ecological and biogeochemical processes in the LEF.

1.2. Literature Review

Climate variability throughout Puerto Rico and Caribbean is strongly associated with global scale wave patterns, tropical cyclones, orographic effects, sea breeze circulations, regional scale wind patterns (primarily the easterly trade winds), and intense solar heating (Taylor et al. 2002). Thus, local scale understanding of climate and precipitation variability depends heavily on the understanding of the regional, synoptic-scale phenomena in the Caribbean (Comarazamy and González 2008; Comarazamy and Gonzalez 2011). The Caribbean has two distinct "rainy" seasons (Taylor et al. 2002; Chen and Taylor 2002; Angeles et al. 2010); the early rainfall season (ERS) begins in late March and extends into July for some locations in the Caribbean with precipitation peaking in May (Angeles et al. 2010). The second rainfall season lasts from August to November with a peak in October, partly associated with hurricanes and tropical storms (Taylor et al. 2002). This period is interrupted beginning in May and lasting until July in some locations by what is commonly called the midsummer dry spell (MSD) (Taylor et al. 2002; Curtis and Gamble 2008; Gamble et al. 2008; Angeles et al. 2010). This annual dry period is expected to intensify in the 21st Century as greenhouse warming is enhanced (Rauscher et al. 2008). The MSD is largely influenced by the intensification and expansion of the North Atlantic Subtropical High (NASH) into the Caribbean region in

May with the eastern Caribbean affected first as the MSD spreads westward (Magaña et al. 1999; Chen and Taylor 2002; Gamble et al. 2008). The eastern and interior Caribbean experiences a dry spell in April, May and June, while the western Caribbean experiences the MSD primarily in June and July (Gamble et al. 2008). The MSD appears to be more pronounced in the western and central Caribbean and on the leeward sides of islands. The bi-modal distribution of Caribbean annual precipitation has been shown to be related to a host of local and regional scale phenomena including vertical wind shear, atmospheric aerosols (e.g. dust), El-Niño-Southern Oscillation (ENSO), and the North Atlantic Oscillation (NAO) (Malmgren et al. 1998; Chen and Taylor 2002; Angeles et al. 2010). The MSD is detectable in eastern Puerto Rico precipitation records, but the signal is muted compared to other regions in the Caribbean basin.

Precipitation over much of Puerto Rico exhibits a bimodal distribution with a dry period separating the wet seasons (Malmgren and Winter 1999; Daly et al. 2003; Comarazamy and González 2008). Spatial precipitation patterns across Puerto Rico are most strongly associated with topography, exposure and direction of the predominant winds, and proximity to the ocean (Daly et al. 2003). Much of the annual precipitation occurs in intense showers from easterly waves and tropical disturbances that occur from May to October with the balance from northerly frontal systems and localized convection (Larsen 2000).

Precipitation increases at a rate of approximately 140% (of island average) per kilometer of elevation (Daly et al. 2003). In addition, the presence of urban areas influences precipitation patterns over Puerto Rico, but to a much lesser extent than topography (Comarazamy and González 2008). Coastal urban regimes have been shown

to enhance convection in the boundary layer in some locations (Shepherd et al. 2010). Thus, San Juan's urban landscape must be considered as a forcing on precipitation processes in the area. In Puerto Rico, the area of highest precipitation is in the Luquillo Mountains (LM) in northeast Puerto Rico (NE PR), where steep terrain, coupled with exposure to predominant northeasterly winds and proximity to the ocean leads to frequent orographic lifting of moisture laden air (Daly et al. 2003). The steady trade winds and steep orographic gradient of the LM and LEF dampen the MSD signal (D. Gamble, 2013, personal communication).

The topography and geography of NE PR provides a unique opportunity to analyze regional scale forcing on precipitation in a tropical montane regime. Due to the remoteness of many tropical locations, reliable precipitation data are sparse. Due to extensive scientific research being conducted in the LM since the 1940s, climate data is available at various sites around the region for long temporal periods. The topographic gradient does present some challenges and limitations. Due to the remote nature of some of the high elevation rain gauges, daily data collection of rainfall can be difficult. Additionally, easterly waves and other strong tropical low pressure systems can disrupt data logging and collection by clogging the rain gauges due to excess leaf litter or by damaging the rain gauge. Typically low-level sites have fewer data continuity issues. Despite these challenges, ample data are available for sites throughout NE PR. Two high elevation sites, one at Pico del Este (Fig.1.1) and the other at El Verde Field Station (Fig. 1.2), are examples of precipitation availability at high elevations.

The spatial distribution of rainfall across Puerto Rico is largely driven by the orientation of mountainous regions with the trade winds. On rainfall maps, the LM are

easily identifiable due to the steep gradient and high precipitation rates at high elevations (Fig. 1.3). The other areas of note are with regard to the localized minima in precipitation on the south and west facing slopes of the Cordillera Central which are oriented east to west across central Puerto Rico. The higher elevations of the LM have high daily rainfall rates year-round and receive over 10 mm of daily precipitation, even during the dry season (DJF) when much of the island experiences less than 5 mm of daily precipitation .

In general, Puerto Rico experiences four important inflection points in annual precipitation (Fig. 1.4). The dry season begins in December and persists through February and March. The ERS begins in late March and extends into July. There is a slight MSD signal in the form of a reduction of rainfall intensity in JJA. However, this signal is muted compared to the central and western Caribbean. There is a secondary peak in precipitation that occurs between August and November. This period is also the wettest period for southern PR as easterly waves and tropical low pressure systems disrupt low-tropospheric circulation and allow for advection of moisture onto the south facing slopes of the LM and the Cordillera Central. There is a high degree of spatial and temporal variability in days with no measureable precipitation (e.g. dry days) across the island (Fig. 1.5). The exception would be at high elevation in the LM where dry day frequency is low and relatively uniform across all seasons.

There is substantial inter-annual precipitation variability in the LM. This is most likely driven by synoptic and regional scale processes as the large scale flow primes the atmosphere for more localized, topographically enhanced processes (i.e. deep, moist convection). Using observed daily precipitation at El Verde Field Station (380 m) as a

representative historical time series of precipitation in the LM, inter-annual seasonal precipitation variability can be assessed.

The three dry seasons with the lowest precipitation have occurred since 2005 (Fig. 1.6). These three years had median rainfall values of zero. The ERS has the highest variability in daily rainfall among all three seasons (Fig. 1.7). Since 2007, the ERS has experienced years with median daily precipitation of 0 mm (2007) and the highest median value of 6.1 mm (2013). There is a large amount of variability in the extreme rainfall events as well with 95^{th} percentile rains as low as 9.6 mm (2014) and as high as 50 mm (2013). The late rainfall season exhibits the least variability which may be attributed to the frequency of organized tropical disturbances (Figure 1.8).

The ecology and biogeochemical cycles of the LM and LEF are highly climate sensitive. As a result, there are large responses in these systems to small climate disturbances, in particular hurricanes and drought. This dissertation provides high resolution precipitation analyses to better understand how precipitation variability has varied historically and is projected to change in the future. This provides critical context for the future of the biota and ecological processes in the LEF and LM.

1.3 Research Objectives

Most climate impact assessments rely on precipitation simulations from Global Climate Models (GCMs). While these data have important applications, higher resolution climate information is desirable, especially on Caribbean islands where complex topography leads to high spatial variability in precipitation. This dissertation addresses the need for high resolution climate information for NE PR in order to resolve the topographic and precipitation gradient that exists in the LM. Chapter 2 identifies the

atmospheric variables that are the largest forcing mechanisms on precipitation in Puerto Rico. This dissertation provides important findings for future climate downscaling studies in the region and provides a more comprehensive model for understanding the processes driving precipitation variability in Puerto Rico. Chapter 2 addresses the following research questions:

- Can artificial neural networks (ANNs) provide a reliable methodological tool to model precipitation in Puerto Rico?
- What atmospheric variables control precipitation across the island?
- What atmospheric levels are most important to precipitation processes at the study sites?
- Does geography, topography, and/or land cover influence which controls are most important?

Chapter 3 examines how atmospheric and precipitation variability in northeastern Puerto Rico has changed over the historical period of record from 1985 – 2014. This chapter addresses the following research questions:

- What are the leading modes of variability in low-tropospheric winds, moisture, and geopotential heights?
- What synoptic types represent environments that are conducive to high/low daily precipitation?
- Are there statistically significant changes in synoptic types over the historical period? Are there any trends in the types of environments that are increasing or decreasing in frequency?

Chapter 4 analyzes how future atmospheric variability, driven by anthropogenic climate change, will alter the hydroclimate of NE PR. A future with a changing hydroclimate has important ramifications for the ecology and biogeochemical cycles of the tropical forests as well as the economic future of the island with regard to municipal water supplies and tourism. The research questions addressed in this chapter are:

- Can a SOM-based empirical-dynamical downscaling approach adequately resolve historical precipitation variability?
- Can this downscaling methodology provide plausible scenarios of future precipitation variability? If so, how is precipitation variability projected to change?
- What synoptic types are projected to increase or decrease in frequency in the future? Do these changes in frequency provide information in how low-tropospheric processes are expected to change?
- Are the future environments properly captured by the historical SOM?
- Do the projected changes in low-tropospheric environments conform to or violate the stationarity assumption implicit in statistical downscaling methodologies?
- How are the low-tropospheric atmospheric variables projected to change in the study area and what can be inferred from these changes about future precipitation variability in NE PR?

1.4. Summary

This dissertation addresses the need for an improved understanding of how global to regional scale atmospheric forcing on precipitation variability in northeastern Puerto

Rico. The topography of the study area provides unique challenges in terms of achieving high resolution climate data and information. In order to address these challenges, the research objectives set forth in the proceeding discussion involve downscaling large-scale atmospheric variability to individual points in NE PR. This provides the highest resolution analysis available for this study. The overarching objectives in subsequent chapters examines the atmospheric controls on precipitation in Puerto Rico and then uses those atmospheric variables to downscale historical and future climate regimes to precipitation in NE PR. A comprehensive understanding of the climatic drivers of precipitation provides important information for the future of eastern Puerto Rico. Future changes in precipitation variability in eastern Puerto Rico can have important impacts on ecology, forest dynamics, water availability, and the islands' economy.

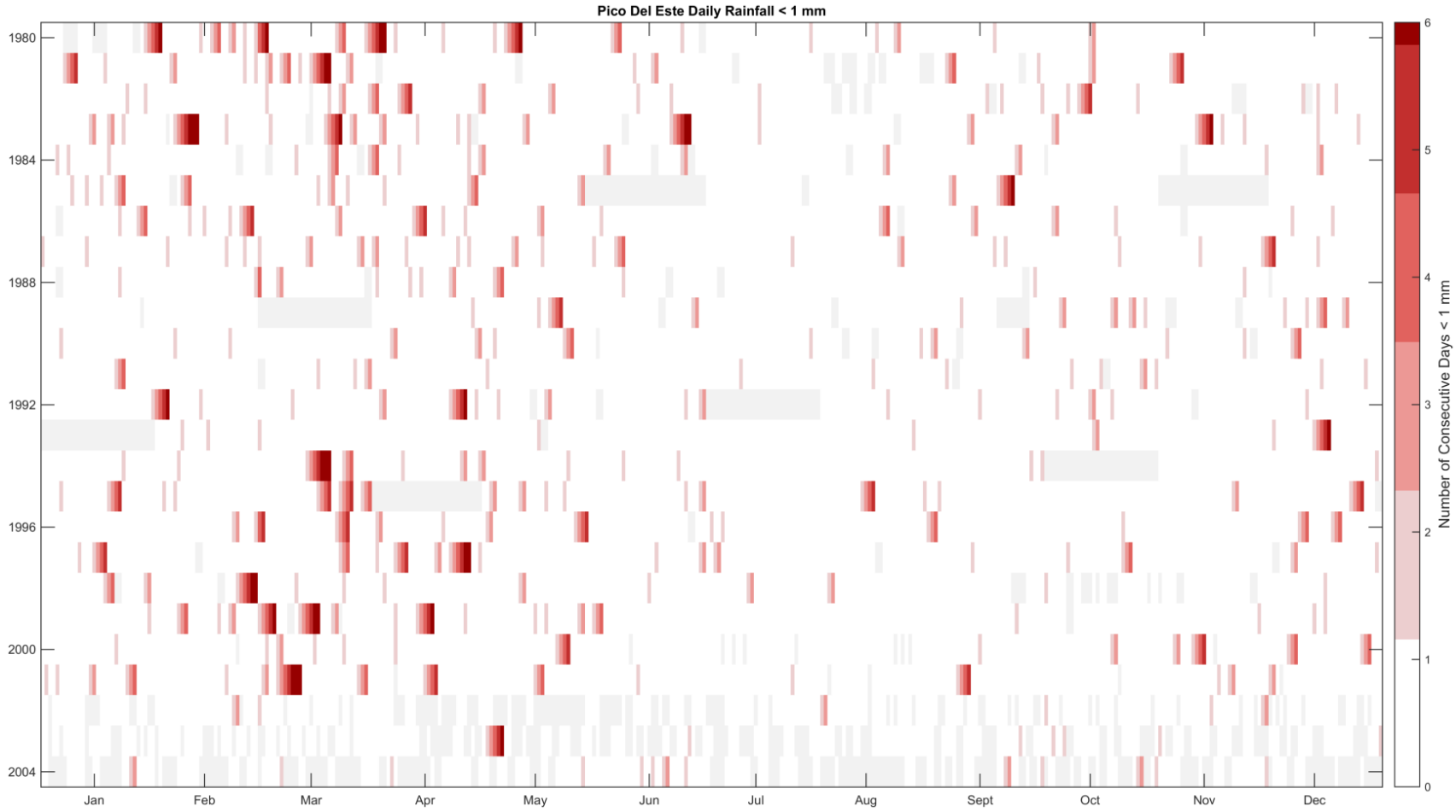

Fig. 1.1 Number of consecutive days with < 1 mm of rainfall at Pico Del Este with missing data represented by gray pixels (Data Source: GHCN-Daily).

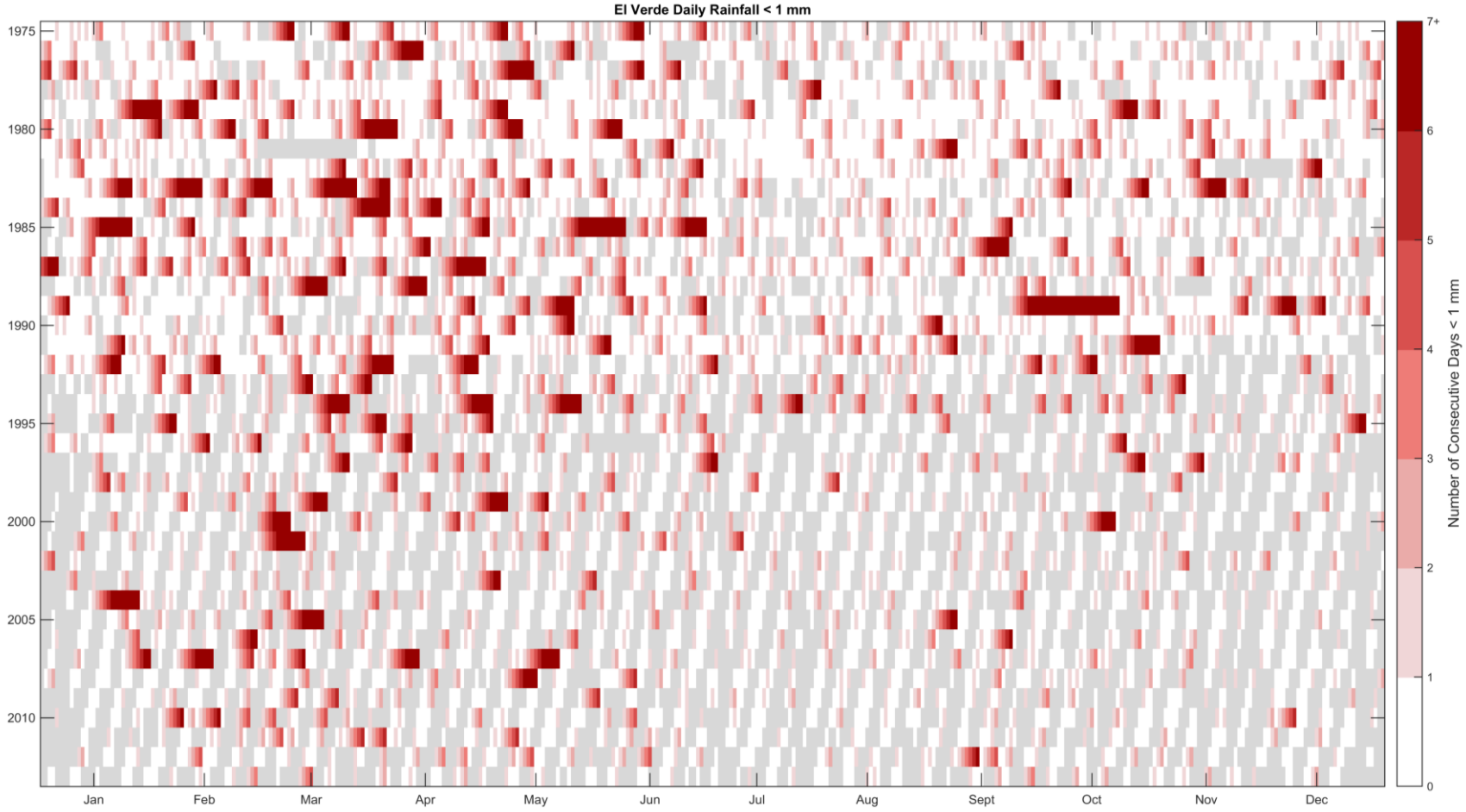

Fig. 1.2 Number of consecutive days with < 1 mm of rainfall at El Verde with missing data represented by gray pixels (Data Source: National Atmospheric Deposition Program).

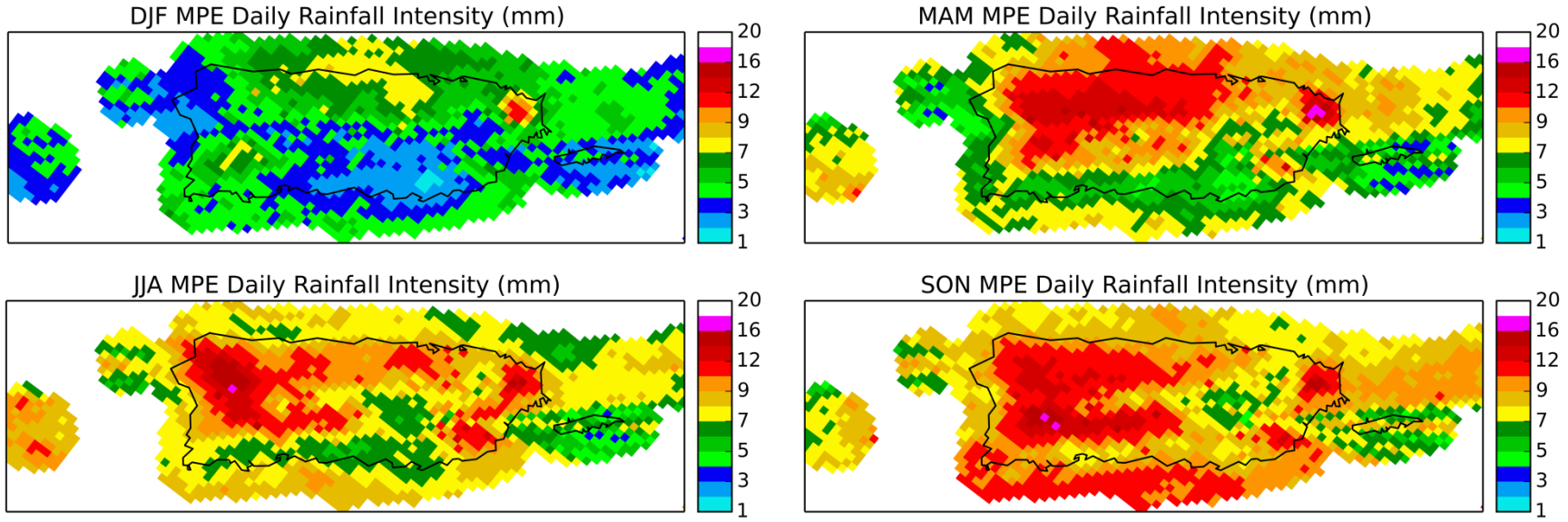

Fig.1.3 Multi-sensor precipitation estimates (MPE) of daily rainfall intensity. MPE integrate rain gauge and satellite rainfall estimates with the radar-only estimates.

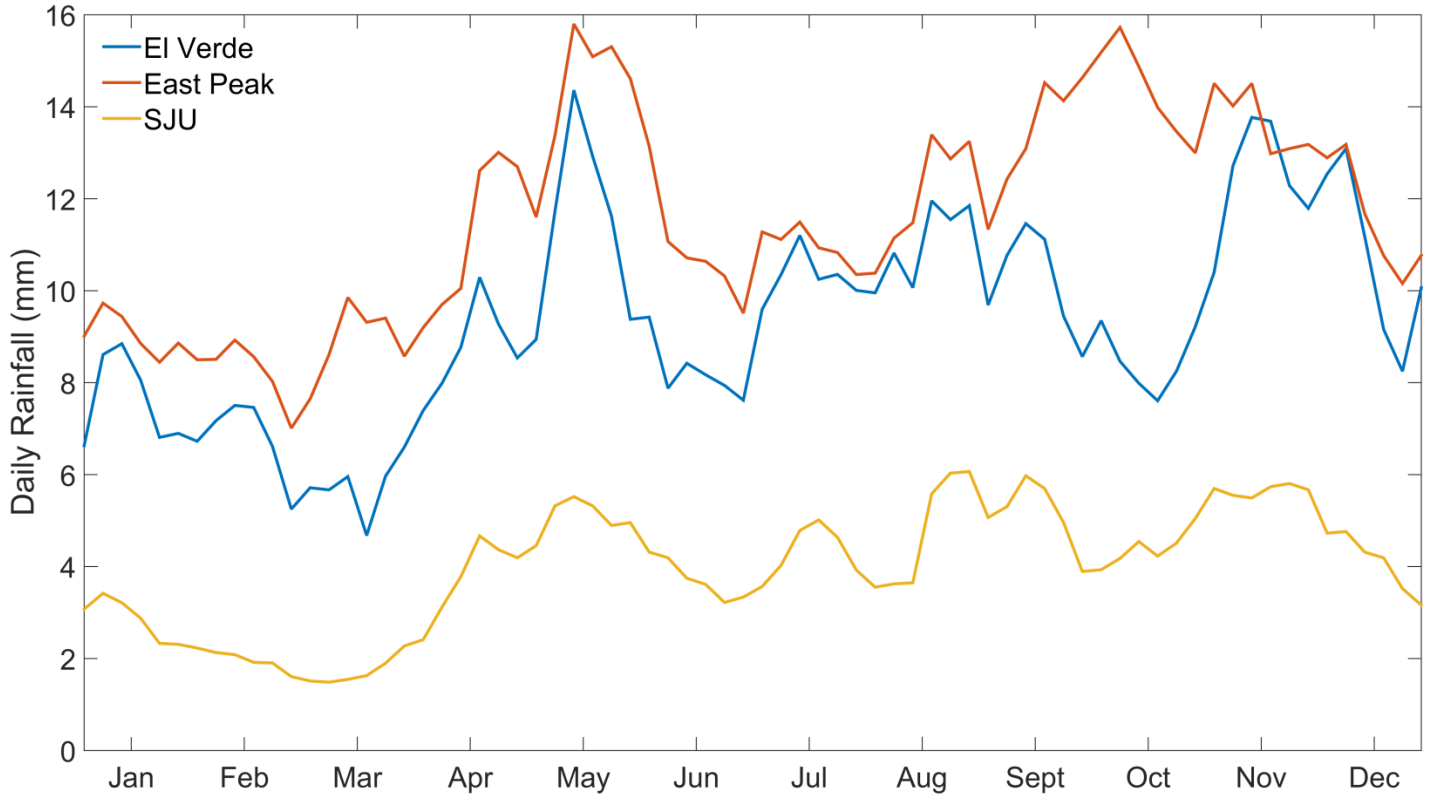

Fig. 1.4 Mean monthly daily rainfall from historical rain gauge observations at three sites in northeast Puerto Rico. San Juan (SJU) represents a low elevation site (10 m). El Verde (380 m) and East Peak (1050 m) represent two high elevation sites.Similar seasonality is represented at all three sites, but the magnitude of daily rainfall increases with increasing elevation.

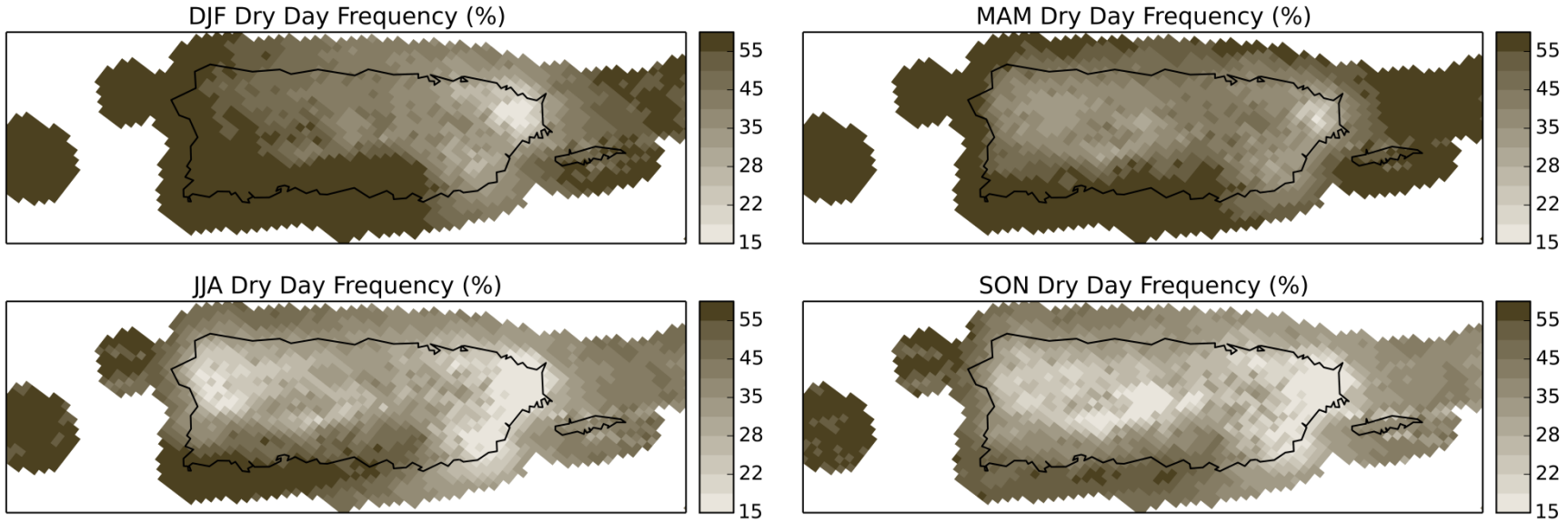

Fig. 1.5 Multi-sensor precipitation estimates (MPE) of dry day frequency. MPE integrate rain gauge and satellite rainfall estimates with the radar-only estimates.

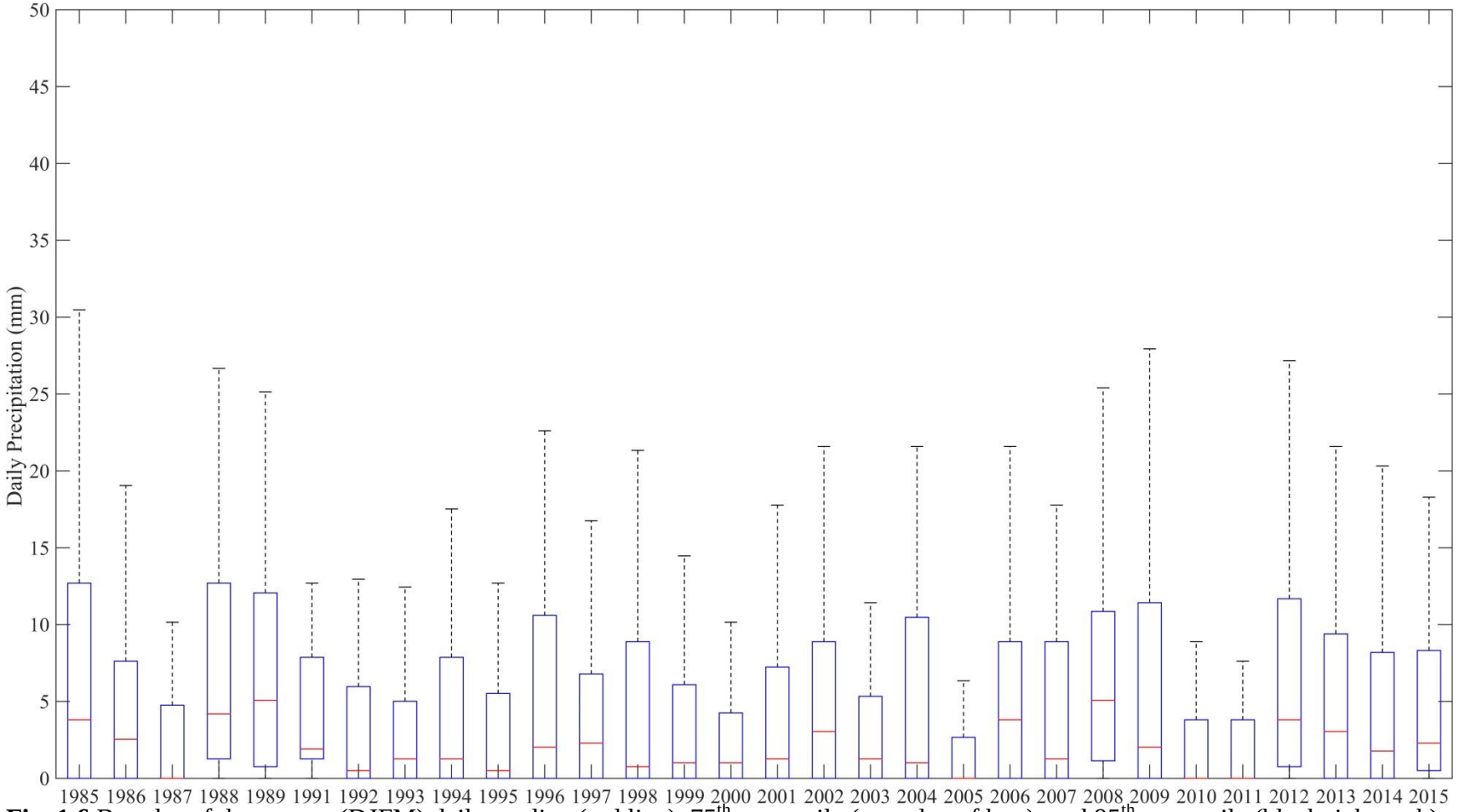

Fig. 1.6 Boxplot of dry season (DJFM) daily median (red line), 75th percentile (top edge of box), and 95th percentile (black tick mark) precipitation values.

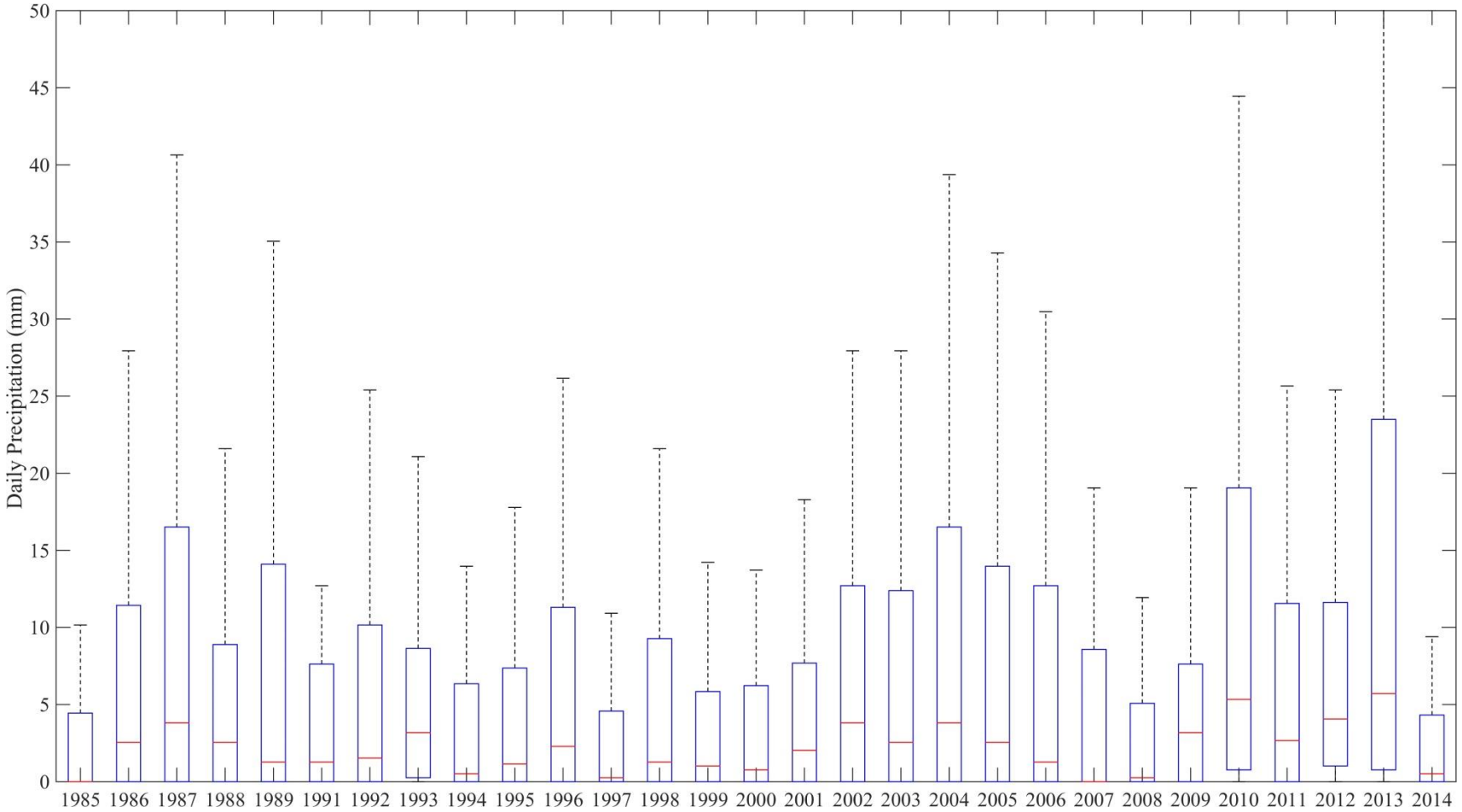

Fig. 1.7 Boxplot of early rainfall season (AMJJ) daily median (red line), 75th percentile (top edge of box), and 95th percentile (black tick mark) precipitation values.

CHAPTER 2

ATMOSPHERIC CONTROLS ON PUERTO RICO PRECIPITATION USING ARTIFICIAL NEURAL NETWORKS[1]

[1]Ramseyer, C.A. and T.L. Mote. 2016. *Climate Dynamics*, in press. DOI:10.1007/s00382-016-2980-3. Reprinted here with permission of the publisher.

Abstract

The growing need for local climate change scenarios has given rise to a wide range of empirical climate downscaling techniques. One of the most critical decisions in these methodologies is the selection of appropriate predictor variables for the downscaled surface predictand. A systematic approach to selecting predictor variables should be employed to ensure that the most important variables are utilized for the study site where the climate change scenarios are being developed. Tropical study areas have been far less examined than mid- and high-latitudes in the climate downscaling literature. As a result, studies analyzing optimal predictor variables for tropics are limited. The objectives of this study include developing artificial neural networks for six sites around Puerto Rico to develop nonlinear functions between 37 atmospheric predictor variables and local rainfall. The relative importance of each predictor is analyzed to determine the most important inputs in the network. Randomized ANNs are produced to determine the statistical significance of the relative importance of each predictor variable. Lower tropospheric moisture and winds are shown to be the most important variables at all sites. Results show inter-site variability in u- and v-wind importance depending on the unique geographic situation of the site. Lower tropospheric moisture and winds are physically linked to variability in sea surface temperatures (SSTs) and the strength and position of the North Atlantic Subtropical High (NASH). The changes forced by anthropogenic climate change in regional SSTs and the NASH will impact rainfall variability in Puerto Rico.

2.1. Introduction

Understanding the controls on rainfall variability is useful for creating climate change scenarios, particularly when employing empirical-dynamical or statistical downscaling techniques. The selection of predictor variables is one of the most important methodological steps in statistical downscaling studies (Winkler et al. 2011). Hewitson and Crane (1996) emphasize that the first assumption of statistical downscaling approaches is the inclusion of the most physically explanatory predictors into the transfer functions. The ideal predictor variable is responsive to climate change, is simulated well by global climate models, has a stable relationship with the predictand, and is sensitive to variability in the predictand (Giorgi et al. 2001; Wilby et al. 2004; Winkler et al. 2011). Sea level pressure (SLP) and geopotential heights and thicknesses have been the most common predictor variables for downscaling studies analyzing local temperature and precipitation (Cavazos and Hewitson 2005). The majority of these studies examine areas in the mid-latitudes and polar regions. Methodological studies identifying appropriate predictor variables for tropical locations have been largely isolated to studies in Mexico (Cavazos and Hewitson 2005; Hewitson and Crane 1992). Atmospheric variables most affecting precipitation and temperature processes are likely to have some regional and local variability. For example, precipitation downscaling studies should consider important factors in precipitation variable including thermodynamics, circulation variables, and moisture content (Cavazos and Hewitson 2005; Trenberth et al. 2003). An appropriate methodological step for downscaling studies in novel study areas is to systematically determine the best predictor variables.

This study aims to produce a methodology for determining the controlling atmospheric variables on rainfall variability in six sites in Puerto Rico (PR). The results will produce a suite of appropriate predictor variables for use in Caribbean downscaling studies. Precipitation falling in northeast PR is critical to the island in multiple facets. Precipitation and streamflow out of the LM are a significant source of municipal water for the San Juan metropolitan area (Crook et al. 2007; Fig. 2.1) The LM contain a range of climate-sensitive ecosystems, including montane cloud forest at the peaks. El Yunque National Forest is located within the LM and is an important economic driver for the island's tourism industry. Changes in precipitation variability have important consequences economically, biologically, and ecologically. The sites examined in the study are primarily located in northeast PR along the steep topographic gradient that culminates at the peaks of the LM with the other sites situated along the southern and western coasts. Analyzing six sites that are uniquely situated both geographically and topographically allows for an analysis of how predictor variables of precipitation vary across the island and up the topographic gradient.

This study employs the use of artificial neural networks (ANN) as it allows for non-linear functions to be established between the atmospheric predictor variables and rainfall at each site. The connections between the input and output in the ANN can be examined to determine the importance of each of the input variables to the predicted output. This provides an approach for determining the controlling atmospheric variables on observed meteorological fields at varying temporal and spatial scales.

2.2. Background

Climate variability throughout the Caribbean region is strongly associated with global scale wave patterns, tropical cyclones, orographic effects, sea breeze circulations, regional scale wind patterns (primarily the easterly trade winds), and intense solar heating (Taylor et al. 2002). Thus, local scale understanding of climate and precipitation variability depends heavily on the understanding of the regional, synoptic-scale phenomena in the Caribbean (Comarazamy and González 2008, 2011).

Precipitation over much of Puerto Rico exhibits a bimodal distribution with a dry period separating the wet seasons (Comarazamy and González 2008; Daly et al. 2003; Malmgren and Winter 1999). Spatial precipitation patterns across Puerto Rico are most strongly associated with topography, exposure and direction of the predominant winds, and proximity to the ocean (Daly et al. 2003). Much of the annual precipitation occurs in intense showers from easterly waves and tropical disturbances that occur from May to October with the balance from northerly frontal systems and localized convection (Larsen 2000).

Precipitation increases at a rate of approximately 140% (of island average) per kilometer of elevation (Daly et al. 2003). In addition, the presence of urban areas influences precipitation patterns over Puerto Rico, but to a much lesser extent than topography (Comarazamy and González 2008). In Puerto Rico, the area of highest precipitation is in the LM in NE PR, where steep terrain, coupled with exposure to predominant northeasterly winds and proximity to the ocean leads to frequent orographic lifting of moisture laden air (Comarazamy and González 2008; Daly et al. 2003; Jury 2009). Perturbations in predominant winds have been shown to spatially redistribute

localized areas of moisture convergence which has a profound effect on precipitation patterns (Comarazamy and González 2011).

Climate downscaling involves taking coarse resolution global climate model (GCM) data and through the use of a dynamical or statistical methodology, producing local, fine-resolution climate information. Dynamical downscaling typically refers to the use of a regional climate model (RCM). These RCMs use nested grids within the GCM grid space to produce increasingly higher resolution simulations. The other group of downscaling techniques can be broadly referred to as statistical downscaling.

ANNs have been used across the physical sciences, including meteorology and climatology, as a powerful, non-linear function approximation and forecasting tool. Due to the ability to produce nonlinear functions, ANNs are well suited to be used as a problem solving tool for climate related problems (Tsonis and Elsner 1992). ANNs have been used for a wide range of applications in the atmospheric sciences including precipitation modeling (Bellerby et al. 2000; Cavazos and Hewitson 2005; Gardner and Dorling 1998; Hall et al. 1999; Haylock et al. 2006; Kuligowski and Barros 1998; Sahai et al. 2000; Schoof and Pryor 2001; Silverman and Dracup 2000; Valverde Ramírez et al. 2005). These studies use ANNs to derive nonlinear relationships between atmospheric controls and precipitation. In this way, ANNs are a tool used for statistical climate downscaling, more specifically, empirical-dynamical downscaling (Winkler et al. 2011). ANNs have been shown to out-perform other statistical techniques for studies analyzing precipitation (Kuligowski and Barros 1998; Valverde Ramírez et al. 2005).

This study uses feed-forward multilayer perceptrons (FFMLP) with backpropagation of error, one type of artificial neural network. The FFMLPs used in the

study contain three layers (Fig. 2.2). The architecture includes an input layer containing a suite of atmospheric variables, a hidden layer, and an output layer that contains the predicted/forecasted precipitation. The input layer receives input data, with one node per input variable. The input layer is connected to the hidden layer through a matrix of weighted connections, referred to as input weights. The hidden layer is linked to the output layer via a matrix of layer weights. The weight matrices are initialized randomly. As training data are introduced to the model, the network "learns" via a training function (Bayesian regularization here). This algorithm seeks to reduce some error statistic (e.g., mean squared error). After each training epoch, the error between the output and the target is back-propagated through the ANN to update the weights for the next training epoch (Cavazos 1999; Cavazos and Hewitson 2005). The remaining data is used for independent testing and/or validation. For further details on FFMLP and the backpropagation of error algorithms, refer to (Hewitson and Crane 1994) and (Maier and Dandy 2000).

2.3. Methodology

a. Data sources

Six study sites in Puerto Rico were chosen representing a range of geographic situations and precipitation climatologies. Three of the sites are located in the vicinity of the LM including Roosevelt Roads (RR), Paraiso (PA), and Pico Del Este (PE). To determine if changes in optimum predictor variables existed up the topographic gradient, each of the three sites is located at different elevations (Fig. 2.1). PE was selected as the high elevation site where the rain gauge is located at an elevation of 1051 m asl. RR provides a low-elevation site at 10 m asl and is situated on the eastern coast of Puerto

Rico. The PA rain gauge is situated at 110 m asl and serves as a mid-elevation site, between PE and RR on the topographic gradient up the LM. The fourth site is located at the San Juan International Airport (SJ) which provides an additional low-elevation site northwest of LM in a highly urbanized area. The two remaining sites are relatively far removed from the other sites with Ponce (PO) on the south coast and Mayaquez (MA) on the west coast (Fig. 2.1).

Daily precipitation data for each site were acquired from the Global Historical Climatology Network-Daily dataset (Menne et al. 2012). All data available from 1980–2009 were obtained. This date range was used as it provided the least missing values for a 30 year period for the six sites. The SJ site is the only site that has a continuous record spanning the 30 year period. The other sites had less than 30 percent missing data. The reanalysis data were pruned to match the availability of the precipitation data at each site. Six-hourly ERA-Interim gridded reanalysis variables were obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) for 12Z for a 0.125° latitude by 0.125° longitude grid situated over Puerto Rico. ERA-Interim is a global atmospheric reanalysis product available from 1979 and updated continuously (Dee et al. 2011). The data assimilation scheme used is a four-dimensional variational analysis (4D-Var) with a 12-hour analysis window. Cycle 31r2 of the ECMWF Integrated Forecast System (IFS) was used in the development of ERA-Interim. Reanalysis data were collected for the study period from 1980–2009. The grid resolution of this product allows for better representation of smaller scale atmospheric circulations and fluctuations in atmospheric variability. This study uses 37 atmospheric predictor variables which include divergence, geopotential heights, surface wind fields, mean sea-level pressure, relative humidity,

specific humidity, vorticity, vertical velocities, u and v winds, and atmospheric thickness at different pressure levels (Table 2.1). These variables were selected following the methodology of Cavazos and Hewitson (2005) with the addition of the 925 hPa level, which was added for increased vertical resolution in the lower troposphere.

b. Methodology

The utility of the networks presented here is to provide a predictive tool of precipitation variability, and ultimately, to determine which atmospheric controls are most important in that precipitation variability. The precipitation and reanalysis data were pre-processed prior to training the ANN. The reanalysis and precipitation data were smoothed with a five-day equally weighted moving average filter. This allowed for the focus to be on the atmospheric modes conducive to precipitation (Hewitson and Crane 1994). Missing data, including the first and last two days, were removed.

The independent (reanalysis atmospheric variables) and the dependent variables (rainfall observations) were normalized prior to training the ANNs. The rainfall data were fitted to the range [0–1] in order to conform to the output provided by the ANN using a log-sigmoid output transfer function. The following algorithm is used to normalize the rainfall data to the range [0–1]:

$$t_n = \frac{Y_n - Y_{min}}{Y_{max} - Y_{min}} \quad (1)$$

where t_n is the normalized value of Y at time n, Y_n is the original rainfall value at time n, and Y_{min} and Y_{max} are the minimum and maximum values of the variable Y. The reanalysis variables were standardized to account for the variation in measurement scales

of each. This involves producing converted reanalysis variables with zero mean and unity standard deviation.

The normalized rainfall data and standardized reanalysis data were assigned as targets and inputs, respectively, in the networks. The MATLAB Neural Network Toolbox was utilized to produce the two-layer FFMLPs in this study. The networks used in this study consisted of three components, an input layer (N=37), a hidden layer, and an output layer. In order to determine the optimal number of hidden neurons in the hidden layer, a suite of ANNs were constructed using a range of neurons in the hidden layer. The hidden layer size was increased at increments of five and ranged from 5–40 neurons. The other model parameters were held constant. The mean square error (MSE) was assessed for each ANN. There are a few considerations that need to be taken into account when deciding the number of hidden neurons. Networks with fewer hidden neurons are preferable due to their decreased computational expense and because the network has fewer connections, and can be easier to interpret. Networks with a large number of hidden neurons can have advantages when the number of input variables is high. If too many neurons are used however, the model can begin to over-fit and become less useful in generalizing. MSE decreased appreciably from the five-node model to the 20-node model. The lowest MSE occurred when having 40 neurons; however, there was a negligible difference in MSE between the 20- and 40-node models. The computational cost was significantly higher for the 40-node network. As a result, the 20-node model architecture was deemed most appropriate for this study (Fig. 2.2). All training networks used in the analysis performed best with 250 epochs (training runs).

Two neural transfer functions are used in the networks. The first transfer function operates on the input data to produce the hidden layer output. For our data, a hyperbolic tangent sigmoid transfer function is appropriate as it creates output in the range [-1 1], which conforms to the range of the standardized reanalysis data. The second transfer function operates on the hidden layer output to create the output layer data. Because the rainfall data only have positive values, the log-sigmoid transfer function was selected as it constrains output to the range [0 1].

For each observation site, 1000 networks were created with each starting with different initial input weights. This step is necessary because any one network may be converging to local minima rather than global minima. A network converging to local minima versus global minima will have different weight matrices. The weight matrices are used in the calculation of variable importance, thus, the local minima weight matrix would lead to incorrect assumptions about variable importance.

Each network divided the input and target data randomly between a training (85%) and test set (15%). The best performing network was selected by determining which network had the lowest MSE. For the optimized network, Pearson's correlation coefficient was calculated on the observed and modeled rainfall data for both the training and test set (Table 2). The MSE of the optimized network was decomposed into variance and bias:

$$\mathrm{MSE} = \sum_{i=1}^{n} (y_i - x_i)^2 \quad (2)$$

$$= \overline{(y_i - x_i)^2} \quad (3)$$

$$= \overline{(y - x)^2} - \overline{[(y - x)]}^2 - \overline{(y - x)}^2 \quad (4)$$

$$= s_{y-x}^2 + (\overline{y} - \overline{x})^2 \quad (5)$$

$$= \text{Variance} + \text{Bias}^2 \quad (6)$$

The variance term in this decomposition is the variance of the errors, and the bias term is the average deviation of the model forecasted rainfall from the observed rainfall. The model forecasts can be improved by eliminating the bias term. Thus, models with a large bias will be improved more than models with a small bias term. The forecasts can be shifted (+/–) by the constant bias term, but leaves the variance unaffected.

The bias term above was discussed as a constant value. However, it can be assumed that bias may be larger for certain portions of the rainfall distribution. In other words, the bias may be dependent on rainfall amount. Assume E(x|y) is the conditional mean (or expected value) of an observed rainfall amount (x), for a given forecasted rainfall amount (y). In order to calculate E(x|y) for our continuous variable, the forecasted rainfall amounts are binned and the average of the observed rainfall is computed for each bin. The resulting means can be plotted as a function of the forecasts and assess the reliability and conditional bias of the model. To account for the sampling variation of the conditional mean, 1000 bootstrapping trials were performed on the conditional mean of the observed rainfall. The sampling variations of the conditional mean can be used to create a reliability plot which visualizes uncertainty at each bin (Marzban 2009).

The input and layer weight matrices of the optimized networks are used to compute the variable importance. Variable importance was determined by following Garson's approach (Garson 1991; Goh 1995). The networks consist of connection weights between the input-hidden layers (IW) and the hidden-output layers (LW). The Garson method involves calculating the product of the IW and LW connection weights

(IHO) for each input and hidden neuron (e.g. $IHO_{A1} = IW_{1,A} \times LW_{1,A}$). With 37 input neurons and 20 hidden neurons, the networks in this study have 740 IHO connection weights.

Next, the absolute value of the IHO weights are summed across each input variable (CW) resulting in 37 overall connection weights (e.g. $CW_{1} = |IHO_{A1}| + |IHO_{B1}|\ldots$) where IHO_{A1} and IHO_{B1} refer to the IHO value for the first input neuron and the first and second hidden neuron, respectively. The overall connection weights are summed and the relative importance (RI) of each input variable is calculated (e.g. $\text{RI}\ (\%) = \frac{CW_1}{\sum_{i=1}^{n} CW_i} \times 100$). Because the absolute value is used in the calculation of the IHO values, the variable importance calculated here is a magnitude, and does not contain information on the direction of the input-output interaction (Olden and Jackson 2002).

A randomization test is performed on each sites' optimized network in an effort to quantitatively determine which IHO connections are statistically significant (Olden and Jackson 2002). Additionally, the statistical significance of the relative importance can be assessed and is represented using p-values. After the optimized network is established and the IHO weights ($IHO_{observed}$) and relative importance ($RI_{observed}$) have been calculated, the rainfall observations are randomly permuted and new neural networks are constructed using the same initial connection weights for the optimized network. It is critical to use the same initial input and layer weights for each randomization. Without this step, important differences between the observed and random weights cannot be separated from the differences that arise from different initial conditions of the weight matrices (Olden and Jackson 2002). In this study, 999 randomized networks are

developed. After each network is trained, the IHO weights (IHO_{random}), relative importance (RI_{random}), and W_{random} are calculated and recorded. The statistical significance can be calculated as the proportion of randomized values whose value is equal to or more extreme than the observed values (e.g. $RI_{random} \geq RI_{observed}$). The p-values from the randomization tests can help determine which IHO, CW, and RI values are statistically different than what would be expected by chance alone. Further details on the randomization test for ANNs can be found in Olden and Jackson (2002).

2.4. Results

a. Network performance

As discussed in the methodology, for every site 1000 networks were created with different network weight initializations for each network. The best performing ANNs were selected based on a range of performance metrics. Table 2.2 shows the summary statistics for each sites' best network. According to the mean squared error, mean absolute error, and the root mean squared performance metrics, the RR and SJ ANNs performed better than the ANNs for PE and PA (Table 2.2). This was expected due to the wider precipitation distribution and greater number of extreme events observed at PE and PA. Smoothed rainfall values of over 20 mm were recorded at PE in 16 percent of all observations while PA surpassed this threshold in seven percent of the observations. SJ, MA, and RR surpassed the 20 mm threshold in two percent or fewer of the observations. These data were included in the training and testing of the network because the authors wanted to ensure that the model was capable of reproducing the variability in the observed precipitation record. Removing the outlier data may result in a model with better performance. However, keeping the data outliers ensures that the variable

importance calculated is based on a model that is capable of predicting rainfall throughout the distribution. The networks show considerable ability at capturing the overall trends in the smoothed rainfall data (Fig. 2.3).

Fig. 2.4 shows the PE reliability plot of the modeled versus measured rainfall. The graphic allows for a visual assessment of the uncertainty and reliability of the predicted rainfall. For reference, a trendline is displayed showing a perfect prediction. The boxplots provide visually represents the sampling distribution of the conditional mean of the observed rainfall, given the predicted rainfall. The horizontal line inside the boxplot represents the median of the conditional mean while the upper and lower edges of the boxes represent the 25^{th} and 75^{th} percentile. A skewed distribution is represented by a horizontal line that is not centrally located in the boxplot. The size of the boxplots represents certainty where small boxes are representative of more certain predictions. The reliability of the predictions can be visually assessed by the proximity of the boxplots to the diagonal. The whiskers represent the extreme data not including the outliers while the outliers are plotted individually. The size of the boxplot is indicative of the uncertainty.

Using PE as an example, moderate predicted rainfall (e.g. 9.2 mm) is accompanied by a narrow range of observed rainfall. This narrow range indicates that there is a high degree of certainty in the predictions. Comparing this to the heaviest predicted rainfall boxplot (50 mm), a larger variation in corresponding observed rainfall is noted. The median of the heavy rainfall boxplot intersects the trendline and represents that the model is predicting rainfall reliably in these amounts. However, due to the spread in the observed rainfall (represented by the size of the boxplot), these predictions are less

certain. The reliability plot also shows that moderate rainfall predictions are reliable and that there is a high degree of certainty about those predictions. For low daily rainfall, the network slightly over-predicts the observed rainfall, and there is a high degree of certainty. Under-estimation of high rainfall and overestimation of low rainfall is common in different methods and has been documented in similar studies (e.g.,Cavazos and Hewitson 2005).

The correlation coefficient between all of the predicted and observed five-day smoothed daily rainfall data points (r_{all}) was 0.84 or greater for all six networks. This indicates that at least 70 percent of the local rainfall can be explained by the atmospheric variables, while 30 percent of the local rainfall is a function of effects not included as input variables. Additionally, neural networks are specifically trained in order to generalize the target data. Neural network training involves using a sub-sample of a continuous function to find a generalized form of that function (Hewitson and Crane 1994). Thus, the models ability to forecast extreme events can be affected. Despite these limitations, the models showed relatively good skill at predicting the onset of a heavy rainfall period. At the PE, there were 1211 rainfall values between 20–40 mm, 223 values between 40–60 mm, and 72 values over 60 mm making it the wettest of the six sites. The mean difference between the observed and predicted precipitation was 6.0 mm, 6.8 mm, and 7.7 mm (0.30 in.) for each bin respectively. The ability for the model to correctly predict the onset of a heavy rainfall event (despite under- or over-predicting the magnitude) ensures that the input variable contribution to the net is relevant for the full precipitation distribution for the site. This is important because the atmospheric controls on precipitation seem to be similar for dry and wet days.

Another metric of assessing the networks performance is to calculate the number of rainfall observations that were predicted to within one mm and five mm of rainfall (Table 2.3). PE had the lowest percentage of predictions within these thresholds with 78.5% of prediction within five mm of the observed rainfall. The four lower elevation sites with less variable rainfall and fewer extreme events had a higher percentage of predictions within these ranges with 89 % or more of predictions within five mm of the observed rainfall.

b. Variable importance in networks

To calculate the relative importance of each of the input variables, Garson's algorithm was employed. The top six contributors to the best networks are listed in Table 2.4. The best networks were randomized using the approach discussed in (Olden and Jackson 2002). After 999 bootstrapping runs, the statistical significance of the importance of the input variables in the original (best) network was assessed. The italicized variables in the table are statistically significant based on $\chi = 0.05$. This indicates that there is a low probability of getting a value that larger or larger than the relative importance of that variable in the original network. The results indicate that 1000 hPa specific humidity is one of the two leading contributors at each of the three sites in or directly adjacent to the LM (RR, PE, and PA). At PE, 1000 hPa specific humidity is also statistically significant.

The low-level u winds (u10, u92, u85) are also important at each of the four northeast sites (Fig. 2.5). The 1000 hPa u-wind is the most important predictor variable at PA and SJ and it also statistically significant. The 925 hPa u-wind is one of the leading six contributors at each of the four northeast sites. 925 hPa u-winds are statistically

significant and the most important variable in the MA network. The MA and PO networks are more controlled by low-level v-winds than the northeast sites. Seasonal models were also produced to evaluate variability in predictor variables between the dry and wet seasons. Low-tropospheric moisture and u-winds were the most common predictors in both seasons. Thus, the vertical profile of u-wind and specific humidity through the lower troposphere from the surface to 700 hPa is driving rainfall in the study area regardless of season. As expected, the dry season models perform better than the yearly models due to decreased precipitation variability and lack of extreme precipitation events during the dry season. The wet season models have slightly decreased performance compared to the dry season or yearly models due to the increase in precipitation variability and extreme events (e.g. easterly waves). Despite a slight decrease in the wet season model performance, these models still capture the variability in the time series. However, the extreme events are under-predicted and cause an increase in the error metrics.

The seasonal models tend to be slightly noisier compared to the yearly models in terms of the top predictors. The seasonal models tend to have less agreement on which levels in the lower troposphere are most important. The primary difference between the seasonal and yearly models was the increased prevalence of 500 hPa geopotential heights as a top predictor in the wet and dry season models. This could be attributed mid-tropospheric intrusions from extratropical frontal systems during the dry season and easterly waves during the wet season.

2.5. Discussion

The most important atmospheric predictor variables for rainfall in Puerto Rico are low tropospheric specific humidity and low tropospheric u-winds. For the southern and western site, low tropospheric v-winds are also an important control. More specifically, 1000 hPa specific humidity (q0) and 10-m u-winds (u0) are the top predictors at the NE PR sites. The PO and MA networks top predictor was 925 hPa u-winds (u9). Additionally, 925 hPa v-winds (v9) are an important predictor in both networks outside of NE PR.

Q0 is the most important variable in the RR and PE networks and is statistically significant in the PE network. The PA, SJ, and PO networks also have q0 as one of the top six contributors. It is important to note that PE is located above the 1000 hPa pressure level. The q0 control on precipitation at PE is likely due to moisture advection up the topographic gradient into the site. Air at 1000 hPa is transported by the easterly trade winds and encounters the LM. Forced ascent up the windward side of the LM provides an additional lifting mechanism and may allow condensation processes to occur even with a lack of another lifting mechanism (e.g. instability). At a more broad scale, high specific humidity likely has two roles in increased precipitation in the LM. High humidity at low-levels leads to decreased atmospheric stability. Additionally, the high moisture content of the air decreases the lifted condensation level (LCL) and allows for initiation of precipitation processes at lower levels.

Other than specific humidity, the low-level u-winds are the other top predictor in the networks. U0 is the top predictor at the PA and SJ sites and is statistically significant. Additionally, u9 is the top predictor at the MA (statistically significant) and PO sites. U9

is one of the top six predictors at the PE and RR sites. The prevalence of the low-level u-winds and low-level specific humidity as the two most important predictors of precipitation in PR in the networks is validated in the literature.

Due to strong ocean-atmosphere coupling in the tropics, low level specific humidity is closely related to SSTs and the low level wind field. Previous research suggests tropical North Atlantic SSTs strongly influences early season Caribbean rainfall variability while equatorial Pacific and Atlantic SSTs influence late season rainfall variability (Taylor et al. 2002). 1000 hPa specific humidity is linked to the magnitude of the near-surface trade winds in the tropical Atlantic. The North Atlantic Subtropical High (NASH) is a large scale driver of these feedbacks as it controls trade wind strength in the region. The intensification of the NASH translates in stronger trade winds and lower SSTs (Gamble and Curtis 2008; Giannini et al. 2000; Granger 1985; Hastenrath 1976).

Variations in low tropospheric and surface trade wind intensity are the primary forcing mechanism in sea-surface temperatures (SST) over the tropical Atlantic (Nobre and Srukla 1996). As winds increase, wind stress and turbulent mixing increase leading to cooler SSTs. This also leads to increased wind shear (assuming no mid-upper tropospheric disturbance) which increases atmospheric stability and caps deep convection (Enfield and Alfaro 1999). Decreases in trade winds leads to increased SSTs, enhanced surface evaporation, and lower tropospheric moisture convergence (Wu and Kirtman 2005, 2011). This also leads to increased atmospheric instability, allowing for greater deep convection and rainfall in the Caribbean (Enfield and Alfaro 1999).

After q0, u0, and u9, the other common variable amongst most of the networks is 700 hPa specific humidity (q7). It is the second most important input variable, and is statistically significant in the PE network. In three other networks q7 is one of the four most important variables. Over this portion of the Caribbean, the trade wind inversion is located at approximately 2300 m (Gutnick 1958). Cumulus and stratocumulus convection is capped at this level, due to subsidence aloft (Schubert et al. 1995). The closest pressure level included in the network input layer was 700 hPa. The importance of q7 in the networks is indicating that the moisture content in the layer from the surface to the trade wind inversion is an important predictor of PR rainfall.

2.6. Conclusions

This study presents a methodology for selecting appropriate predictor variables to be used in climate downscaling studies. FFMLP neural networks were tested and randomized for six sites in Puerto Rico. The results of this study indicate that 1000 hPa specific humidity, 10-m and 925 hPa u-winds are the largest controls on precipitation in Puerto Rico. Increased low-tropospheric moisture destabilizes the lower troposphere and leads to lower lifted condensation levels. Decreased low-tropospheric and surface u-winds enhance evaporation off the Caribbean Sea and tropical North Atlantic (increasing 1000 hPa specific humidity) and increase low-tropospheric moisture convergence. Decreases in low-tropospheric winds leads to low wind shear environments which allows for effective trade wind cumulus convection. The notable difference between the northeast sites and the PO and MA sites was the slight increase in importance of the low-level v-component of the wind. This is due to the orientation of the sites to nearby topography. Our results indicate that site elevation had little influence on the controlling

atmospheric variables in the networks. The highest elevation site (PE) in northeast PR had similar controls as the nearby low elevation sites.

Although not explicitly resolved in this study, the results of this study highlight the importance of SSTs on precipitation variability in the study area. Due to the strong ocean-atmosphere coupling in the tropics, SST variability is linked to low tropospheric moisture and wind. These changes are coupled with variability in NASH strength and position, which drive trade wind strength. Future changes in Caribbean and tropical North Atlantic SSTs as well as the strength and position due to anthropogenic climate change will force changes in precipitation variability in Puerto Rico. The results from this study suggest that future climate scenarios of precipitation variability in PR and the Caribbean using statistical downscaling methodologies should use low tropospheric specific humidity and winds as predictor variables.

2.7. References

Bellerby, T., M. Todd, D. Kniveton, and C. Kidd, 2000: Rainfall Estimation from a Combination of TRMM Precipitation Radar and GOES Multispectral Satellite Imagery Through the Use of an Artificial Neural Network. *J Appl Meteorol,* **39,** 2115–2128.

Cassano, J. J., P. Uotila, A. H. Lynch, and E. N. Cassano, 2007: Predicted changes in synoptic forcing of net precipitation in large Arctic river basins during the 21st century. *Journal of Geophysical Research: Biogeosciences*, **112,** 1– 20.

Cavazos, T., 1999: Large-Scale Circulation Anomalies Conducive to Extreme Precipitation Events and Derivation of Daily Rainfall in Northeastern Mexico and Southeastern Texas. *J Clim,* **12,** 1506–1523.

Cavazos, T., and B. C. Hewitson, 2005: Performance of NCEP-NCAR Reanalysis Variables in Statistical Downscaling of Daily Precipitation. *Clim Res,* **28,** 95–107.

Comarazamy, D. E., and J. E. González, 2008: On the Validation of the Simulation of Early Season Precipitation on the Island of Puerto Rico Using a Mesoscale Atmospheric Model. *J Hydrometeorol,* **9,** 507–520.

——, 2011: Regional Long-Term Climate Change (1950–2000) in the Midtropical Atlantic and its Impacts on the Hydrological Cycle of Puerto Rico. *J Geophys Res*, **116**.

Crook, K., F. N. Scatena, and C. M. Pringle, 2007: Water Withdrawn From the Luquillo Experimental Forest, 2004. . F. S. U.S. Department of Agriculture, International Institute of Tropical Forestry, Ed., 26.

Daly, C., E. H. Helmer, and M. Quiñones, 2003: Mapping the climate of Puerto Rico, Vieques and Culebra. *Int J Clim*, **23,** 1359–1381.

Dee, D., and Coauthors, 2011: The ERA-Interim Reanalysis: Configuration and Performance of the Data Assimilation System. *Q J R Meteorolog Soc*, **137,** 553–597.

Enfield, D. B., and E. J. Alfaro, 1999: The Dependence of Caribbean Rainfall on the Interaction of the Tropical Atlantic and Pacific Oceans. *J Clim*, **12,** 2093–2103.

Gamble, D. W., and S. Curtis, 2008: Caribbean Precipitation: Review, Model and Prospect. *Prog Phys Geogr*, **32,** 265–276.

Gardner, M. W., and S. R. Dorling, 1998: Artificial Neural Networks (the Multilayer Perceptron)—A Review of Applications in the Atmospheric Sciences. *Atmos Environ*, **32,** 2627–2636.

Garson, D. G., 1991: Interpreting Neural-Network Connection Weights. *Artif Intell Expert*, **6,** 47–51.

Giannini, A., Y. Kushnir, and M. A. Cane, 2000: Interannual Variability of Caribbean Rainfall, ENSO, and the Atlantic Ocean. *J Clim*, **13,** 297–311.

Giorgi, F., and Coauthors, 2001: Emerging Patterns of Simulated Regional Climatic Changes for the 21st Century due to Anthropogenic Forcings. *Geophys Res Lett*, **28,** 3317–3320.

Goh, A., 1995: Back-Propagation Neural Networks for Modeling Complex Systems. *Artif Intell Eng*, **9,** 143–151.

Granger, O. E., 1985: Caribbean Climates. *Prog Phys Geogr*, **9,** 16–43.

Gutnick, M., 1958: Climatology of the trade-wind inversion in the Caribbean. *Bull Am Meteorol Soc*, **39,** 410–420.

Hall, T., H. E. Brooks, and C. A. Doswell III, 1999: Precipitation Forecasting Using a Neural Network. *Weather Forecasting*, **14,** 338–345.

Hastenrath, S., 1976: Variations in Low-Latitude Circulation and Extreme Climatic Events in the Tropical Americas. *J Atmos Sci*, **33,** 202–215.

Haylock, M. R., G. C. Cawley, C. Harpham, R. L. Wilby, and C. M. Goodess, 2006: Downscaling Heavy Precipitation over the United Kingdom: a Comparison of Dynamical and Statistical Methods and their Future Scenarios. *Int J Clim*, **26,** 1397–1415.

Hewitson, B., and R. Crane, 1996: Climate Downscaling: Techniques and Application. *Clim Res*, **7,** 85–95.

Hewitson, B. C., and R. G. Crane, 1992: Large-Scale Atmospheric Controls on Local Precipitation in Tropical Mexico. *Geophys Res Lett*, **19,** 1835–1838.

Hewitson, B. C., and R. G. Crane, 1994: *Neural Nets: Applications in Geography: Applications for Geography.* Vol. 29, Springer Science & Business Media.

Hewitson, B. C., and R. G. Crane, 2006: Consensus between GCM climate change projections with empirical downscaling: precipitation downscaling over South Africa. *Int J Clim*, **26,** 1315–1337.

Jones, P. W., 1999: First- and Second-Order Conservative Remapping Schemes for Grids in Spherical Coordinates. *Mon Weather Rev*, **127,** 2204–2210.

Jury, M. R., 2009: An Intercomparison of Observational, Reanalysis, Satellite, and Coupled Model Data on Mean Rainfall in the Caribbean. *J Hydrometeorol*, **10,** 413–430.

Kuligowski, R. J., and A. P. Barros, 1998: Localized Precipitation Forecasts from a Numerical Weather Prediction Model Using Artificial Neural Networks. *Weather Forecasting*, **13,** 1194–1204.

Larsen, M. C., 2000: Analysis of 20th Century Rainfall and Streamflow to Characterize Drought and Water Resources in Puerto Rico. *Phys Geogr*, **21,** 494–521.

Maier, H. R., and G. C. Dandy, 2000: Neural networks for the Prediction and Forecasting of Water Resources Variables: A Review of Modelling Issues and Applications. *Env Modell Software*, **15,** 101–124.

Malmgren, B. A., and A. Winter, 1999: Climate Zonation in Puerto Rico Based on Principal Components Analysis and an Artificial Neural Network. *J Clim*, **12,** 977–985.

Marzban, C., 2009: Performance Measures and Uncertainty. *Artificial Intelligence Methods in the Environmental Sciences*, S. Haupt, A. Pasini, and C. Marzban, Eds., Springer Netherlands, 49–75.

Menne, M. J., I. Durre, R. S. Vose, B. E. Gleason, and T. G. Houston, 2012: An overview of the global historical climatology network-daily database. *J Atmos Ocean Technol*, **29,** 897–910.

Nobre, P., and J. Srukla, 1996: Variations of Sea Surface Temperature, Wind Stress, and Rainfall over the Tropical Atlantic and South America. *J Clim*, **9,** 2464–2479.

Olden, J. D., and D. A. Jackson, 2002: Illuminating the "black box": A Randomization Approach for Understanding Variable Contributions in Artificial Neural Networks. *Ecol Model*, **154,** 135–150.

Sahai, A., M. Soman, and V. Satyan, 2000: All India Summer Monsoon Rainfall Prediction using an Artificial Neural Network. *Clim Dyn*, **16,** 291–302.

Schoof, J. T., and S. C. Pryor, 2001: Downscaling Temperature and Precipitation: A Comparison of Regression-based Methods and Artificial Neural Networks. *Int J Clim*, **21,** 773–790.

Schubert, W. H., P. E. Ciesielski, C. Lu, and R. H. Johnson, 1995: Dynamical Adjustment of the Trade Wind Inversion Layer. *J Atmos Sci*, **52,** 2941–2952.

Silverman, D., and J. A. Dracup, 2000: Artificial Neural Networks and Long-Range Precipitation Prediction in California. *J Appl Meteorol*, **39,** 57–66.

Skific, N., and J. Francis, 2012: Self-Organizing Maps: A Powerful Tool for the Atmospheric Sciences.

Taylor, M. A., D. B. Enfield, and A. A. Chen, 2002: Influence of the tropical Atlantic versus the tropical Pacific on Caribbean rainfall. *J Geophys Res: Ocean*, **107,** 10-11-10-14.

Trenberth, K. E., A. Dai, R. M. Rasmussen, and D. B. Parsons, 2003: The Changing Character of Precipitation. *Bull Am Meteorol Soc*, **84,** 1205–1217.

Tsonis, A. A., and J. B. Elsner, 1992: Nonlinear Prediction as a way of distinguishing chaos from random fractal sequences. *Nat*, **358,** 217–220.

Valverde Ramírez, M. C., H. F. de Campos Velho, and N. J. Ferreira, 2005: Artificial neural network technique for rainfall forecasting applied to the São Paulo region. *J Hydrol*, **301,** 146–162.

Wilby, R., S. Charles, E. Zorita, B. Timbal, P. Whetton, and L. Mearns, 2004: Guidelines for use of climate scenarios developed from statistical downscaling methods.

Winkler, J. A., and Coauthors, 2011: Climate Scenario Development and Applications for Local/Regional Climate Change Impact Assessments: An Overview for the Non-Climate Scientist. *Geogr Compass*, **5,** 275–300.

Wu, R., and B. P. Kirtman, 2005: Roles of Indian and Pacific Ocean air–sea coupling in tropical atmospheric variability. *Clim Dyn*, **25,** 155–170.

——, 2011: Caribbean Sea rainfall variability during the rainy season and relationship to the equatorial Pacific and tropical Atlantic SST. *Clim Dyn*, **37,** 1533–1550.

Table 2.1 ECMWF ERA-Interim reanalysis variables used in the network input layer.

Atmosphere Level	Circulation	Moisture	Thickness
Surface/1000 hPa	SLP (slp) Geopotential Height (z0) U and V winds (u0, v0) Vorticity (vo0) Divergence (d0)	SH (q0) RH (rh0)	500 – 1000 hPa (th1)
925 hPa	Geopotential Height (z9) U and V winds (u9, v9) Divergence (d9) Vertical Velocity (w9)	SH (q9)	
850 hPa	U and V winds (u8, v8) Divergence (d8) Geopotential Height (z8) Vertical Velocity (w8)	SH (q8)	500 – 850 hPa (th8)
700 hPa	U and V winds (u7, v7) Divergence (d7) Geopotential Height (z7) Vertical Velocity (w7)	SH (q7) RH (rh7)	
500 hPa	U and V winds (u5, v5) Divergence (d5) Geopotential Height (z5) Vorticity (vo5)	SH (q5)	
200 hPa	Divergence (d2) Geopotential Height (z2)		

Table 2.2 Performance statistics from the best performing networks for each site from 1980 – 2009 including mean-squared error (MSE), root mean-squred error (RMSE), variance (VAR), bias, mean absolute error (MAE), Pearson's Coefficient for the test set (r_{test}) and for all observations (validation and test sets; r_{all}), and the coefficient of determination for the test set (R^2_{test}) and for all observations (R^2_{all}). MSE, MAE, RMSE, VAR, and BIAS are calculated on the full time series.

Location	RMSE (mm)	MSE (mm^2)	MAE (mm)	VAR (mm^2)	BIAS (mm)	r_{test}	r_{all}	R^2_{test}	R^2_{all}
RR	2.50	6.24	1.70	6.18	-0.25	0.80	0.93	0.71	0.86
PE	5.14	26.40	3.86	26.39	-0.11	0.81	0.90	0.70	0.80
PA	4.49	20.14	3.25	20.12	-0.18	0.76	0.84	0.59	0.70
SJ	2.64	6.98	1.82	6.96	-0.13	0.77	0.86	0.61	0.73
MA	3.21	10.33	2.18	10.31	-0.16	0.78	0.86	0.58	0.74
PO	2.34	5.45	1.35	5.34	-0.34	0.81	0.93	0.68	0.86

Table 2.3 Percentage of daily smoothed rainfall values where the difference between observed and predicted rainfall was within 1 mm and 5 mm.

	Predicted ± 1 mm (%)	Predicted ± 5 mm (%)
RR	45.6	94.1
PE	18.8	72.5
PA	24.7	78.9
SJ	43.3	93.6
PO	60.5	95.2
MA	40.8	89.2

Table 2.4 The most important input variables for the best performing network as calculated using Garson's algorithm for determining variable relative importance (RI). Italicized variables represent those RI values that are statistically significant based on $\chi = 0.05$ when compared to the randomized networks.

RR	PE	PA	SJ	MA	PO
q10	*q10*	*u10*	*u10*	*u92*	u92
u85	*q70*	q10	z50	v92	q70
u92	u85	w92	u92	w92	rh7
q70	v92	q92	q70	w85	v92
u70	*rh7*	v92	q10	u10	u85
w85	u92	u92	u85	v85	q10

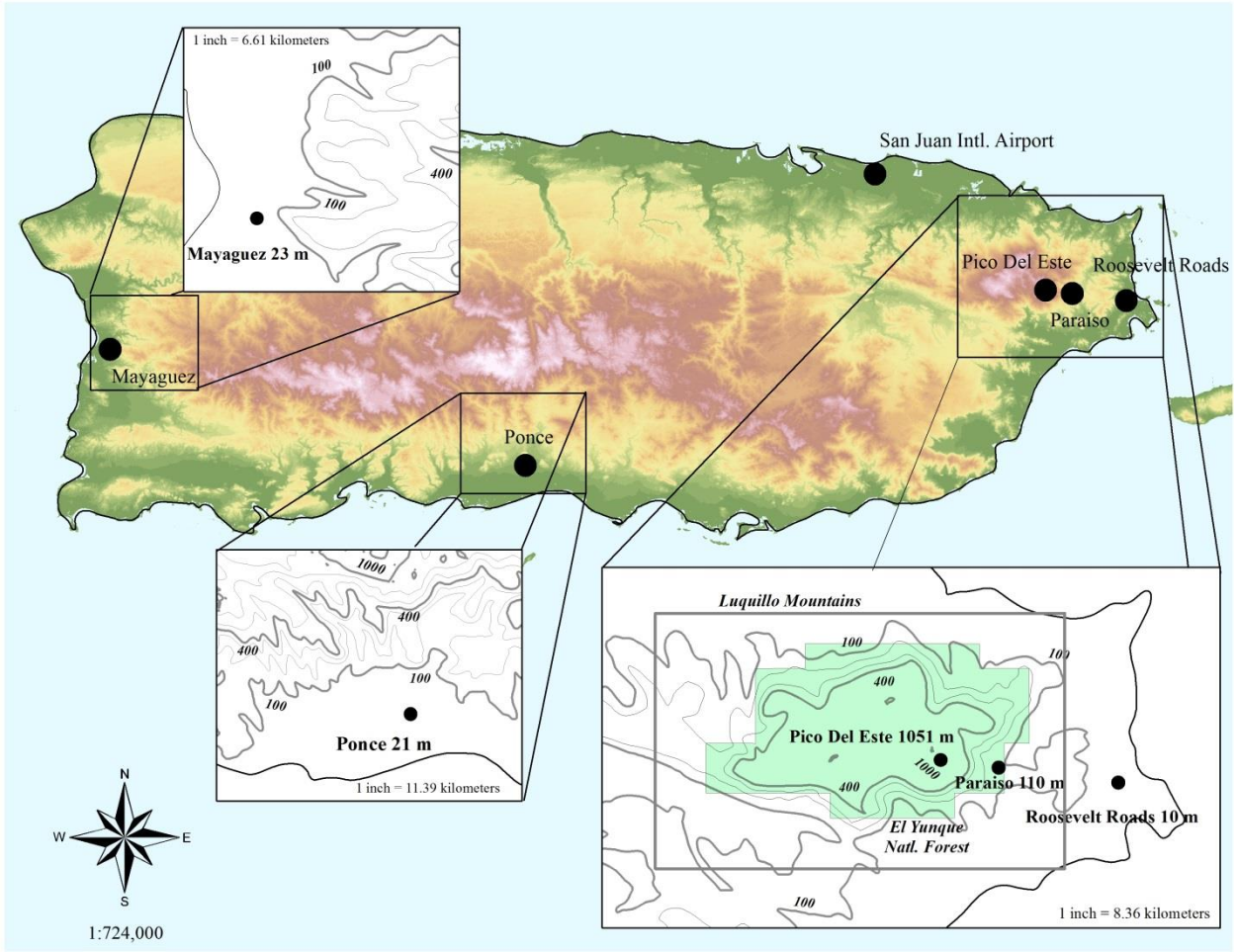

Fig. 2.1 Puerto Rico relief map showing the locations of the study sites. The inset maps show local contour maps for selected sites. The green polygon in the lower-right inset map denotes the boundary of the El Yunque National Forest.

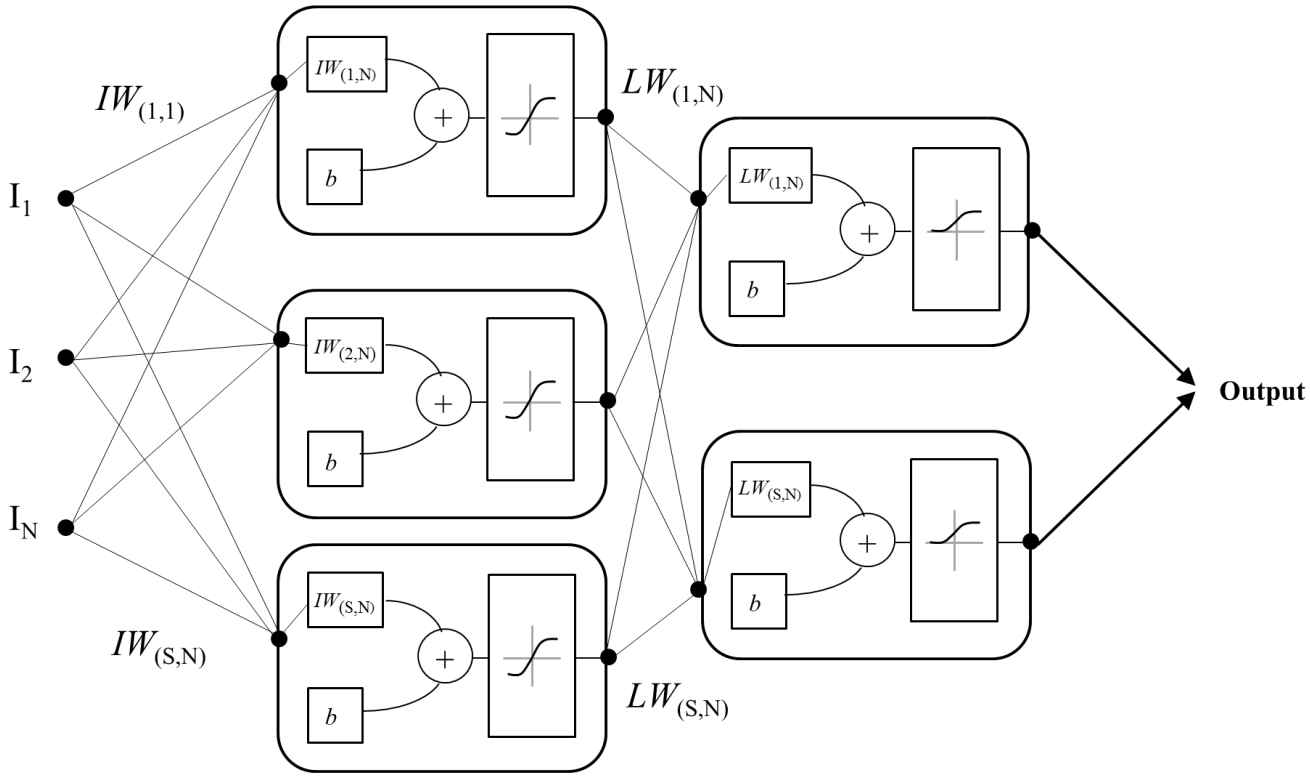

Fig. 2.2 Schematic of the architecture of the multilayer perceptron used in this study where the number of input weights (IW) and layer weights (LW) are a function of the number of inputs (N=37) and the number of hidden nodes (S=20).

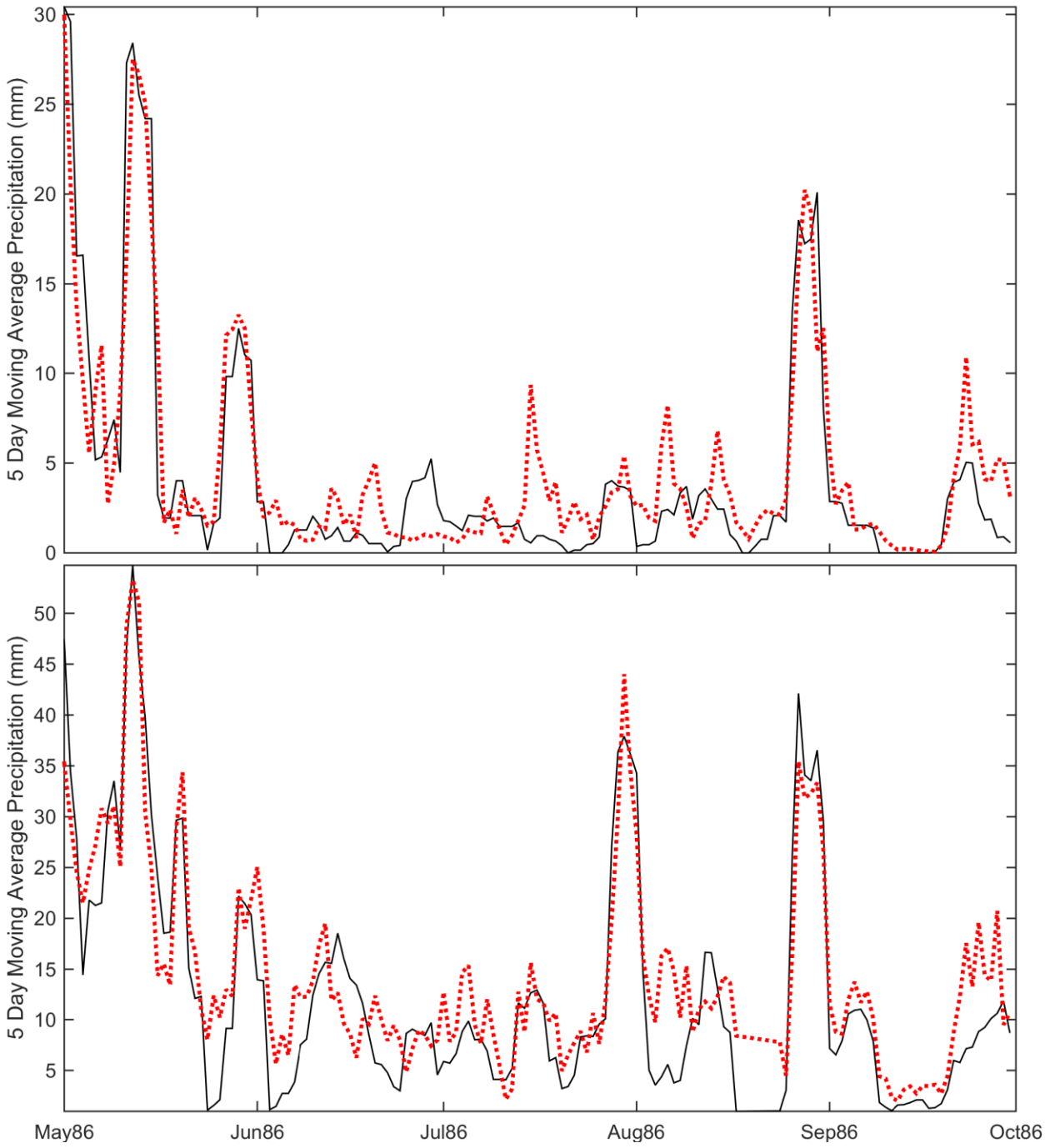

Fig. 2.3 Time series of predicted (red) and observed (black) smoothed rainfall for SJ (top) and PE (bottom) during the wet months from May 1986 – October 1986.

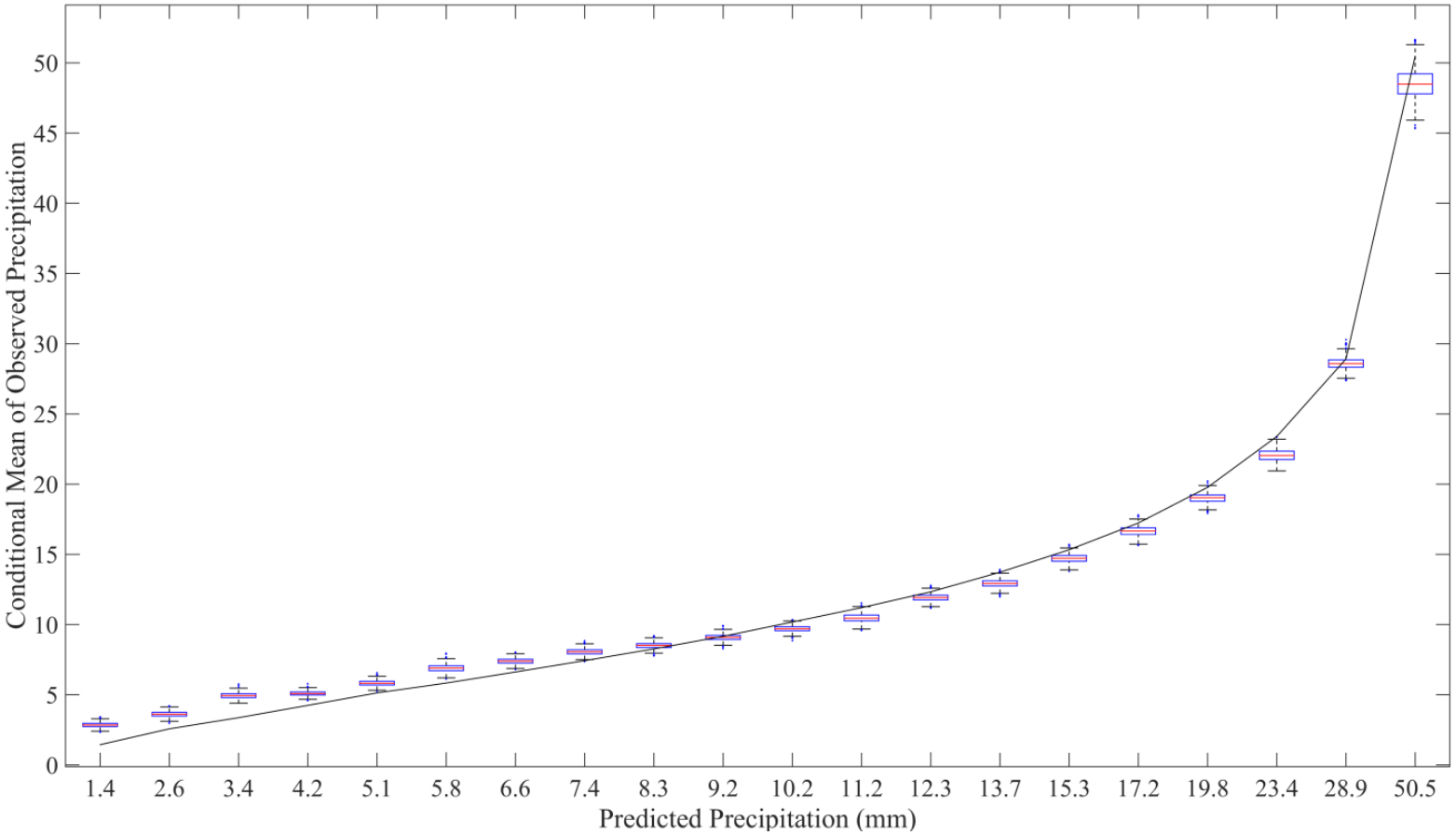

Fig. 2.4 Reliability plot for PE showing the conditional mean of the observed precipitation and the predicted precipitation. A reference line is added showing an optimal 1:1 relationship

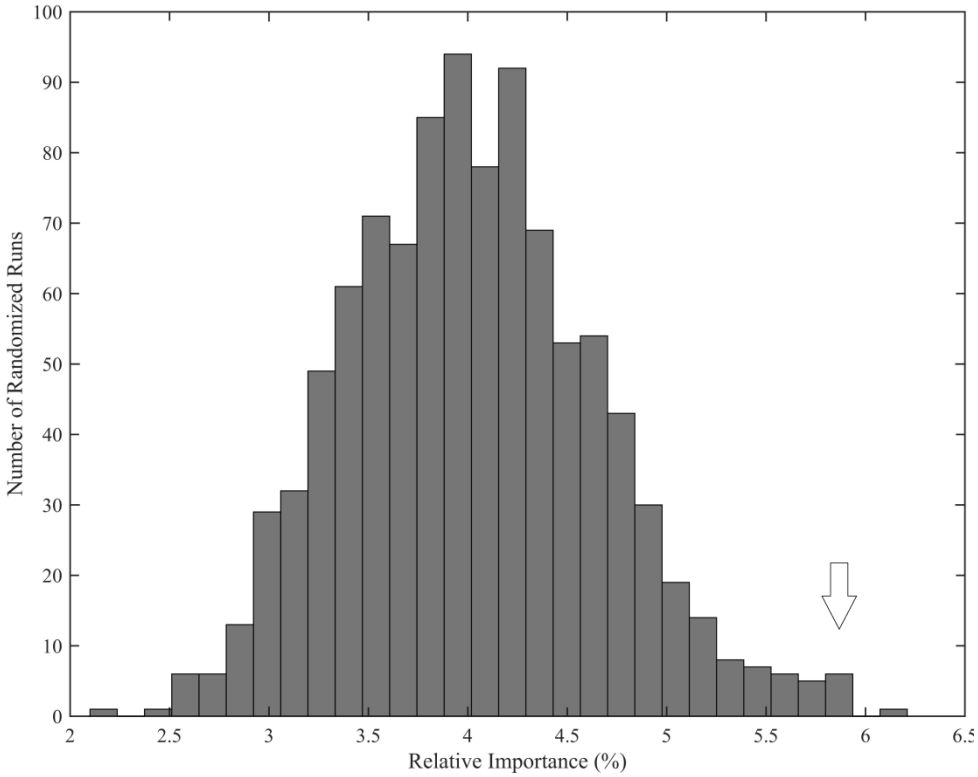

Fig. 2.5 The relative importance (%) of q0 for all randomized networks for PE. The arrow indicates the relative importance for the observed network, which in this example is statistically significant.

CHAPTER 3

ANALYZING REGIONAL CLIMATE FORCING ON HISTORICAL PRECIPITATION VARIABILITY IN NORTHEAST PUERTO RICO

[1]Ramseyer, C.A. and T.L. Mote. To be submitted to *International Journal of Climatology*

Abstract

The tropical forests of Northeast Puerto Rico (NE PR) and the Luquillo Mountains (LM) are a large repository for biodiversity and have an important role in regional biogeochemical processes. Precipitation is a key driver of the productivity of these sensitive ecosystems. This study analyzes historical precipitation variability from 1985-2014 along the topographic gradient of the LM. Self-organizing maps (SOMs) are used to produce synoptic modes of variability from ERA-Interim low-tropospheric moisture and circulation variables over the Caribbean and tropical north Atlantic. These modes of variability are linked to precipitation variability in NE PR. A probability density function of observed precipitation is calculated for each node (e.g. mode of variability) in the SOM. Changes in node frequency over the study period are calculated and evaluated for statistical significance. Results indicate that dry precipitation environments are represented by nodes with high 1000 – 700 hPa wind shear with low levels of 700 hPa moisture. Wet days in the study region are associated with moist low-tropospheric environments with low wind shear. Node frequency shifts over the period indicate a shift in dry season environments with lower 700 hPa moisture. Also, the early rainfall season (ERS) has experienced an increase in node frequency that represents a moister lower troposphere with lower bulk wind shear. Trends in extreme events indicate that dry season/ERS heavy rainfall events have decreased/increase in magnitude. Dry day frequency has decreased in the last 15 years. These results allude to changes in precipitation variability towards more frequent low precipitation events and seasonal changes to the magnitude of extreme rainfall events. Thus, despite little change in total

annual or seasonal rainfall over the study period, precipitation variability appears to be changing with deviations in synoptic environments.

3.1. Introduction and Background

The ecosystems of the tropical forests in Northeast Puerto Rico (NE PR) are sensitive to perturbations in climatic variability, particularly precipitation. For example, prolonged drought can alter rainfall and throughfall nutrient fluxes (e.g., Heartsill-Scalley et al. 2007) and reduce streamflow, creating barriers for migratory species (e.g., Pringle 1997; Benstead et al. 1999; Covich et al. 2006). These effects may become more frequent in tropical rainforests if climate change increases drought frequency (Larsen 2000; Giannini et al. 2001; Covich et al. 2003, Crook et al. 2009). Moreover, balancing water needs, between ecological integrity and extraction by expanding human development outside of NE PR, is an increasing problem due to low water storage capacity, the flashy nature of storm flows, and a high rate of loss from pipes (Pringle and Scatena 1998). A water budget for the El Yunque National Forest in NE PR developed in 2004, showed that on a typical day, 70 percent of water generated within the forest is diverted into municipal water supplies before reaching the ocean – up from an estimated 54% in 1994 (Crook et al. 2007). Due to the climate sensitivity of ecological and human systems in NE PR to changes in daily precipitation, a comprehensive understanding of regional climate forcing on daily precipitation variability is critical. This study uses self-organizing maps to determine the regional modes of variability in low-tropospheric moisture, winds, and geopotential heights and determine precipitation patterns in each synoptic type. This allows for examination of the regional environments that lead to wet/dry periods. Change in frequency in these regimes has impacts on precipitation variability in NE PR.

Synoptic climatological approaches involve the classification of GCM output into "main modes of atmospheric circulation," "synoptic types," or "synoptic patterns," which

are then related to surface climate variables. These approaches are sometimes considered a subset of statistical downscaling approaches (empirical-statistical or statistical-dynamical).

Many synoptic typing approaches use Empirical Orthogonal Function (EOF) analysis or Principal Component Analysis (PCA) to reduce dimensionality of atmospheric data and extract the primary modes of variability, typically from sea level pressure or geopotential height fields. The EOFs or PCs are then related to surface climate fields through multiple linear regression, canonical correlation analysis, or maximum covariance analysis (e.g., Vimont et al. 2010; Widmann et al. 2003; Mote 1998), sometimes combined with cluster analysis (e.g., Huth 2001).

One synoptic-typing approach that has become increasingly common over the past decade is self-organizing maps (SOMs; Lynch et al. 2006; Sheridan and Lee 2011). SOMs produce results similar to most other synoptic typing methodologies. SOM methodology is aimed at identifying and classifying nodes or clusters of large-scale atmospheric conditions that lead to the variability observed at a particular point or location (Sheridan and Lee 2011). SOMs attempt to find nodes in a measurement space that are representative of the observations and are able to describe the multidimensional distribution function of the data set. Given a multidimensional matrix of data points (observations or model output), the SOM places nodes within the data space (Hewitson and Crane 2002). The resultant distribution of the nodes will be representative of the multi-dimensional distribution function of the input data and the nodes will be closer together in data rich areas (Hewitson and Crane 2002). The number of nodes is ultimately decided by the user (Sheridan and Lee 2011; Sauter and Venema 2011). The nodes retain

the full dimensionality of the observed input data through the defining of an associated reference vector that is representative of the synoptic scale weather patterns (Reusch et al. 2005). A successful SOM will produce generalized patterns that are representative of the input data (Reusch et al. 2005).

SOMs offer advantages compared to other downscaling and synoptic classification techniques. SOMs offer a larger number of patterns that are easier to understand and interpret because the nodes exist on a continuum instead of discrete patterns, like those produced by EOF or PCA analysis (Sheridan and Lee 2011). The result allows for an understanding of transitional zones between atmospheric nodes, which can be helpful when examining subtle climate variability that may occur in the tropics and subtropics. The other distinctive advantage of SOMs is the capability to utilize multiple visualization techniques, which allow for more coherent and clear depiction of the SOM array (Sheridan and Lee 2011). This SOM approach is not without its disadvantages, which include the usual statistical downscaling issues (e.g., assumption of stationarity). Other issues include user error selecting the correct predictor variables and user error in the selection of the number of nodes used (Sheridan and Lee 2011).

One of the more common climatological applications of SOMs is for downscaling precipitation (Sheridan and Lee 2011). Some examples include applications in Texas and Mexico (Cavazos 2000), over the Canadian Arctic (Cassano and Cassano 2010), South Africa (Hewitson and Crane 2006) and the Asian monsoon region (Chattopadhyay et al. 2008). The SOM methodology has been used for precipitation over islands (e.g., Sobie and Weaver 2012) and on ice sheets and glaciers (Schuenemann and Cassano 2010; Reusch et al. 2005), which include areas with steep environmental gradients. This

technique has also been used to analyze climate zones over Puerto Rico (Malmgren and Winter 1999).

3.2 Methodology

a. Data

Precipitation data were acquired for El Verde, PR from the National Atmospheric Deposition Program (NADP). The El Verde rain gauge is located at an elevation of 380 m and provides daily data from 1985 – present. To provide a 30 year analysis, only the data up to 2014 are used. El Verde is located on the north facing slopes of the LM and is within the bounds of the El Yunque National Forest.

Daily precipitation data were acquired for the San Juan International Airport were acquired from the Global Historical Climatology Network-Daily dataset (Menne et al. 2012). These data were acquired for the period 1980 – 2014. Precipitation data from San Juan provides a low elevation, urban site to compare the relationship between precipitation and synoptic type.

This study uses mean daily atmospheric fields constructed from 6-hourly ERA-Interim reanalysis data from 1980 – 2014 (Dee et al. 2011). These data are available on a 1° × 1° global grid. The atmospheric variables used in this study are specific humidity, u-wind, v-wind, and geopotential heights and 1000 and 700 hPa. These variables have been shown to be the most important predictor variables of precipitation in Puerto Rico (Ramseyer and Mote 2016). Additionally, humidity has been identified as a critical predictor variable for all precipitation empirical downscaling methodologies (Cavazos and Hewitson 2005). Surface variables are important to include in regions where

orographic forcing plays a significant role as it helps improve the characterization of the atmospheric states (Hewitson and Crane 2006).

b. Methodology

This empirical downscaling methodology involves applying SOMs to define the range of atmospheric states for a localized domain to the target location. For each of the modes of variability defined by the SOM output, probability density functions (PDFs) of the rainfall distribution at the target location are constructed.

The SOM in this study is produced using the SOM Toolbox 2.0 which is available in MATLAB and is translated from the SOM_PAK software package (Kohonen et al. 1996). The closest ERA-Interim grid point to the El Verde rain gauge is used as the center point of a 10° × 10° domain grid, and the SOM is trained using the reanalysis atmospheric data from the 100 grid points in the domain. Using eight atmospheric variables, there is an 800-element vector that describes each atmospheric state (SOM node). The SOM is trained with over 13,000 daily data vectors. A 9 × 11 map size is used which allows for 99 atmospheric states. As the dimensionality of the input climatic data increases, larger map sizes are needed to most accurately represent the data space. This map size allows for a broad range of states to be represented including relatively lower frequency atmospheric states. These states could become more frequent in the future and are important to be captured by the SOM. Smaller maps would lead to increased generalization of the atmospheric states resulting in a loss of some of the less frequent states which are of interest in this study.

Prior to training, the input data were normalized by setting the variance to unity, and the mean to zero. The training parameters for the SOM were an initial update radius

of 7 reduced to 1 by the end of the training and a learning rate of 0.05. The SOM trained for 1000 iterations. At the end of training, each daily vector can be mapped to the best-matching unit (BMU). The BMU is determined by calculating the distance between the daily data vector and the node vectors and determining for which node vector that distance is smallest. Thus, each day in the time series uniquely maps to one of the atmospheric states. If new data are made available (e.g. GCM output), they can be mapped to the existing node space created by the SOM and its BMU can be determined.

In order to produce precipitation PDFs representative of each SOM node, all of the days that are assigned to a given node are determined and daily precipitation values for those days are acquired. This methodology follows Hewitson and Crane (2006), and is conducted for all 99 SOM nodes. Precipitation observations are sorted from low to high and include all occurrences of zero. A spline is fit to the sorted precipitation data. The spline curve is sampled by dividing the PDF into 100 equal units to account for nodes of differing numbers of observations. These 100 interpolated values are equivalent to the PDF for the particular node. For both rain gauge locations, 99 PDFs describe the rainfall distribution for each of the atmospheric states produced from the SOM.

To analyze shifts in synoptic type during the study period, the data are separated into two epochs, an early epoch from 1981 – 1990 and a late epoch from 2001 – 2010. Node frequency is calculated for each epoch and the change in frequency is calculated and evaluated for its statistical significance. The statistical significance in the difference in node frequency between the two epochs was calculated by producing a binomial distribution (Cassano et al. 2007; Skific and Francis 2012) . This test statistic is produced as a range in a 95% confidence interval:

$$= \pm 1.96 \sqrt{\frac{p_1(1-p_1)}{\frac{n}{3}} + \frac{p_2(1-p_2}{\frac{n}{3}}} \quad (1)$$

Where p_1 and p_2 are the expected node frequencies ($p = \frac{1}{99}$) and n is the number of days in each epoch (3650). To account for persistence of synoptic type, the degrees of freedom are approximated as $\frac{n}{3}$. An analysis of the persistence of synoptic type of the study domain indicates that the atmospheric variables used as input variables in the SOM remain in a similar pattern for three days. This was determined empirically by calculating node sequences of the same or adjacent nodes. This test is conducted to analyze annual changes as well as the seasonal changes between epochs.

3.3. Results

Several SOMs were produced before an optimal SOM was achieved. The SOM selected for analysis was had a map size of 9 x 11. This allowed for less frequent extreme environments to be represented as nodes. This map size also would allow for more than 100 observations per node if the observations were to be evenly distributed across the nodes. The SOM selected for analysis was trained using an intial neighborhood function radius of 5, an intial learning rate of 0.05, and a training length of 10,000 epochs.

The SOM produces 99 nodes which can be used to represent discrete modes of variability in the 1000 and 700 hPa specific humidity (q), u- and v- winds, and geopotential heights (z). After the SOM has completed training, the best matching unit (BMU) can be calculated for each daily 800-dimension vector. Daily vectors with matching BMUs can be composited to produce a mean climate state of each mode of variability. Internodal variability can be assessed by calculating the mean state for each mode. The spatial mean of each composite node map is calculated to easily differentiate

atmospheric variability between nodes. These nodes respresent the full spectrum of atmospheric variabililty for the study area.

High and low precipitation nodes are determined by taking the median rainfall value from the precipitation PDF for each of the 99 nodes. The low/high precipitation nodes are identified by taking the 12 nodes with the highest/lowest median rainfall. We assume that median rainfall provides a better representation of a "normal" value than the mean due to the propensity for low-frequency, extreme rainfall events (e.g. tropical cyclones) at the study sites which can skew mean values in the PDFs.

The ERA-Interim atmospheric data are composited by taking the mean state of each variable for every node. This approach allows for analysis of winds, moisture, and geopotential heights at the top and bottom of the lower troposphere for each synoptic mode of variability. When analyzing composited data for all indices in the 12 highest and lowest precipitation nodes, there is a reduction (~20%) in moisture at 1000 and 700 hPa in the dry versus wet nodes (Fig. 3.2 and 3.3). However, there are low precipitation nodes with higher specific humidity and wet nodes with lower specific humidity. This demonstrates that dry periods can occur in relatively moist environments during the early and late rainfall seasons.

The wind speed and wind direction spatial mean maps are derived from the composited u- and v-winds for each node (Fig. 3.2 and 3.3).. Wind speeds are generally higher in wet nodes; however, there are examples of wet conditions occurring under light winds. Wind direction variability in high and low precipitation nodes was distinctly different. 1000 and 700 hPa winds for the high precipitation nodes were all ENE to E ranging from 65 – 98°. This indicates that winds in the low-troposphere during the

highest precipitation days have an easterly component to them. Dry nodes feature a higher degree of variability in wind direction. 1000 hPa winds are NE to SE ranging from 35 – 137° while 75% of dry node 700 hPa winds have a westerly component. There are also instances where 700 hPa winds are northwesterly possibly suggesting influence from extratropical features. Thus, increasing deviations away from easterly low-tropospheric winds is indicative of lower precipitation in NE PR.

Due to the apparent difference in wind fields, wind speed and wind direction were used to calculate 1000 – 700 bulk wind shear (BWS). BWS is calculated by first determining meridional and zonal shear. Meridional shear (MS) is the difference in the v-wind at two atmospheric levels:

$$MS = |V_{1000} - V_{700}| \quad (1)$$

Vertical shear (VS) is calculated as the difference in the u-wind at two atmospheric levels:

$$VS = |U_{1000} - U_{700}| \quad (2)$$

BWS is then calculated as the vector:

$$BWS = \sqrt{MS^2 + VS^2} \quad (3)$$

Spatially averaged BWS ranged from 4.9 – 7.7 m/s for low precipitation nodes (Fig. 3.4) and 2.4 – 5.0 m/s for high precipitation nodes (Fig. 3.5). The increased BWS observed in low precipitation nodes is driven by both directional and speed shear. Unlike BWS, the geopotential height and thicknesses are highly variable across both the high and low precipitation nodes.

a. Epochal Changes

Epochal changes in relevant metrics are presented in Table 3.1. The values listed are spatial means of the nodes with statistically significant changes in frequency between epochs. Analysis of the dry season (DJFM) shows a shift to nodes with reduced lower tropospheric moisture with little change in wind shear and geopotential height. Despite the propensity for drier atmospheric conditions, precipitation metrics are relatively similar.

The ERS is defined as the period from April – July. The nodes with increasing frequency during the ERS months represent synoptic types with slightly higher moisture, lower 1000 hPa geopotential heights, and lower wind shear (~20%). This translates into higher median and mean daily precipitation as well as an increase frequency of heavy precipitation events. Additionally, there is a decreased frequency of dry days.

Analysis of the late rainfall season (LRS) from August – November exhibits increasing trends in synoptic types with slightly higher 700 hPa moisture and lower geopotential heights throughout the lower troposphere. Near surface moisture and wind shear across the increasingly frequent nodes were relatively unchanged compared to the nodes becoming less frequent in the latter epoch. Due to the minimal difference between the atmospheric circulation and moisture between the increasing and decreasing nodes, only slight changes in precipitation variability was observed. The increasingly frequent nodes are slightly higher variability as exhibited by a decrease in median daily precipitation but an increase in mean daily precipitation.

Changes in synoptic atmospheric variability are consistent with trends noted in the ERS and LRS. Increased moisture, lower geopotential heights, higher 1000 – 700 hPa thicknesses and lower wind shear (~ 30% reduction) are observed in the increasingly

frequent nodes as compared to the decreasingly frequent nodes. This shift in synoptic type frequency favors environments leading to slightly higher daily precipitation and a small increase in the frequency of high precipitation events.

b. Historical Precipitation Variability

To assess changes in extreme precipitation, percentile ranks were calculated to the entire dataset as well as a subset of days including only those days with measurable precipitation (non-zero days). Analysis of the interannual variability of the 95th percentile exhibits little trend in dry season and LRS precipitation (Fig. 3.6). However, the 95th percentile rank for the ERS has steadily increased over the study period. Median rainfall has increased slowly since the 1990s during the dry season, showed little trend in the LRS, and has been highly variable since 2000 during the ERS with a general increasing trend (Fig. 3.7).

In order to further analyze precipitation variability and extreme precipitation events, datasets including only precipitation events (daily precipitation > 0) were constructed and different trends emerge. When the term "event" or precipitation is used in following discussions it is referring to analysis on only those days with measurable precipitation and excludes dry days. This allows for a better representation of how rainfall events are changing and does not include the influence of dry day frequency, which is considered separately.

The amount of precipitation falling during the extreme precipitation events (95th percentile) during the dry season has decreased since 2000 (Fig. 3.8). In other words, the extremes have become less extreme during the dry season. Extreme events have increased/decreased in magnitude during the ERS/LRS, but these changes are not as large

as those observed during the dry season. Decreasing median rainfall is observed during dry season precipitation events, following the trend of the extreme rainfall events (Fig. 3.9). LRS median rainfall events have steadily increased since 2000. Most interestingly, ERS median rainfall events have decreased, contrary to the trend noted in the extreme events.

3.4 Discussion

A weak, positive linear trend in the San Juan and El Verde precipitation observations is observed. However, the fit of this line is not statistically significant. This indicates that the trend in yearly precipitation throughout the study period is largely unchanged. There is, however, high interannual yearly and seasonal precipitation variability throughout the period of observation (Figs. 3.6 – 3.9). This study attempts to elucidate trends in rainfall variability through the historical record.

El Verde has experienced a relatively stable LRS precipitation regime over the last 30 years with median and 95th percentile rainfall having relatively low variability with little trend through the period. Median LRS precipitation events have increased in magnitude over the past 15 years with a 1.5 mm increase over that period. This is accompanied by an 8 mm decrease in the amount of rainfall during extreme precipitation events.

The dry season has experienced up to a 30 mm decrease in 95th percentile rainfall and a 1 –1.5 mm decrease in median rainfall during precipitation events over the last 15 years. Over this period, dry day frequency has decreased (Fig. 3.10). Despite lower magnitude median and extreme rainfall events the number of days with precipitation has increased. When analyzing all precipitation observations (including dry days) there is

little discernible trend in median or extremes. This indicates an increase the frequency of days with measurable rainfall but with lower amounts of precipitation.

The ERS has seen the most change in precipitation variability over the period of analysis. Analysis of precipitation events indicates an increase in rainfall during extreme events with a decrease in median rainfall during precipitation events. This is accompanied by a steady decrease in dry day frequency over the last 30 years. There is no trend in total precipitation during the ERS through the study period. However, the analysis presented here shows that ERS rainfall variability has changed. El Verde has experienced an increase in the frequency of days with measureable rainfall while the extreme rainfall events have increased in magnitude. However, due to an increase in frequency of low magnitude precipitation events, the average (represented by the median) amount of rainfall falling on days with precipitation has decreased. ERS precipitation variability shows the greatest variability and the largest changes in variability over the historical period.

The results of this study identify the regional scale lower-tropospheric processes leading to the observed precipitation variability in NE PR. As two sites at different elevations were examined and provide similar findings, these two study sites provide realizations of the climatic forcing responsible for precipitation variability in NE PR. Results from the SOM indicate that lower-tropospheric wind fields and moisture variability are linked to precipitation variability. In particular, BWS appears to be the strongest indicator of precipitation variability in the nodes.

High daily precipitation is most likely when lower-tropospheric moisture (particularly at 700 hPa) is elevated and BWS is low. While there is variability in the

amount of moisture in these nodes, wind shear is uniformly low across all of the high precipitation nodes. Conversely, the strongest wind shear regimes coupled with low 700 hPa specific humidity lead to the lowest daily precipitation. Wind shear appears to be the dominant control on precipitation in Puerto Rico due to the inefficiency of trade wind cumuli development in high wind shear regimes. Additionally, high wind shear environments are not conducive to organized tropical systems.

Low precipitation environments are also related to drier moisture in the lower troposphere, especially at 700 hPa. As the Saharan Air Layer (SAL) is advected westward across the north Atlantic, it typically resides near 700 hPa over the eastern Caribbean (Prospero and Carlson 1972). This hot, dry elevated layer of air strengthens the existing trade wind inversion and increases atmospheric stability. The increased stability at this level acts as a capping mechanism on convection and vertical development of trade wind cumuli. The SAL could be responsible for some of the dryness detected in the low precipitation nodes as indicated by dry air being advected into the study area from the east.

There are linkages between the trends detected in precipitation observations and changes in synoptic variability over the historical period analyzed. This is particularly evident in the ERS as synoptic nodes that have increased in frequency have higher specific humidity and lower wind shear that those nodes that have decreased in frequency (Table 3.1). This has produced environments conducive to more extreme rainfall events and a lower frequency of dry days during the ERS. The dry season has experienced a decrease in both median and extreme event intensity. This is coincident with increased frequency of synoptic environments that are drier at 1000 and 700 hPa. The LRS has

exhibited the most stable precipitation regime and that is reflected in minor changes in synoptic moisture and circulation variability. Generally speaking, the historical record exhibits an increase in frequency of synoptic moister low-tropospheric environments with moister, slightly lower wind shear environments except during the dry season.

3.5 Conclusions

The general trend in total yearly precipitation in NE PR indicates a weakly, statistically insignificant positive increase. There is considerable interannual variability in the seasonal, monthly, and annual precipitation record. Thus, despite no trend in total precipitation, precipitation variability appeared to be changing. This study analyzes the historical trend in synoptic variability in low-tropospheric moisture, winds, and geopotential heights and produces a downscaling methodology to relate the synoptic types to precipitation in NE PR. Additionally, trends in precipitation events are analyzed to evaluate changes in variability.

The results of this study identify BWS as a primary forcing mechanism on precipitation in NE PR. The lowest precipitation days are associated with synoptic conditions with high wind shear in the lower troposphere. Wind shear in the dry nodes appears to be primarily driven by directional shear due to intrusions by the NASH. Moisture is also an important driver of precipitation variability, and results show that high levels of specific humidity exist throughout the entire depth of the lower troposphere during the wettest periods. However, moisture is not a limiting factor on precipitation as there are examples of low precipitation with high levels of moisture present. These nodes also were coincident with high wind shear which limits precipitation despite the high specific humidity.

There has been a decrease in dry day frequency across all seasons. Despite a higher frequency of precipitation days, the dry season is experiencing a decrease in median and extreme precipitation intensity. This is coincident with an increase in synoptic types with lower moisture in the lower troposphere. ERS extreme precipitation events have increased in intensity and appear to compensate for a decrease in the median event. An increase in days with low precipitation amounts appears to be contributing to this trend. There has been an increase in synoptic environments with higher moisture and lower wind shear during the ERS.

This study finds that precipitation variability seems to be changing despite little trend in total yearly precipitation. Changes in synoptic environments are at least partly to blame as dry season environments are getting drier and ERS environments are becoming wetter with decreased wind shear. Moving forward, changes in future precipitation variability in NE PR will be largely dependent on changes in the lower tropospheric wind shear regime and changes in moisture near 700 hPa. Shifts in climate regimes to higher frequency of synoptic types with increased wind shear and decreased moisture at 700 hPa would have large impacts on the ecology and forest dynamics of NE PR due to the higher incidence of dry spells.

3.6 References

Benstead, J. P., J. G. March, C. M. Pringle, and F. N. Scatena, 1999: Effects of a low-head dam and water abstraction on migratory tropical stream biota. *Ecological Applications*, **9**, 656–668.

Cassano, J.J., P. Uotila, and A. Lynch, 2007: Predicted changes in synoptic forcing of net precipitation in large Artic river basins during the 21st Century. *J Geophys Res Biogeosciences*, **112(4)**, 1 – 20, doi:10.1029/2006JG000332.

Cassano, E. N., and J. J. Cassano, 2010: Synoptic forcing of precipitation in the Mackenzie and Yukon River basins. *Int J Climatol*, **30**, 658–674.

Cavazos, T., 2000: Using self-organizing maps to investigate extreme climate events: An application to wintertime precipitation in the Balkans. *J Clim*, 1718–1732.

Cavazos, T., and B. C. Hewitson, 2005: Performance of NCEP-NCAR Reanalysis Variables in Statistical Downscaling of Daily Precipitation. *Clim Res*, **28,** 95–107.

Chattopadhyay, R., A. K. Sahai, B. N. Goswami, 2008: Objective identification of nonlinear convectively coupled phases of monsoon intraseasonal oscillation: implications for prediction. *J Atmos Sci*, **65**, 1549–1569.

Covich, A. P., T. A. Crowl, and F. N. Scatena, 2003: Effects of extreme low flows on freshwater shrimps in a perennial tropical stream. *Freshwater Biology*, **48**, 1199–1206.

Covich, A. P., T. A. Crowl, and T. Heartsill-Scalley, 2006: Effects of drought and hurricane disturbances on headwater distributions of palaemonid river shrimp (Macrobrachium spp.) in the Luquillo Mountains, Puerto Rico. *Journal of the North American Benthological Society*, **25**, 99–107.

Crook, K. E. , F. N. Scatena, and C. M. Pringle: 2007. Water withdrawal from the Luquillo Experimental Forest 2004. General Technical Report IIT-GTR-34, International Institute of Tropical Forestry, Pacific Northwest Research Station, 40 p.

Crook, K. E., C. M. Pringle, and M. C. Freeman, 2009: A method to assess longitudinal riverine connectivity in tropical streams dominated by migratory biota. *Aquatic Conservation* , **19**, 714–723.

Dee, D., and Coauthors, 2011: The ERA-Interim Reanalysis: Configuration and Performance of the Data Assimilation System. *Q J R Meteorolog Soc*, **137,** 553–597.

Giannini, A., Y. Kushnir, and M. A. Cane, 2001: Seasonality in the impact of ENSO and the North Atlantic high on Caribbean rainfall. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, **62**, 143–147.

Heartsill-Scalley T., Scatena F.N., Estrada C., McDowell W.H., Lugo A.E., 2007: Disturbance and long-term patterns of rainfall and throughfall nutrient fluxes in a subtropical forest in Puerto Rico. *J Hydrol*, **33**, 472–485.

Hewitson, B. C., and R. G. Crane, 2002: Self-organizing maps: applications to synoptic climatology. *Clim Res*, **22**, 13–26.

Hewitson, B. C., and R. G. Crane, 2006: Consensus between GCM climate change projections with empirical downscaling: precipitation downscaling over South Africa. *Int J Climatol*, **26,** 1315–1337.

Huth, R., 2001: Disaggregating climatic trends by classification of circulation patterns. *Int J Climatol*, **21**, 135–153.

Kohonen, T., Hynninen, J., Kangas, J., and J. Laaksonen. SOM_PAK: The Self-Organizing Map Program Package**.** Technical Report A31, Helsinki University of Technology, Laboratory of Computer and Information Science, FIN-02150 Espoo, Finland, 1996.

Larsen, M. C., 2000: Analysis of 20th century rainfall and streamflow to characterize drought and water resources in Puerto Rico. *Phys Geog*, **21**, 494–521.

Lynch, A. H., P. Uotila, and J. J. Cassano, 2006: Changes in synoptic weather patterns in the polar regions in the 20th and 21st centuries, Part 2: Antarctic. *Int J Climatol*, **26**, 1181–1199.

Malmgren, B., and A. Winter, 1999: Climatic zonation in Puerto Rico based on multivariate statistical analysis and an artificial neural network. *J Clim*, 12, 977–985.

Mote, T., 1998: Mid-tropospheric circulation and surface melt on the Greenland ice sheet. Part II: Synoptic climatology. *Int J Climatol*, **18**, 131–146.

Pringle, C. M., 1997: Exploring how disturbance is transmitted upstream: going against the flow. *Journal of the North American Benthological Society*, **16**, 425–438.

Pringle, C. M. and F. N. Scatena, 1998: Freshwater resource development: Case studies from Puerto Rico and Costa Rica, *Tropical managed ecosystems: New persepctives on sustainability.* U. Hatch and M. E. Swisher, Eds. Oxford University Press. 114–221.

Prospero, J.M. and T.N. Carlson, 1972: Vertical areal distributions of Saharan dust over the western equatorial North Atlantic Ocean. *J Geophys Res*, **77**, 5255–5265.

Ramseyer, C.A. and T. Mote, 2016: Atmospheric controls on Puerto Rico precipitation using artificial neural networks. *Clim Dyn*, doi: 10.1007/s00382-016-2980-3.

Reusch, D.B., R. B. Alley, and B. C. Hewitson, 2005: Relative performance of self-organizing maps and principle component analysis in pattern extraction from synthetic climatological data. *Polar Geog*, **29**, 188–212.

Sauter, T., V. Venema, 2011: Natural three-dimensional predictor domains for statistical precipitation downscaling. *J Clim*, 24, 6132–6145.

Schuenemann, K., and J. J. Cassano, 2010: Changes in synoptic weather patterns and Greenland precipitation in the 20th and 21st centuries 2: Analysis of 21st Century atmospheric changes using self-organizing maps. *J Geophys Res–Atmos*, **115**, D05108.

Sheridan, S. C., and C. C. Lee, 2011: The self-organizing map in synoptic climatological research. *Prog Phys Geog*, 35, 109–119.

Skific, N. and J. Francis, 2012: Self-organizing maps: a powerful tool for the atmospheric sciences, applications of self-organizing maps. In *Applications of self-organizing maps* (ed. M. Johnsson). InTech.

Sobie, S.R., and A. J. Weaver, 2012: Downscaling of Precipitation over Vancouver Island using a Synoptic Typing Approach. *Atmosphere-Ocean*, 50, 176–196.

Vimont, D. J., D. S. Battisti, and R. L. Naylor, 2010: Downscaling Indonesian precipitation using large-scale meteorological fields. *Int J Climatol*, 30, 1706–1722.

Widmann, M. M., C. S. Bretherton, and E. P. Salathe, 2003: Statistical precipitation downscaling over the Northwestern United States using numerically simulated precipitation as a predictor. *J Clim*, **16,** 799 – 816.

Winkler, J. A., and Coauthors, 2011: Climate Scenario Development and Applications for Local/Regional Climate Change Impact Assessments: An Overview for the Non-Climate Scientist. *Geogr Compass*, **5,** 275–300.

Table 3.1 Statistically significant increasing (+) and decreasing (-) node composite grid point averages for the dry season (DS), early rainfall season (ERS), late rainfall season (LRS), and all months (All).

	DS (+)	DS (-)	ERS (+)	ERS (-)	LRS (+)	LRS (-)	All (+)	All (-)
1000 hPa *q* (g/kg)	13.1	14.3	16.3	15.6	16.4	16.2	16.0	15.2
700 hPa *q* (g/kg)	2.4	2.8	4.7	4.1	5.2	4.6	4.5	3.4
1000 hPa *z* (m)	138	136	131	139	118	138	127	141
700 hPa *z* (m)	3149	3157	3174	3174	3163	3180	3168	3173
BWS (kts)	10.1	10.0	7.5	9.1	7.0	7.0	6.3	9.6
Dry Days (%)	38	38	30	38	31	29	31	36
Days > 20 mm (%)	9	10	17	11	15	13	14	12
EV median daily rainfall (mm)	1.6	1.6	3.4	1.6	2.8	3.1	2.5	2.4
EV mean daily rainfall (mm)	7.8	7.4	10.2	7.8	9.5	8.7	8.9	8.5

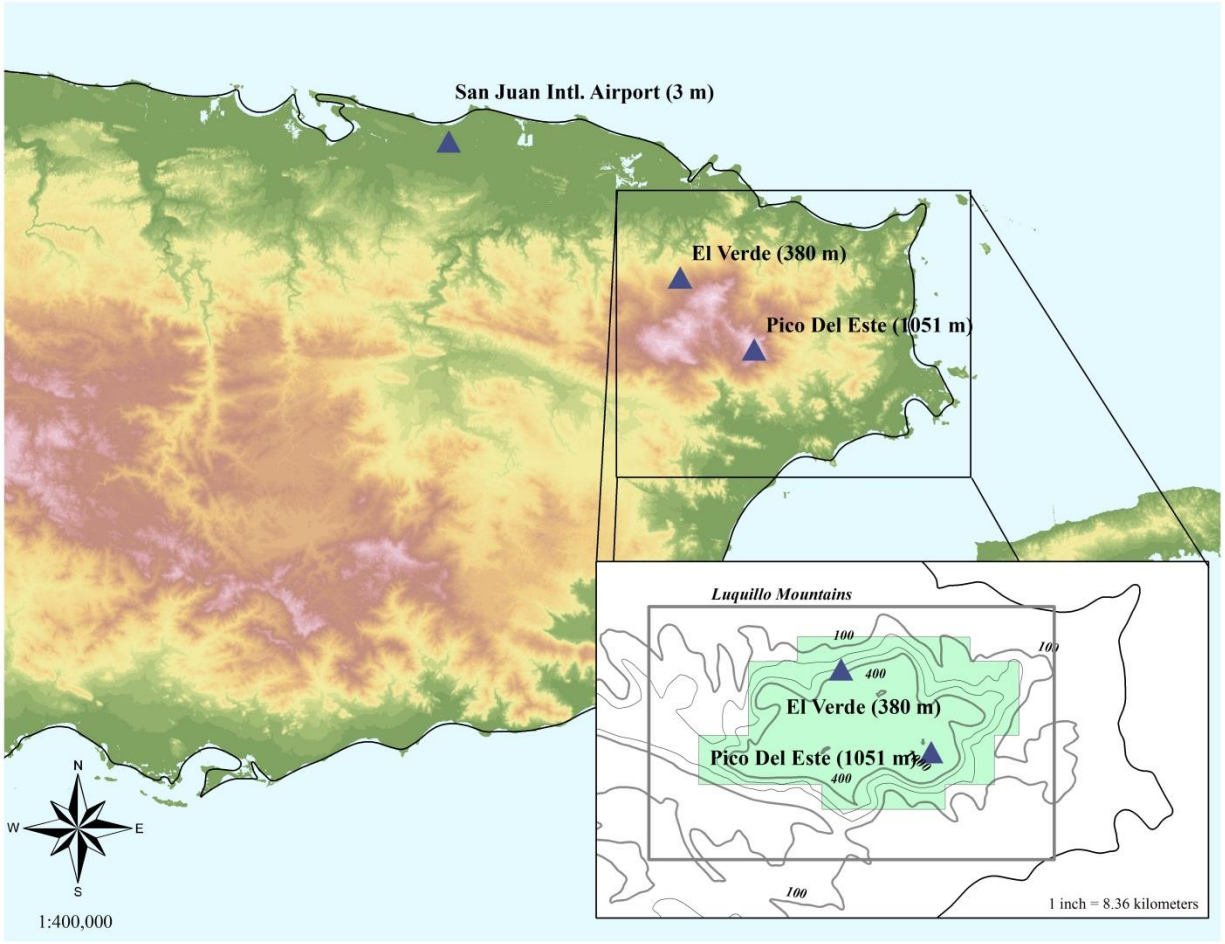

Fig. 3.1 Map of eastern Puerto Rico and a subset map highlighting the Luquillo Mountains and the El Yunque National Forest (green shading).

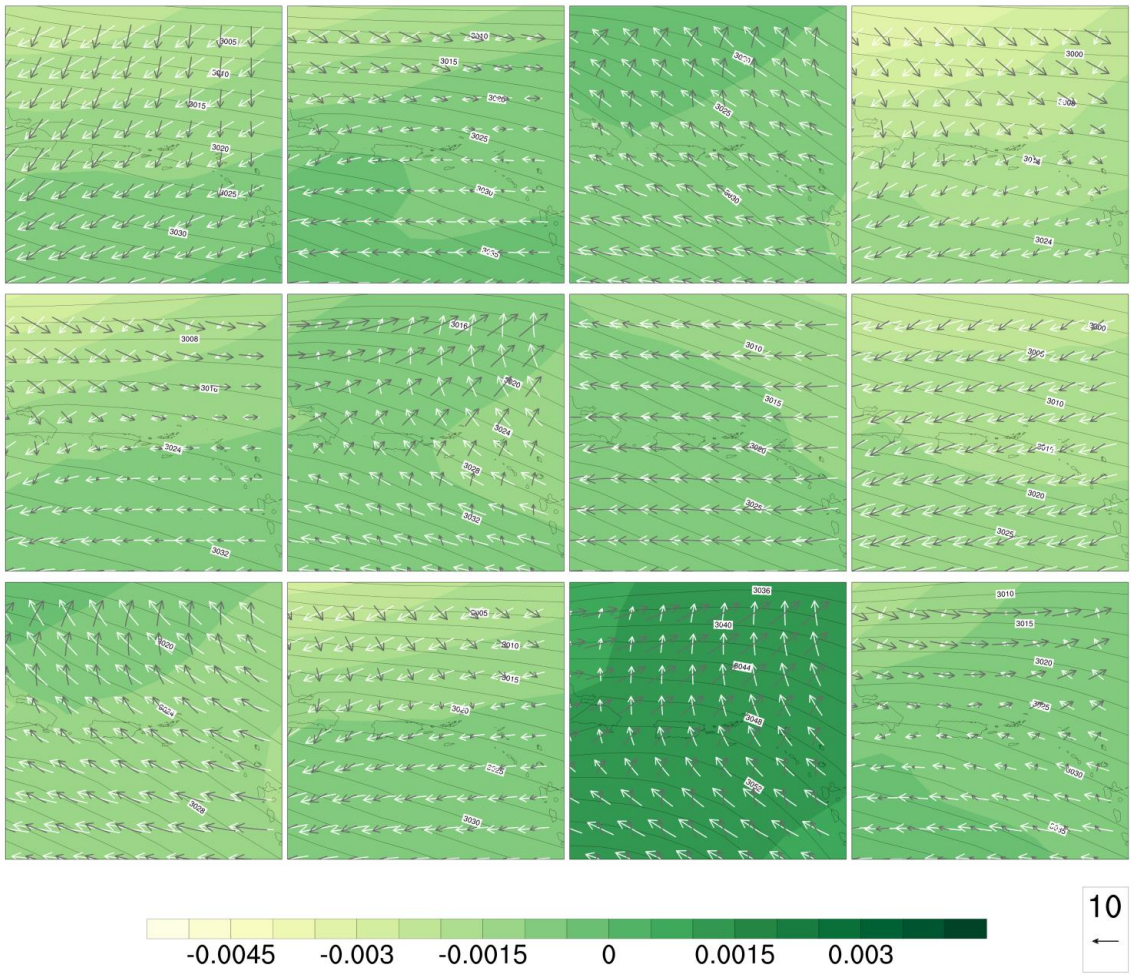

Fig. 3.2 Composite maps of 700 hPa specific humidity (green shading), 700 hPa winds (grey vectors), 1000 hPa winds (white vectors), and 700 hPa geopotential heights (contours) for the nodes associated with the twelve lowest precipitation PDFs.

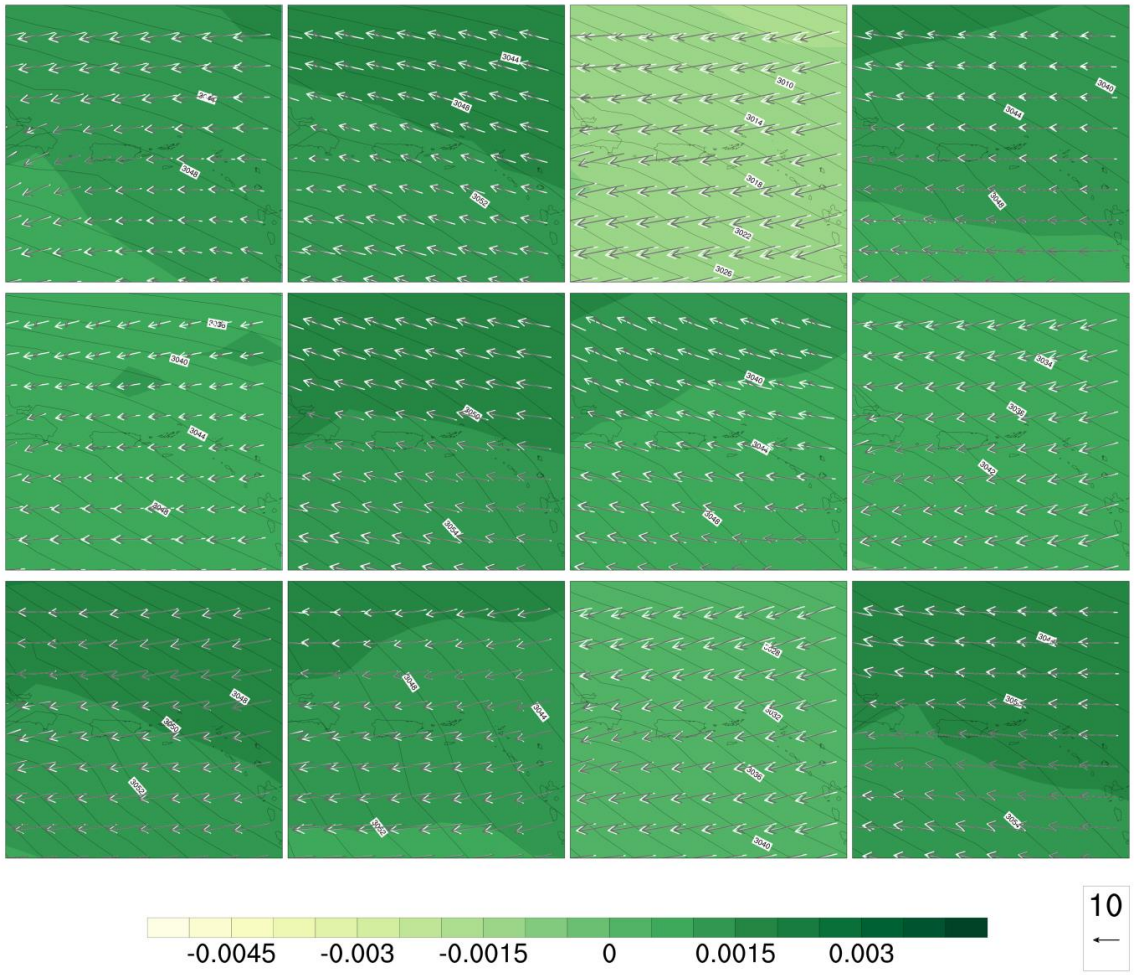

Fig. 3.3 Composite maps of 700 hPa specific humidity (green shading), 700 hPa winds (grey vectors), 1000 hPa winds (white vectors), and 700 hPa geopotential heights (contours) for the nodes associated with the twelve highest precipitation PDFs.

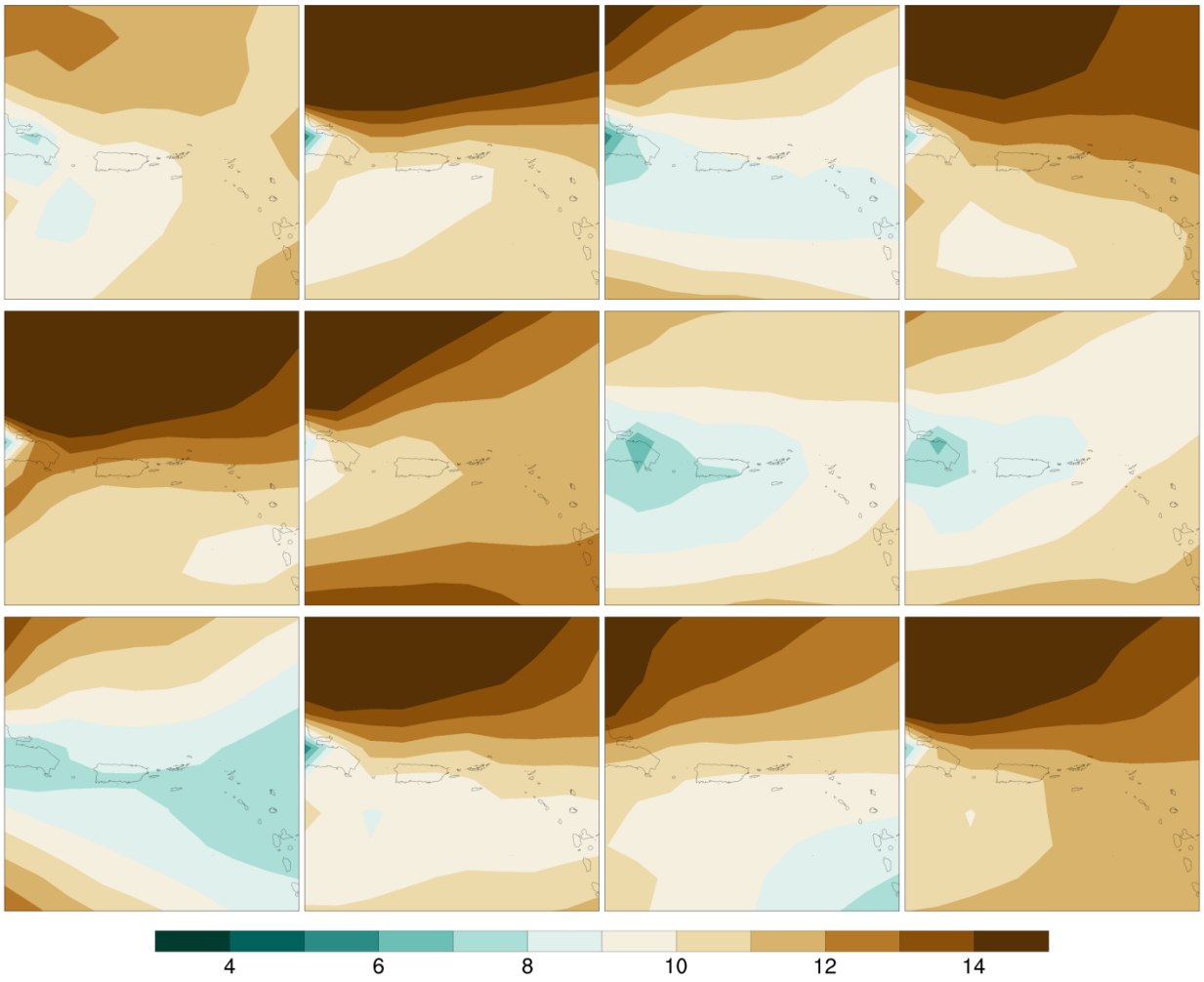

Fig. 3.4 Composite maps of 1000 – 700 hPa bulk wind shear for the twelve nodes associated with the lowest precipitation PDFs.

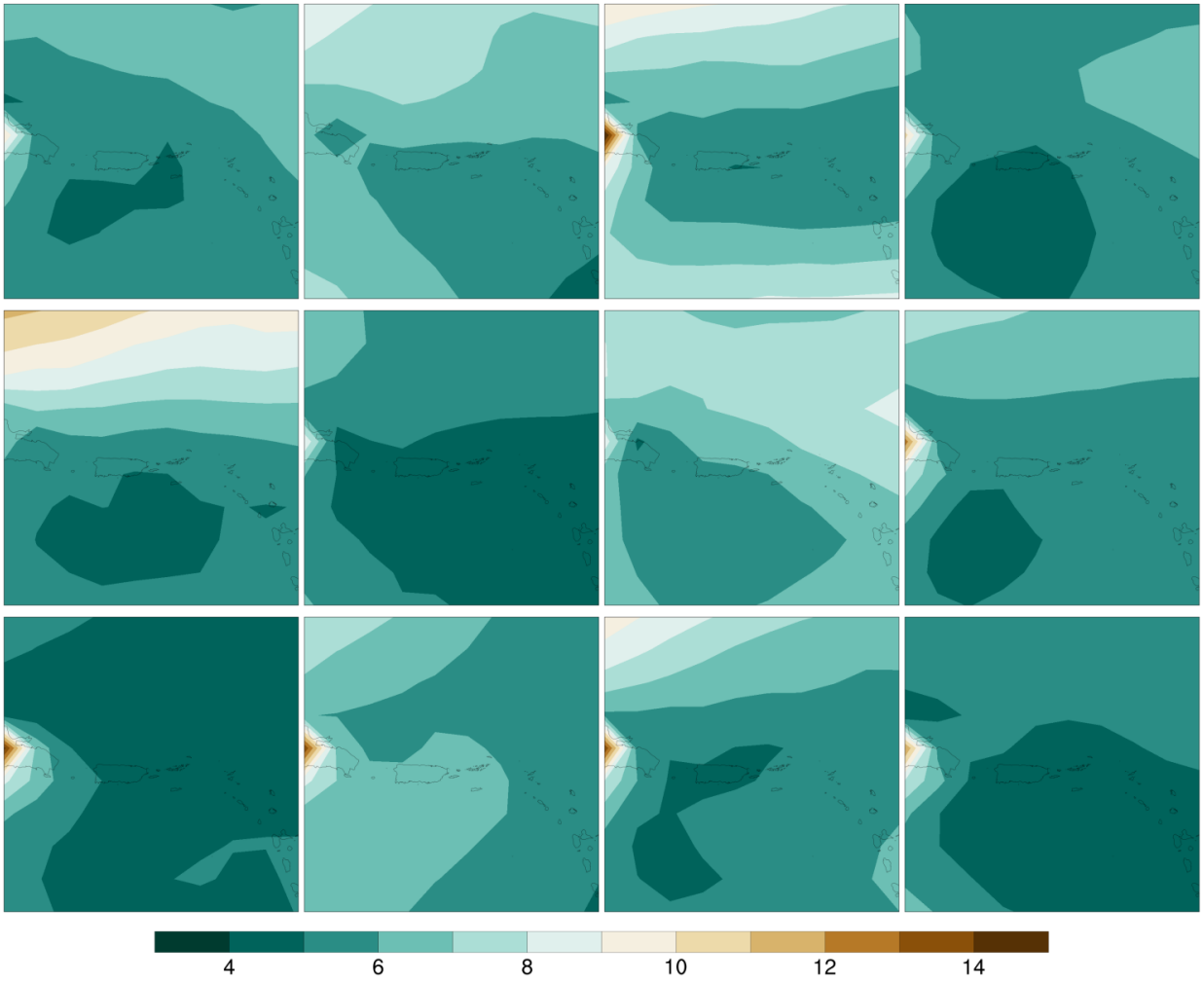

Fig. 3.5 Composite maps 1000 – 700 hPa bulk wind shear for the twelve nodes associated with the highest precipitation PDFs.

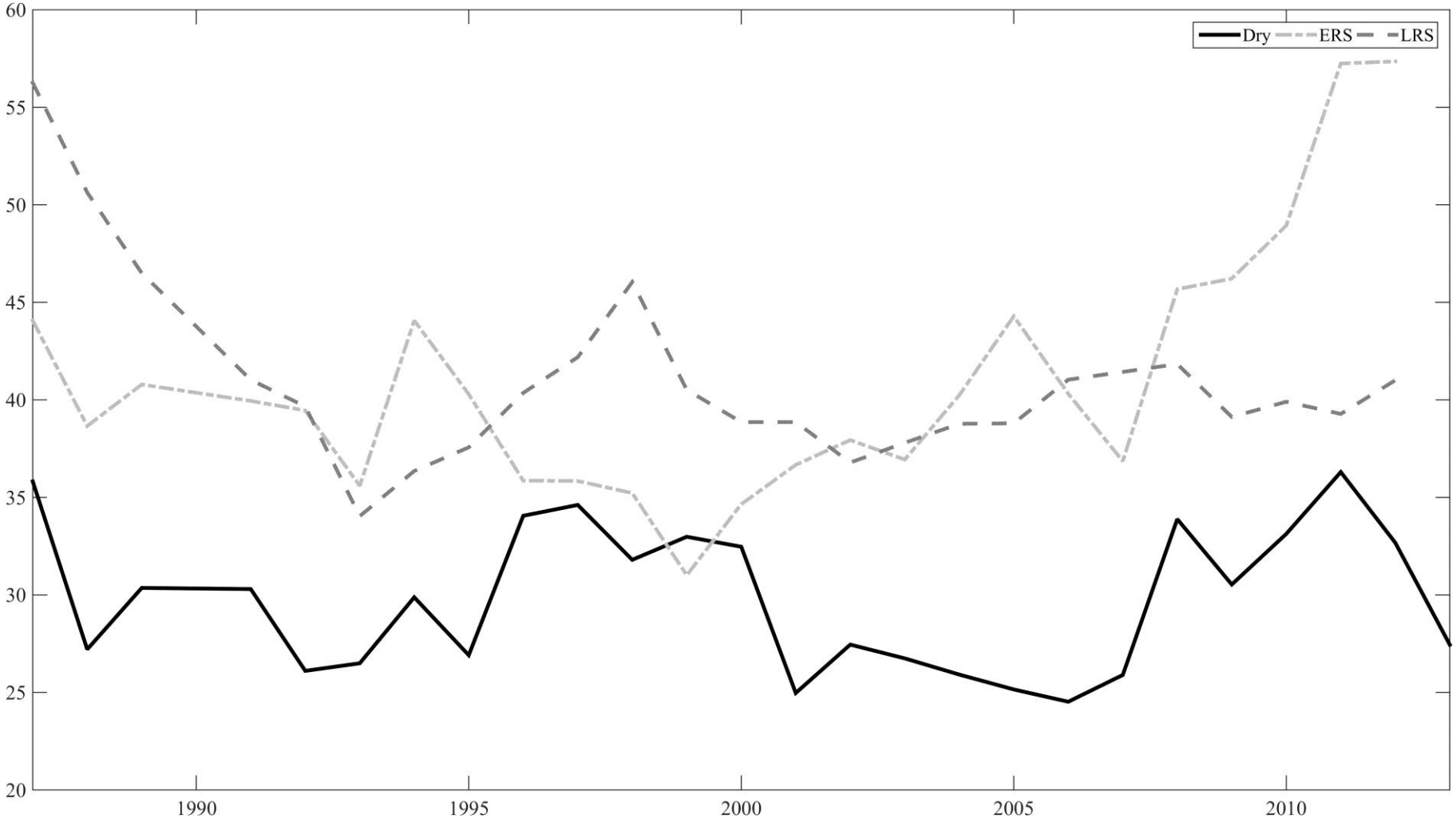

Fig. 3.6 Seasonal analysis of daily 95th percentile precipitation at El Verde.

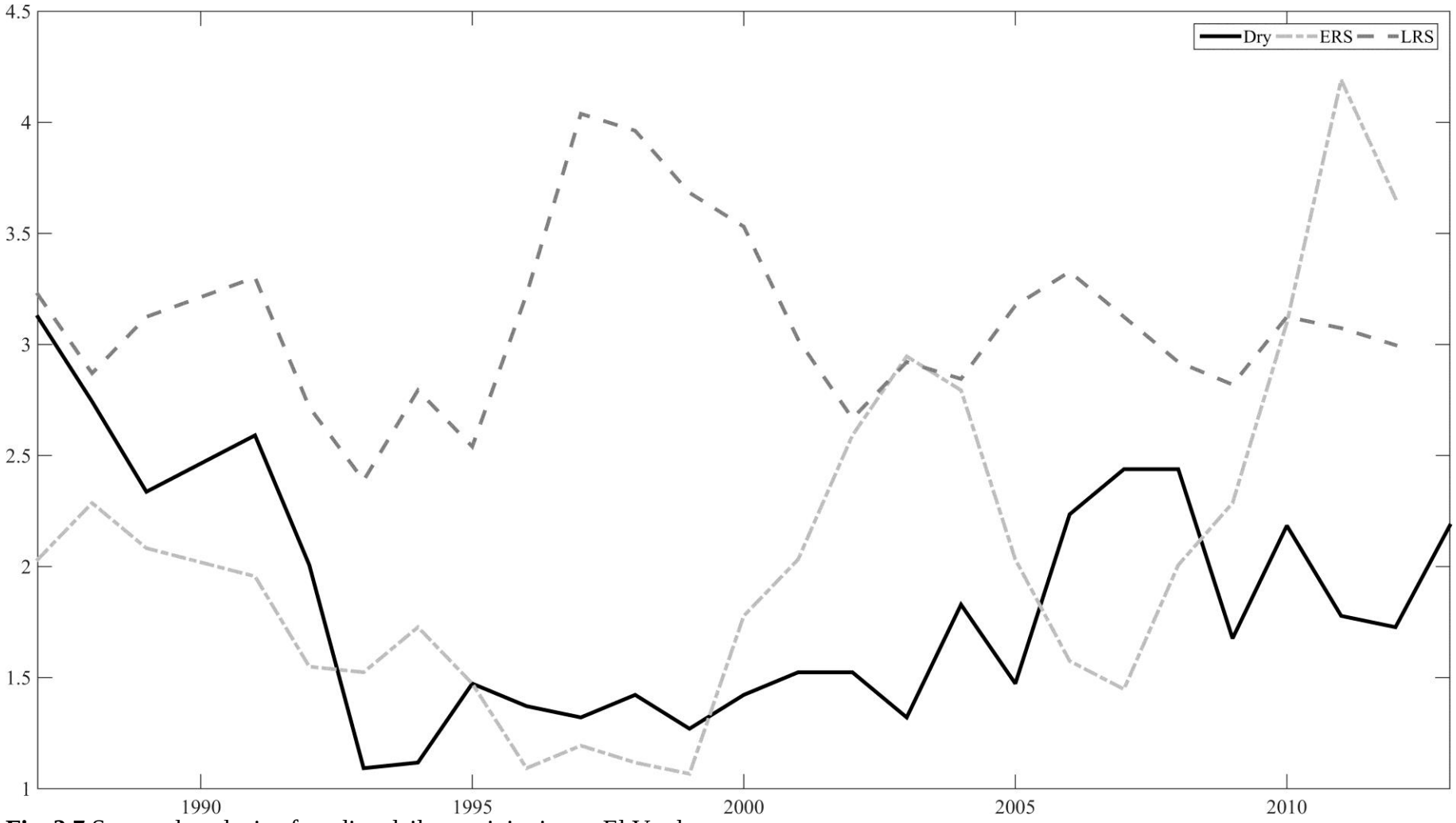

Fig. 3.7 Seasonal analysis of median daily precipitation at El Verde.

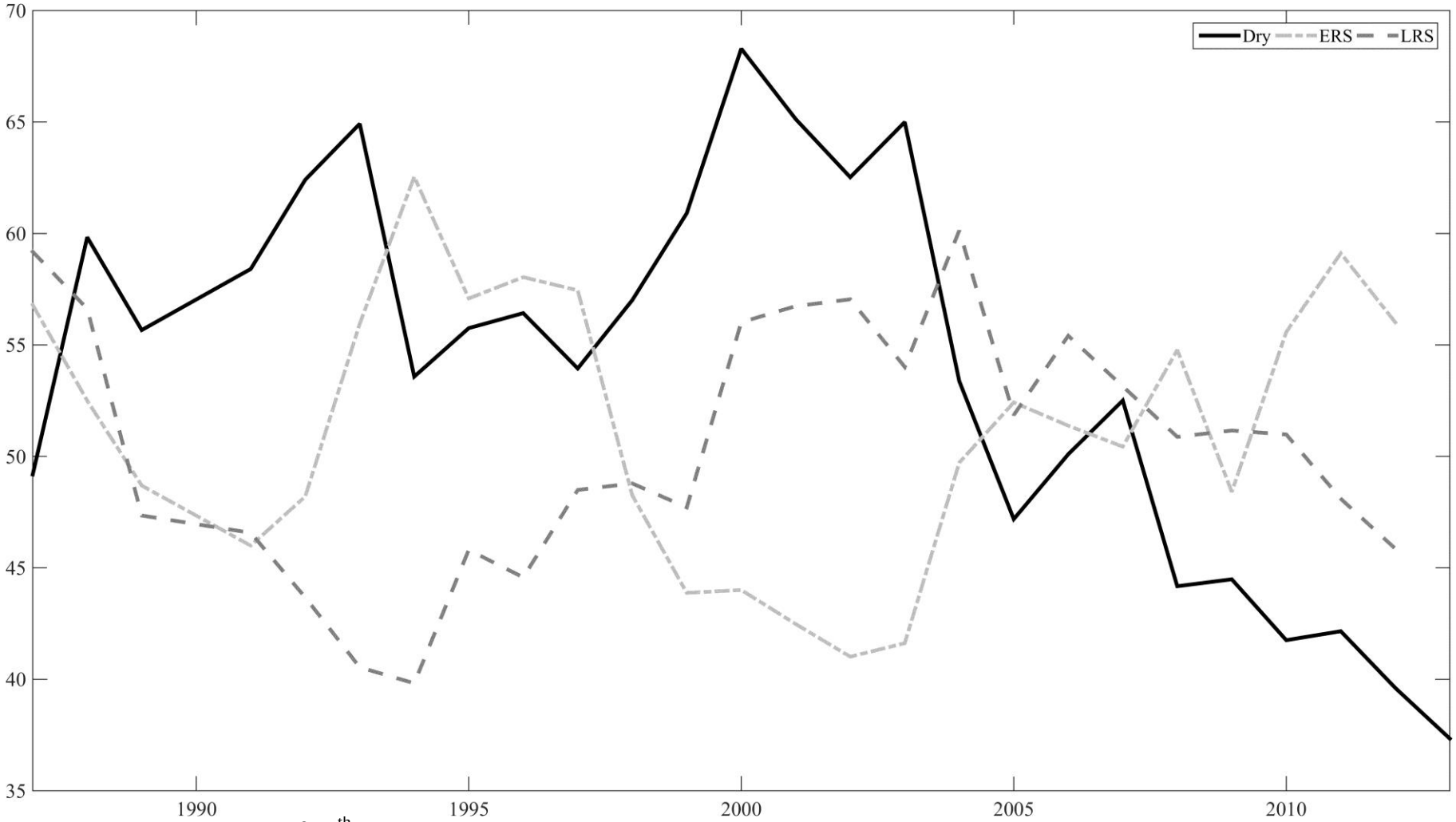

Fig. 3.8 Seasonal analysis of 95th percentile precipitation events at El Verde.

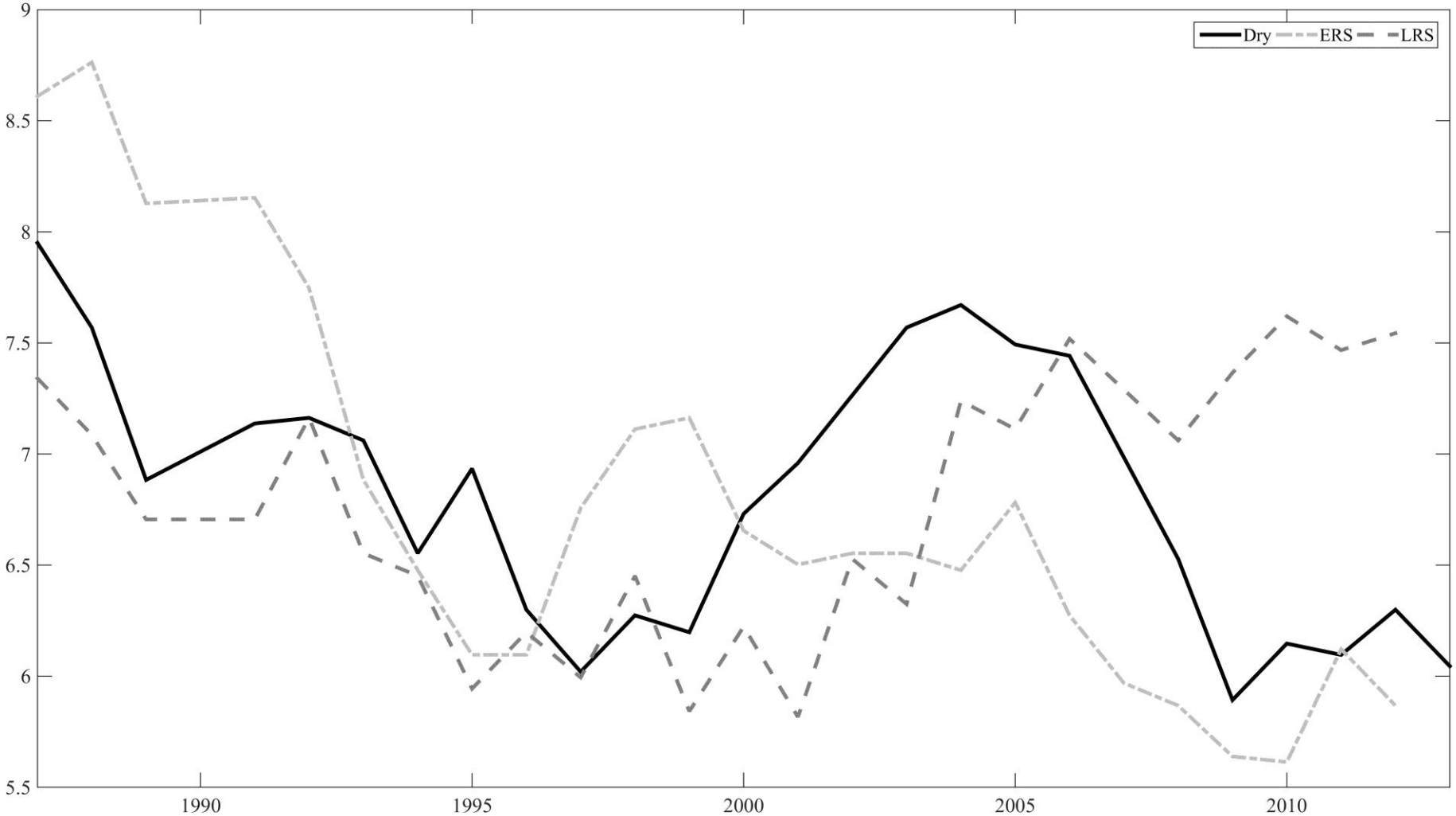

Fig. 3.9 Seasonal analysis of median precipitation events at El Verde.

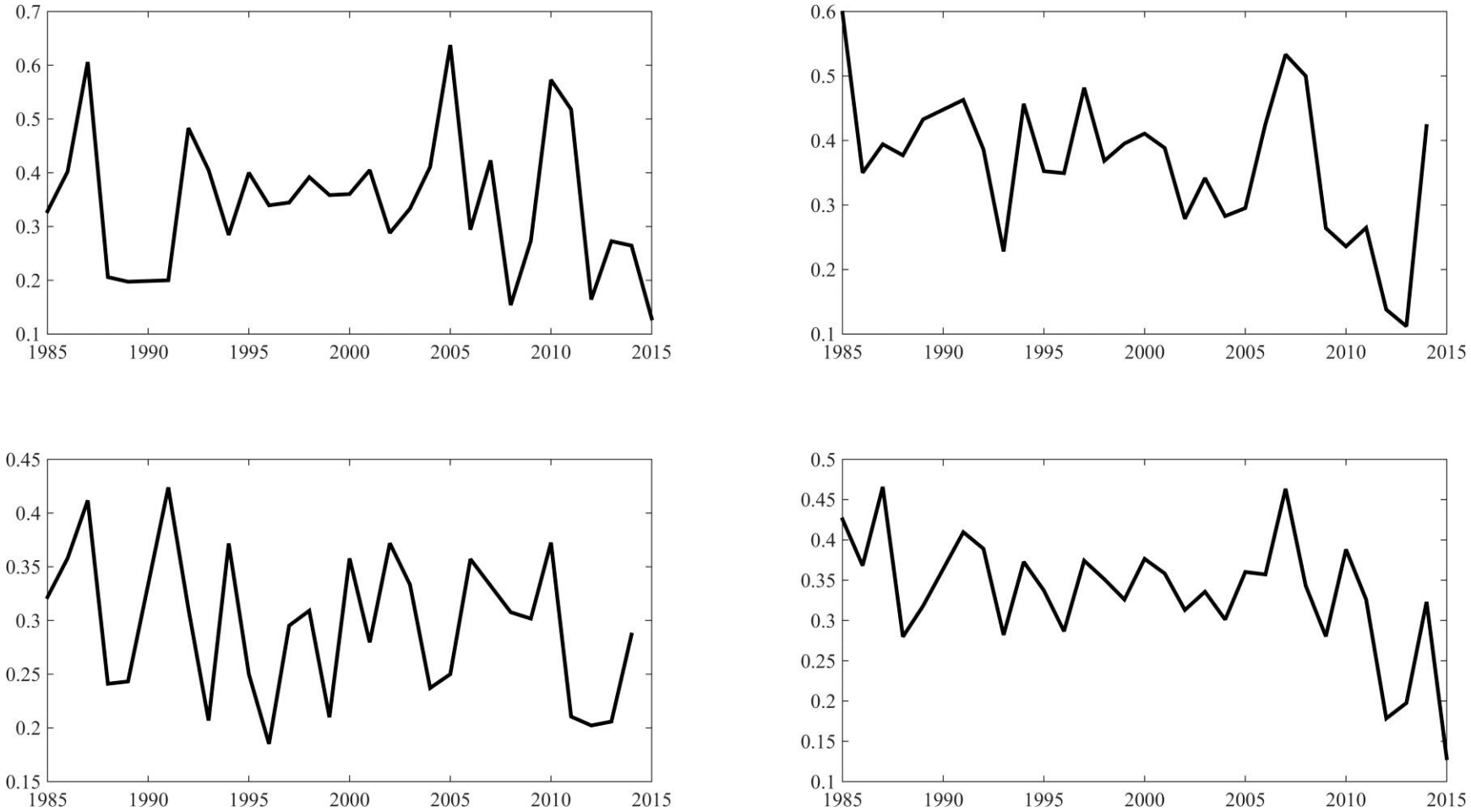

Fig. 3.10 Frequency of dry days for the dry season (upper-left), early rainfall season (upper-right), late rainfall season (lower-left), and all months (lower-right).

CHAPTER 4

FUTURE CHANGES IN LOWER TROPOSPHERIC ATMOSPHERIC VARIABLES AND PRECIPITATION VARIABILITY IN THE CARIBBEAN [1]

[1]Ramseyer, C.A. and T.L. Mote. To be submitted to *Journal of Climate*.

Abstract

The Luquillo Mountains (LM) of Northeast Puerto Rico NE PR have a steep precipitation gradient due to the variable topography, thus, producing reliable climate change projections for this region has been difficult. This study utilizes a downscaling methodology involving self-organizing maps to produce future scenarios of precipitation variability at rain gauges in NE PR. A suite of daily climate variables from ERA-Interim reanalysis are used to train a self-organizing map representing a continuum of synoptic types over the eastern Caribbean region. A probability density function (PDF) for each site's daily precipitation observations is produced for each type and can be randomly sampled to produce representative time series given each day's synoptic type. Historical CMIP5 simulations are mapped to the historical self-organizing map to determine each GCM's ability to reproduce the distribution of synoptic types. The best performing CMIP5 GCMs are used as an ensemble to produce future simulations of atmospheric and precipitation variability. By randomly sampling the precipitation PDFs for the synoptic type time series produced by the GCM ensemble, a modeled precipitation time series is produced. The future simulations show significant shifts in the distribution of synoptic types leading to a highly non-stationary future climate. Geopotential heights are projected to increase at both 1000 and 700 hPa with higher magnitude deviations at 700 hPa. The CMIP5 simulations have consilience in higher specific humidity at 1000 hPa and decreases in moisture at 700 hPa. Additionally, the GCMs indicate increasing 1000 – 700 hPa wind shear across the study domain. The future climate occupies a different node space and is representative of a completely new climate regime. Thus, the SOM based statistical downscaling methodology underestimates the impacts of anthropogenic climate

change of future precipitation variability. However, these research findings provide conclusive proof of the atmospheric changes driving the decreased precipitation projected by GCMs and other downscaling studies. The future synoptic environment in the eastern Caribbean likely includes a warmer lower troposphere with higher geopotential heights and pressure. Additionally, future synoptic regimes will have moister conditions near the surface as SSTs increase. However, 700 hPa specific humidity is likely to decrease due to greater intrusion from the Saharan Air Layer (SAL) and the North Atlantic Subtropical High (NASH).

4.1. Introduction and Background

Ecological and biogeochemical cycles in the tropical forests of Luquillo Mountains (LM) in Northeast Puerto Rico (NE PR) to disturbances and climatic variability, particularly precipitation. Dry spells can alter throughfall nutrient fluxes (e.g., Heartsill-Scalley et al. 2007) and reduce streamflow, which can cause widespread disruption to stream ecology (e.g., Pringle 1997, Benstead et al. 1999, Covich et al. 2006). If changes in the future climate system of the region leads to increased precipitation variability and increased dry spell frequency, the ecological and biogeochemical cycles of NE PR may become increasingly at risk of irreversible changes (Larsen 2000; Giannini et al. 2001; Covich et al. 2003, Crook et al. 2009). Future changes in precipitation variability would also have substantial impacts on water availability for the eastern half of the island. A water budget for the El Yunque National Forest in NE PR indicated that 70 percent of water generated within the forest is diverted into municipal water supplies, particularly to the San Juan Metropolitan Area (Crook et al. 2007).

Future precipitation changes are likely to be significant and are likely more important than temperature change for wet tropical ecosystems (Mulligan and Burke 2005). The response of the ecosystem to these precipitation changes will be important for the future of the forests of NE PR. This study analyzes climate scenarios for NE PR to determine how atmospheric variables that force localized precipitation variability are projected to change. These high resolution climate analyses will provide valuable context for future changes to ecological and biogeochemical processes in the LM.

This study analyzes future changes in atmospheric variables related to precipitation in NE PR. These projections of future climate are used to deduce precipitation variability based on an empirical-dynamical downscaling methodology utilizing self-organizing maps. Low-tropospheric specific humidity and winds have been identified as leading forcing mechanisms on precipitation on NE PR precipitation (Ramseyer and Mote 2016). The need for higher resolution projections of climatic variability are needed to confirm or refute precipitation projections from global climate models (GCMs). GCMs predict robust decreases in precipitation in the Caribbean and adjacent land areas (Neelin et al. 2006). CMIP5 models have shown good agreement that summertime precipitation will be reduced in the Caribbean (Maloney et al. 2014). The topography of NE PR presents unique modeling challenges as the elevation gradient leads to a sharp rainfall gradient. Due to the coarse resolution of GCMs, this gradient is not resolved. Thus, additional information is needed to determine if NE PR precipitation trends will follow that of the GCMs or does additional mechanical lift provided by the topography lead to mitigation of the dry trend projected for this region.

The projected response of climate in future projections on regional and local scales is more complex than for resolution of synoptic and global scale processes, especially on shorter temporal scales (Maloney et al. 2014). Of regional interest, Rauscher et al. (2008) found that the midsummer dry spell (MSD) in Mexico and Central American is projected to begin earlier and be stronger in magnitude. The MSD has been discussed as a climate feature evident in much of the Caribbean Basin and is driven by a number of synoptic scale forcing mechanisms including the expansion of the North Atlantic Subtropical High (NASH; Magana et al. 1999, Gamble et al 2008). CMIP3 and

CMIP5 models have also been found to capture the spatial and temporal aspects of the MSD (Rauscher et al. 2008, Sheffield et al. 2013). This indicates that CMIP5 models are able to reproduce the synoptic circulation patterns of the Caribbean basin.

A CMIP5 ensemble projection most pronounced drying is projected to occur in the Caribbean during July and August (Maloney et al. 2014). Additionally, the MSD is projected to strengthen in both the RCP 4.5 and 8.5 emissions simulations. In most Caribbean locations, this is primarily the result of a higher magnitude decrease in early summer rainfall relative to the late summer rainfall peak (Maloney et al. 2014). CMIP5 models have also been shown to adequately represent other intraseasonal variability, in particular the Madden-Julian Oscillation (Sheffield et al. 2013). One aspect of the Caribbean climate that CMIP5 models to have large uncertainty and intermodal disagreement is in future projections of North Atlantic hurricane frequency (Maloney et al. 2014).

4.2. Data and Methods

a. Historical Climate Data

Precipitation data were acquired for El Verde, PR from the National Atmospheric Deposition Program (NADP). The El Verde rain gauge is located at an elevation of 380 m and provides daily data from 1985 – present. To provide a 30 year analysis, only the data up to 2014 are used. El Verde is located on the north facing slopes of the LM and is within the bounds of the El Yunque National Forest.

Daily precipitation data were acquired for the San Juan International Airport were acquired from the Global Historical Climatology Network-Daily dataset (Menne et al. 2012). These data were acquired for the period 1980 – 2014. Analyzing precipitation data

from San Juan provides a low elevation, urban site to compare the relationship between precipitation and synoptic type.

This study uses mean daily atmospheric fields constructed from 6-hourly ERA-Interim reanalysis data from 1980 – 2005 (Dee et al. 2011). These data are available on a 1° × 1° global grid. The atmospheric variables used in this study are specific humidity, u-wind, v-wind, and geopotential heights and 1000 and 700 hPa. These variables have been shown to be important predictor variables of precipitation in Puerto Rico (Ramseyer and Mote 2016). Humidity has been identified as a critical predictor variable for all precipitation empirical downscaling methodologies (Cavazos and Hewitson 2005). Surface variables are important to include in regions where orographic forcing plays a significant role as it helps improve the characterization of the atmospheric states (Hewitson and Crane 2006).

b. General Circulation Model Data

General circulation model (GCM) data were acquired from the fifth phase of the Coupled Model Intercomparison Project (CMIP). CMIP is a standard experimental framework for global coupled ocean-atmosphere general circulation models produced in collaboration with the Intergovernmental Panel on Climate Change (IPCC) fifth annual report (IPCC 2014). The CMIP5 data allows for ensemble analyses of emissions feedbacks on the climate system. Historical runs of these models allows for validation studies to determine optimal GCM ensembles for individual study domains and methodologies. Future simulations are available for four different emissions scenarios, termed Representation Concentration Pathways (RCPs). They describe four possible

climate futures, all of which are deemed possible depending on greenhouse gas emissions over the next few decades (Fig. 4.1).

CMIP5 data were acquired from the World Data Center for Climate (WDCC) for those modeling studies that had available daily specific humidity, u-wind, v-wind, and geopotential heights data at 1000 and 700 hPa. Additionally, all model data downloaded were available for the historical, RCP 4.5, and RCP 8.5. There were 23 modeling experiments that had all of these data available. RCP 4.5 and RCP 8.5 are the most common pathways available and allowed for the most model members to be analyzed. RCP 4.5 represents a "middle-of-the-road" emissions scenario while RCP 8.5 represents an acceleration of current global emissions output leading to large emissions concentration in the atmosphere (Fig. 4.1). Analyzing these two RCPs allows for a range of scenarios of the future climate circulation.

c. Methodology

A novel validation methodology was developed in order to determine which CMIP5 members provided an optimal ensemble. This involved evaluating each GCMs ability to reproduce the distribution of synoptic types observed when mapping the ERA-Interim reanalysis data to the SOM. Each GCM has a unique model grid and resolution, which makes them incompatible with mapping to the ERA-Interim SOM. Thus, the data for each GCM were remapped to the ERA-Interim grid. Climate data operators (CDO) remapping command-line functions were utilized to conduct first order conservative remapping of the GCM grids (Jones 1999). Data were spatially constrained to the same 10 x 10 grid that was used as input to train the ERA-Interim SOM (Fig. 4.2). This enabled each GCM to be mapped to the SOM and for the best matching unit (BMU) to be

calculated. The BMU is the 800-dimensional vector representing one of the 99 nodes that is closest to the input 800-dimensional vector. GCM temporal coverage periods varied, thus, only the ERA-Interim node distribution that corresponds with each GCMs temporal domain is analyzed.

A time series of downscaled precipitation data can be produced by using the PDFs and the BMU time series. For a particular day, the BMU determines which PDF represents the precipitation distribution from which a random value will be extracted. A random number generator is used to produce a value (*r*) from zero to one. This value is then scaled by multiplying by the number of interpolated values in the PDF, in this study, 100. The scaled *r* value dictates where the PDF is sampled from. That value is recorded as the daily precipitation value and the process is repeated for all steps in the time series. This methodology has been shown to produce rainfall events of the correct magnitude but slightly underestimates the number of days with measurable rainfall in regions with high rainfall (Hewitson and Crane 2006).

When an atmospheric state occurs for multiple days in a row and precipitation occurs on day 1, it is more likely that the subsequent days will have precipitation. To account for this persistence, if rainfall occurs on day 1 then $\sqrt{r}$ is used for day 2 in order to nudge *r* toward the higher end of the PDF. This has been shown to provide ample compensation for persistence in areas with high rainfall (Hewitson and Crane 2006).

This methodology is repeated for the entire length of the time series. In order to account for sampling variability of the PDFs, 100 different time series are constructed. For each time series, the monthly mean, monthly median, and monthly standard deviation are calculated. The monthly time series that most closely matches the median monthly

time series is selected and the daily form is used as a representative time series of daily values (Hewitson and Crane 2006).

The SOM downscaling methodology was validated by producing time series from historical CMIP5 GCM atmospheric fields. Historical GCM simulations were acquired for 23 CMIP5 models (Table 4.1). The CMIP5 GCM specific humidity, geopotential height and wind fields at 700 and 1000 hPa were mapped to the SOM produced using the ERA-Interim reanalysis data. This involves mapping each 8000-dimensional (100 grid points x 8 atmospheric variables) daily vector from the GCM to the closest-matching node vector in the SOM. From this, a daily time series of the best-matching node (e.g. 1–99) is produced for each GCM for their respective historical temporal periods. In order to account for different time series lengths for the GCMs, a node frequency was calculated. The GCM node frequencies were then compared to the ERA-Interim node frequencies. In order to produce an optimal ensemble of GCMs, those models that best reproduced the ERA-Interim node frequency distribution were used. The performance of the GCMs was determined by summing the difference between the reanalysis node frequency and the GCM node frequency. Due to the high number of nodes needed in this study, node differences were calculated by also including the three closest node vectors. For example, the vector in node 1 is most similar to node vectors 2, 10, and 11. These four nodes were considered together when calculating the GCMs ability to reproduce node 1. This validation methodology identifies the GCMs that best reproduce the daily regional circulation patterns. The seven GCMs with the best performance were used to provide an ensemble of future atmospheric conditions.

4.3 Results and Discussion

a. Downscaling Model

The SOM produced in this study consists of 99 nodes. This allows for the representation of extreme events in the geopotential height, specific humidity, and wind shear fields. This also provides for the representation of the most extreme precipitation probability density functions (PDFs). Figure 4.3 visualizes an example of an extremely wet and dry precipitation PDF for El Verde. Generally speaking, PDFs produced for San Juan were similar for the nodes but with lower magnitude precipitation. The El Verde node PDFs with a higher percentage of days with no measureable precipitation were also representative of dry San Juan PDFs. This alludes to the regional scale climate forcing being similar for both sites but with El Verde receiving higher precipitation due to the addition of some local forcing not represented in the regional atmospheric variable fields (e.g. orographic ascent).

The ERA-Interim data are used to produce a downscaled time series of historical precipitation which are then matched to days with observed rainfall at El Verde and San Juan. The downscaled precipitation model shows adequate ability to capture the general trend in daily precipitation from month-to-month, but slightly underestimates rainfall events thus leading to underestimated monthly precipitation. The most extreme example of this is that the model tends to underestimate daily precipitation in November by nearly 2.5 mm. This is likely due to some increased influence from variability not captured by the predictor variables used here (e.g. tropical disturbances, easterly waves, localized surface pressure anomalies).

The model predicts dry day frequency to within 5% of the observed dry day percentage in all three season (Table 4.2). Additionally, the model predicts heavy precipitation (> 20 mm) days to within 2% of the observed events for all seasons. The model adequately reproduces the mean and median daily rainfall of all three seasons. The ability to reproduce average and extreme daily rainfall frequency as well as the accuracy to predict total seasonal precipitation indicates that the downscaling model incorporates variables that have strong controls on precipitation.

When applying the downscaled model to the GCM future simulations, the results indicate minimal changes in precipitation for El Verde and San Juan. Seasonal changes in daily precipitation are relatively unchanged with a small increase in four day periods with no rainfall. There is little variability among the GCM ensemble members. These findings are generally inconsistent with the patterns in modeled precipitation fields from the CMIP5 GCMs which show decreases in precipitation across the Caribbean. There have also been recent statistical downscaling projects which have used transfer functions to produce high resolution precipitation fields from GCM precipitation projections (e.g. Hjimans et al. 2005, Hayhoe 2013). These downscaled products also show varying degrees of drying across the Caribbean and Puerto Rico. There is reason to further question the results from the SOM downscaling methodology as precipitation is predicted to be stationary.

b. CMIP5 Future Projections of Low-Tropospheric Variables

In order to determine the validity of the downscaled precipitation projections produced, the lower-tropospheric variables are investigated to analyze shifts in moisture, wind patterns, and geopotential heights. Our previous study has shown that these lower-

tropospheric variables are the largest controls on precipitation in Puerto Rico (Ramseyer and Mote 2016). Understanding how these variables are projected to change can provide further evidence as to whether the downscaled precipitation data are reasonable.

One of the primary obstacles of statistical downscaling is the issue of non-stationarity of future climate regimes. Shifts in future distributions of the synoptic types can be indicative of the magnitude of non-stationarity. When mapping the CMIP5 GCM vectors to the SOM, large increases in frequency of nodes in nodes representing extreme synoptic environments can indicate a high degree of non-stationarity. Despite being mapped to these nodes as being the best-matching, the CMIP5 vectors may be more representative of new synoptic types that are not represented in the historical SOM.

Fig. 4.4 provides a visual representation of the change in node frequency between the end-century CCSM4 RCP 8.5 model simulation and the historical node distribution. Most of the nodes at the edge of the map experience varying degrees of decreased frequency. Generally speaking, the nodes with the largest increase in frequency are located in the middle of the map. Node 1 is projected to have the largest decrease in frequency while nodes 37 and 69 are projected to have the largest frequency increase. SOMs used in climate research are typically trained with one atmospheric variable which allows for easy interpretation as the edges of the map typically represent the most extreme environments. The SOM trained in this study includes eight climate variables, which introduces additional complexity in interpreting map geometry. In order to better interpret the changes in map frequency, the spatial mean of the atmospheric variables is determined and is visualized as an overlay on the node frequency change map. 1000 hPa geopotential heights tend to increase in magnitude with increasing row number (Fig. 4.5).

Nodes with larger changes in frequency do not exhibit any notable trends with 1000 hPa geopotential heights.

700 hPa geopotential heights generally increase with increasing row number (Fig. 4.6). Node 1 (largest decrease in frequency) represents the environment with the lowest 700 hPa heights. The nodes projected to increase in frequency are generally in the middle and upper-half of the distribution. Additionally, those nodes representing the lower half of the distribution of 700 hPa geopotential heights are mostly projected to decrease in frequency. Future node frequency indicates a preference for nodes with moderate to higher geopotential heights at the top of the low-troposphere. Meanwhile, those nodes representing the low geopotential height environments are projected to be less frequent in the future.

There does not appear to be a clear progression of 1000 hPa specific humidity across the node space (Fig. 4.7). Nodes representing high levels of near-surface moisture are distributed throughout the map. Node 1 represents the driest node while node 69 (largest projected frequency increase) represents one the highest specific humidity environments. Those nodes projected to decrease in the future predominantly represent the lower half of the 1000 hPa specific humidity distribution.

Spatial means of 700 hPa specific humidity are highly variable across the node space (Fig. 4.8). One of the most notable features is node 1 represents the fourth driest environment. However, node 37 (highest frequency increase) represents the fifth driest 700 hPa environment. From this analysis alone, the future node distribution does not show a clear preference for a moister or drier 700 hPa level. It appears that those nodes

projected to increase in frequency tend to be nodes with higher magnitude 1000 hPa specific humidity but that is not the case at 700 hPa.

1000 – 700 hPa bulk wind shear (BWS) is highly variable across node space (Fig. 4.9). Node 1 represents the second lowest BWS environment while node 69 has a BWS environment that is representative of the upper end of the distribution. Nodes with BWS environments at the lower end of the distribution are projected to be predominantly decreasing in the future.

Figures 4.10 – 4.15 provide further analysis on the atmospheric characteristics of the nodes and allows for some interpretation of the types of environments that are projected to increase/decrease in frequency. Node 1 experiences the largest decrease in frequency in the future and is representative of an extreme environment with low BWS, moisture, and geopotential heights. The two nodes that are projected to have the largest increases in frequency in the future do not show any consistent trends between them.

To assess non-stationarity of nodes, the difference in spatial mean for each node is calculated from the projected CCSM4 atmospheric fields and compared to the historical node spatial mean. Large deviations would indicate that although the nodes represent the closest synoptic type, the new environments are representative of a new climate regime that is unobserved in the historical SOM.

1000 hPa geopotential height node spatial means are 20 – 40 m higher across all nodes in the future simulation (Fig. 4.10). The increases are largest in magnitude at the upper-edge and upper-right corner of the SOM where geopotential height increases are close to 40 m. 700 hPa geopotential height node spatial means are 54 – 72 m higher across all nodes in the future simulation (Fig. 4.11). The highest deviations occur at the

bottom edge of the SOM. Across all nodes, the deviations in 700 hPa heights are at least twice as large as those at 1000 hPa. The 1000 hPa specific humidity node anomalies are all positive and increases range from 1.1 to 3.0 grams per kilogram (Fig. 4.12). The largest anomalies tend to be on the edges of the map. The 700 hPa specific humidity anomalies exhibit a much different trend as both positive and negative anomalies exist across the node space (Fig. 4.13). Most of the anomalies are negative, indicating 700 hPa specific humidity is drier for the days mapping to those nodes in the future. The positive anomalies are restricted to the bottom of the node space. Positive and negative node anomalies are projected in the 1000 – 700 hPa BWS field (Fig. 4.14). Most of the decreases in wind shear are projected in nodes near the edges of the SOM and are much less frequent than nodes with positive BWS anomalies. Negative BWS anomalies are as large as -0.87 m/s while positive anomalies of up to 1.7 m/s are projected.

The spatial mean anomalies derived from the CCSM4 data indicate there is a high degree of non-stationarity in the atmospheric variables analyzed. This is evident from the substantial deviations across all nodes for each atmosphere variable. The consequence of this is poor predictive ability in models employing certain downscaling methodologies in this study region. The SOM-based downscaling approach utilized in this study works well in regions where future climate regimes are more or less within the bounds of the current synoptic modes of variability (i.e. Hewitson and Crane 2006). In the mid-latitudes, there is a high degree of atmospheric variability throughout an annual cycle, providing a wide range of synoptic types. There is less atmospheric variability in the Caribbean, which leads to a narrow range of historical realizations of the climate system. The changes to atmospheric variability projected by the GCM model fall well outside of

the bounds of the current climate. As a result, when the GCM input vectors are mapped to the historical SOM, it calculates the closest historical node vector. In reality, as shown here, while those nodes may be the closest match, they do not represent the future climate regimes. The future climate is shifted into an entirely new node space. Thus, when developing transfer functions or other models relating future climate to historical circulation, the results will have limited skill as it is unable to resolve how extreme the new environments as there is no precedent in the training period.

In addition to evaluating node frequency shifts between the historical and future, analysis of the atmospheric fields across the study domain can also provide further evidence to non-stationarity. Additionally, analyzing these fields can provide evidence of future precipitation variability. The CMIP5 ensemble members were analyzed using two temporal periods and two relative concentration pathways (RCP). RCP 4.5 and 8.5 were used for this study as they are typically most available and represent an extreme emissions scenario (RCP 8.5) and an emissions scenario that is indicative of a leveling off of emissions in the 21st century (RCP 4.5). A mid-century temporal period is used for 2046 – 2065 while an end-century period extends from 2080 – 2099. These twenty year temporal periods capture interannual and interdecadal variability.

c. CMIP5 Future Projection of Extreme Environments

Monthly percentiles (5th, 50th, and 95th) were calculated for all grid points in the GCM ensemble members. These percentiles were then compared to the climatological monthly percentiles derived from ERA-Interim. This allows for analysis on how climatic variability is projected to change in the future. By mid-century, 1000 hPa geopotential heights are projected to exceed the historical 95th percentile from 15 – 50%, with the

highest exceedance at the end of the ERS (Fig. 4.15). There is little difference in the two mid-century RCPs. Similar trends are noted in the mid-century 700 hPa geopotential heights but with increased daily exceedance (20 – 75%) of the climatological 95th percentile (Fig. 4.16). RCP 8.5 has a uniform 20% increase in exceedance of the extreme threshold over the RCP 4.5 ensemble simulation.

Mid-century 1000 hPa specific humidity is expected increase and exceeds the climatological 95th percentile with more frequency (Fig. 4.17). RCP 8.5 represents about a 15 – 20% increase in exceedance over the 4.5 simulation. The increases peak in the late ERS and early LRS at around 40% for the RCP 8.5 ensemble simulation. The 700 hPa simulation has much lower instances of exceeding the 95th percentile; in fact, it is roughly 5% when averaged over all months (Fig. 4.18). Mid-century changes in BWS are similar in magnitude with only about a 5 – 10% frequency in monthly exceedance of the 95th percentile (Fig. 4.19). Of note in the BWS projections, a definitive peak occurs in June, at the onset of the MSD. It appears that both the 700 hPa specific humidity and BWS variables are less sensitive to the differences in RCP 4.5 and 8.5 as there is little difference between the two.

For the end-century simulations of 1000 hPa geopotential heights, the trend is similar to the mid-century simulation but there is an increase in the frequency in which the historical 95th percentile is exceeded (Fig. 4.20). Exceedance peaks in August and October at 50% (RCP 4.5) and 60% (RCP 8.5). Of all the variables analyzed, 700 hPa geopotential heights are projected to exceed the historical 95th percentile with the most frequency (Fig. 4.21). From July to September, nearly every day in the 20 year period exceeds the climatological 95th percentile in the RCP 8.5 scenario. 1000 hPa specific

humidity follows a similar pattern as seen in the mid-century simulations, however, the frequency increases for the RCP 8.5 but stays relatively similar for the RCP 4.5 ensemble (Fig. 4.22). 700 hPa specific humidity follows a similar trend as seen during the mid-century simulations with a slight increase in frequency of exceedance in most months with a peak of around 20% in October (Fig. 4.23). The frequency of extreme dry environments at 700 hPa is also examined (Fig. 4.24). At the end of the ERS and beginning of the LRS, 700 hPa specific humidity is drier than the historical 5th percentile 8 – 15% of days in the end-century period. This indicates that the extremes are likely to become more extreme for the future 700 hPa moisture regime. The BWS peak observed in the mid-century ensemble simulations is evident again in the end-century ensemble (Fig. 4.25). The end-century peak indicates nearly 25% of June days in the future will exceed the historical 95th BWS percentile.

The mid-century ensemble projection of median wind shear indicates higher shear across all months in both the RCP 4.5 (not shown) and 8.5 simulations (Fig. 4.26). The largest increases in shear are during June and October where a 90% increase is projected across the southern part of the study domain. In all months, higher moisture and geopotential heights are projected at 1000 hPa (Fig. 4.27). The largest increases in specific humidity are indicated during the dry season while the largest increases in geopotential heights are projected for the late rainfall season. Increases in specific humidity are around 15% in November, December, and January in the northern part of the domain for RCP 8.5 ensemble with slightly lower increases in the RCP 4.5 simulations. Geopotential heights are have the highest anomalies in September and October with increases as high as 22 m in both RCP ensembles.

Analysis of the 700 hPa specific humidity and geopotential height mid-century departures indicate a drying trend with increasing geopotential heights (Fig. 4.28). The signal is slightly different in the dry season where small decreases in moisture are projected with moisture increases in the northern part of the domain. Starting in March, drying occurs at every grid point and increases in magnitude through the remainder of the ERS and the beginning of the LRS. The strongest drying appears in April and May with decreases as high as 25% and this magnitude is projected in both RCP ensembles. The moisture increase in the northern part of the domain is slightly larger in the RCP 4.5 ensemble with December increases as large as 50%. The largest increases in geopotential heights occur in the LRS with September increases as high as 34 m (RCP 8.5) and 36 m (RCP 4.5). Increases of similar magnitude are noted throughout the LRS.

The end-century RCP 4.5 and 8.5 ensembles of wind shear are spatially and temporally consistent with the mid-century projections (Fig. 4.29). Wind shear is projected to substantially increase across all months with increases over 90% across much of the domain during the ERS with a peak in May and June as well as a secondary peak in October. The drying trend at 700 hPa observed in the mid-century projection is apparent in the end-century simulations as well (Fig. 4.30). However, the end-century projections indicate a slightly less dry scenario. Additionally, the areas in the mid-century projections where moisture was increasing appear to be even moister in the end-century scenario. 700 hPa geopotential heights are projected to be even higher in the end-century scenario with increases as high as 60 m in September. This trend is also noted in the 1000 hPa height fields for September as increases are as high as 30 m (Fig. 4.31). The spatial

and temporal characteristics of the end-century projections of 1000 hPa specific humidity are similar to the mid-century projections but with higher magnitude increases.

4.4 Conclusions

The study addresses the need for a more comprehensive understanding of the how climate and precipitation variability are projected to change in the future in NE PR. The SOM downscaling model developed in this study performed adequately in reproducing the distribution of daily historical rainfall. A validation study was performed on CMIP5 GCMs to determine which models best represented the distribution of historical synoptic types of lower-tropospheric environments. This identified seven models that appropriately reproduced the distribution of synoptic variability. The SOM downscaling model projected little change in future precipitation which is contrary to the GCM predicted precipitation and other downscaling studies in the region (Hijmans et al. 2005, Hayhoe 2013). The magnitude of change forecasted indicates a shift in all synoptic modes of variability identified in the historical record. In fact, it could be argued the climate predictor variables used here occupy an entirely new node space. Thus, using a SOM based downscaling approach has shortcoming in this study area as the model greatly underestimates how extreme the new synoptic environment is projected to be. There are too few historical training vectors that adequately represent the projected future climate. The deviations in low-tropospheric environments projected by the future CMIP5 model ensembles are best represented historically by regimes that have been linked with decreased precipitation at El Verde (Chapter 3).

A thorough analysis of future changes in node frequency and the deviations in low-tropospheric environments indicates a substantial shift in the entire node space.

Substantial changes in atmospheric variability are projected. Geopotential heights are projected to increase throughout the low-troposphere with 700 hPa heights having higher magnitude deviations than near the surface. This indicates a warmer low-troposphere with higher pressure. Additionally, this could add to the increasing evidence in the literature for increased influence from the NASH on Caribbean climate variability (e.g. Rauscher et al. 2008, Diem 2013, Ryu and Hayhoe 2014). Moisture is projected to increase near the surface but decrease at 700 hPa. This is likely indicative of an increasing influence of the NASH and/or the SAL. This drying at 700 hPa coupled with the increasing 1000 – 700 hPa thicknesses represent a stronger trade wind inversion and increased atmospheric stability. The CMIP5 ensemble projects that 1000 – 700 hPa wind shear will likely increase in the future. This tends to agree with recent modeling studies (Lee et al. 2011, Maloney et al. 2014), which concluded that wind shear was likely to increase in the main development region of the tropical North Atlantic. It is also plausible that intensification and spatial deviations in the mean state of the NASH could further increase pressure gradients and wind shear in the Caribbean.

Future climate simulations suggest large deviations in all variables in all three seasons examined. The direction of these deviations indicates a shift to higher frequency of dry conditions. However, the changes appear to be most significant during the ERS. In the future, the ERS looks to be significantly drier at 700 hPa with substantial increases in low tropospheric wind shear. It is likely that the climate features currently thought to be responsible for the Caribbean MSD are projected to become larger forcing mechanisms, including the NASH (Gamble et al. 2008). This finding is consistent with other studies

which indicate the MSD may become stronger and longer-lasting in the future (Rauscher et al. 2008).

A stronger, longer MSD coupled with increases in BWS and decreases in 700 hPa moisture suggests higher frequency of dry spells in the forests of NE PR. This will lead to greater disruption to biogeochemical and ecological processes and higher frequency of disturbance. Forest composition is likely to change and adapt to the new climate regime. Additionally, the municipal water supplies in the San Juan Metropolitan Area are likely to become increasingly stressed in a future, drier climate.

4.5 References

Benstead, J. P., J. G. March, C. M. Pringle, and F. N. Scatena, 1999: Effects of a low-head dam and water abstraction on migratory tropical stream biota. *Ecological Applications*, **9**, 656–668.

Covich, A. P., T. A. Crowl, and F. N. Scatena, 2003: Effects of extreme low flows on freshwater shrimps in a perennial tropical stream. *Freshwater Biology,* **48**, 1199–1206.

Covich, A. P., T. A. Crowl, and T. Heartsill-Scalley, 2006: Effects of drought and hurricane disturbances on headwater distributions of palaemonid river shrimp (Macrobrachium spp.) in the Luquillo Mountains, Puerto Rico. *Journal of the North American Benthological Society*, **25**, 99–107.

Crook, K. E. , F. N. Scatena, and C. M. Pringle: 2007. Water withdrawal from the Luquillo Experimental Forest 2004. General Technical Report IIT-GTR-34, International Institute of Tropical Forestry, Pacific Northwest Research Station, 40 p.

Crook, K. E., C. M. Pringle, and M. C. Freeman, 2009: A method to assess longitudinal riverine connectivity in tropical streams dominated by migratory biota. *Aquatic Conservation* , **19**, 714–723.

Diem, J.E., 2013: Comments on "Changes to the North Atlantic subtropical high and its role in the intensification of summer rainfall variability in the southeastern United States". *J Clim*, **26**, 679 – 682.

Gamble, D. W., and S. Curtis, 2008: Caribbean Precipitation: Review, Model and Prospect. *Prog Phys Geogr*, **32,** 265– 276.

Gamble, D. W., D. B. Parnell, and S. Curtis, 2008: Spatial variability of the Caribbean mid-summer drought and relation to north Atlantic high circulation. *Int J Climatol*, **28**, 343–350.

Giannini, A., Y. Kushnir, and M. A. Cane, 2001: Seasonality in the impact of ENSO and the North Atlantic high on Caribbean rainfall. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, **62**, 143–147.

Hayhoe, K., 2013: Quantifying key drivers of climate variability and change for Puerto Rico and the Caribbean. Agreement number G10AC00582. Caribbean Landscape Conservation Cooperative.

Heartsill-Scalley T., Scatena F.N., Estrada C., McDowell W.H., Lugo A.E., 2007: Disturbance and long-term patterns of rainfall and throughfall nutrient fluxes in a subtropical forest in Puerto Rico. *J Hydrol*, **33**, 472–485.

Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones and A. Jarvis, 2005: Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol*, **25**, 1965–1978.

IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

Larsen, M. C., 2000: Analysis of 20th century rainfall and streamflow to characterize drought and water resources in Puerto Rico. *Phys Geog*, **21**, 494–521.

Lee, S.-K., D. B. Enfield, and C. Wang, 2011: Future impact of differential interbasin ocean warming on Atlantic hurricanes. *J Clim*, **24**, 1264–1275.

Magana, V., J. A. Amador, and S. Medina, 1999: The midsummer ~ drought over Mexico and Central America. *J Clim*, **12**, 1577–1588

Maloney, E. D., and Coauthors, 2014: North American climate in CMIP5 experiments. Part III: Assessment of twenty-first-century projections. *J Clim*, **27**, 2230–2270, doi:10.1175/JCLI-D-13-00273.1.

Mulligan, M., and S. M. Burke, 2005: FIESTA: Fog Interception for the Enhancement of Streamflow in Tropical Areas, 174 pp.

Neelin, J. D., M. Munnich, H. Su, J. E. Meyerson, and C. E. Holloway, 2006: Tropical drying trends in global warming models and observations. *Proc Natl Acad Sci USA*, **103**, 6110–6115, doi:10.1073/pnas.0601798103.

Pringle, C. M., 1997: Exploring how disturbance is transmitted upstream: going against the flow. *Journal of the North American Benthological Society*, **16**, 425–438.

Ramseyer, C.A. and T. Mote, 2016: Atmospheric controls on Puerto Rico precipitation using artificial neural networks. *Clim Dyn*, doi: 10.1007/s00382-016-2980-3.

Rauscher, S. A., F. Giorgi, N. S. Diffenbaugh, and A. Seth, 2008: Extension and intensification of the meso-American midsummer drought in the twenty-first century. *Clim Dyn*, **31**, 551–571.

Ryu, J.-H. and K. Hayhoe, 2014: Understanding the sources of Caribbean precipitation biases in CMIP3 and CMIP5 simulations. *Clim Dyn*, **42**, 3233– 3252, doi: 10.1007/s00382-014-2351-x.

Sheffield, J. and Coauthors, 2013: North American climate in CMIP5 experiments. Part II: Evaluation of historical simulations of intraseasonal to decadal variability. *J Clim*, **26**, 9247–9290

Table 4.1 CMIP5 GCMs analyzed and validated with the models selected for use in future ensembles underlined.

Model Name	Responsible Agency
BCC-CSM 1.1	Beijing Climate Center (China)
BCC-CSM 1.1(m)	BCC (China)
CanESM2	Canadian Centre for Climate Modeling and Analysis
CNRM-CM5	Centre National de Recherches Météorologiques, Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (France)
CMCC-CM	Centro Euro-Mediterraneo sui Cambiamenti Climatici (Italy)
CMCC-CMS	CMCC (Italy)
ACCESS 1.0	Commonwealth Scientific and Industrial Research Organization/Bureau of Meteorology (Austrailia)
ACCESS 1.3	CSIRO (Austrailia)
IPSL-CM5B-LR	Institut Pierre Simon Laplace (France)
IPSL-CM5A-MR	IPSL (France)
FGOALS-g2	Institue of Atmospheric Physics, Chinese Academy of Sciences (China)
MIROC5	Atmosphere and Ocean Reseach Institute, National Institute for Environmental Studies (Japan)
MIROC-ESM-CHEM	MIROC (Japan)
HadGEM-CC	Met Office Hadley Centre (UK)
MRI-CGCM3	Meteorological Research Institute (Japan)
MPI-ESM-LR	Max Planck Institute for Meteorology (Germany)
MPI-ESM-MR	MPI-M (Germany)
NorESM1-M	Bjerkes Centre for Climate Research, Norwegian Meteorological Institute (Norway)
GFDL-ESM2G	Geophysical Fluid Dynamics Laboratory (USA)
GFDL-ESM2M	GFDL (USA)
GFDL-CM3	GFDL (USA)
CCSM4	National Center for Atmospheric Research (USA)

Table 4.2 Various metrics used to demonstrate the ability of the SOM downscaling model to predict historical seasonal precipitation.

Precipitation Metric	Observed	Modeled
Dry Season mean daily (mm)	7.1	6.7
Dry Season median daily (mm)	1.8	1.8
Dry Season dry days (%)	37.1	34.3
Dry Season daily frequency > 20 mm (%)	9.3	9.6
ERS mean daily (mm)	8.8	7.6
ERS median daily (mm)	1.5	2.0
ERS dry days (%)	36.7	32.0
ERS daily frequency > 20 mm (%)	13.7	11.6
LRS mean daily (mm)	10.0	9.0
LRS median daily (mm)	3.0	2.8
LRS dry days (%)	29.7	27.8
LRS daily frequency > 20 mm (%)	15.5	14.5

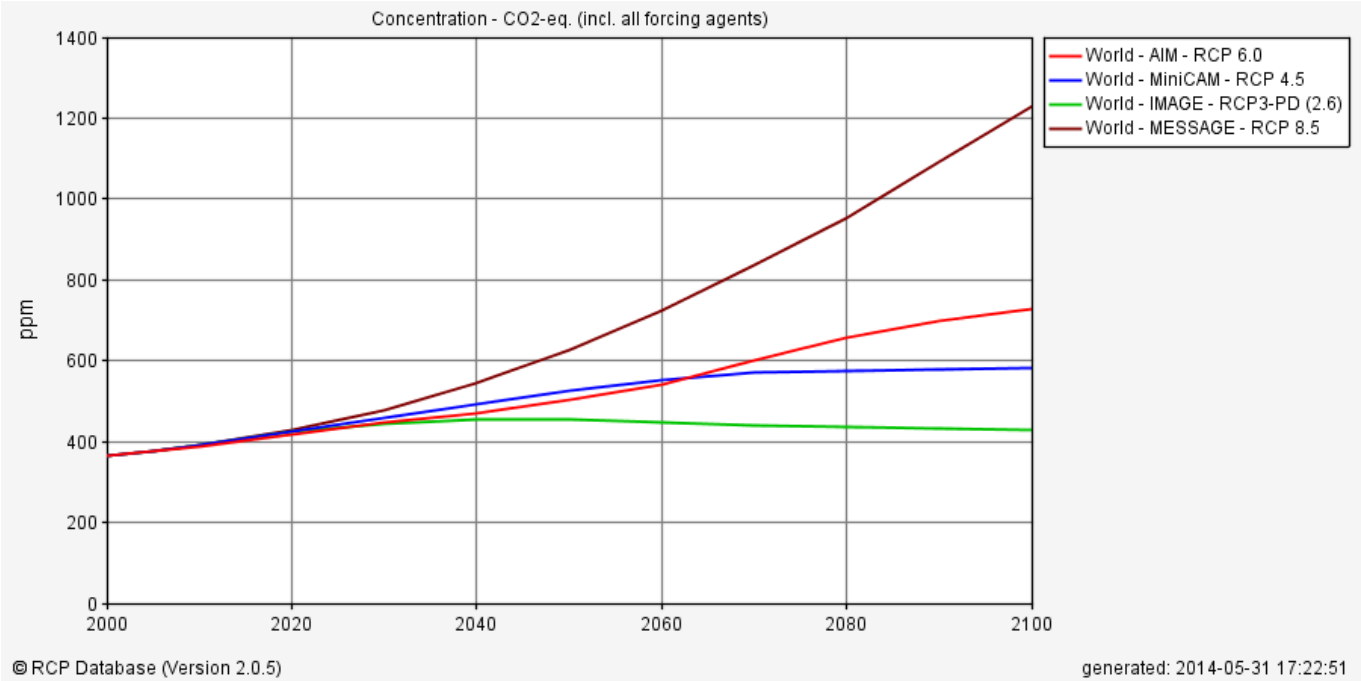

Fig. 4.1 A graphic representation of the representative concentration pathways utilized in the CMIP5 modeling framework. This image was generated using the RCP Database online GUI.

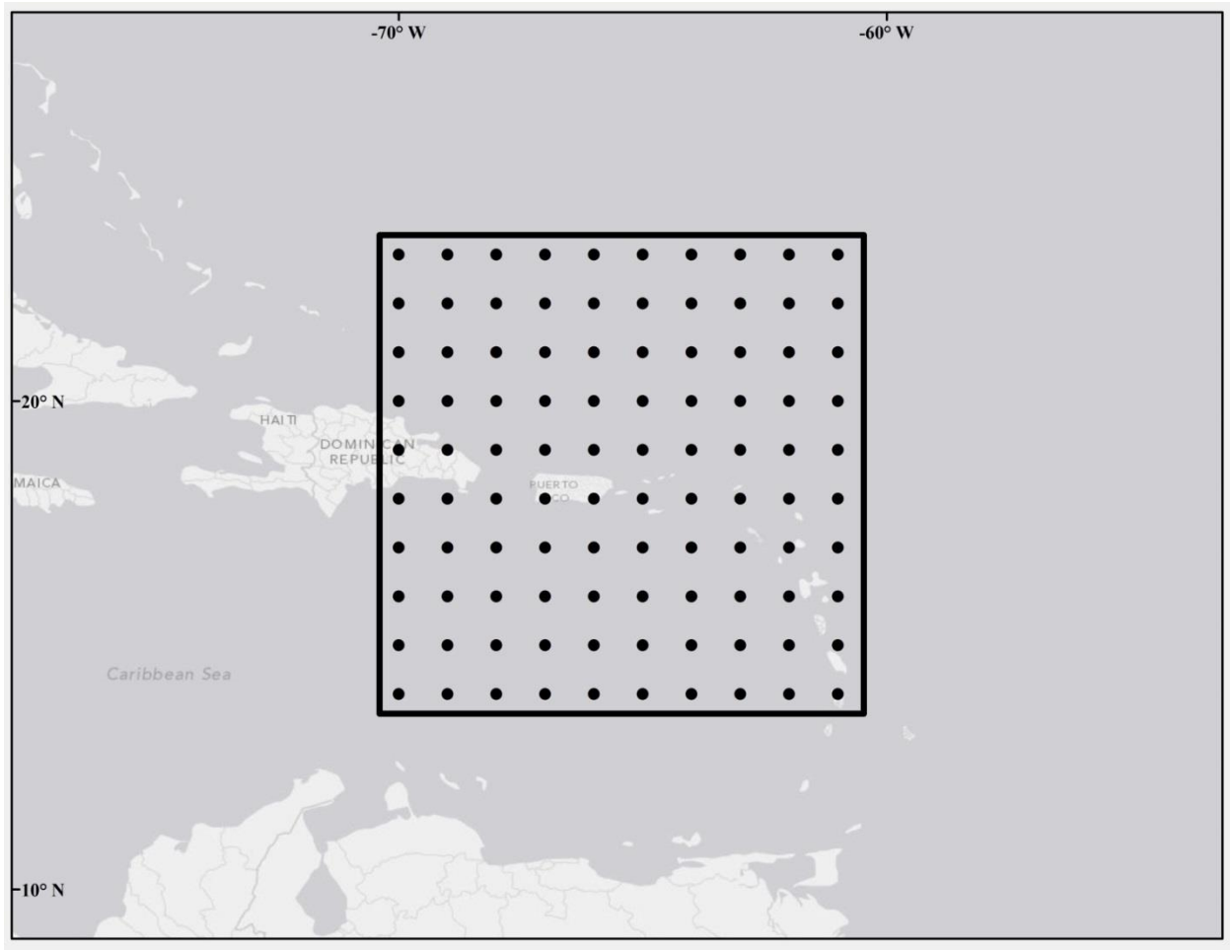

Fig. 4.2 A map showing the extent of the spatial domain and the ERA-Interim grid points used. The CMIP5 models used in the study were regridded to provide data at these grid points.

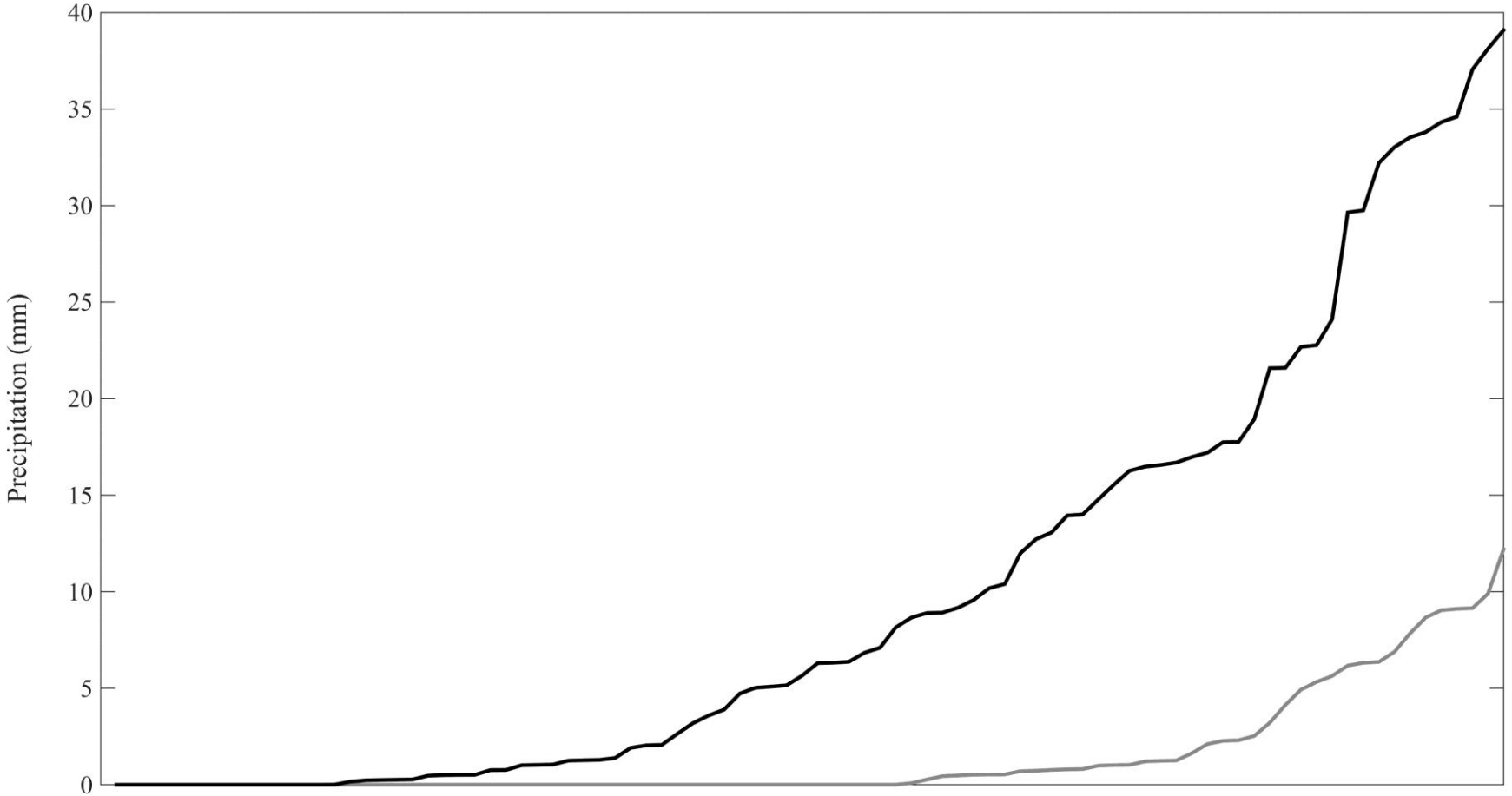

Fig. 4.3 An example of a high precipitation node PDF (black line) and a low precipitation node PDF (grey) at El Verde. The x-axis represents different realizations for the two nodes.

Fig. 4.4 Map of frequency shifts between historical nodes and the node distribution from the end-century CCSM4 RCP 8.5 simulation.

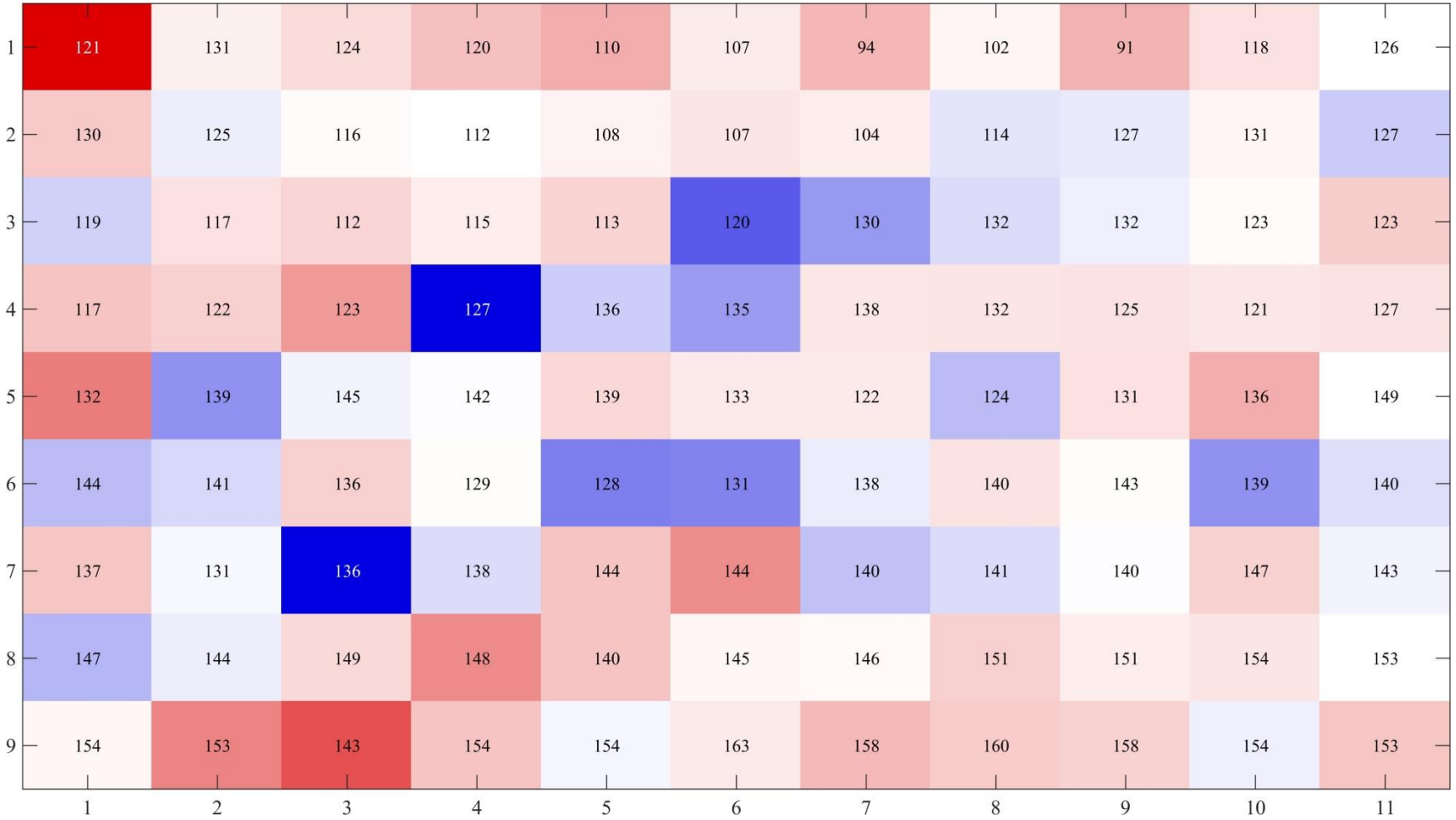

Fig. 4.5 Map of changes in node frequency as projected by the end-century CCSM4 RCP 8.5 simulation (shading) and the historical grid point spatial average of 1000 hPa geopotential heights (numeric). Blue shading indicates an increased frequency in the future while red shading indicates a decrease in frequency in the future.

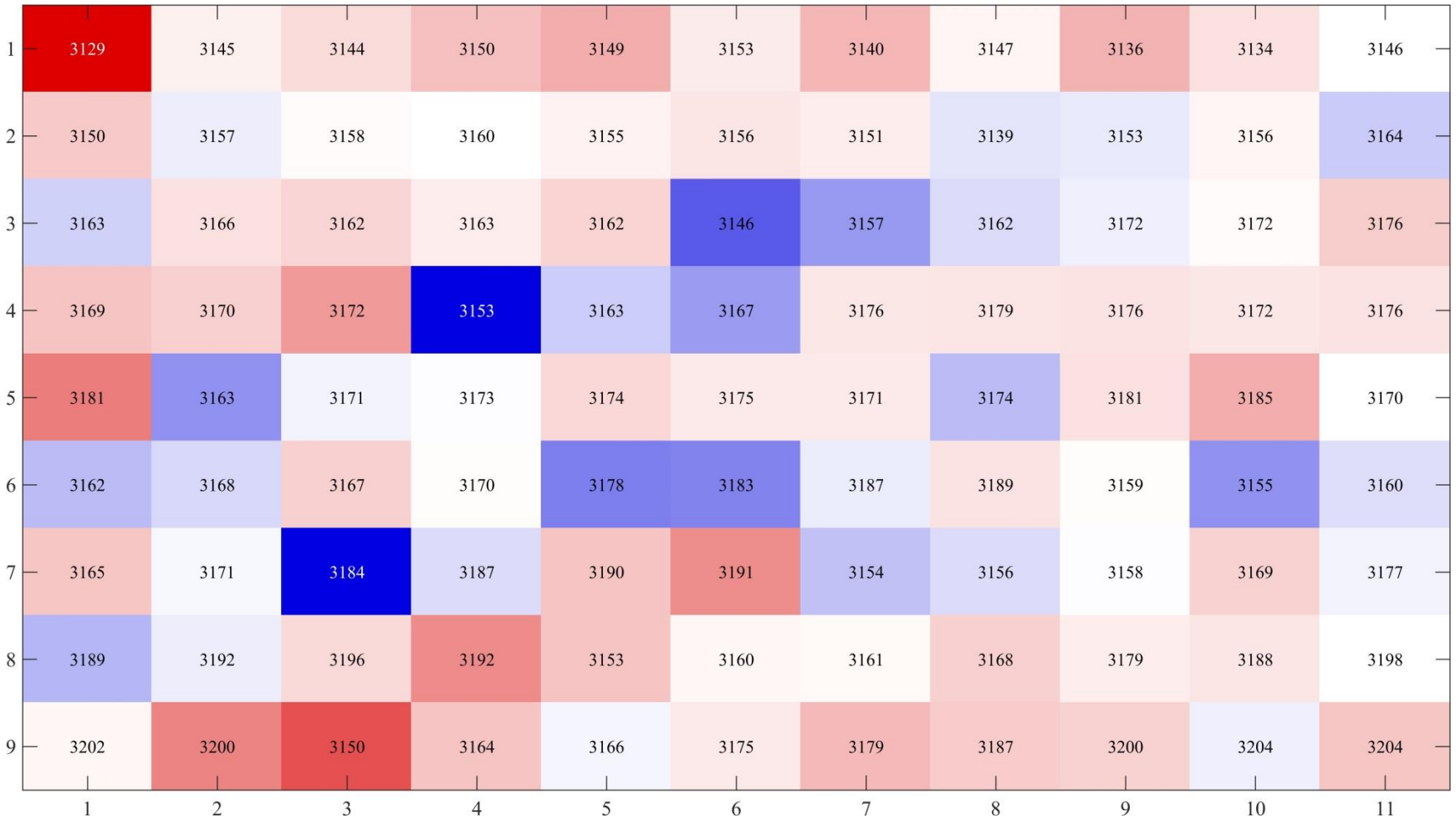

Fig. 4.6 Map of changes in node frequency as projected by the end-century CCSM4 RCP 8.5 simulation (shading) and the historical grid point spatial average of 700 hPa geopotential heights (numeric). Blue shading indicates an increased frequency in the future while red shading indicates a decrease in frequency in the future.

Fig. 4.7 Map of changes in node frequency as projected by the end-century CCSM4 RCP 8.5 simulation (shading) and the historical grid point spatial average of 1000 hPa specific humidity (numeric). Blue shading indicates an increased frequency in the future while red shading indicates a decrease in frequency in the future.

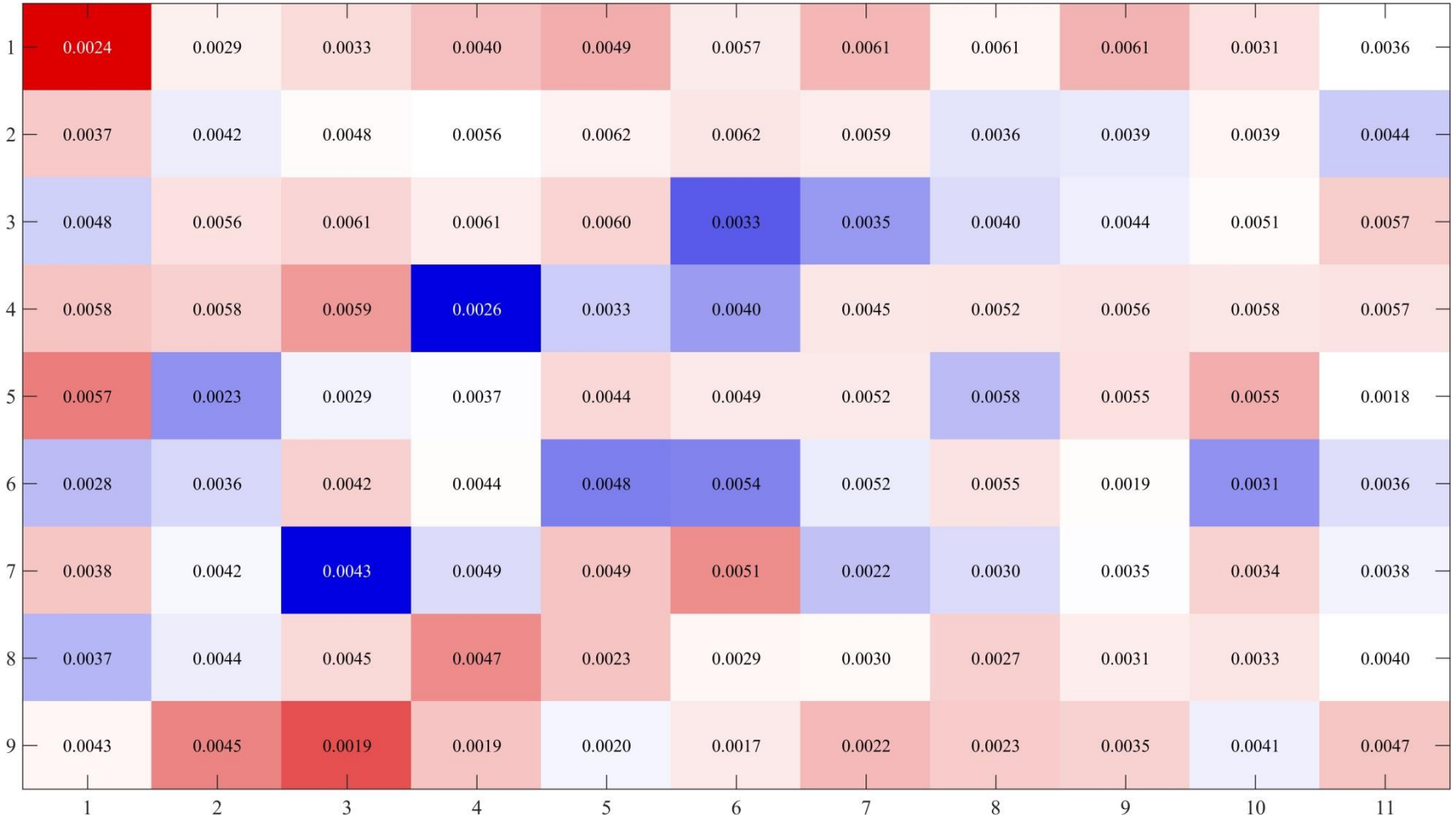

Fig. 4.8 Map of changes in node frequency as projected by the end-century CCSM4 RCP 8.5 simulation (shading) and the historical grid point spatial average of 700 hPa specific humidity (numeric). Blue shading indicates an increased frequency in the future while red shading indicates a decrease in frequency in the future.

Fig. 4.9 Map of changes in node frequency as projected by the end-century CCSM4 RCP 8.5 simulation (shading) and the historical grid point spatial average of 1000 – 700 hPa bulk wind shear (numeric). Blue shading indicates an increased frequency in the future while red shading indicates a decrease in frequency in the future.

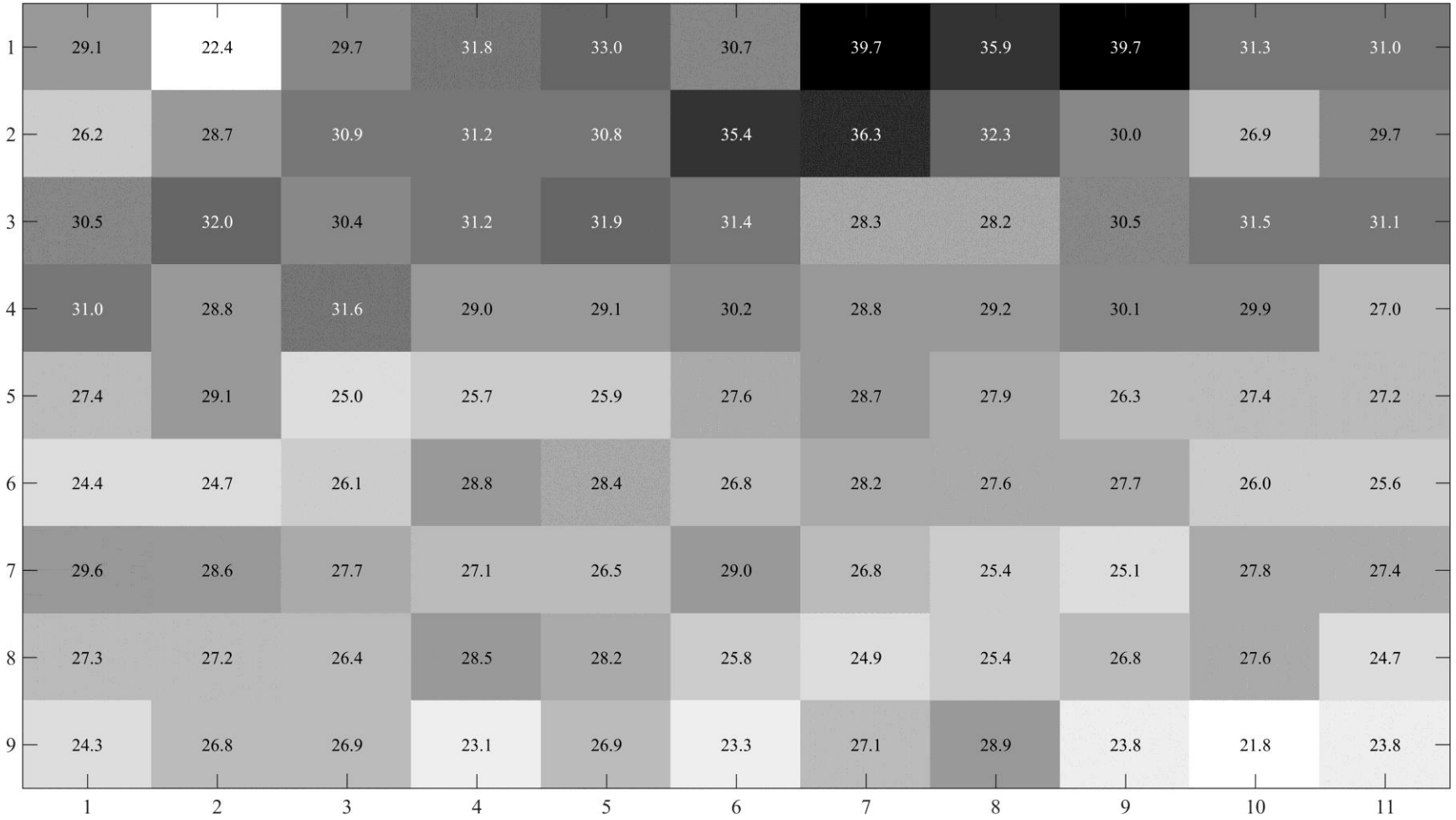

Fig. 4.10 Map of 1000 hPa geopotential height node anomalies (m) as projected by the end-century CCSM4 RCP 8.5 simulation (shading and numeric). Darker shading indicates higher anomalies.

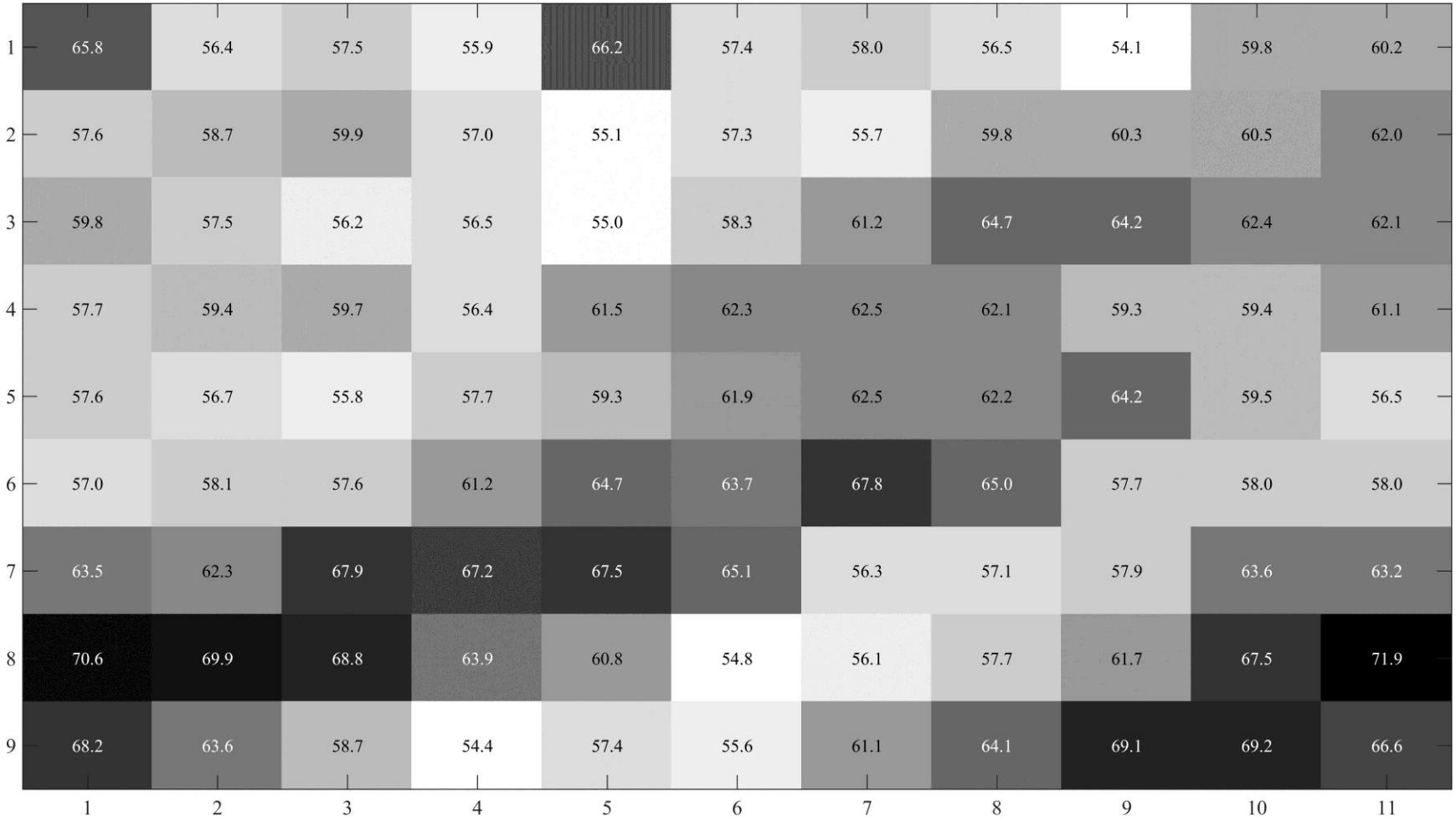

Fig. 4.11 Map of 700 hPa geopotential height node anomalies (m) as projected by the end-century CCSM4 RCP 8.5 simulation (shading and numeric). Darker shading indicates higher anomalies.

Fig. 4.12 Map of 1000 hPa specific humidity node anomalies (kg kg^{-1}) as projected by the end-century CCSM4 RCP 8.5 simulation (shading and numeric). Darker shading indicates higher anomalies.

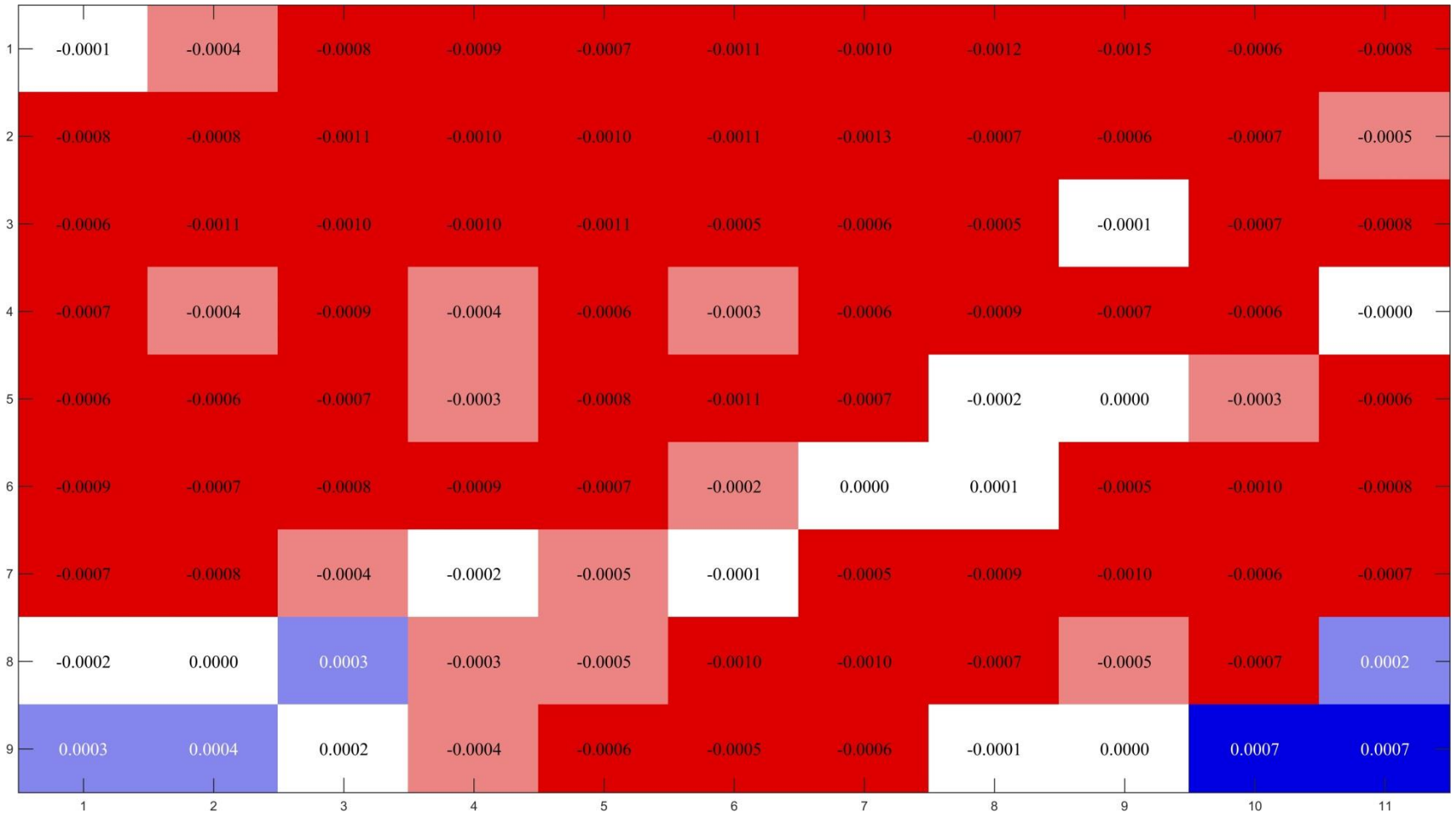

Fig. 4.13 Map of 700 hPa specific humidity node anomalies (kg kg^{-1}) as projected by the end-century CCSM4 RCP 8.5 simulation (shading and numeric). Red shading indicates negative anomalies and blue shading indicates positive anomalies.

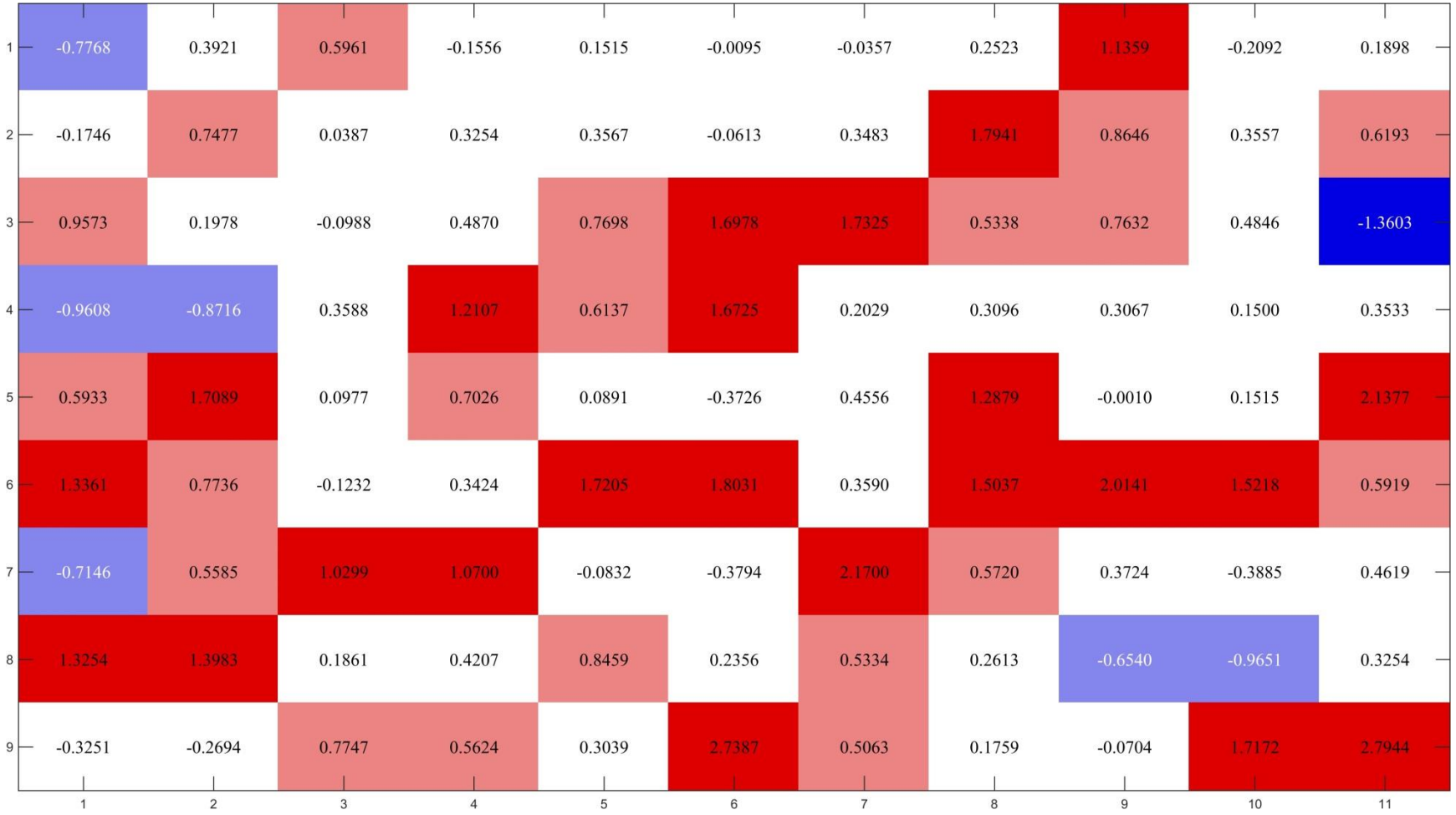

Fig. 4.14 Map of bulk wind shear node anomalies (kts) as projected by the end-century CCSM4 RCP 8.5 simulation (shading and numeric). Red shading indicates negative anomalies and blue shading indicates positive anomalies.

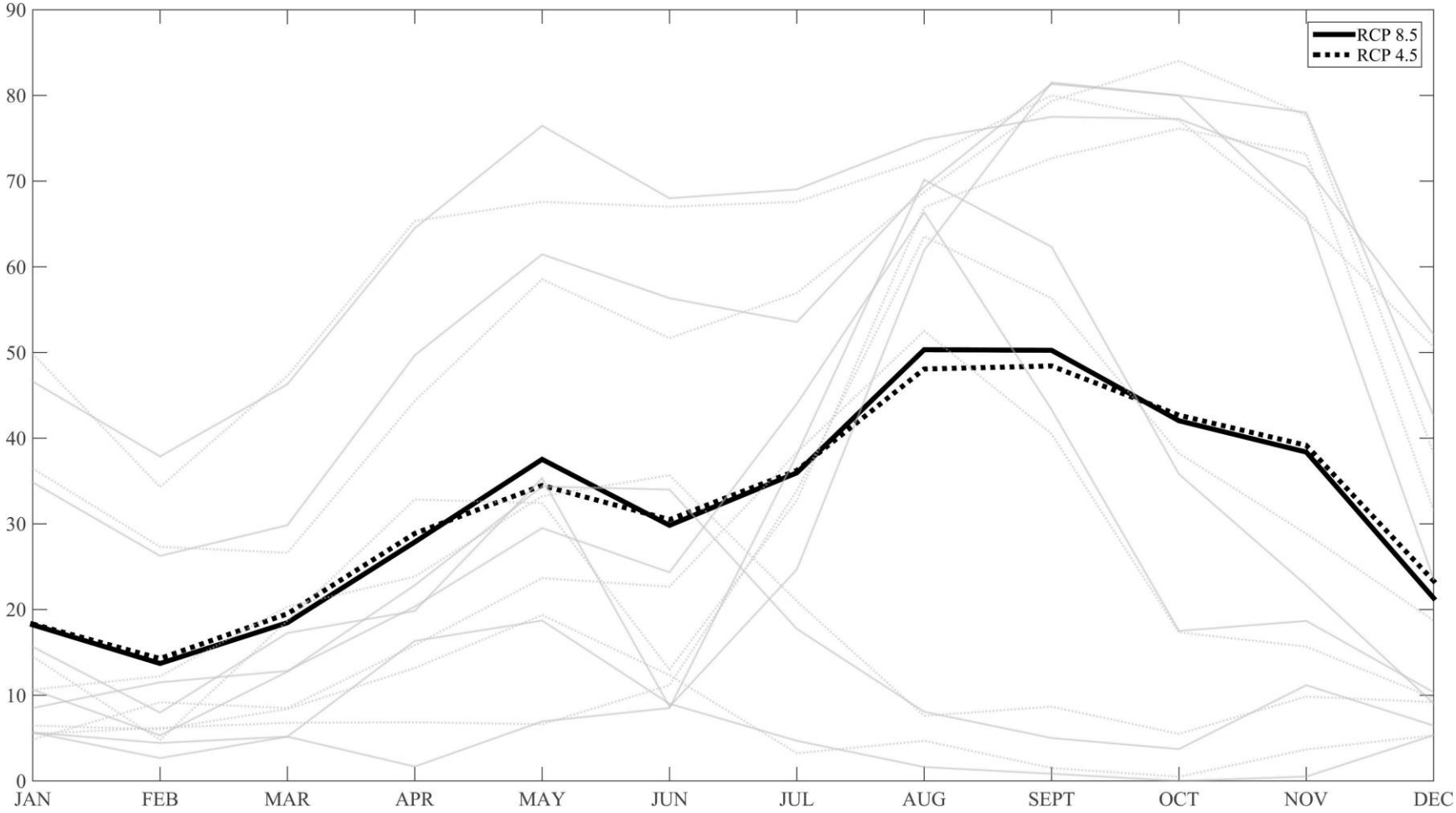

Fig. 4.15 Monthly mid-century ensemble projection of 1000 hPa geopotential heights grid point average exceeding the climatological 95th percentile for the RCP 4.5 (dashed black line) and RCP 8.5 ensembles (solid black line).

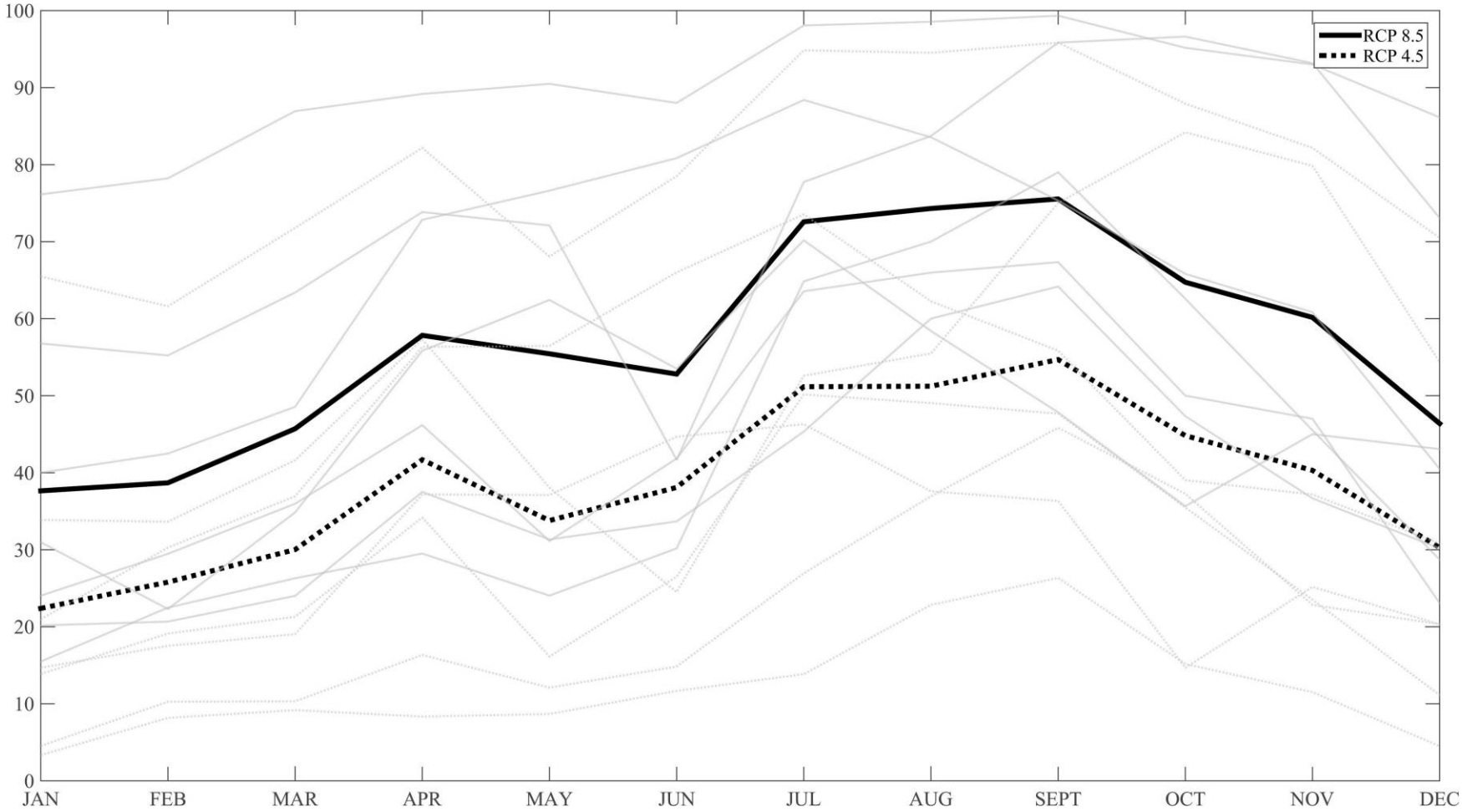

Fig. 4.16 Monthly mid-century ensemble projection of 700 hPa geopotential heights grid point average exceeding the climatological 95th percentile for the RCP 4.5 (dashed black line) and RCP 8.5 ensembles (solid black line).

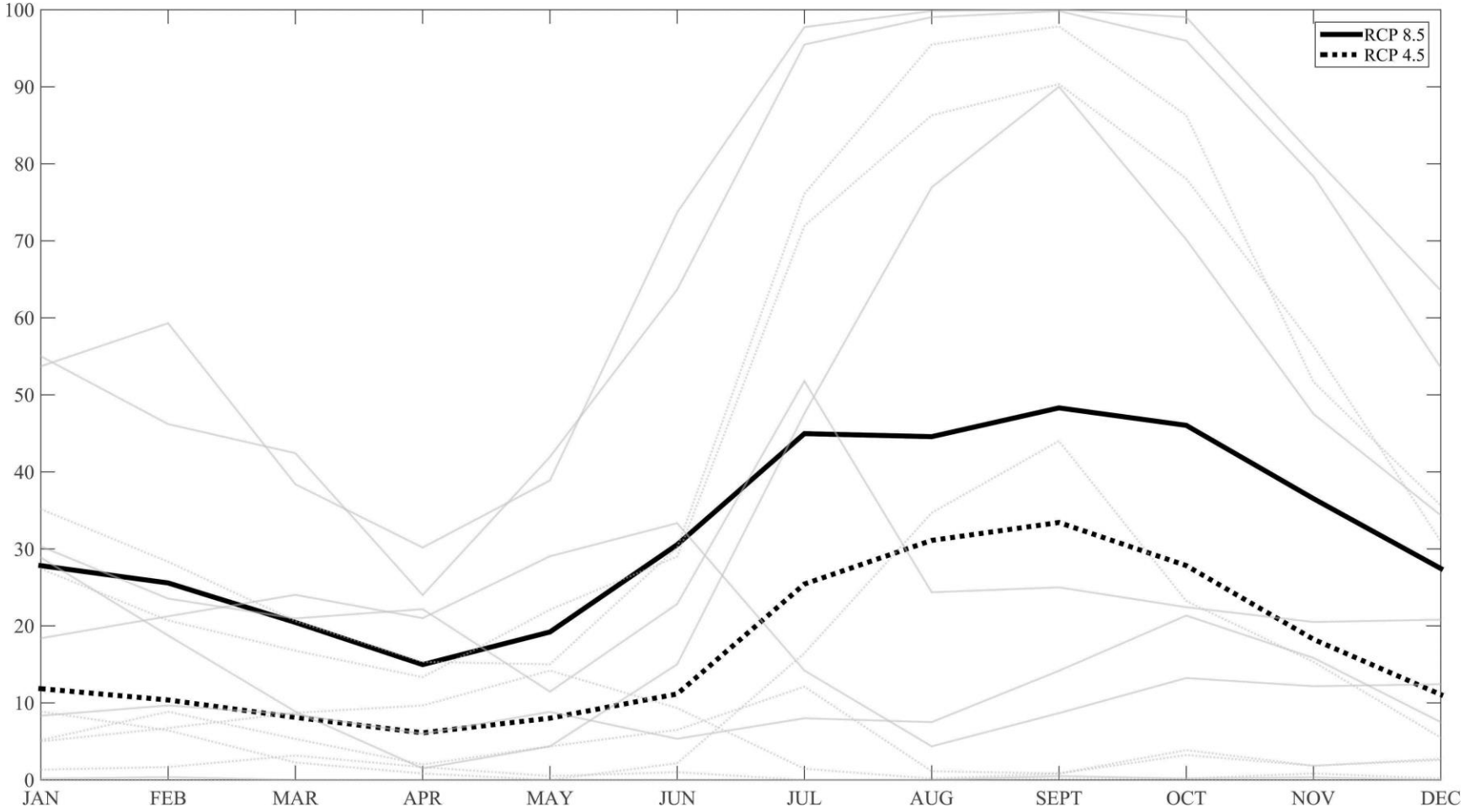

Fig. 4.17 Monthly mid-century ensemble projection of 1000 hPa specific humidity grid point average exceeding the climatological 95th percentile for the RCP 4.5 (dashed black line) and RCP 8.5 ensembles (solid black line).

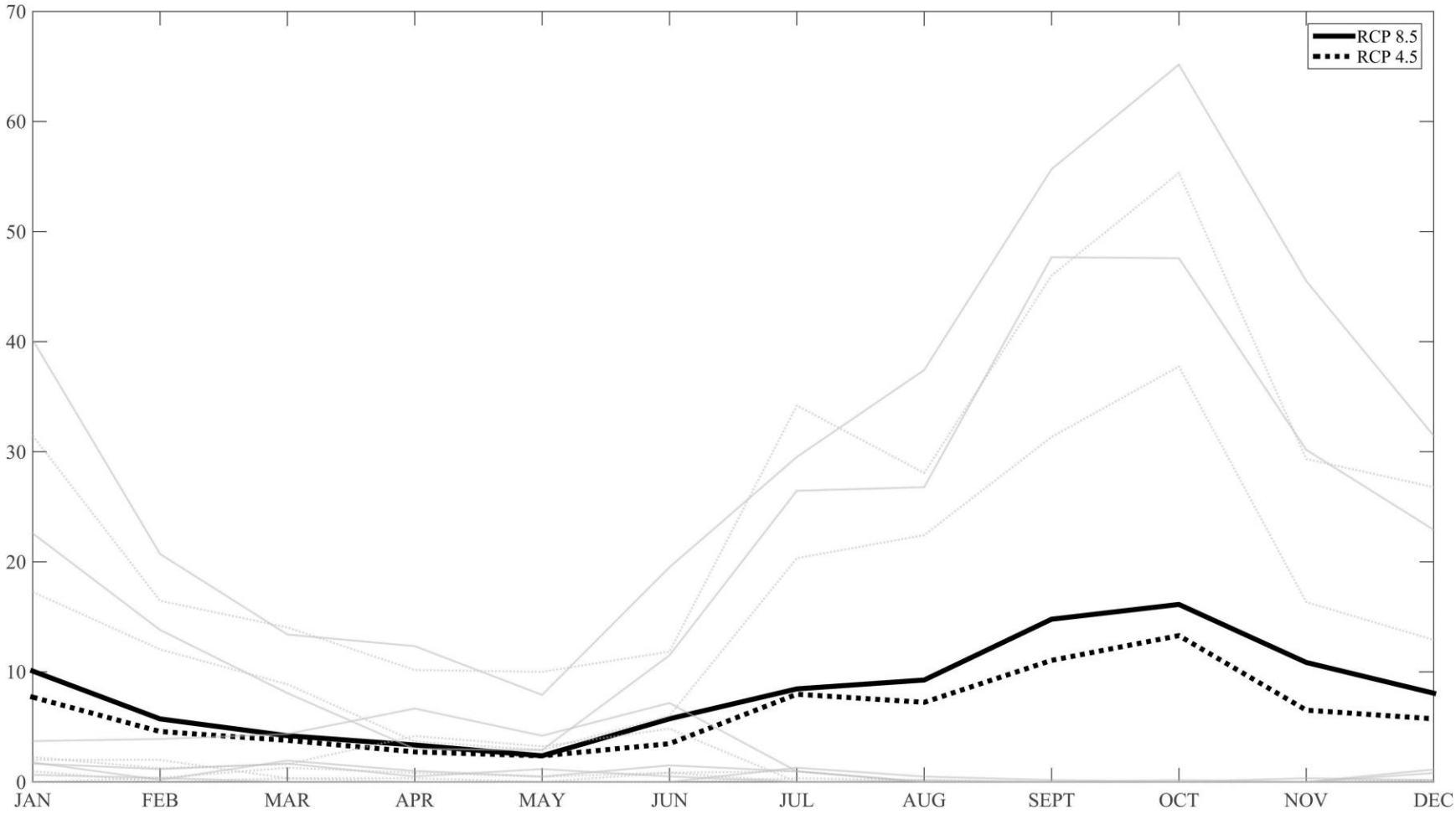

Fig. 4.18 Monthly mid-century ensemble projection of 700 hPa specific humidity grid point average exceeding the climatological 95th percentile for the RCP 4.5 (dashed black line) and RCP 8.5 ensembles (solid black line).

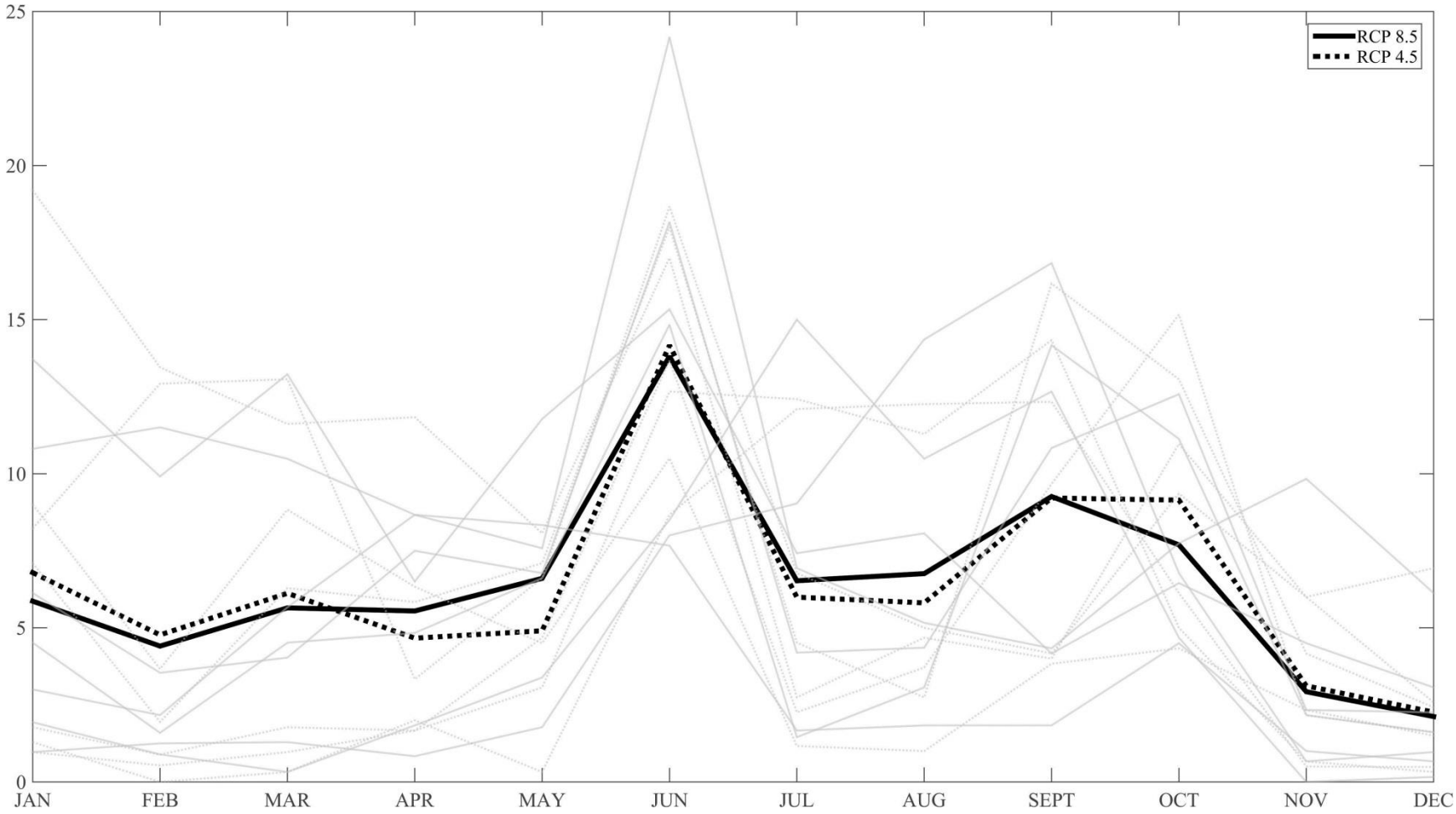

Fig. 4.19 Monthly mid-century ensemble projection of bulk wind shear grid point average exceeding the climatological 95% percentile for the RCP 4.5 (dashed black line) and RCP 8.5 ensembles (solid black line).

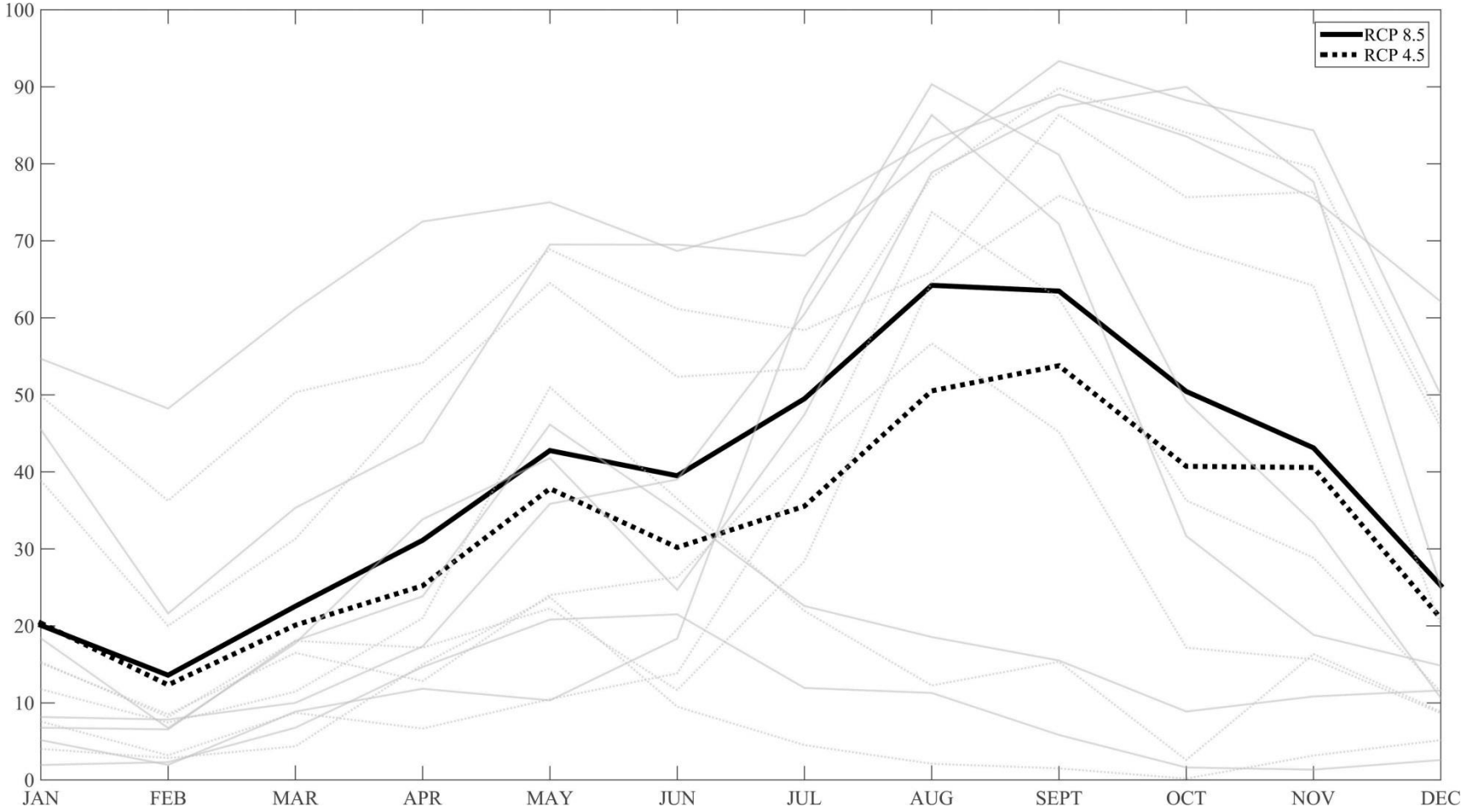

Fig. 4.20 Monthly end-century ensemble projection of 1000 hPa geopotential heights grid point average exceeding the climatological 95% percentile for the RCP 4.5 (dashed black line) and RCP 8.5 ensembles (solid black line).

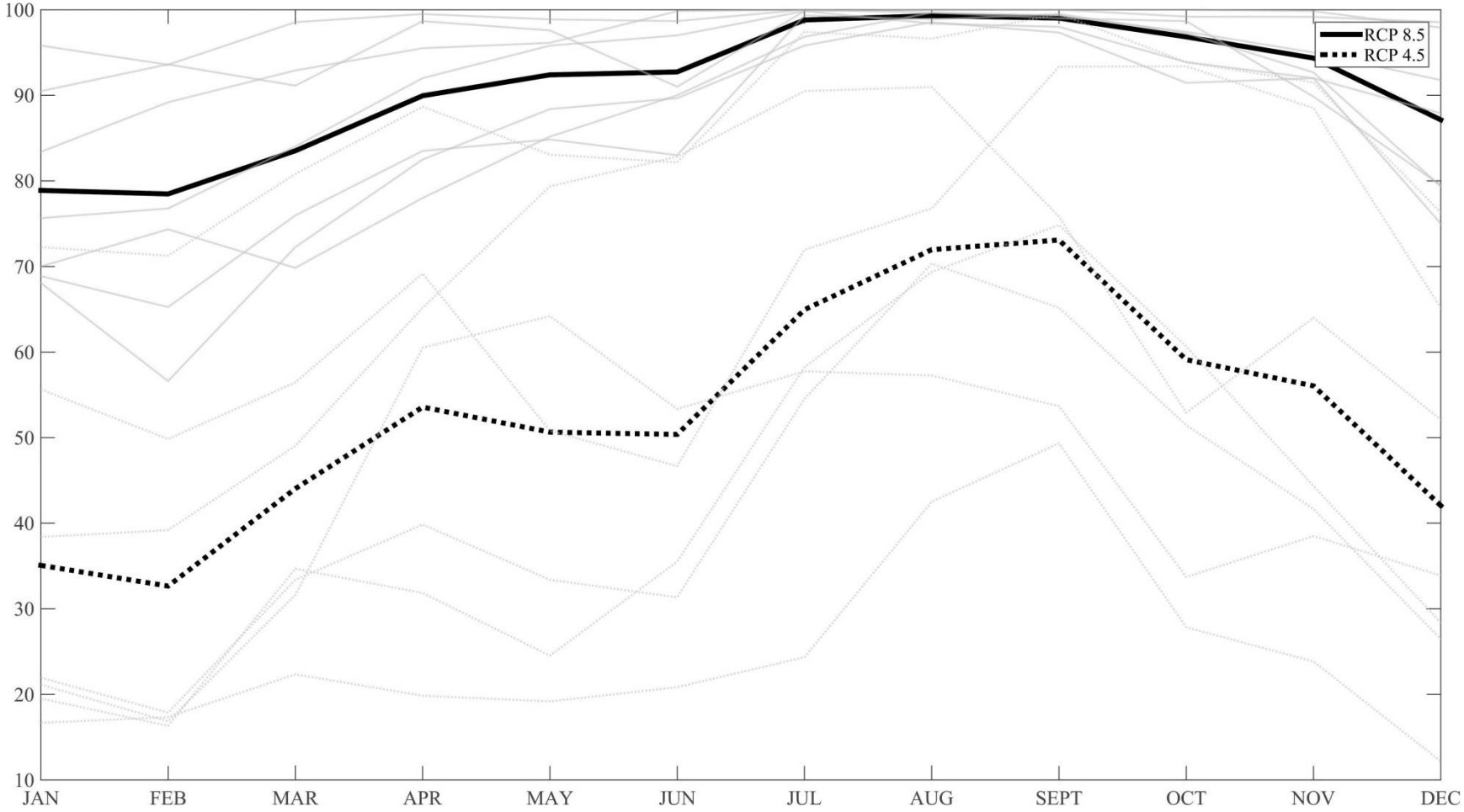

Fig. 4.21 Monthly end-century ensemble projection of 700 hPa geopotential heights grid point average exceeding the climatological 95th percentile for the RCP 4.5 (dashed black line) and RCP 8.5 ensembles (solid black line).

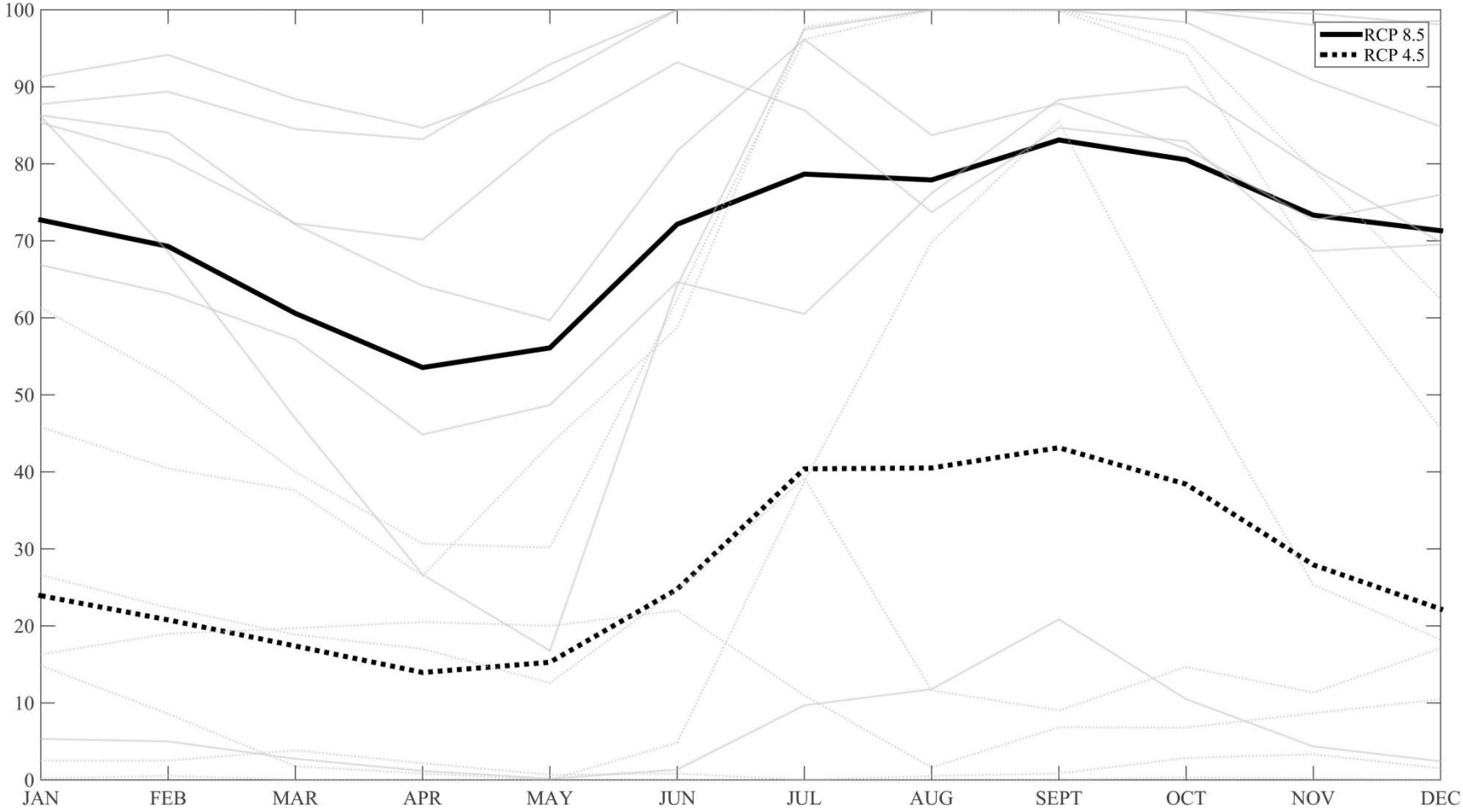

Fig. 4.22 Monthly end-century ensemble projection of 1000 hPa specific humidity grid point average exceeding the climatological 95th percentile for the RCP 4.5 (dashed black line) and RCP 8.5 ensembles (solid black line).

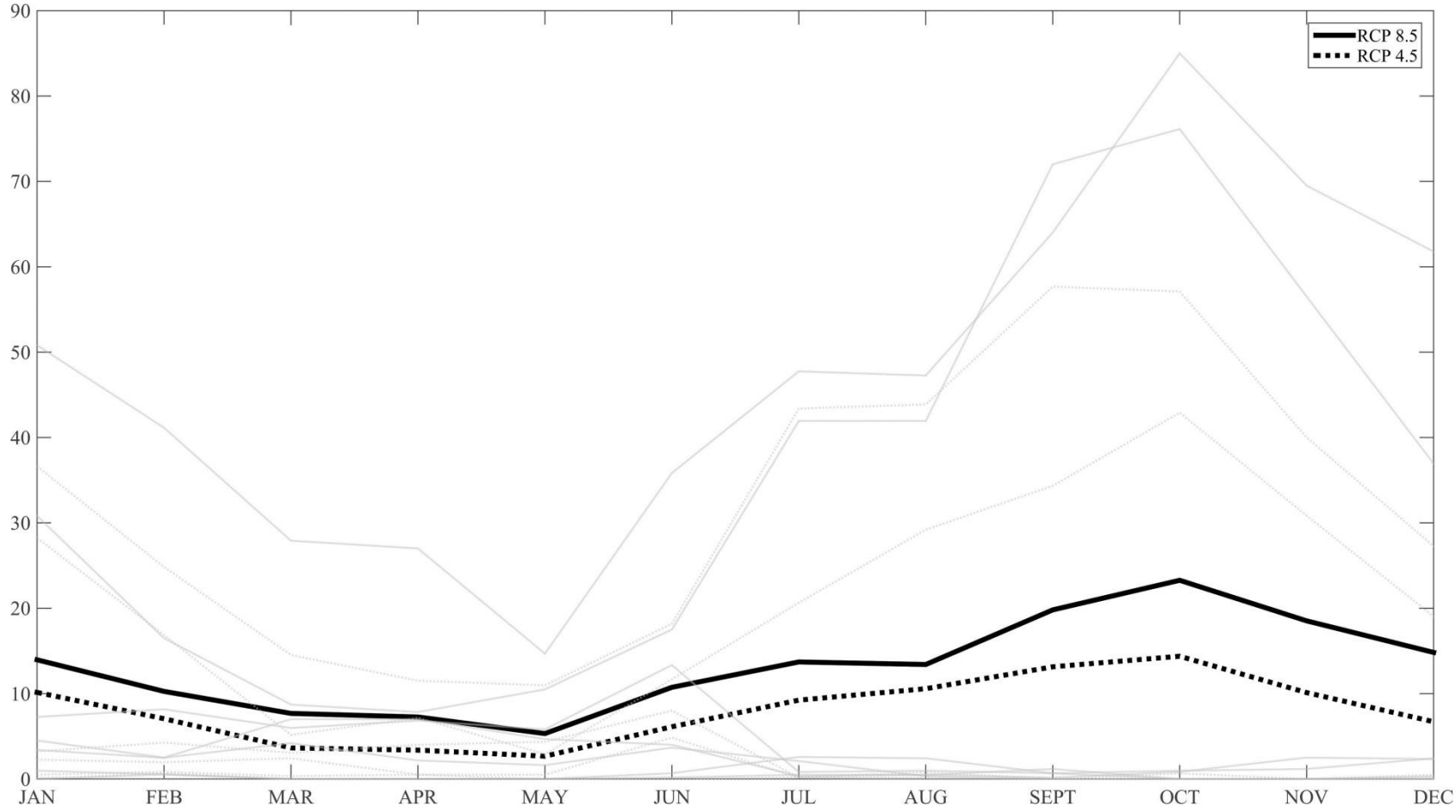

Fig. 4.23 Monthly end-century ensemble projection of 700 hPa specific humidity grid point average exceeding the climatological 95th percentile for the RCP 4.5 (dashed black line) and RCP 8.5 ensembles (solid black line).

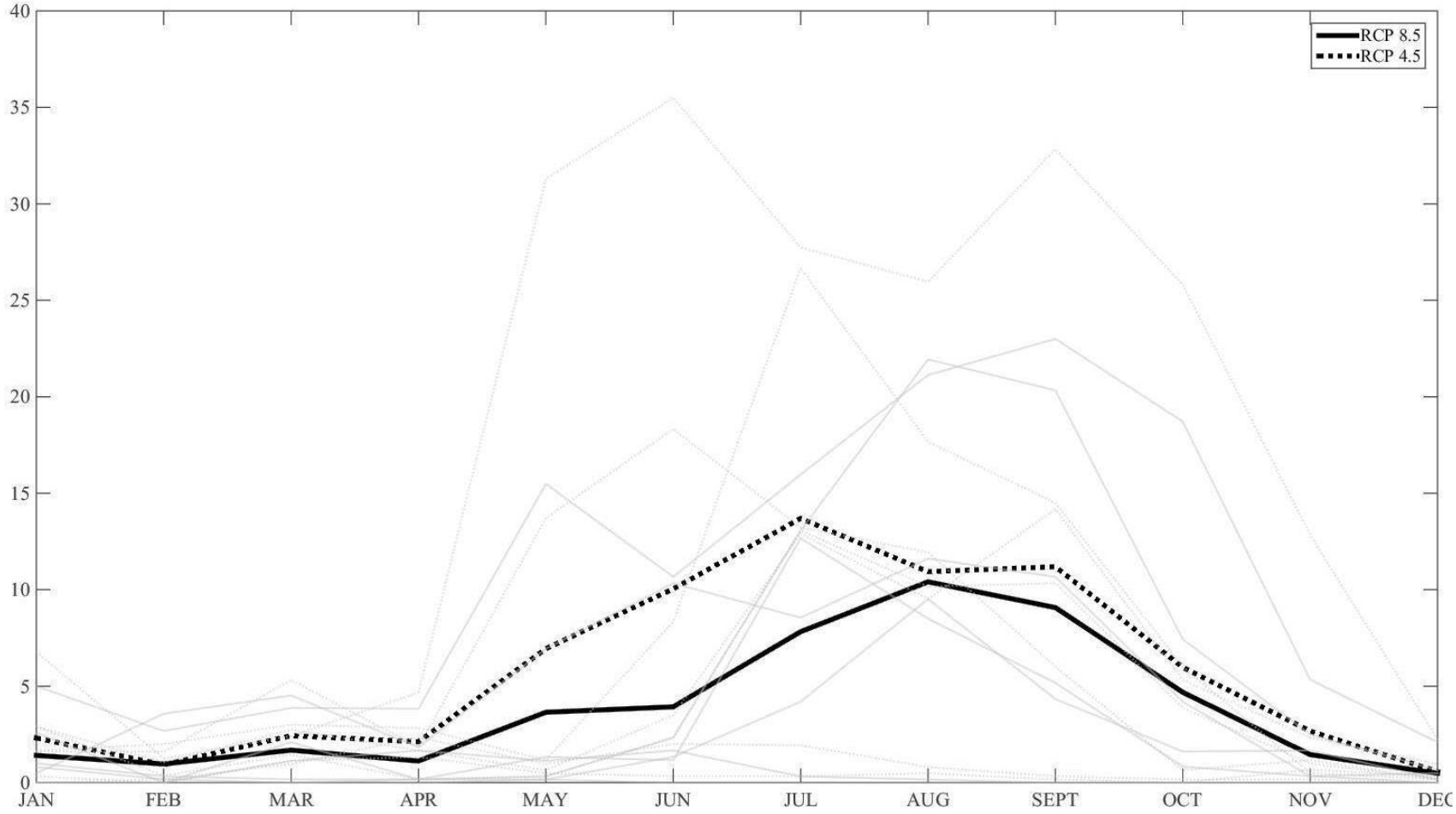

Fig. 4.24 Monthly end-century ensemble projection of 700 hPa specific humidity grid point average below the climatological 5th percentile for the RCP 4.5 (dashed black line) and RCP 8.5 ensembles (solid black line).

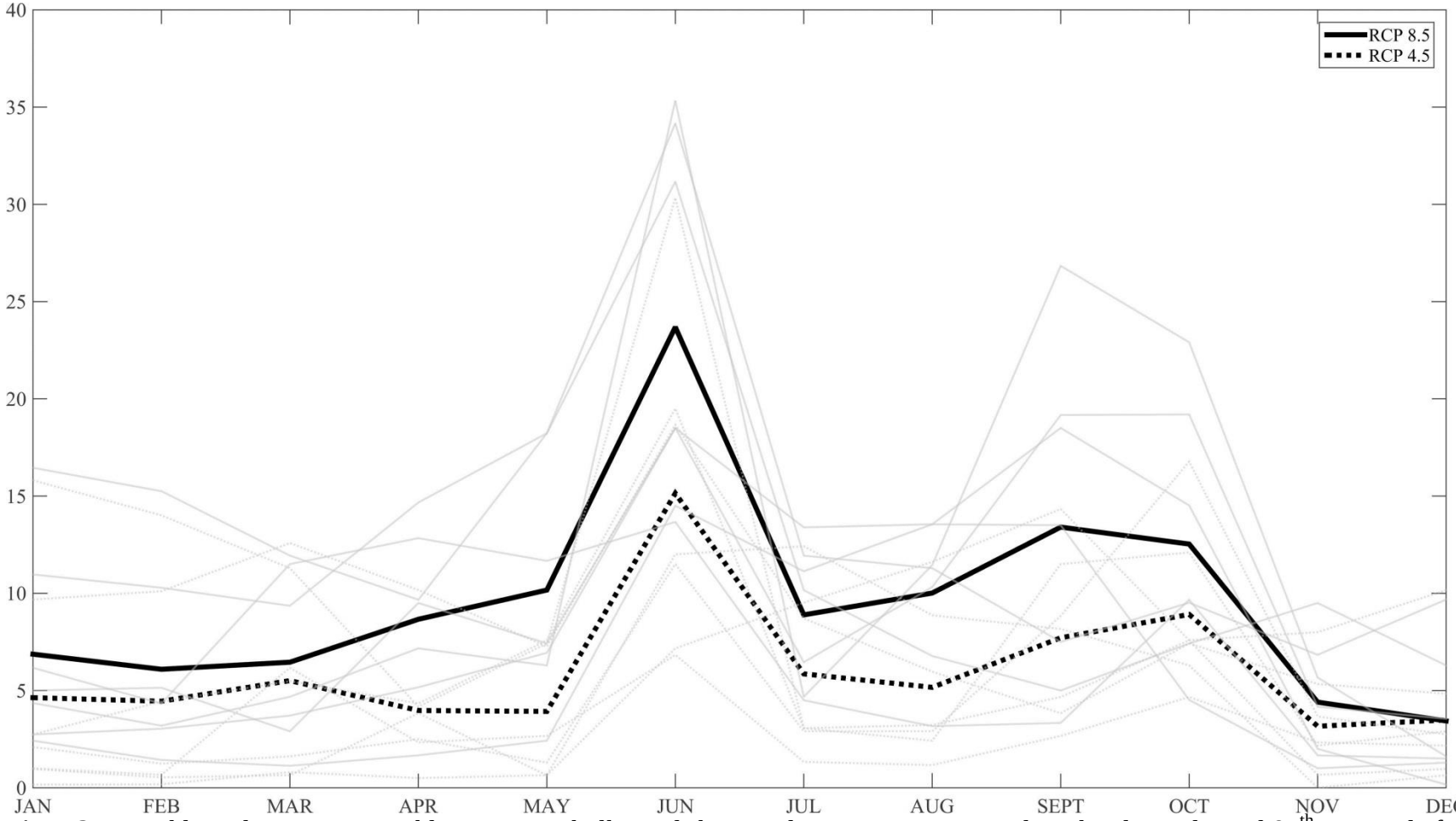

Fig. 4.25 Monthly end-century ensemble projection bulk wind shear grid point average exceeding the climatological 95th percentile for the RCP 4.5 (dashed black line) and RCP 8.5 ensembles (solid black line).

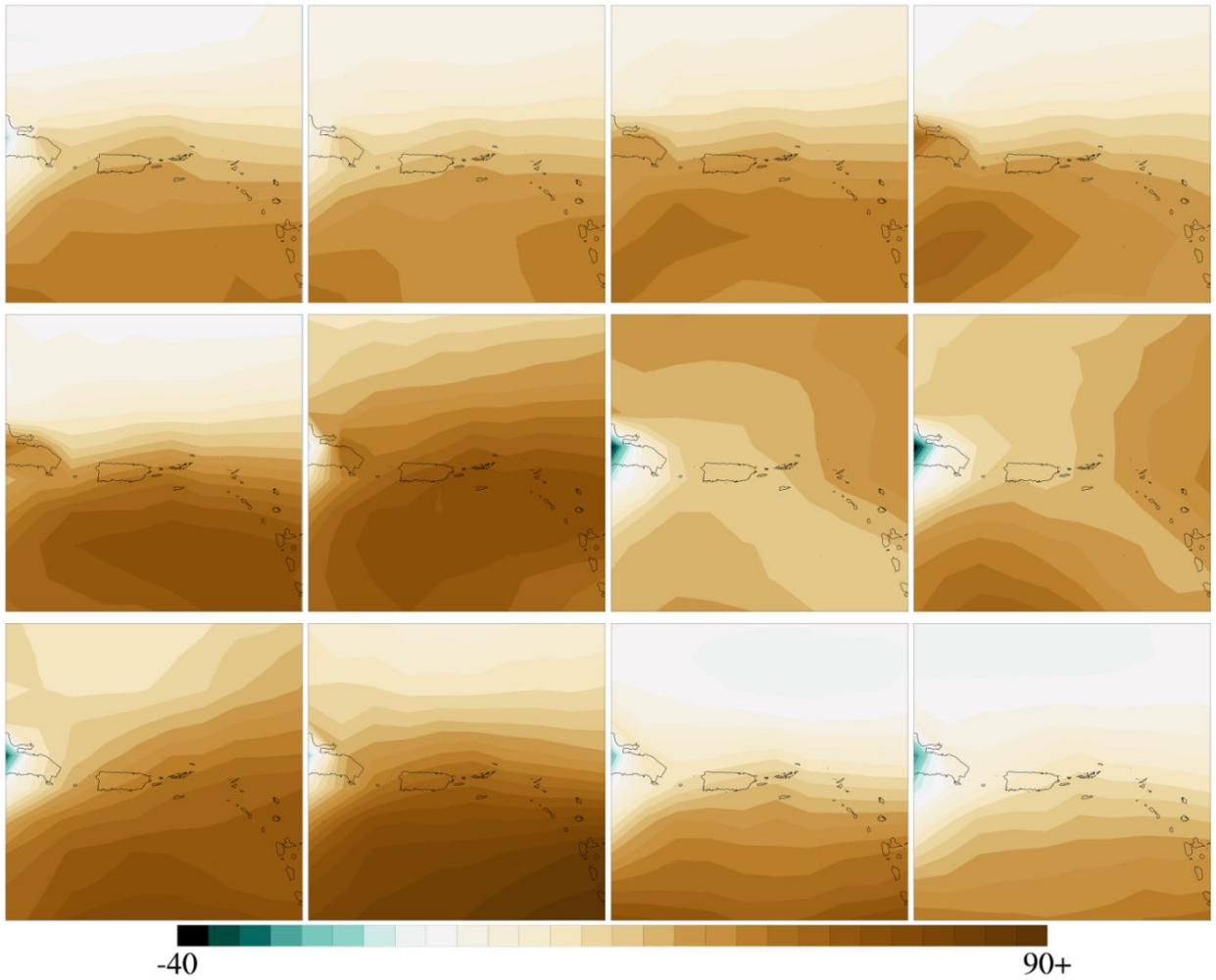

Fig. 4.26 Monthly bulk wind shear deviations from climatology as projected by the mid-century RCP 8.5 model ensemble. January is visualized in the upper-left panel with December in the lower-right panel.

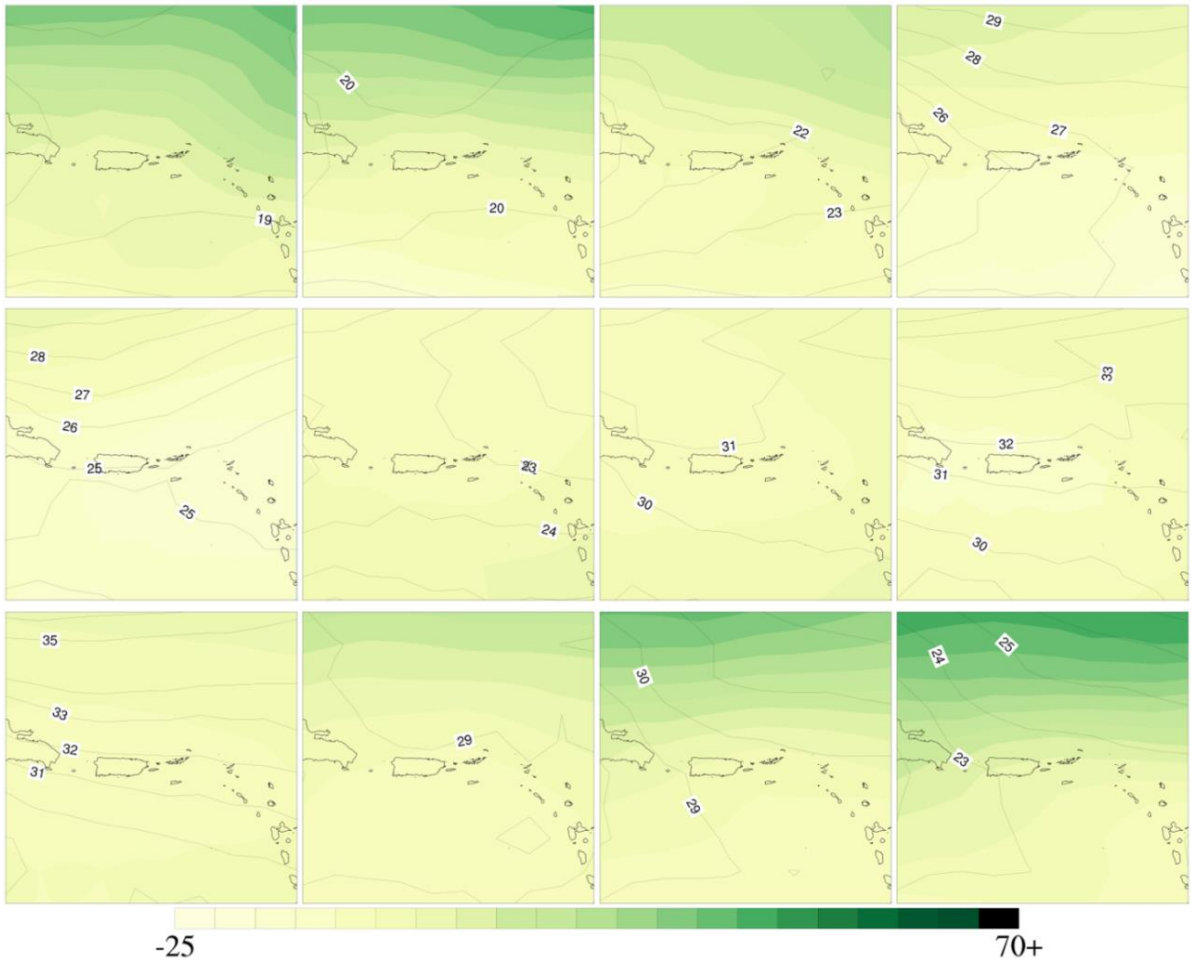

Fig. 4.27 Monthly 700 hPa specific humidity (green-shading) and geopotential height (contour lines) deviations from climatology as projected by the mid-century RCP 8.5 model ensemble. January is visualized in the upper-left panel with December in the lower-right panel.

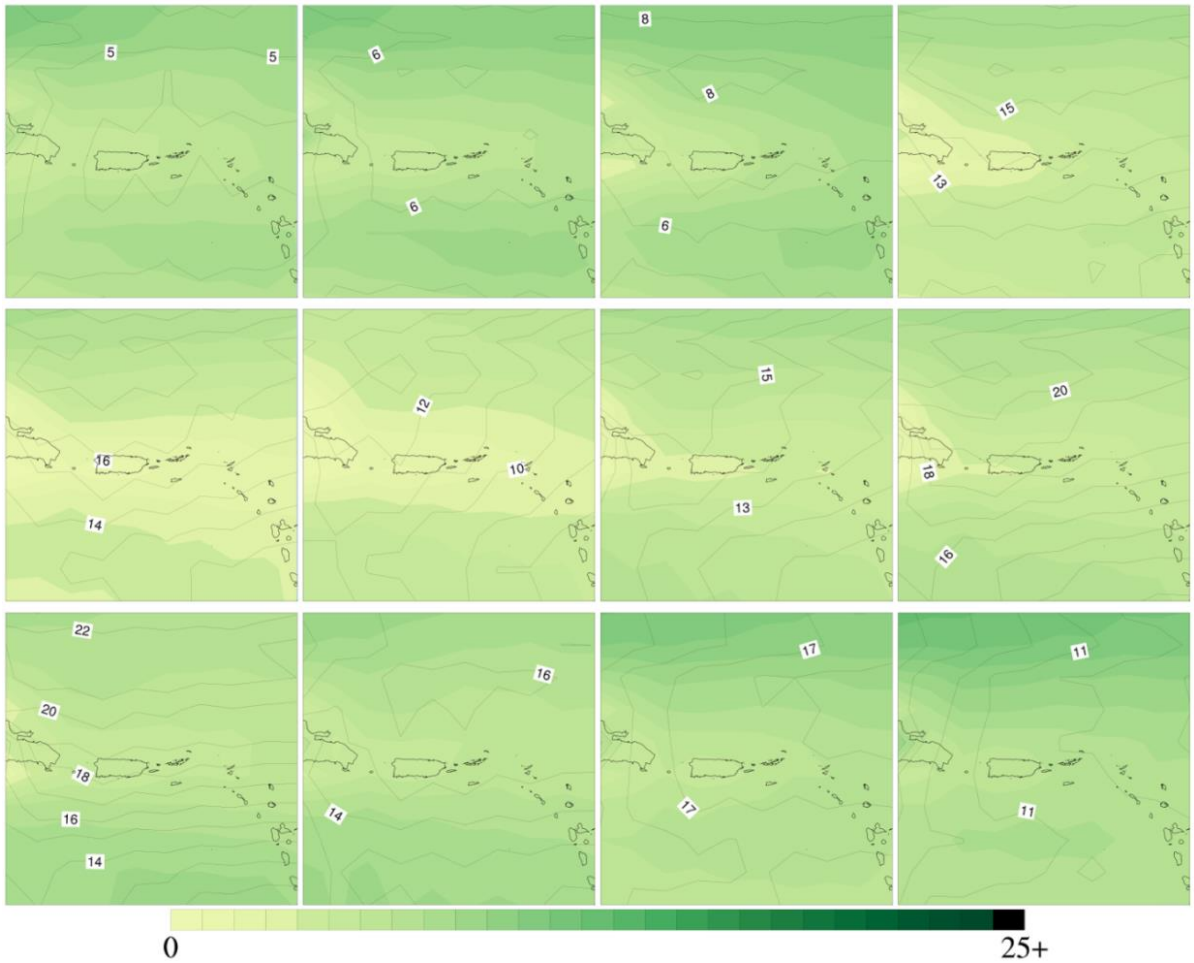

Fig. 4.28 Monthly 1000 hPa specific humidity (green-shading) and geopotential height (contour lines) deviations from climatology as projected by the mid-century RCP 8.5 model ensemble. January is visualized in the upper-left panel with December in the lower-right panel.

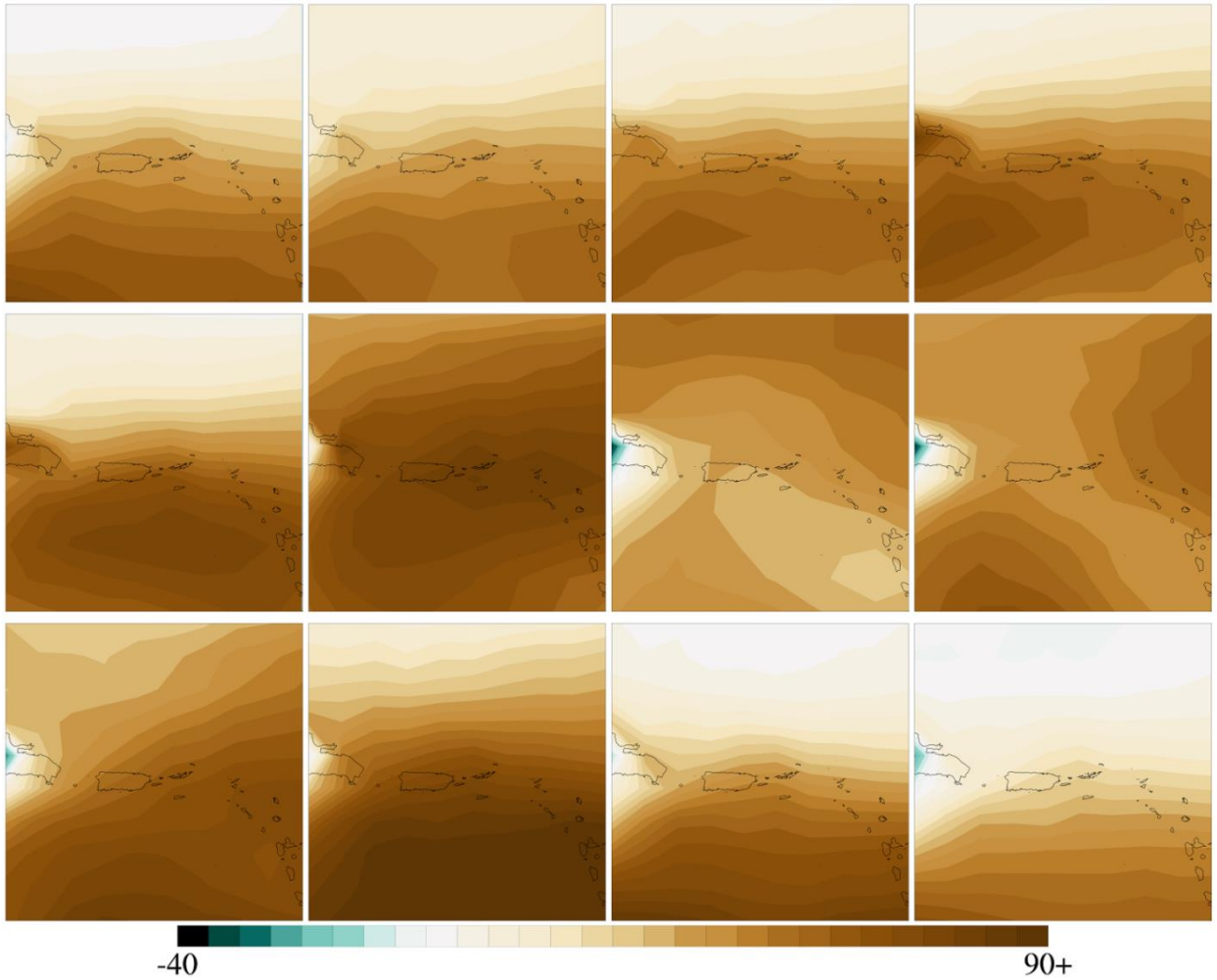

Fig. 4.29 Monthly bulk wind shear deviations from climatology as projected by the end-century RCP 8.5 model ensemble. January is visualized in the upper-left panel with December in the lower-right panel.

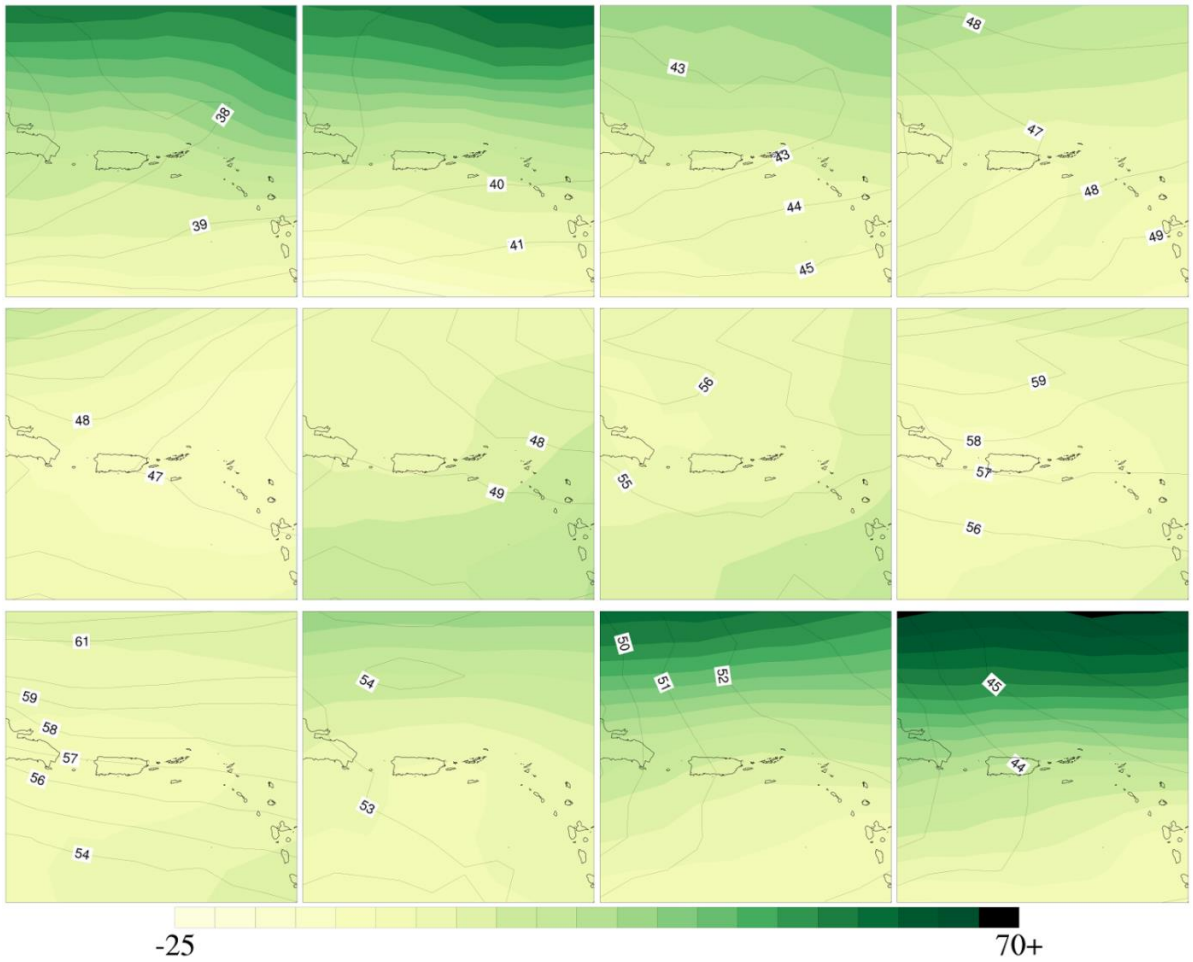

Fig. 4.30 Monthly 700 hPa specific humidity (green-shading) and geopotential height (contour lines) deviations from climatology as projected by the end-century RCP 8.5 model ensemble. January is visualized in the upper-left panel with December in the lower-right panel.

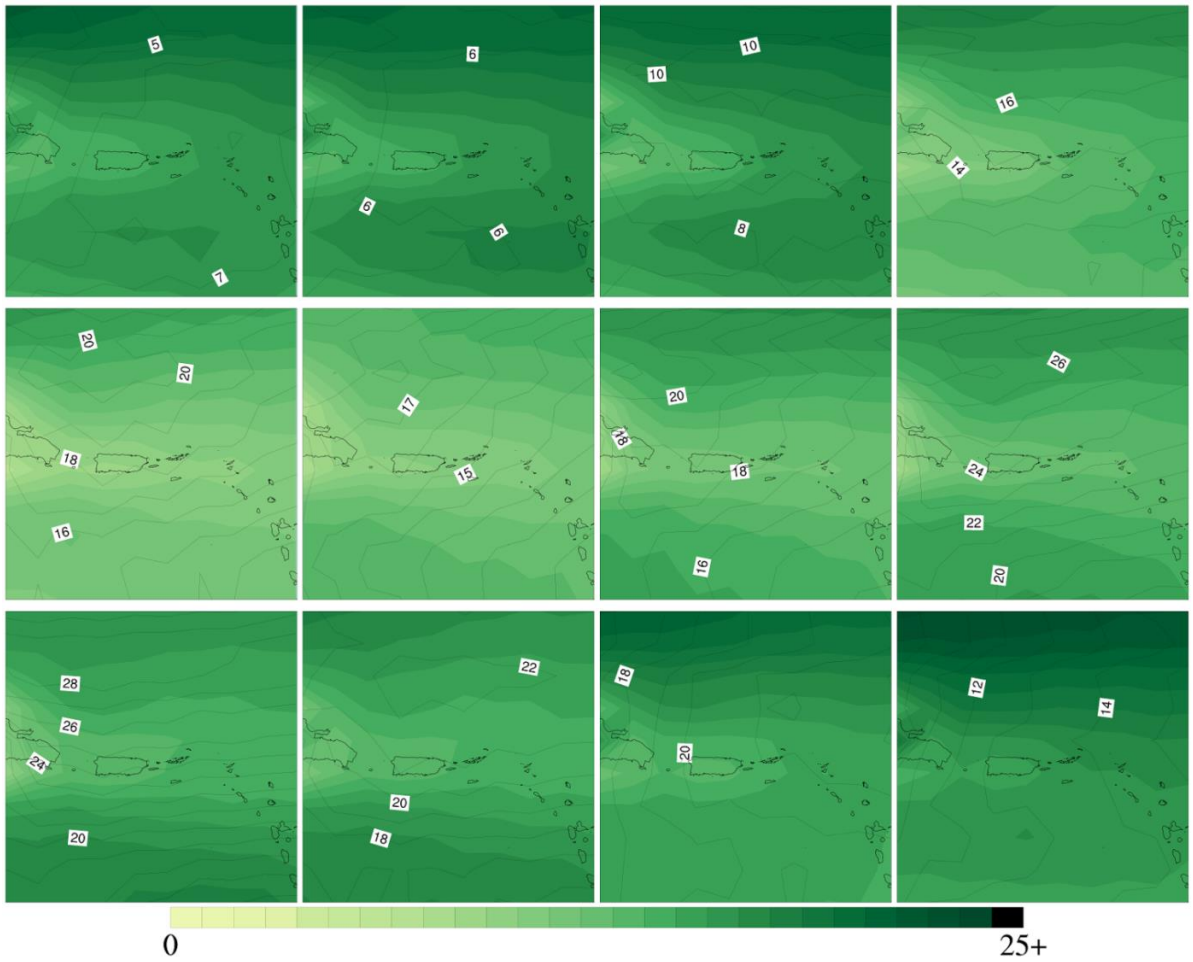

Fig. 4.31 Monthly 1000 hPa specific humidity (green-shading) and geopotential height (contour lines) deviations from climatology as projected by the end-century RCP 8.5 model ensemble. January is visualized in the upper-left panel with December in the lower-right panel.

CHAPTER 5

SUMMARY AND CONCLUSIONS

5.1 Overview

NE PR is a landscape defined by the LM which is host to diverse tropical forests. The ecosystems located in the mosaic of stands are sensitive to small shifts in climate, most notably precipitation variability. Climatic variability leads to changes in biotic structure and biogeochemistry and it is likely that an increase in drought and extreme events will accelerate the rate of change and result in altered ecosystem compositions and productivity (Covich et al. 1991, Silver 1998, Lugo et al. 2012, Willig et al. 2012).

San Juan and the surrounding areas are highly dependent on precipitation for the municipal water supply. A significant portion of water flowing out of the LM is diverted for municipal water purposes (Crook et al. 2007). Increased precipitation variability will likely cause greater stress on the San Juan Metropolitan water supply as drought frequency increases and populations grow. Additionally, the El Yunque National Forest in the LM attracts millions of visitors a year and provides an important economic service for the island. Shifts in forest composition and ecosystems could lead to the area becoming less frequented by visitors.

This dissertation addresses the need for a more comprehensive, higher resolution understanding of the past, present, and future hydroclimate of NE PR, more specifically, the LM. The major objectives of this dissertation are to model current precipitation to better understand the atmospheric controls on rainfall variability (Chapter 2). Chapter 3

identifies the leading modes of variability in low-tropospheric environments and downscale those modes to precipitation in NE PR. Lastly, Chapter 4 uses a SOM downscaling technique to project future changes in atmospheric and precipitation variability in NE PR.

5.2 Summary

Chapter 2 uses artificial neural networks to model precipitation from 37 atmospheric variables at six sites located around Puerto Rico with an emphasis on the LM in northeastern Puerto Rico. This study finds that precipitation variability on the island is primarily driven by variations in moisture and winds between 1000 – 700 hPa. This finding is important for understanding the key drivers of future precipitation variability and for future downscaling methodologies in the region. This study provides quantitative justification for the use of low-tropospheric specific humidity and winds as the best predictor variables for use in downscaling methodologies in the Caribbean.

Chapter 3 uses self-organizing maps to identify the most common synoptic patterns in low-tropospheric moisture, geopotential height, and wind fields. As these are identified as the atmospheric variables with the largest control on precipitation in the study area, the nodes in the SOM are related to precipitation variability in NE PR. This study finds that the SOM nodes leading to the driest conditions represent environments with high 1000 – 700 wind shear and low 700 hPa specific humidity. While wind shear impacts on Atlantic tropical storm development is well discussed in the literature, the findings of this study indicate that increases in wind shear have important impacts on smaller scales as well.

Chapter 4 addresses the need for a more comprehensive understanding of the how climate and precipitation variability are projected to change in the future in NE PR. The study uses a SOM-based downscaling methodology and adequately reproduces historical precipitation variability. However, when future climate simulations CMIP5 GCMs are mapped to the nodes, the SOM downscaling model performs poorly and produces results inconsistent with GCM precipitation data and other regional downscaling studies. The cause of this is a high degree of non-stationarity in the climate data. Future climatic modes of variability are projected to occupy an entirely new data space, and thus, when GCM data are mapped to the historical nodes, the extreme nature of the future environments are greatly underestimated or not represented at all in the historical data space. The future low-tropospheric environment in the study region is projected to be warmer and drier at 700 hPa which will lead to increase atmospheric stability. Compounding this, 1000 – 700 hPa wind shear is projected to increase in magnitude. These appear to be the driving mechanisms behind the drying trend estimated in GCM precipitation projections. While all three season appear to follow a drying trend, the ERS appears to be the most impacted with the highest magnitude deviations in precipitation.

Our conclusion indicating a future regional climate conducive for drier conditions at El Verde finds consilience with future GCM and downscaled precipitation data in the region. Projections of future monthly rainfall at El Verde indicate a drying trend in all months with the largest decreases in July (Fig. 5.1). Spatially, these trends are consistent throughout the Caribbean with peak decreases in precipitation during July (Fig. 5.2). This dissertation shows that this is likely due to large drying trends at 700 hPa and increases in

wind shear through the lower troposphere. These are likely manifestations of a stronger, longer-lasting MSD for the region.

5.3 Conclusions

The findings from this dissertation play an important role in providing a comprehensive understanding of the interactions between atmospheric variability and precipitation in the LM and NE PR. The results presented here indicate that an increase in precipitation variability has occurred through the historical record. Additionally, CMIP5 models indicate that future climate regimes in the study region will occupy an entirely new data space. Historical climatic variability is unable to adequately represent the new environments. Thus, the future climate of the Caribbean is highly non-stationary. However, the results of this dissertation provide sufficient evidence to support the trends suggested in GCM predicted precipitation. It appears plausible that despite the topographic lift provided by the LM, a substantial shift in precipitation is to be expected. The new synoptic low-tropospheric environment suggests a drier future for the LM. The most substantial changes are projected for the ERS and become increasingly influenced by current large-scale drying forcing mechanisms (i.e. NASH, SAL).

The 2015 ERS may have provided a harbinger of future climate for NE PR. Analysis of the deviations in May 2015 700 hPa specific humidity shows a substantial decrease in moisture. This appears to be to most significant forcing mechanisms for the 2015 drought. This persistent dry tongue was likely the result of some combination of the expansion of the NASH and a moderately strong SAL. This hot, dry air layer led to a statically stable low-troposphere and limited deep, moist convection and produced the driest May on record at El Verde. The social and ecological implications from this

drought were widespread as biogeochemical cycles were severely altered and water rationing was frequent in San Juan throughout the ERS and into the LRS. The ecological implications of this drought will become increasingly evident throughout the next couple of years as extensive research is ongoing at the Luquillo Long-Term Ecological Research Site and at the International Institute of Forestry. The results of this dissertation suggest that events like these will become more frequent in the future and it is of critical importance to understand the lasting impacts this increased precipitation variability will have on the future of NE PR.

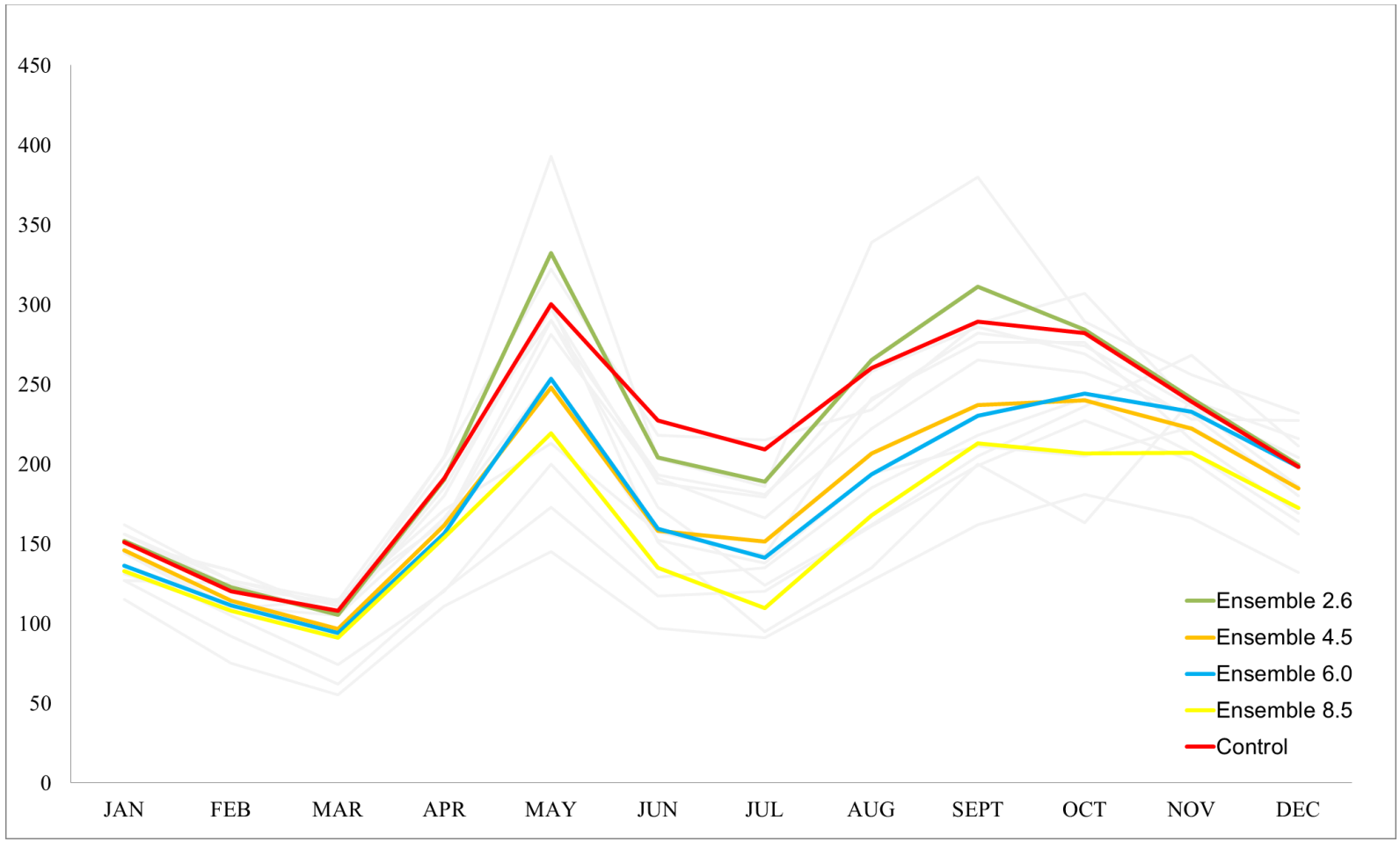

Fig. 5.1 Monthly downscaled precipitation anomalies for El Verde as simulated by a WorldClim ensemble. Ensemble members are shown in light gray.

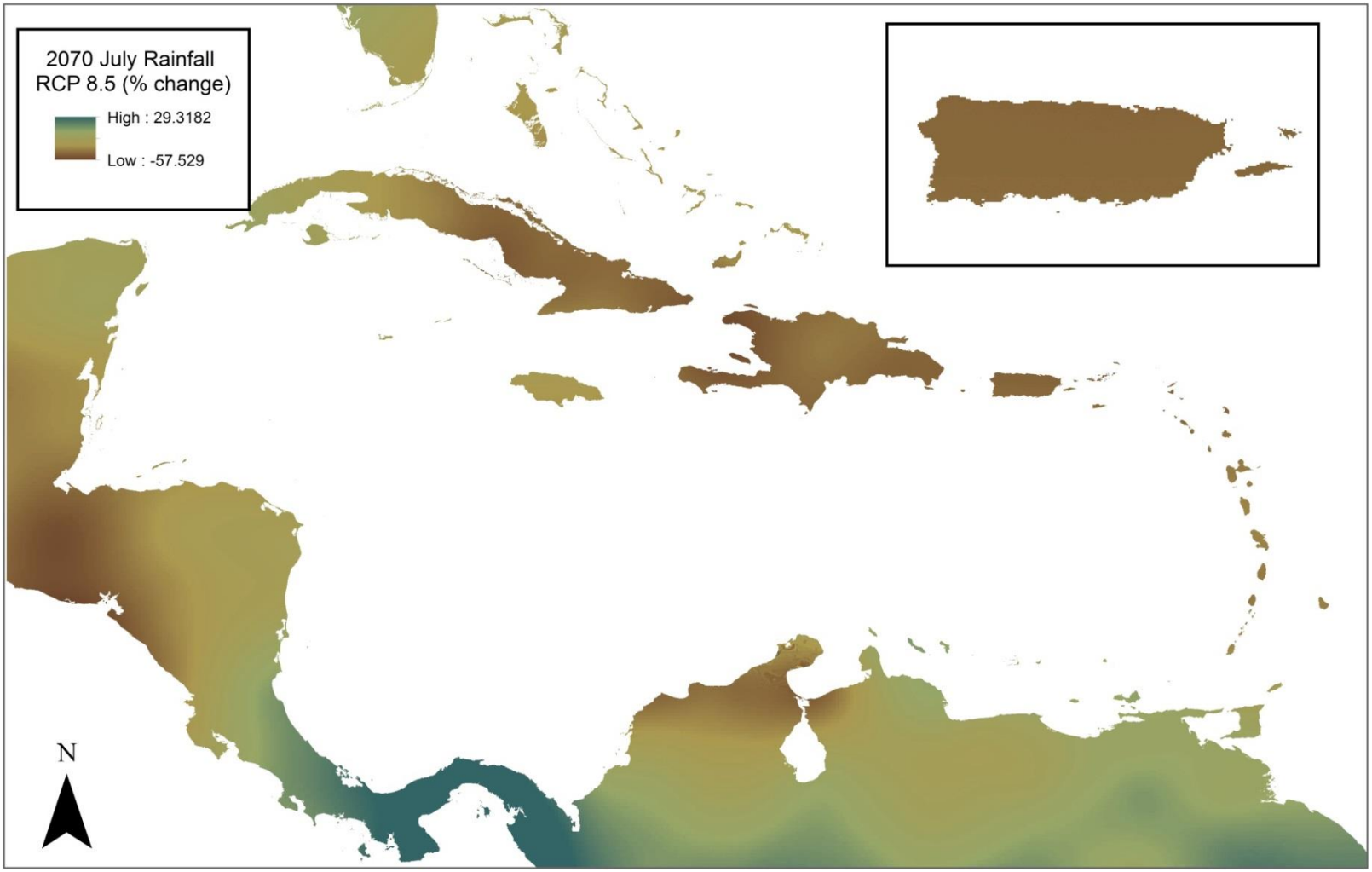

Fig. 5.2 Downscaled July 2070 precipitation for the Caribbean Basin as simulated by a WorldClim ensemble (Hijmans et al. 2005).

REFERENCES

Angeles, M.E., J. E. González, N. D. Ramírez-Beltrán, C. A. Tepley, and D. E. Comarazamy, 2010: Origins of the Caribbean rainfall bimodal behavior. *J Geophys Res*, **115**, D11106, doi: 10.1029/2009JD012990.

Bellerby, T., M. Todd, D. Kniveton, and C. Kidd, 2000: Rainfall Estimation from a Combination of TRMM Precipitation Radar and GOES Multispectral Satellite Imagery Through the Use of an Artificial Neural Network. *J Appl Meteorol*, **39,** 2115–2128.

Benstead, J. P., J. G. March, C. M. Pringle, and F. N. Scatena, 1999: Effects of a low-head dam and water abstraction on migratory tropical stream biota. *Ecological Applications*, **9**, 656–668.

Cassano, J. J., P. Uotila, A. H. Lynch, and E. N. Cassano, 2007: Predicted changes in synoptic forcing of net precipitation in large Arctic river basins during the 21st century. *J Geophys Res: Biogeosciences*, **112,** 1–20.

Cassano, E. N., and J. J. Cassano, 2010: Synoptic forcing of precipitation in the Mackenzie and Yukon River basins. *Int J Climatol*, **30**, 658–674.

Cavazos, T., 1999: Large-Scale Circulation Anomalies Conducive to Extreme Precipitation Events and Derivation of Daily Rainfall in Northeastern Mexico and Southeastern Texas. *J Clim*, **12,** 1506–1523.

——, 2000: Using self-organizing maps to investigate extreme climate events: An application to wintertime precipitation in the Balkans. *J Clim*, 1718–1732.

Cavazos, T., and B. C. Hewitson, 2005: Performance of NCEP-NCAR Reanalysis Variables in Statistical Downscaling of Daily Precipitation. *Clim Res*, **28,** 95–107.

Chattopadhyay, R., A. K. Sahai, B. N. Goswami, 2008: Objective identification of nonlinear convectively coupled phases of monsoon intraseasonal oscillation: implications for prediction. *J Atmos Sci*, **65**, 1549–1569.

Chen, A., and M. A. Taylor, 2002: Investigating the link between early season Caribbean rainfall and the El Niño + 1 year. *Int J Clim*, **22**, 87–106.

Comarazamy, D. E., and J. E. González, 2008: On the Validation of the Simulation of Early Season Precipitation on the Island of Puerto Rico Using a Mesoscale Atmospheric Model. *J Hydrometeorol*, **9,** 507–520.

——, 2011: Regional Long-Term Climate Change (1950–2000) in the Midtropical Atlantic and its Impacts on the Hydrological Cycle of Puerto Rico. *J Geophys Res*, **116**, doi:10.1029/2010JD015414.

Covich, A., A. T. Crowl, A. H. Johnson, D. Varza, and D. Certain, 1991: Post-Hurricane Hugo increases in atyid shrimp abundances in a Puerto Rican montane stream. *Biotropica* , **23**,448–454.

Covich, A. P., T. A. Crowl, and F. N. Scatena, 2003: Effects of extreme low flows on freshwater shrimps in a perennial tropical stream. *Freshwater Biology*, **48**, 1199–1206.

Covich, A. P., T. A. Crowl, and T. Heartsill-Scalley, 2006: Effects of drought and hurricane disturbances on headwater distributions of palaemonid river shrimp (Macrobrachium spp.) in the Luquillo Mountains, Puerto Rico. *Journal of the North American Benthological Society*, **25**, 99–107.

Crook, K. E. , F. N. Scatena, and C. M. Pringle: 2007. Water withdrawal from the Luquillo Experimental Forest 2004. General Technical Report IIT-GTR-34, International Institute of Tropical Forestry, Pacific Northwest Research Station, 40 p.

Crook, K. E., C. M. Pringle, and M. C. Freeman, 2009: A method to assess longitudinal riverine connectivity in tropical streams dominated by migratory biota. *Aquatic Conservation* , **19**, 714–723.

Curtis, S., and D. W. Gamble, 2008: Regional variations of the Caribbean mid-summer drought. *Theor Appl Climatol*, **94**, 25–34.

Daly, C., E. H. Helmer, and M. Quiñones, 2003: Mapping the climate of Puerto Rico, Vieques and Culebra. *Int J Clim*, **23**, 1359–1381.

Dee, D., and Coauthors, 2011: The ERA-Interim Reanalysis: Configuration and Performance of the Data Assimilation System. *Q J R Meteorolog Soc*, **137,** 553–597.

Diem, J.E., 2013: Comments on "Changes to the North Atlantic subtropical high and its role in the intensification of summer rainfall variability in the southeastern United States". *J Clim*, **26**, 679 – 682.

Enfield, D. B., and E. J. Alfaro, 1999: The Dependence of Caribbean Rainfall on the Interaction of the Tropical Atlantic and Pacific Oceans. *J Clim*, **12,** 2093–2103.

Gamble, D. W., D. B. Parnell, and S. Curtis, 2008: Spatial variability of the Caribbean mid-summer drought and relation to north Atlantic high circulation. *Int J Climatol*, **28**, 343–350.

Gardner, M. W., and S. R. Dorling, 1998: Artificial Neural Networks (the Multilayer Perceptron)—A Review of Applications in the Atmospheric Sciences. *Atmos Environ*, **32,** 2627–2636.

Garson, D. G., 1991: Interpreting Neural-Network Connection Weights. *Artif Intell Expert*, **6,** 47–51.

Giannini, A., Y. Kushnir, and M. A. Cane, 2000: Interannual Variability of Caribbean Rainfall, ENSO, and the Atlantic Ocean. *J Clim*, **13,** 297–311.

——, 2001: Seasonality in the impact of ENSO and the North Atlantic high on Caribbean rainfall. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, **62**, 143–147.

Giorgi, F., and Coauthors, 2001: Emerging Patterns of Simulated Regional Climatic Changes for the 21st Century due to Anthropogenic Forcings. *Geophys Res Lett*, **28,** 3317–3320.

Goh, A., 1995: Back-Propagation Neural Networks for Modeling Complex Systems. *Artif Intell Eng*, **9,** 143–151.

Granger, O. E., 1985: Caribbean Climates. *Prog Phys Geogr*, **9,** 16–43.

Gutnick, M., 1958: Climatology of the trade-wind inversion in the Caribbean. *Bull Am Meteorol Soc*, **39,** 410–420.

Hall, T., H. E. Brooks, and C. A. Doswell III, 1999: Precipitation Forecasting Using a Neural Network. *Weather Forecasting*, **14,** 338–345.

Hastenrath, S., 1976: Variations in Low-Latitude Circulation and Extreme Climatic Events in the Tropical Americas. *J Atmos Sci*, **33,** 202–215.

Hayhoe, K., 2013: Quantifying key drivers of climate variability and change for Puerto Rico and the Caribbean. Agreement number G10AC00582. Caribbean Landscape Conservation Cooperative.

Haylock, M. R., G. C. Cawley, C. Harpham, R. L. Wilby, and C. M. Goodess, 2006: Downscaling Heavy Precipitation over the United Kingdom: a Comparison of Dynamical and Statistical Methods and their Future Scenarios. *Int J Clim*, **26,** 1397–1415.

Heartsill-Scalley T., Scatena F.N., Estrada C., McDowell W.H., Lugo A.E., 2007: Disturbance and long-term patterns of rainfall and throughfall nutrient fluxes in a subtropical forest in Puerto Rico. *J Hydrol*, **33**, 472–485.

Hewitson, B., and R.G. Crane, 1996: Climate Downscaling: Techniques and Application. *Clim Res*, **7,** 85–95.

——, 1992: Large-Scale Atmospheric Controls on Local Precipitation in Tropical Mexico. *Geophys Res Lett*, **19,** 1835–1838.

——, 1994: *Neural Nets: Applications in Geography: Applications for Geography.* Vol. 29, Springer Science & Business Media.

——, 2002: Self-organizing maps: applications to synoptic climatology. *Clim Res*, **22**, 13–26.

——, 2006: Consensus between GCM climate change projections with empirical downscaling: precipitation downscaling over South Africa. *Int J Clim*, **26,** 1315–1337.

Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones and A. Jarvis, 2005: Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol*, **25**, 1965–1978.

Huth, R., 2001: Disaggregating climatic trends by classification of circulation patterns. *Int J Climatol*, **21**, 135–153.

IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

Jones, P. W., 1999: First- and Second-Order Conservative Remapping Schemes for Grids in Spherical Coordinates. *Mon Weather Rev*, **127,** 2204–2210.

Jury, M. R., 2009: An Intercomparison of Observational, Reanalysis, Satellite, and Coupled Model Data on Mean Rainfall in the Caribbean. *J Hydrometeorol*, **10,** 413–430.

Kuligowski, R. J., and A. P. Barros, 1998: Localized Precipitation Forecasts from a Numerical Weather Prediction Model Using Artificial Neural Networks. *Weather Forecasting*, **13,** 1194–1204.

Larsen, M. C., 2000: Analysis of 20th century rainfall and streamflow to characterize drought and water resources in Puerto Rico. *Phys Geog*, **21**, 494–521.

Lee, S.-K., D. B. Enfield, and C. Wang, 2011: Future impact of differential interbasin ocean warming on Atlantic hurricanes. *J Clim*, **24**, 1264–1275.

Lugo, A. E., R. B. Waide, M. R. Willig, A. T. Crowl, F. N. Scantena, J. Thompson, W. L. Silver, W. H. McDowell, and N. V. L. Brokaw. 2012. Ecological Paradigms for the Tropics Old Questions and Continuing Challenges. *A Caribbean forest tapestry: the multidimensional nature of disturbance and response.* N.V.L. Brokaw, A.T. Crowl, A. E. Lugo, W. H. McDowell, F. N. Scantena, R. B. Waide, and M. R. Willig, Eds. Oxford University Press, New York, NY.

Lynch, A. H., P. Uotila, and J. J. Cassano, 2006: Changes in synoptic weather patterns in the polar regions in the 20th and 21st centuries, Part 2: Antarctic. *Int J Climatol*, **26**, 1181–1199.

Magaña, V., J. A. Amador, and S. Medina, 1999: The midsummer drought over Mexico and Central America. *J Clim*, **12**, 1577–1588.

Maier, H. R., and G. C. Dandy, 2000: Neural networks for the Prediction and Forecasting of Water Resources Variables: A Review of Modelling Issues and Applications. *Env Modell Software*, **15,** 101–124.

Malmgren, B. A., A. Winter, A. and D. Chen, 1998: El Nino-Southern oscillation and North Atlantic oscillation control of Caribbean climate. *J Clim*, **11**, 2713–2717.

Malmgren, B., and A. Winter, 1999: Climatic zonation in Puerto Rico based on multivariate statistical analysis and an artificial neural network. *J Clim*, **12**, 977–985.

Maloney, E. D., and Coauthors, 2014: North American climate in CMIP5 experiments. Part III: Assessment of twenty-first-century projections. *J Clim*, **27**, 2230–2270, doi:10.1175/JCLI-D-13-00273.1.

Marzban, C., 2009: Performance Measures and Uncertainty. *Artificial Intelligence Methods in the Environmental Sciences*, S. Haupt, A. Pasini, and C. Marzban, Eds., Springer Netherlands, 49–75.

Menne, M. J., I. Durre, R. S. Vose, B. E. Gleason, and T. G. Houston, 2012: An overview of the global historical climatology network-daily database. *J Atmos Ocean Technol*, **29,** 897–910.

Mote, T., 1998: Mid-tropospheric circulation and surface melt on the Greenland ice sheet. Part II: Synoptic climatology. *Int J Climatol* , **18**, 131–146.

Mulligan, M., and S. M. Burke, 2005: FIESTA: Fog Interception for the Enhancement of Streamflow in Tropical Areas, 174 pp.

Neelin, J. D., M. Munnich, H. Su, J. E. Meyerson, and C. E. Holloway, 2006: Tropical drying trends in global warming models and observations. *Proc Natl Acad Sci USA*, **103**, 6110–6115, doi:10.1073/pnas.0601798103.

Nobre, P., and J. Srukla, 1996: Variations of Sea Surface Temperature, Wind Stress, and Rainfall over the Tropical Atlantic and South America. *J Clim*, **9,** 2464–2479.

Olden, J. D., and D. A. Jackson, 2002: Illuminating the "black box": A Randomization Approach for Understanding Variable Contributions in Artificial Neural Networks. *Ecol Model*, **154,** 135–150.

Pringle, C. M., 1997: Exploring how disturbance is transmitted upstream: going against the flow. *Journal of the North American Benthological Society*, **16**, 425–438.

Pringle, C. M. and F. N. Scatena, 1998: Freshwater resource development: Case studies from Puerto Rico and Costa Rica, *Tropical managed ecosystems: New persepctives on sustainability.* U. Hatch and M. E. Swisher, Eds. Oxford University Press. 114–221.

Prospero, J.M. and T.N. Carlson, 1972: Vertical areal distributions of Saharan dust over the western equatorial North Atlantic Ocean. *J Geophys Res*, **77**, 5255–5265.

Ramseyer, C.A. and T. Mote, 2016: Atmospheric controls on Puerto Rico precipitation using artificial neural networks. *Clim Dyn*, doi: 10.1007/s00382-016-2980-3.

Rauscher, S. A., F. Giorgi, N. S. Diffenbaugh, and A. Seth, 2008: Extension and intensification of the meso-American midsummer drought in the twenty-first century. *Clim Dyn*, **31**, 551–571.

Reusch, D.B., R. B. Alley, and B. C. Hewitson, 2005: Relative performance of self-organizing maps and principle component analysis in pattern extraction from synthetic climatological data. *Polar Geog*, **29**, 188–212.

Ryu, J.-H. and K. Hayhoe, 2014: Understanding the sources of Caribbean precipitation biases in CMIP3 and CMIP5 simulations. *Clim Dyn*, **42**, 3233– 3252, doi: 10.1007/s00382-014-2351-x.

Sahai, A., M. Soman, and V. Satyan, 2000: All India Summer Monsoon Rainfall Prediction using an Artificial Neural Network. *Clim Dyn*, **16,** 291–302.

Sauter, T., V. Venema, 2011: Natural three-dimensional predictor domains for statistical precipitation downscaling. *J Clim*, 24, 6132–6145.

Schoof, J. T., and S. C. Pryor, 2001: Downscaling Temperature and Precipitation: A Comparison of Regression-based Methods and Artificial Neural Networks. *Int J Clim*, **21,** 773–790.

Schubert, W. H., P. E. Ciesielski, C. Lu, and R. H. Johnson, 1995: Dynamical Adjustment of the Trade Wind Inversion Layer. *J Atmos Sci*, **52,** 2941–2952.

Schuenemann, K., and J. J. Cassano, 2010: Changes in synoptic weather patterns and Greenland precipitation in the 20th and 21st centuries 2: Analysis of 21st Century atmospheric changes using self-organizing maps. *J Geophys Res–Atmos*, **115**, D05108.

Sheffield, J. and Coauthors, 2013: North American climate in CMIP5 experiments. Part II: Evaluation of historical simulations of intraseasonal to decadal variability. *J. Climate*, **26**, 9247–9290

Shepherd J.M., M. Carter, M. Manyin, D. Messen, S. Burian, 2010: The impact of urbanization on current and future coastal precipitation: a case study for Houston. *Environment and Planning B: Planning and Design*, **37(2)**, 284 – 304.

Sheridan, S. C., and C. C. Lee, 2011: The self-organizing map in synoptic climatological research. *Prog Phys Geog*, 35, 109–119.

Silver, W. L. 1998: The potential effects of elevated CO2 and climate change on tropical forest biogeochemical cycling. *Climatic Change*, **39**, 337–361.

Silverman, D., and J. A. Dracup, 2000: Artificial Neural Networks and Long-Range Precipitation Prediction in California. *J Appl Meteorol*, **39,** 57–66.

Skific, N., and J. Francis, 2012: Self-Organizing Maps: A Powerful Tool for the Atmospheric Sciences.

Sobie, S.R., and A. J. Weaver, 2012: Downscaling of Precipitation over Vancouver Island using a Synoptic Typing Approach. *Atmosphere-Ocean*, 50, 176–196.

Taylor, M. A., D. B. Enfield, and A. A. Chen, 2002: Influence of the tropical Atlantic versus the tropical Pacific on Caribbean rainfall. *J. Geophys. Res.*, **107**, 3127., doi:10.1029/2001JC001097.

Trenberth, K. E., A. Dai, R. M. Rasmussen, and D. B. Parsons, 2003: The Changing Character of Precipitation. *Bull Am Meteorol Soc*, **84,** 1205–1217.

Tsonis, A. A., and J. B. Elsner, 1992: Nonlinear Prediction as a way of distinguishing chaos from random fractal sequences. *Nat*, **358,** 217–220.

Valverde Ramírez, M. C., H. F. de Campos Velho, and N. J. Ferreira, 2005: Artificial neural network technique for rainfall forecasting applied to the São Paulo region. *J Hydrol*, **301,** 146–162.

Vimont, D. J., D. S. Battisti, and R. L. Naylor, 2010: Downscaling Indonesian precipitation using large-scale meteorological fields. *Int J of Climatol*, 30, 1706–1722.

Widmann, M. M., C. S. Bretherton, and E. P. Salathe, 2003: Statistical precipitation downscaling over the Northwestern United States using numerically simulated precipitation as a predictor. *J. of Clim.*, **16,** 799 – 816.

Wilby, R., S. Charles, E. Zorita, B. Timbal, P. Whetton, and L. Mearns, 2004: Guidelines for use of climate scenarios developed from statistical downscaling methods.

Willig, M. R., C. P. Bloch, A. Covich, C. A. S. Hall, D. J. Lodge, A. E. Lugo, W. L. Silver, R. B. Waide, L. A. Walker, and J. K. Zimmerman, 2012: Long-Term Research in the Luquillo Mountains: synthesis and foundations for the future. *A Caribbean Forest Tapestry: The Multidimensional Nature of Disturbance and* Response, N.V.L Brokaw, A.T. Crowl, A. E. Lugo, W. H. McDowell, F. N. Scatena, R. B. Waide, and M. R. Willig, Eds., Oxford University Press, New York, NY, 361–441.

Winkler, J. A., and Coauthors, 2011: Climate Scenario Development and Applications for Local/Regional Climate Change Impact Assessments: An Overview for the Non-Climate Scientist. *Geogr Compass*, **5,** 275–300.

Wu, R., and B. P. Kirtman, 2005: Roles of Indian and Pacific Ocean air–sea coupling in tropical atmospheric variability. *Clim Dyn*, **25,** 155–170.

——, 2011: Caribbean Sea rainfall variability during the rainy season and relationship to the equatorial Pacific and tropical Atlantic SST. *Clim Dyn*, **37,** 1533–1550.