2016

Potential predictability of precipitation over the continental United States

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

POTENTIAL PREDICTABILITY OF PRECIPITATION OVER THE CONTINENTAL UNITED STATES

by

DANIEL JOSEPH SHORT GIANOTTI

B.S., Harvey Mudd College, 2003

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2016

Approved by

First Reader

Bruce T. Anderson, PhD
Professor of Earth & Environment

Second Reader

Guido D. Salvucci, PhD
Professor of Earth & Environment

Third Reader

Michael Dietze, PhD
Associate Professor of Earth & Environment

Acknowledgments

I'd first like to thank Bruce Anderson and Guido Salvucci who created these projects, supported me financially and educationally, and offered endless amounts of time and wisdom throughout my time at Boston University. You are a perfect advising team, offering big picture perspective and detailed focus, methodology and motivation, wordsmithing and structure, and high standards. I hope to work with both of you in the future. I'd also like to thank Mike Dietze for some truly inspirational head trips into the world of models, and Dara Entekhabi for welcoming me into his group — I am enjoying the new view from 685 km.

I'd also like to thank the faculty of Boston University's Department of Earth and Environment — the department is a wonderful place, and you all work hard to make it better all the time. Our illustrious director of graduate studies, Lucy Hutyra, has given me a lot of patient direction (always with a faux stern glance), for which I am much obliged, and Nora Watson has put up with far too many late registration forms from me. The E&E staff may be the most impressive collection of go-getters in Boston — thank you for your tireless quest against institutional entropy.

I'd like to thank L. Lamport and D. Knuth for their assistance in typesetting; Nils Napp for a healthy supply of probability textbooks and spicy cuisine; and Dmitriy Kogan, DD for rescuing all of my files from oblivion countless times since 1999, often from his guru's perch atop the San Gabriels.

I'd particularly like to thank my favorite Professor and collaborator for her constant efforts to help me be the person I want to be; Æ for encouraging a slower path towards graduation, filled with much more wonder and learning; and Ramona for all of the introspective morning walks.

POTENTIAL PREDICTABILITY OF PRECIPITATION OVER THE CONTINENTAL UNITED STATES

DANIEL JOSEPH SHORT GIANOTTI

Boston University, Graduate School of Arts and Sciences, 2016

Major Professor: Bruce T. Anderson, Professor of Earth & Environment

ABSTRACT

Our ability to predict precipitation on climate time-scales (months–decades) is limited by our ability to separate signals in the climate system (cyclical and secular) from noise — that is, variability due to processes that are inherently unpredictable on climate time-scales. This dissertation describes methods for characterizing "weather" noise — variability that arises from daily-scale processes — as well as the *potential predictability* of precipitation on climate time-scales. In each method, we make use of a climate-stationary null model for precipitation and determine which characteristics of the true, non-stationary system cannot be captured by a stationary assumption. This un-captured climate variability is potentially predictable, meaning that it is due to climate time-scale processes, although those processes themselves may not be entirely predictable, either practically or theoretically.

The three primary methods proposed in this dissertation are

1. A stochastic framework for modeling precipitation occurrence with proper daily-scale memory representation, using variable order Markov chains and information criteria for order selection.

2. A corresponding method for representing precipitation intensity, allowing for memory in intensity processes.

3. A semi-parametric stochastic framework for precipitation which represents intensity and occurrence without separating the processes, designed to handle the issues that arise from estimating likelihoods for zero-inflated processes.

Using each of these methods, potential predictability is determined across the contiguous 48 United States. Additionally, the methods of Chapter 4 are used to determine the magnitude of weather and climate variability for the "historical runs" of five global climate models for comparison against observational data.

It is found that while some areas of the contiguous 48 United States are potentially very predictable (up to $\sim 70\%$ of interannual variability), many regions are so dominated by weather noise that climate signals are effectively masked. Broadly, perhaps 20–30% of interannual variability may be potentially predictable, but this ranges considerably with geography and the annual seasonal cycle, yielding "hot spots" and "cold spots" of potential predictability. The analyzed global climate models demonstrate a fairly robust representation of weather-scale processes, and properly represent the ratio of weather-to-climate induced variability, despite some regional errors in mean precipitation totals and corresponding variability.

Contents

List of Tables

List of Figures

List of Symbols and Abbreviations

AIC_c	(Corrected) Akaike Information Criterion
AO	Arctic Oscillation
CPC	Climate Prediction Center
CPC-1/4°, CPCQ	Climate Prediction Center Unified Gauge-based Analysis of Daily Precipitation over the Continental United States at 1/4° resolution
CPC-1°, CPC1	Climate Prediction Center Unified Gauge-based Analysis of Daily Precipitation over the Continental United States at 1° resolution
DJF	December-January-February (inclusive)
ENSO	El Niño/Southern Oscillation
GCM	Global Climate Model
GHCN	Global Historical Climatology Network
GHG	Greenhouse gas
h	Kernel bandwidth parameter
IPCC	Intergovernmental Panel on Climate Change
JJA	June-July-August (inclusive)
KA	Kernel-auto-regressive model
MAM	March-April-May (inclusive)
MJO	Madden-Julian Oscillation
NAO	North Atlantic Oscillation
NCEP	National Centers for Environmental Prediction
NOAA	National Oceanic and Atmospheric Administration
OMC	Occurrence Markov Chain model

PDO	Pacific Decadal Oscillation
PP	Potential predictability
PP_{89}	Potential predictability of 89-day (seasonal) precipitation totals
SON	September-October-November (inclusive)
SST	Sea Surface Temperature
WG	Weather Generator
WPC	Weather Prediction Center
ZL	Zero-lag model
$\sigma^2_{89,obs}$	Interannual variance of 89-day observed precipitation totals
$\sigma^2_{89,sim}$	Interannual variance of 89-day simulated precipitation totals

Chapter 1

Introduction

Our species runs on rain. Water feeds our agriculture and energy systems, our industrial and domestic processes, our ecosystems and transportation networks. Water establishes our climate and the locations of our towns and cities. Our access to water determines our resilience to disease, famine, drought, floods, and conflict; our location on a very unevenly-distributed map of precipitation determines our available freshwater resources. Some 4,000 cubic kilometers of water are extracted from lakes, streams, and aquifers each year for human use (Black and King, 2009), making water management one of humanity's largest infrastructure systems.

Our natural resources are similarly driven by rain. Beyond precipitation's role in the formation of valleys, glaciers, soils, forests, and deserts, roughly 85% of all species and an even larger fraction of the Earth's biomass live in terrestrial or freshwater systems (Costello et al., 2012; Groombridge and Jenkins, 2002; Kallmeyer et al., 2012), meaning that the biosphere is primarily rain-fed. The distribution of species — from fish to flowers — depends on persistent hydroclimatic conditions, and our understanding of ecosystem health hinges on our understanding of precipitation patterns.

By this point in the twenty-first century, we have come to understand much about precipitation and its patterns, mechanisms, and impacts. We broadly know where rain comes from, how much it rains in any location, how much variability there is from day-to-day and year-to-year, and how likely it is to rain today, tomorrow, or next week. The work in this dissertation delves deeper into precipitation to answer questions about prediction and the limits that may constrain our ability to make forecasts about the future.

1.1 Historical background

Phenomenological records of precipitation and its presumed causes can be found in the early archives of most societies [for example, 6th century India (Mihira and Panditbhushan, 1946), 1st century China (Needham, 1986), and 4th century B.C.E. Greece (Wills and Wills, 2003; Aristotle and Webster, 2006)]. Discussion of measurement systems and early rain gauges suggest the existence of quantitative records from the fourth century B.C.E. in India, allegedly for the sake of a water-resource-based taxation system (Biswas, 1967). Near-contiguous records of precipitation have been kept in Korea since the invention of a standardized rain-gauge in 1441 and a first national network in 1442 (Cho et al., 2015), and dedicated long-term meteorological observatories have been in place since at least the late eighteenth century (1874 in Calcutta: Baird, 1879, pg 130). Additional rain gauges developed in Italy in the 1630s and England in the 1660s (the first "tipping-bucket" rain gauge) and their derivatives were exported widely across Europe and the European colonies, leading to much of the early modern academic literature on meteorology (Biswas, 1967).

From these collections of observations, a field of study was born. Discussion of trends and variability in precipitation from standardized observation networks were already discussed in the 1870s[1] (Baird, 1879). Today, a multitude of meteorological station networks exist globally, involving some tens of thousands of weather stations, ranging from automated amateur backyard devices to large-scale observatories with full-time staff meteorologists. In the United States this includes roughly 1,600 stations used for local long-term

[1] From Baird (1879), pg 122:

> The report of the New York Meteorological Observatory in Central Park, for the past year, by Daniel Draper, contains the usual tables from the self-recording instruments, and also a discussion on the two following questions:
>
> 1st. "Has there been in late years a change in the rainfall of New York City or its vicinity, affecting seriously its water supply?" It was found by the observations used that there had been an increase of rain until 1869, and after that year a steady decrease.
>
> 2d. "Does the rainfall of New York still diminish, will it continue to do so, and does this variation occur in the early or latter portion of the year?" By the observations used it appears that the rainfall of the city will most probably continue to decrease by fluctuations for several years to come, and that the variations are nearly the same in the two portions of the year.

climatological purposes, 10,000 stations used for short term weather forecasting, 1,000 used for local aviation conditions, 100 used as a high-caliber calibration reference, and 1,500 used for global-scale intercomparison (National Oceanic and Atmospheric Administration, a).

These networks have provided us with spatio-temporal views of the distribution of precipitation, as well as a means to establish the likelihood and magnitudes of extreme events (Easterling et al., 2000; Kunkel et al., 2003), ties between precipitation and global circulation metrics (Ropelewski and Halpert, 1986; McCabe and Dettinger, 1999; Barlow et al., 2001; Solomon et al., 2011; Becker et al., 2011), and changes in precipitation amounts and patterns on decadal and longer time-scales (Groisman and Easterling, 1994; Karl et al., 1995; Karl and Knight, 1998; Kunkel et al., 1999; Frich et al., 2002; Kunkel et al., 2003; Groisman and Knight, 2004).

Modern probabilistic estimation of future precipitation has been attempted since at least the 1870s in the US (Greely, 1891; Gore and Thomson, 1909; Ward, 1919; National Weather Service), following the centralization of reported weather conditions to the Smithsonian Institute by telegraph in 1849 and the founding of the US Weather Bureau in 1870 (National Weather Service), following the lead of colonial-era Britain (Bloxam, 1875; Raulin, 1875; Strachan, 1875; Trachan, 1877; Cox, 1878). Modern short-term forecasts use physical models of the atmosphere as well as statistical models (Glahn, 1982), and have steadily improved over time with innovation in scientific understanding, modeling capability, and computational resources (see Figure 1.1). Similarly, based on our understanding of how atmospheric circulation drives water vapor transport coupled with an understanding of how slowly-varying Earth system processes like sea surface temperatures can drive atmospheric circulation, we now use both physical and statistical models to estimate future precipitation regimes far beyond the range of daily-scale numerical weather prediction (Kirtman et al., 2013).

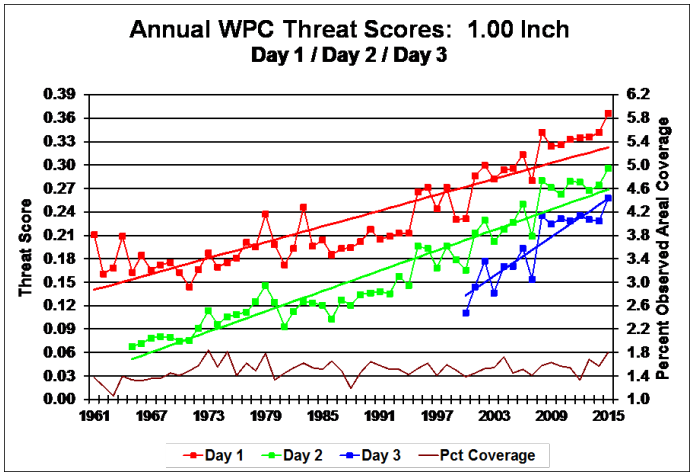

Figure 1.1: The equal threat skill — a metric of forecast quality — of the occurrence of one inch of precipitation or more over the United States from 1961–2015, as forecast by the National Oceanic and Atmospheric Administration's (NOAA) Weather Prediction Center (WPC; formerly the Hydrometeorological Prediction Center). Higher skill scores correspond to more accurate forecasts, and it is apparent that the WPC's forecast success has been consistently improving over time, to the point that present-day forecasts with three-day lead time are more accurate than forecasts with a single-day lead time in the 1960s–1980s. Reproduced from National Oceanic and Atmospheric Administration (b).

1.1.1 Modern lens and motivation

A large body of research has suggested broad changes in precipitation regimes in the future, due to geological and physical processes [volcanism (Ciais et al., 2013), ocean mixing (Aoki et al., 2013), weathering (Ciais et al., 2013), etc.], hydrological feedbacks from the terrestrial surface (Hartmann et al., 2013), and biological drivers of these hydrogeochemical systems (Ciais et al., 2013), all of which display variability across a very broad range of time-scales. As humans have large impacts on nearly all of these systems (Waters et al., 2016), there are additionally a multitude of anthropogenic drivers of changes in the hydrologic system (Myhre et al., 2013). These potential changes are significant in that water resource reliability has a direct impact on human livelihoods through social, health, and economic means (National Climate Assessment and Development Advisory Committee, 2013). Additionally, our ability to predict hydrologic changes drives our ability to mitigate or adapt to them on the time-scale of civil and political infrastructure development.

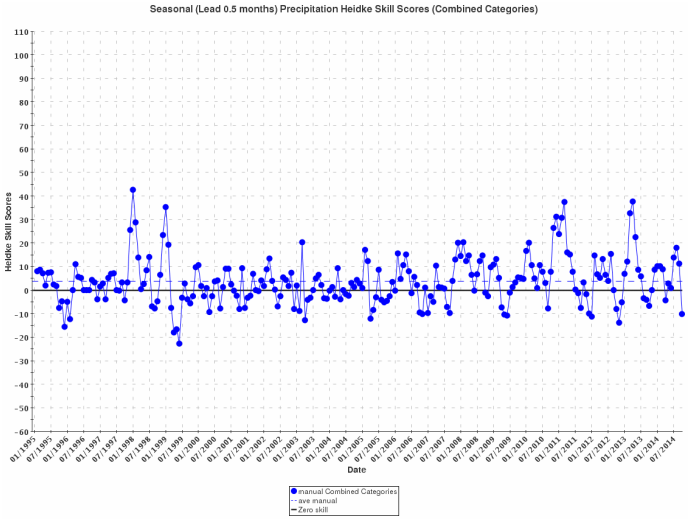

Figure 1.2: Heidke skill scores over time for seasonal precipitation with a 0.5 month lead from the NOAA NCEP Climate Prediction Center from 1995–2014. Reproduced from National Oceanic and Atmospheric Administration (c). Note the low skill (a skill of 0 corresponds to the skill of a random forecast) and lack of improvement over time as opposed to Figure 1.1.

Figure 1.2 shows the Heidke skill scores for seasonal precipitation forecasts (average/above-average/below-average) over the United States with a half-month lead time from 1995–2014 (National Oceanic and Atmospheric Administration, c). The skill scores are on average slightly positive, but nearly as often negative (100 corresponds to a perfect forecast, -50 corresponds to a perfectly incorrect forecast, and 0 corresponds to a random forecast), suggesting very little skill. Additionally, there seems to be no significant improvement in forecast skill over the last two decades, in contrast to the positive trends in Figure 1.1. This suggests a question for climate researchers:

- **How predictable is precipitation on climate time-scales?**

1.1.2 Potential predictability

While the state of the atmosphere is a great predictor for precipitation for the next hours-to-weeks, beyond that time-scale it is inherently internally unpredictable, and thus may as well be stochastic from a prediction perspective (Lorenz, 1973). However, there is still

some clear correlation between precipitation and the state of the broader Earth system at these time scales through slowly-varying climate modes, oscillations, and secular trends, which serve as boundary conditions on the atmosphere.

Our ability to predict those boundary conditions determines our ability to predict characteristics of precipitation at longer "climate-scale" lead-times[2]. Depending on how far in the future we are trying to predict, we may think of those as more or less deterministic versus stochastic.

Broadening our view a bit, we might ask how noisy precipitation is on time-scales beyond the "weather scale" of meteorological forecasts. Since the atmosphere is internally stochastic at climate-scale lead times, we can think of long-term prediction in a signal/ noise context. To re-state this idea, if we treated all slowly-varying sources of climate variability as deterministic and predictable, our predictions would still be constrained by the variability that comes from basic atmospheric dynamics and processes that lead to variability at shorter time scales. We will at times in this dissertation refer to this high-frequency (sub-seasonal) variability as "weather noise." The following chapters attempt to characterize this weather noise and quantify its magnitude relative to the total variability in precipitation at climate time scales. In each case, a weather model is constructed and fit to observational data, using the inherently flawed assumption that the background climate state is stationary. This flawed assumption allows us to estimate upper bounds on the amount of climate-scale variability that we would expect to see in precipitation data given a stationary climate. The degree to which our weather models fail to fully capture climate-scale variability will determine the *potential predictability* (PP) of the observed precipitation, the limit on our ability to predict future precipitation. These limits have not been fully quantified previously (DelSole and Tippett, 2007) and have even been proposed by some to be indeterminable (National Research Council, 2010).

[2]In some senses, the secular trend in energy stored in the atmosphere, land, and ocean due to GHG accumulation may end up being the *most* predictable of these drivers of variability (at least on the annual-to-decadal scale, and perhaps longer as well), and so much of the research effort into estimating changes in precipitation on those time scales has been focused on climate change studies (Allen and Ingram, 2002; Trenberth, 2011).

To determine the potential predictability, we will need to appropriately model our observed data using a weather model framework. This leads us to ask:

- **What is the magnitude of weather-scale variability in precipitation data?**
- **How can we best model daily precipitation in a stationary weather-scale context as a null model for climate variability?**

1.1.3 Weather generator models

Weather generator models are commonly used in the earth and environmental sciences as stochastic representations of meteorological variables of interest. The "classic" weather generator is the Richardson-type stochastic weather generator (Richardson, 1981), which models precipitation occurrence as a Markov-chain (most simply as a first-order Markov chain with constant transition probabilities throughout the year), intensity as independent draws from a positive-valued probability distribution (classically exponential, but often gamma), and further meteorological variables as conditioned on precipitation. The Markov representation of precipitation occurrence is significantly older than Richardson, though (Buishand, 1978; Lowry and Guthrie, 1968; Gabriel and Neumann, 1962), and the notion of stochastic precipitation models dates back at least to Quetelet[3] (Quetelet, 1852).

Markov chain models are common in precipitation studies due to their ability to represent the persistence (or auto-correlation, or "memory") in a series of observations, which was found to be necessary by at least the early 20th century (Gold, 1929; Cochran, 1938; Blumenstock, 1942). There has been a great deal of work further developing Markov chain-based representations of precipitation, significantly by Richard Katz (with numerous co-authors) from the 1970s–1990s (Katz, 1977, 1985; Katz and Parlange, 1998), including work developing the notion that simple stationary stochastic models of daily precipitation can properly represent interannual variability, but seemingly only at daily accumulation

[3]Quetelet could even be given credit for determining first-order Markov transition probabilities for daily precipitation occurrence, although he pre-dated Markov's work (Markov, 1906) by more than 50 years, so his probabilities are expressed in the language of geometric series (Quetelet, 1852).

scales. The "overdispersion" of observed precipitation at longer time-scales relative to these models suggested that the daily-scale models were not fully capturing all of the variability-generating processes, and that some additional lower-frequency processes were necessary for proper stochastic representation of precipitation on seasonal or longer time-scales (Katz and Parlange, 1993). This in turn lead to the notion of segregating high-frequency "weather-scale" processes (those which can be well-represented by simple stationary weather generator models) from low-frequency "climate-scale" processes (those which cannot), and determining the potential predictability of precipitation using these models (Katz and Parlange, 1998; Madden et al., 1999).

This overdispersion phenomenon hearkens back to numerous earlier empirical studies of hydrological phenomena with persistence on multiple time-scales, for example that of Hurst (1951) and Mandelbrot and Wallis (1968), both of which led to the development of the theoretically rich research areas of fractional Brownian motion (Mandelbrot, 1971; Koutsoyiannis, 2002), self-similarity theory (Lovejoy and Mandelbrot, 1985; Tyralis and Koutsoyiannis, 2011), and Lévy processes (Menabde and Sivapalan, 2000). Today, these ideas are commonly expressed in multiple methodological frameworks, particularly time-series analysis (Box-Jenkins and extensions), spectral theory (Molini et al., 2010), Bayesian hierarchical models (Cooley et al., 2007), and numerous machine learning paradigms (Hall et al., 1999; Kleiber et al., 2012), all of which are mainstays of the field of stochastic hydrology.

In 1999, Madden et al. used a daily one-lagged model for precipitation occurrence and an AR(1) model for precipitation intensity to determine potential predictability for a precipitation time series. In this work, simulations using the model in a weather-generator-type setting represented the variability that could be generated by high-frequency weather variations — unpredictable on climate time-scales, and thereby the "noise" component of a signal-to-noise ratio. The potential predictability was then determined as the fraction of variability which was *not* due to these weather variations. The work in this dissertation follows this path, but with methodological improvements to the representation of weather-

scale noise.

1.2 Outline of dissertation

The intent of this dissertation is to answer the following research questions:

- **How can we best model daily precipitation as a null model for climate variability studies?**
- **Can we quantify the magnitude of the stochastic component of precipitation on climate time-scales?**
- **How predictable is precipitation beyond the window of atmospherically-driven numerical weather predication?**

These questions will be answered through the methods and results of four linked studies:

Chapter 2 builds upon the methods of Madden et al. (1999) to select optimal stochastic representations of precipitation occurrence processes. These representations are then used to determine the magnitude of the stochastic component of precipitation occurrence across the contiguous 48 United States.

Chapter 3 uses the occurrence framework of Chapter 2 and proposes a complementary stochastic framework for precipitation intensity to represent the weather-scale variability of precipitation across the contiguous 48 United States. The potential predictability of precipitation determined from this method is then decomposed into predictability due to occurrence and intensity processes, demonstrating both constructive and destructive interference in predictability signals in precipitation totals.

Chapter 4 proposes a new methodological framework for representing persistence in precipitation time series to improve upon the separated occurrence /intensity paradigm of weather generator models using kernel density methods. This method is particularly desirable in its applicability to both station-based data (which is significantly

zero-inflated), and more coarse gridded or remotely-sensed data products (which tend to show little or no memory in occurrence processes).

Chapter 5 applies the methods of Chapter 4 to observational data as well as the output from corresponding historical runs from five Global Climate Models (GCMs). Weather-scale variability, climate-scale variability, and potential predictability are then compared between the data sets to determine whether GCMs accurately apportion variability between high-frequency and low-frequency processes.

Chapter 6 summarizes the findings of Chapters 2–5, and proposes future directions of research building upon these results.

Chapter 2

Modeling Precipitation Occurrence*

2.1 Abstract

We present a generalizable method for establishing the potential predictability for seasonal precipitation occurrence using rain gauge data. Using this method, we are able to provide an observationally-based upper limit for potential predictability for 774 weather stations in the contiguous 48 United States. We find that the potentially-predictable fraction varies seasonally and spatially, and that on average 30% of year-to-year seasonal variability is potentially explained by predictable climate processes. Potential predictability is generally highest in winter, appears to be enhanced by orography and land-surface coupling, and is lowest (stochastic variance is highest) along the Pacific coast. These results depict "hot" spots of climate variability, for use in guiding regional climate forecasting and in uncovering processes driving climate. Identified "cold" spots are equally useful in guiding future studies as predictable climate signals in these areas will likely be undetectable.

2.2 Introduction

It has long been known that the Earth's atmosphere is not entirely predictable in practical or theoretical terms (Lorenz, 1973; Leith, 1978), but many climate processes do display a predictability that we can exploit in creating forecasts, discovering trends, and uncovering mechanisms of climate variability (Lorenz, 1973; Leith, 1978; Kirtman and Schopf, 1998).

*This chapter is previously published as:
Daniel J. Gianotti, Bruce T. Anderson, and Guido D. Salvucci. What Do Rain Gauges Tell Us about the Limits of Precipitation Predictability? *Journal of Climate*, 26(15):5682–5688, aug 2013.

These forecasts of climate hinge on determining future probability distributions conditioned on the present and past (Palmer, 2006; National Research Council, 2010; DelSole and Tippett, 2007) and are based on uncovering variability due to low-frequency processes that have persistence beyond the chaotic time-scale of weather (Leith, 1978; Palmer, 2006; National Research Council, 2010). Many of these low-frequency processes (El Niño/Southern Oscillation, Madden-Julian Oscillation, long-term trends) are regularly used in forecasting climate variables on seasonal to decadal time scales (Palmer, 2006; National Research Council, 2010), while other potential predictors may remain undiscovered and unutilized.

The chaotic nature of atmospheric dynamics limits the predictability of climate variables, although the limits have not been quantified (DelSole and Tippett, 2007) and have even been proposed to be indeterminable (National Research Council, 2010). In this paper we estimate these limits — termed the potential predictability (PP) — for a single climate variable, seasonal precipitation occurrence, with the method generalizable to other variables.

Precipitation occurrence is of interest for a number of reasons, in particular (a) variability in the number of rain events is a driver of extreme hydrological events (droughts and flooding) (Sun et al., 2006), (b) the major climate trends in US precipitation have been shown to be due to an increase (or decrease) in the number of rain events as much as by major changes in precipitation intensity (Karl and Knight, 1998), and (c) our tools for decadal and century-scale forecasting (coupled global climate models) have had difficulty representing the actual variability of precipitation occurrence (the "constant drizzle" problem (Sun et al., 2006; Dai, 2006)).

Ferguson et al. (2011) have shown that — at least for general circulation model ensemble data — the slowly-varying, boundary-induced precipitation signal is generally non-stationary over the 20^{th} century, while the stochastic component of precipitation variability is stationary. This suggests the idea of separating variance due to short-memory (high-frequency) processes and longer (low-frequency) oscillations associated with potentially predictable processes. Because of this, if we model precipitation as an inter-annually-sta-

tionary process, we would expect to represent the high-frequency ("stochastic") variability well, but not represent any of the low-frequency (potentially predictable) variability; differences between the (modeled) stochastic variability and the (observed) total variability should give an estimate of the potentially predictable variability (Katz and Parlange, 1998; Madden and Shea, 1978; Wang et al., 2006).

2.3 Methods

One means of determining an upper bound for the predictability of a particular variable is to establish what fraction of the variability is due to long-term versus short-term processes (Katz and Parlange, 1998). Here we establish this fraction using variable-complexity stochastic models trained on the historic record to simulate precipitation occurrence (Katz, 1977). Although not all low-frequency variability in the historic record is predictable, we can estimate what portion of the signal is attributable to (non-predictable) short-term dynamics through comparison with an entirely high-frequency data generating process. Our models only incorporate short memory information (and match the observed short timescale statistics identically), so that no low-frequency (climate) variability is introduced. To determine the appropriate complexity of the models, we use the corrected Akaike Information Criterion (AIC_c) to optimize the trade-off between goodness-of-fit and parsimony/complexity (Hurvich and Tsai, 1989). We then use the models to simulate 1,000 realizations of each station's roughly 100-year historic record, which form a distribution of purely stochastic simulated precipitation records for a given station.

To compare the observed record with the simulated records, we use the number of days of precipitation within a centered seasonal (89-day) window. Although any window could be selected, the seasonal window establishes a predictability metric of appropriate scale for use in assessing local droughts (Ferguson et al., 2011), agricultural growing periods (D'Odorico and Ridolfi, 2000), or winter snowpack (Bohr and Aguado, 2001) and still allows us some insight into the seasonal cycle of PP throughout the year.

Potential predictability of observed precipitation frequency at a given station is defined as the portion of variability that cannot be explained by the stochastic evolution of daily precipitation events — an unexplained variance — usually as a normalized ratio. Although sometimes expressed as a signal-to-noise ratio (Ferguson et al., 2011; Leith, 1978) or a signal-to-forecast-error ratio (Somerville, 1987), we use the unexplained variance fraction definition (Boer, 2004; Rowell, 1998):

$$\mathrm{PP}_{89} = \frac{\sigma^2_{89,obs} - \left\langle \sigma^2_{89,sim} \right\rangle}{\sigma^2_{89,obs}} \tag{2.1}$$

where PP_{89} is the potential predictability for a seasonal 89-day window centered on calendar day t, $\sigma^2_{89,obs}$ is the variance (between years) in the number of rainy days within an 89-day window of t (± 44) in the observed record, and $\left\langle \sigma^2_{89,sim} \right\rangle$ is the mean variance (between years) in the number of rainy days within an 89-day window of t in the 1,000 simulated records. The numerator represents the variance not explained by the stationary, stochastic model and is expressed as a fraction of the total (observed) variance from the historical record. Positive values of PP_{89} correspond to values of $\sigma^2_{89,obs}$ in the positive tail of the $\sigma^2_{89,sim}$ distribution; the higher the value of PP_{89}, the less likely it is that the observed variance is due the integration of random short-memory processes (i.e. is due to chance).

Our models and analysis are based on the United States Historical Climatology Network (USHCN) data set (Williams et al., 2006) using daily precipitation values through December 31, 2009. Of the 1,218 available weather stations in the USHCN, a subgroup of 774 stations is selected to meet longevity and completeness criteria (see Supplementary Materials). All stations are required to have at least 80 years of continuous data (Jan 1, 1930–Dec 31, 2009) and cannot have more than 5% of daily values missing. The longest period ending on Dec 31, 2009 with at least 95% complete data is selected. The threshold for precipitation occurrence is 0.254 mm (0.01 in); climatology of precipitation occurrence from these data is shown in Figure 2.1a–d.

Markov chain models are well-suited for stochastic precipitation occurrence modeling

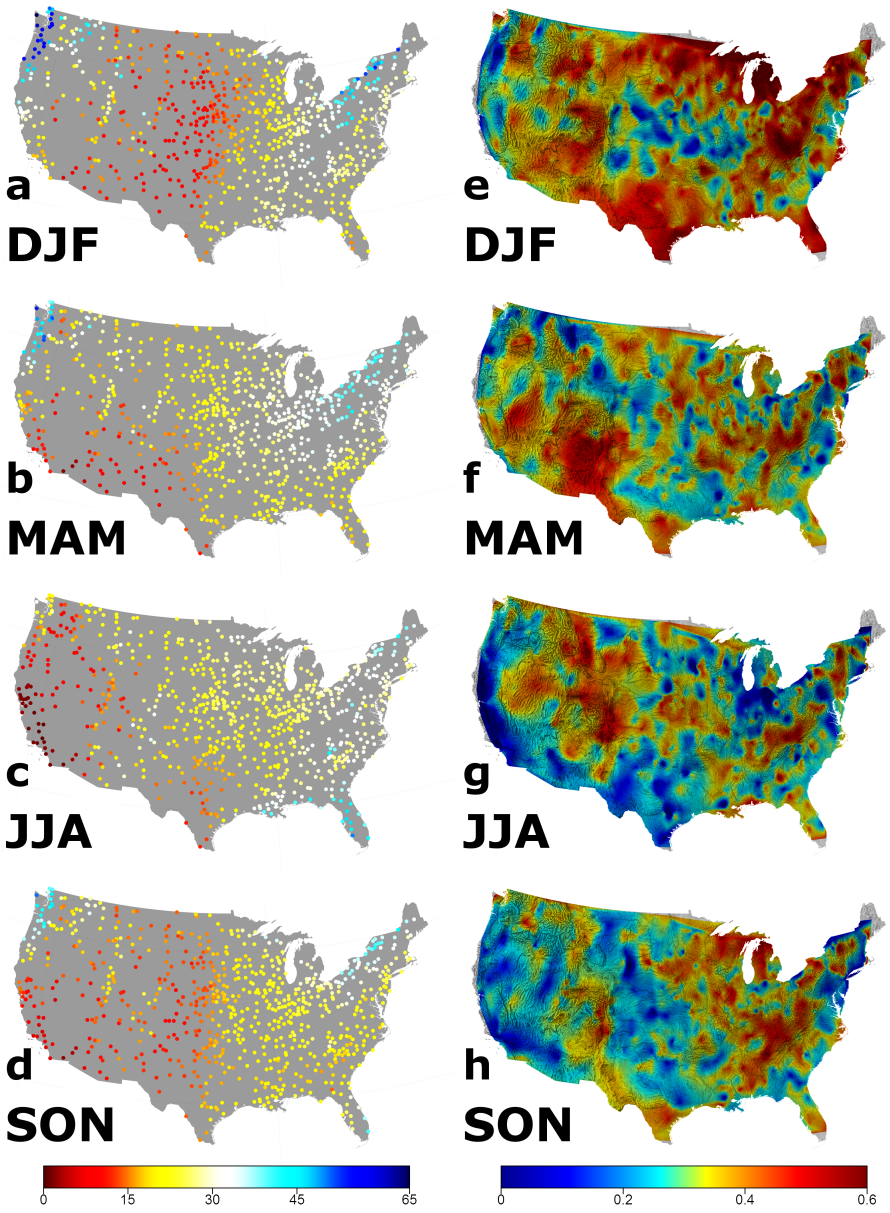

Figure 2.1: Climatology and potential predictability (PP) of precipitation occurrence. (a)–(d) show the mean number of days with precipitation (> 0.254 mm or 0.01 in) for 89-day seasonal windows centered on Jan 15, Apr 15, Jul 15, and Oct 15, respectively. Circles show the location of the 774 weather stations used in the study. (e)–(h) show PP as calculated by $(\sigma^2_{89,obs} - \langle\sigma^2_{89,sim}\rangle)/\sigma^2_{89,obs}$. The Appalachians, southern Rockies, and Great Lakes regions show the highest PP, while high wintertime stochastic variability in the Northwest and Pacific coast lead to lower-than-national-average PP. See Figure A.4 for timings of minimum and maximum PP and Figure A.6 for spatial significance.

(Katz, 1977), and model complexity can be selected by allowing for a variable chain order, m. Each state of the Markov chain is represented by the day of the year and occurrence pattern over the previous m days (Wilks and Wilby, 1999). The transition probabilities — the probabilities of precipitation on a certain day given the last m days' history — are found empirically from the historic record for each possible occurrence pattern on each day of the year.

Our method for determining seasonal PP values is shown schematically in Figure A.1. Models are created to match the empirical historic record for each of 774 selected weather stations across the US; for each station, the order m is established by using the AIC_c (see Supplementary Materials) and is allowed to vary from day-to-day (Hurvich and Tsai, 1989; Katz and Parlange, 1993). The AIC_c is used because it allows for an un-biased optimization of the trade-off between model complexity/parsimony (using a higher order in determining the transition probabilities is more complex and introduces more variability) and model likelihood (higher order models increase likelihood-based fit). We also allow pooling of our data set (using multiple neighboring days' observations for fitting model transition probabilities — see Supplementary Materials), again using the AIC_c to determine appropriate pooling windows. While daily chain orders are selected and used for model fitting and simulation, the mean chain order of each station for each season is shown in Supplementary Figure A.2.

Once the pooling size, daily chain orders, and transition probabilities are determined, the stochastic model for that station is complete, and simulations can be created where each day's occurrence is determined probabilistically based on the previous m days' occurrence history. We simulate 1,000 realizations of each station's roughly 100-year historic record ($\sim 3 \times 10^{10}$ individual days of precipitation occurrence in total) which then form a distribution of purely stochastic simulated precipitation records for a given station.

To analyze the spatial patterns of predictability, we grid the $\mathrm{PP}_{89}(t)$ vector ($t \in \{1, 2, \ldots, 365\}$) at 0.0775° resolution using natural neighbor interpolation (Sibson, 1981). To create clusters we use the k-means algorithm (MacQueen, 1967) with an L2 norm

(squared Euclidean distance) to group the clusters so that moving any one grid point to another cluster increases the mean intra-cluster distance.

2.4 Results

The maps shown in Figure 2.11e–h indicate areas of high and low PP by season, with considerable variability observed in both magnitude and spatial pattern throughout the year. This figure is intended to serve as a "road map" for areas with high and low likelihoods for success in interannual-to-multidecadal-scale modeling of precipitation occurrence, corresponding to high and low PP values, respectively.

In all of the Figure 2.1e–h maps there are clear correlations between high PP and orographic features, most notably in the southern Rockies and Appalachians. This suggests that orography may play a role in amplifying not just spatial but temporal precipitation variability (and thereby PP), presumably by amplifying temporal variations in local water vapor convergence (storm tracks and convection). The Pacific coast is noteworthy in its relatively low PP, due primarily to very low observed variability in the summer and fall (Figure 2.1g–h) and relatively high stochastic variability in the fall through spring (Figure 2.1e,g–h, Figure A.3). The mean PP value across all stations and days of the year is 0.3 (30% of year-to-year seasonal variability is potentially explained by predictable climate processes), although at some stations the seasonal mean PP exceeds 0.7. All PP values above 0.24 are significant at the 0.05 confidence level; those below 0.24 are either significant (but low) or indistinguishable from zero (see Supplementary Materials).

It is important to note that low PP does not necessarily mean that predictable processes have little influence in a region. Rather, it suggests that the predictable portion of the interannual variance in number of rainy days in a season is small relative to the high-frequency-driven stochastic variability that the area experiences.

To be able to attribute climate processes to their effects on specific climate variables, we hope to find common regional, seasonal behavior patterns. To this end, we performed

a k-means cluster analysis based on an evenly-gridded set of PP values and plotted the resulting regional patterns (Figure 2.2). Although implicitly dependent on spatial patterns to some degree through the gridding process, the clustering is otherwise independent of location and based only on similarities in the seasonal pattern of PP. The overall 774 station mean pattern shows higher predictability in the late fall and winter, and lower predictability in the summer and early fall, confirming an earlier result by Rowell (1998).

Cluster 1 has higher than spatial-mean predictability in summer, lower in fall and winter, which can also be thought of as a "flattening" of the seasonal predictability cycle. This group of stations has the highest relative summer PP, which agrees well with previous studies on the strong JJA soil-moisture convective precipitation feedbacks in the western Great Plains and Ozarks/Gulf Coast (Koster et al., 2003), driven, at least in part, by sea surface temperature (SST) variability and atmospheric water vapor transport from the Gulf of Mexico (Wang et al., 2010). In addition, Koster et al. (2004) suggest that this feedback — in which local soil moisture enhances local evaporation and precipitation — should happen in transition areas from wet-to-dry climates, further supporting a soil moisture explanation for enhanced summertime predictability in the western portion of Cluster 1.

Cluster 2 displays an amplification of the spatial average predictability cycle, with high late-fall and winter PP. The generally high PP in the Great Lakes and Finger Lakes regions suggests a PP amplification during the lake-effect snow period from October through mid-March (Niziol et al., 1995), with the available water supply likely amplifying the predictability of temporal patterns in atmospheric water vapor. Although the mechanisms are likely different, Texas and Florida have similar seasonal patterns to the Great Lakes region due to their strong correlation with tropical Pacific SSTs from December through March (Boer, 2004; Ropelewski and Halpert, 1986).

Cluster 3 — located mostly in the Southwest — has a seasonal predictability cycle that generally follows the spatial average with the exception of late winter, when predictability is significantly higher. Although coastal southern California has, on average throughout

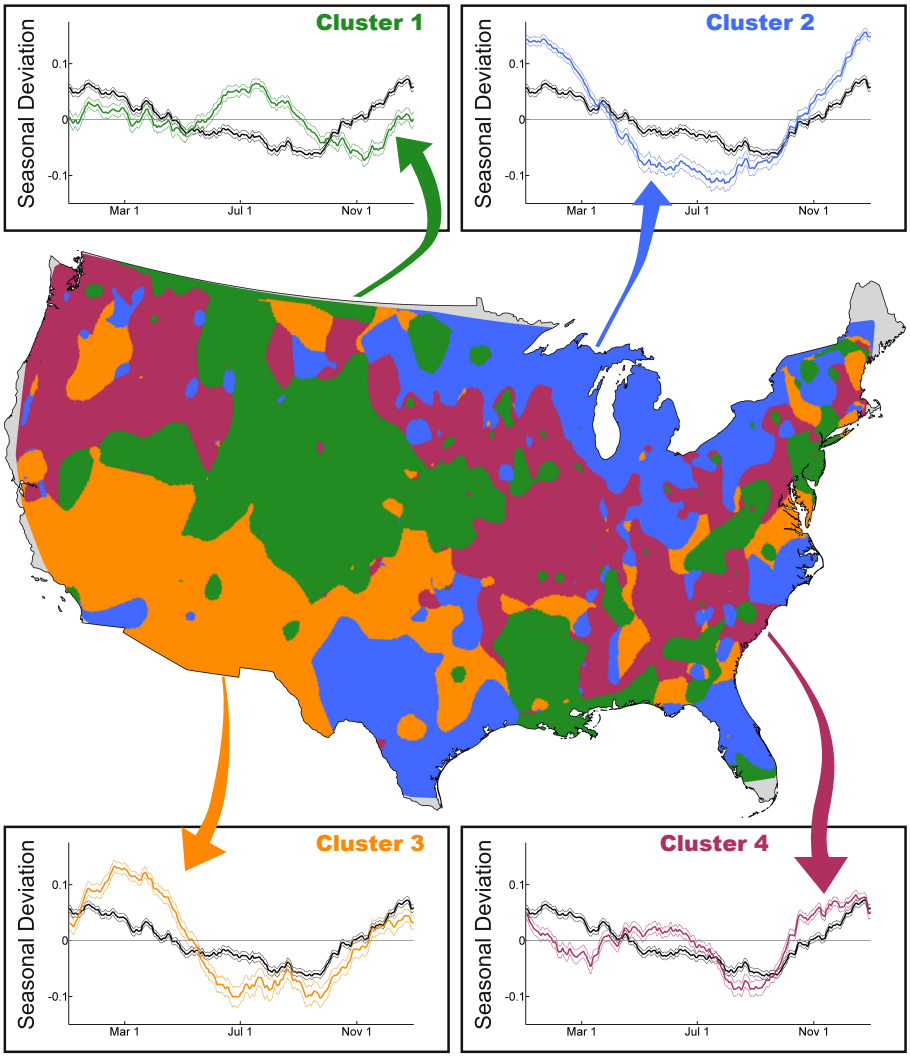

Figure 2.2: Cluster analysis based on potential predictability seasonal time series. Regions determined by an interpolated L2-norm clustering algorithm performed on the seasonal predictability curves. Subplots show the mean seasonal anomaly curve ±1.96 SE over all 774 stations (black) and the mean anomaly curve ±1.96 SE for the cluster (color). Clusters were determined using k-means on a gridded PP field, and cluster curves are the mean of the station values in that region (rather than the mean of the gridded values).

the year, a much lower predictability than the Colorado Plateau to the East (Figure 2.2), the region as a whole lies along the ENSO precipitation anomaly track (Cayan et al., 1999), and as such the predictability peaks for stations in this region are representative of the typical late-winter ENSO precipitation timings (McCabe and Dettinger, 1999).

Cluster 4 is best-described as a flattening of the mean PP cycle from January through June with a PP increase in October and November. Although wintertime precipitation in the Pacific Northwest has been shown to be strongly affected by ENSO variability, we find that the predictable variability in precipitation occurrence is minimal compared to winter in the Southwest, in agreement with previous results (Feldl and Roe, 2011). As previously mentioned, this does not mean that ENSO does not affect Northwest precipitation occurrence, but that the predictable fraction of the variance is small relative to the stochastic component (which is in fact higher in the Northwest than any other region — see Figure A.3).

2.5 Summary

All PP is due to the variability of processes leading to precipitation, and the timing/ magnitude of the PP response is a function of interactions between PP drivers (SSTs (Cayan et al., 1998), anthropogenic forcing (Zhang et al., 2007)), feedbacks (soil-moisture (Koster et al., 2003, 2004); orography), and unpredictable dynamics. Rigorous attribution of PP-related precipitation processes to oscillations, trends, and feedbacks is likely best targeted towards areas of high PP — the Appalachians, Rockies, and Great Lakes, for instance — using Figure 2.1 as a guide for region selection and Figure 2.2 as a guide for timings and likely processes. Further, the method for determining PP discussed here is algorithmic and generalizable to climate variables other than precipitation occurrence, which suggests a high-level approach to climate forecasting: (a) establish the magnitudes of predictable and stochastic variances to determine PP, (b) look for processes that drive variability in high PP regions to close the gap between current model skill and PP limits,

(c) use the findings from regional process modeling to update global models and forecasts, and (d) monitor areas for changes in PP over time that might signify changes in predictable climate processes.

Chapter 3

The potential predictability of precipitation occurrence, intensity, and seasonal totals over the continental United States*

3.1 Introduction

Precipitation data is very noisy due to many non-linear processes and thresholds in both the physics of precipitation and our ability to measure it (Mandelbrot and Wallis, 1968; Pruppacher and Klett, 1997). At our common measurement time-scales (e.g., daily), spatially-based, physical process models dominate our precipitation forecasting methodologies (Lynch, 2008), whereas the chaotic behavior of the atmosphere beyond a few weeks renders these methods nearly useless for forecasts at seasonal scales or longer (Lorenz, 1965). In time-series analyses of precipitation point-data, the perceived deterministic time-scale switches; there is significant persistence in aggregated seasonal and annual totals (particularly when tied to slowly changing atmospheric indices, Katz and Parlange, 1998), but daily data are generally considered random variables due to low daily persistence (Katz and Parlange, 1998) as well as the complex spatial behavior of precipitation (Olsson and Niemczynowicz, 1996; Gong et al., 2003).

As we move from daily data to longer aggregated time-scales, we see that precipitation data becomes less chaotic and less stationary (Mandelbrot and Wallis, 1968) — that is,

*This chapter is previously published as:
Daniel J. Short Gianotti, Bruce T. Anderson, and Guido D. Salvucci. The Potential Predictability of Precipitation Occurrence, Intensity, and Seasonal Totals over the Continental United States. *Journal of Climate*, 27(18):6904–6918, 2014.

the probability distributions for precipitation random variables are not the same from one month, season, or year to the next. This is perhaps most easily envisioned as the superposition of multiple precipitation patterns at multiple time-scales: diurnal evaporation and moisture convergence (Meisner and Arkin, 1987), the annual cycle (Meisner and Arkin, 1987), correlations with intra-seasonal to multi-annual climate oscillations [e.g. Madden-Julian Oscillation (Becker et al., 2011), El Niño/Southern Oscillation (ENSO, Ropelewski and Halpert, 1986), Arctic Oscillation/North Atlantic Oscillation (AO/NAO, Hurrell, 1995), and Pacific Decadal Oscillation (PDO, Mantua and Hare, 2002)], and empirical trends over the course of our data record (Karl and Knight, 1998). This non-stationarity is what fundamentally allows us to make successful predictions of precipitation despite the chaotic regime of atmospheric dynamics, as evidenced by the typically positive forecast skill scores on seasonal time-scales (National Oceanic and Atmospheric Administration Climate Prediction Center, 2013).

We are interested in establishing the degree to which precipitation is potentially predictable at these climatic time-scales [i.e. longer than the chaotic noise time-scale on the order of days to weeks (Lorenz, 1965; Leith, 1978)]. Throughout this paper, we will use the terms "high-frequency" and "weather noise" to refer to processes with a time-scale on the order of days, and "low-frequency" and "climate signal" to refer to processes — other than the annual cycle — with time-scales longer than the weather noise. The characterization of high-frequency processes as noise rather than signal is simply due to this study's focus on climate-scale processes.

Potential predictability (PP) can be thought of as an upper-bound on forecast skill or a partitioning of a time-series into signal and noise. In the case of this study, the noise is due to short-persistence (on the order of days) processes in the Earth system, and the signal is due to a number of forcings, feedbacks, and slowly-changing boundary conditions that lead to persistence beyond our noise time-scale. This upper-bound can provide us with useful information, including 1) the broad limits of our ability to confidently forecast longer-term climate behaviors, 2) where and when our present forecasts are limited by an

understanding of the system versus an absence of useful climate signals, and 3) a metric to test whether or not our current forecast models display signals and noise in the same manner as the real climate system.

The first point is particularly critical as unexpected droughts and floods damage infrastructure in US towns, cities, and farmlands; and water resource reliability impacts human livelihoods through social, health, and economic means (National Climate Assessment and Development Advisory Committee, 2013). It is important to note that these impacts are as much an issue of the predictability of precipitation as its magnitude. For example, the ability of modern human infrastructure to provide for quite similar lives in the deserts of Arizona and the rain-forests of Washington relies on a proper use of climatological data to predict future conditions. This prediction is obviously less effective if stationarity is improperly assumed, or if the data used to make predictions is insufficient for the time-scale of infrastructure planning (Milly et al., 2008). Beyond the direct human impacts, increased precipitation variability has been found to affect ecosystem services — for example through changes in plant community composition and ecosystem carbon cycling processes (Knapp et al., 2002) — and our ability to mitigate or adapt to these impacts are limited by our ability to predict them.

Since day-to-day precipitation is unpredictable on climatic time-scales (i.e, beyond the weather noise time-scale of days to weeks), we will think of it as a stochastic process with some amount of intrinsic variability. From the forecasting perspective, we might equate precipitation variability with uncertainty and think that large variances must be related to low predictability. While it is certainly true that stochastic weather variability increases uncertainty, any climate signal — which we can think of as the variability of the weather noise distribution through time (e.g., Mandelbrot and Wallis, 1968) — also increases precipitation variability. From a climate perspective, precipitation variance is a combination of (stochastic) weather noise and (potentially predictable) climate signal. If we think of separating our precipitation variance into a weather component and a climate component, the latter forms the basis for our climatic potential predictability, and large values of climate

variability relative to weather variability suggest high PP. As an example, if the statistics of precipitation were stationary from year to year, there could be interannual variability due to weather; however, there would be no climate-driven variability and thus no opportunity for prediction based on climate indicators (sea surface temperatures, atmospheric modes, land-atmosphere coupling). The degree of non-stationarity in the statistics of precipitation is thus one measure of potential predictability. In this paper, we will specifically assess how well (or poorly) stationary statistics can describe the interannual variability of precipitation.

Potentially predictable climate signals can manifest as deviations from stationary precipitation characteristics spanning years to centuries. There has been considerable previous investigation into precipitation variability on these time-scales, much of which focuses on partitioning natural variability and trend components. For instance, mean daily precipitation increased over the contiguous U.S. by roughly 10% over the twentieth century (Karl and Knight, 1998). Precipitation occurrence in the contiguous U.S. has generally increased over the last century as well (Karl and Knight, 1998; Groisman et al., 1999; Higgins and Kousky, 2013; Pal et al., 2013), with some locations showing an additional 15 days per year of precipitation over the last half of the twentieth century (Higgins et al., 2007). At the same time, dry spell lengths have also increased, leading to more wet days in wet periods and longer droughts in dry periods (Groisman and Knight, 2008; Pal et al., 2013), as well as possible shifts in the timing of the seasonal precipitation cycle (Pryor and Schoof, 2008; Pal et al., 2013). Roughly half of the increase in the number of wet days has been attributed to the PDO and AO (Higgins et al., 2007); and modeling studies have linked up to 40% of the variability in annual precipitation in the U.S. Southeast, Southwest, and Mid-Atlantic regions to sea-surface temperatures (SSTs), and found significant amplifications of variance in the Central Plains due to land-atmosphere feedbacks (Koster et al., 2000).

Much of the increase in twentieth century precipitation is due to increases at the extreme high end of the distribution (Karl and Knight, 1998). Models suggest that heavy precipitation should increase with trends in increased atmospheric temperatures, perhaps

with an increase in dry days as well (Groisman and Knight, 2005; Trenberth et al., 2003), and that Probable Maximum Precipitation should increase following the Clausius-Clapeyron relation (Kunkel et al., 2013; Sun et al., 2007). Using maximum daily precipitation, this increase in heavy precipitation has been found to hold globally, but U.S. precipitation has shown a less coherent relationship (Westra et al., 2013). Changes in heavy precipitation have been tied to increased convective thunderstorm activity (except in the Great Lakes and Mid-Atlantic — Changnon, 2001), extratropical cyclones (Great Lakes and Northeast — Kunkel et al., 2012), frontal systems (Central Plains — Kunkel et al., 2012), and the influence of tropical cyclones (Southeast and Cornbelt — Knight and Davis, 2009; Kunkel et al., 2012). Much of the increase in heavy precipitation in the U.S. appears to have occurred after 1970 (Groisman and Knight, 2005), although intensities of similar magnitudes were observed at the same frequency around the turn of the twentieth century, suggesting a background of significant natural climate variability (Kunkel et al., 2003). As part of our analysis in this paper, we will assess the degree to which these observed changes in both intensity and occurrence processes can be said to contribute to precipitation's variability and potential predictability.

Given its importance, significant research has addressed the issue of isolating predictable characteristics of climate-scale precipitation variance, using both process models (Koster et al., 2004; Moron et al., 2006; Jia and DelSole, 2011) and observational approaches (Madden et al., 1999; Anderson et al., 2009; Robertson et al., 2009; Gianotti et al., 2013). PP is generally established from signal-to-noise ratios (Ferguson et al., 2011), variance partitioning (Koster et al., 2004; Jia and DelSole, 2011; Boer, 2009), or through the divergence of states from initial model conditions (Collins, 2002). Precipitation is found to have lower typical PP than air temperatures (Collins, 2002; Boer, 2009; Feng et al., 2011) — often less than 30% of the total variance is deemed potentially predictable (Koster et al., 2000; Boer, 2009; Westra and Sharma, 2010; Feng et al., 2011) — and the processes driving PP in a certain location tend to be dominated by either ocean processes or land-feedbacks (Koster et al., 2000). Prior work has separated the role of precipitation occurrence and intensity,

and occurrence has been found to demonstrate more potential predictability than intensity in the tropics (Moron et al., 2007; Robertson et al., 2009) and Chile (Verbist et al., 2010), perhaps due to the stronger spatial coherence of occurrence processes (Moron et al., 2006).

In this study, we assess PP for U.S. seasonal precipitation totals using purely observation-based data as opposed to modeling studies which may have artifacts due to model specification. Studies of the U.S. have found higher PP in the West (Yilmaz and DelSole, 2010; Feng et al., 2011), with possible feedbacks in the Central Plains (Koster et al., 2004) using gridded precipitation data. Using station data, Gianotti et al. (2013) found the seasonal PP for rain gauge precipitation occurrence to be somewhat higher than other studies found for total precipitation, 30% on average across the U.S., with some locations higher than 70% in certain seasons.

The primary components of this study and this paper are 1) a strategy for creating "Stationary Stochastic Weather Models" to model precipitation under the assumption of stationary weather statistics (no climate variability), 2) the decomposition of precipitation variability into intensity and occurrence components, 3) simulation of precipitation data using the SSWMs as a null hypothesis of no potentially predictable climate signal against which to compare the observed precipitation data, 4) calculation of potential predictability from the observed data, and 5) discussion of spatial features of the calculated potential predictability values. Section 2 presents an overview of the data and statistical methods employed to quantify the potential predictability of climate-induced variations in seasonal-mean precipitation amounts, occurrences, and intensities across the U.S. The temporal and spatial characteristics of potentially predictable signals are then presented in Sections 3 and 4. Discussion of the results and implications for understanding the role climatic factors versus short-persistence processes play in generating these events follow in Section 5.

3.2 Methods

3.2.1 Stochastic Modeling

To establish potential predictability, we compare the interannual variability of observed data with the variability of annually-stationary models designed to represent the observed data as well as possible within a stationary framework. We will refer to these models as "Stationary Stochastic Weather Models" (SSWMs) to emphasize their use in simulating artificial weather data with no interannual variability aside from that created by "weather" processes with persistence on the order of days.

Daily precipitation has been modeled in the past, often as a chain-based process with some assumed parametric or non-parametric intensity distribution (Stern and Coe, 1984; Katz, 1977; Wilks, 1999; Wang et al., 2006; Robertson et al., 2006; Schoof and Pryor, 2008; Hanson and Vogel, 2008). Although generally fit to observational data, the simplest models are prone to under-representation of observed interannual variability (Gregory et al., 1993; Katz and Parlange, 1998). To the degree that this is due to misrepresentation of high-frequency variance, it can be corrected by using higher chain-order models and/or more complex intensity distributions.

When considering occurrence, we can envision a single probability distribution for intensity with a large density at zero to account for dry days (Salvucci and Song, 2000; Robertson et al., 2009), or we can think of it as two processes: an occurrence process and an intensity process, which will not necessarily be statistically independent. This separation is potentially useful in cases where some physical processes may be tied more directly to either occurrence or intensity — for example if soil moisture is connected to occurrence through convective triggering or SSTs are connected to intensity through synoptic convergence.

To construct our SSWMs, we use seasonally-varying variable-order Markov chains to model daily precipitation occurrence and continuous parametric density functions (conditioned on the chain of occurrence patterns) to model intensity. To properly represent

the high-frequency variability, we allow the occurrence models to vary their complexity via their chain-order and use the corrected Akaike Information Criterion (AIC_c; Hurvich and Tsai, 1989) to select the appropriate order to optimize the goodness-of-fit/parsimony trade-off.

Our observational precipitation data is drawn from the U.S. Historical Climate Network (USHCN) dataset beginning on January 1, 1900 and ending on December 31, 2009 (Williams et al., 2006). We focus on station-based data in order to ensure that results are not biased by interpolation algorithms, scale selection, changes in station locations, or coverage gaps as they may be with gridded data (Hamlet and Lettenmaier, 2005). From the 1218 available weather USHCN stations, 774 are selected that have at least 80 years of data with less than 5% of precipitation values missing — the longest period meeting these criteria is used.

Our occurrence models are variable-order Markov chains with daily-varying transition probabilities and chain orders (see Gianotti et al., 2013). At each station, a data-pooling window is selected using the AIC_c to improve transition probability estimates. Pooling sizes for occurrence are typically between 15 and 35 days, and all days within that window (across all years) are used in estimating the transition probabilities for the central day's occurrence model. For each day of the year, parameters are fit to the empirical data for chain orders from zero to five days, and the appropriate model for each day is selected, again using the AIC_c. Pooling sizes are fixed for a given station, while chain orders can vary from day to day. All combinations of pooling sizes from 1 to 91 days are tested using all chain orders from 0 to 5 days for each day of the year. The AIC_c is used over the similar Bayesian Information Criterion (Schwarz, 1978) because the AIC_c has a weaker penalty for overfitting, potentially allowing more variability in the models, which leads to more conservative confidence intervals and PP estimates.

As an example, in Ajo, AZ (pooling size 17 days) on June 18, the optimal chain order is zero. This means that the modeled probability of precipitation for June 18 is just the climatological frequency of occurrence in the pooling window (June 10 – June 26) around

June 18. In contrast, in Selma, AL (pooling size 23 days) on June 18, the selected chain order is two, meaning that there are four different modeled probabilities of precipitation occurrence depending on whether it rained on both the 16th and 17th, neither the 16th nor 17th, on just the 16th, or on just the 17th. Each of those four probabilities is determined from the climatological frequency of occurrence following the matching two-day history for all days in the observed record within the pooling window of June 7 – June 29.

Our modeled intensities are drawn from parameterized probability distributions fit to the first two moments of the observed data. To allow for bi-modal precipitation distributions with always-finite variances, the intensity models take the form of a five-parameter gamma-gamma probability distribution mixture model:

$$f(x) = \lambda \frac{\beta_1^{\alpha_1} x^{\alpha_1 - 1} e^{-\alpha_1 \beta_1}}{\Gamma(\alpha_1)} + (1 - \lambda) \frac{\beta_2^{\alpha_2} x^{\alpha_2 - 1} e^{-\alpha_2 \beta_2}}{\Gamma(\alpha_2)}, \tag{3.1}$$

where $0 \leq \lambda \leq 1$, α_1 and α_2 are shape parameters, β_1 and β_2 are rate parameters, and $\Gamma(\cdot)$ is the gamma function. The gamma-gamma distribution was selected from among a number of 1–5 parameter models (exponential, gamma, Dagum, Weibull, kappa, log-normal, and mixtures of each) by the AIC_c as the best trade-off between parsimony and goodness of fit. As with the occurrence process, a model is constructed for each combination of station, day of year, pooling size, and chain order (conditioning on the previous m days' occurrence, rather than intensity), the data pooling size is selected using AIC_c, and then the chain order is selected also using AIC_c. This process is fairly computationally intensive, requiring over 100 million non-linear five-dimensional constrained optimization solutions, and yields an intensity model for each day of the year at each station. The 365 daily models for occurrence and intensity, each of which is interannually stationary, combine to make our SSWM for a given location. Selected chain orders for occurrence are generally zero to three days, while they are almost exclusively zero (no memory) for intensity. This suggests that, given 100 years of observational data — even allowing data pooling (effectively increasing observations by an order of magnitude) for more robust parameterization — intensity

estimates are not substantially improved by conditioning on the past few days' occurrence patterns.

We then create stochastically-generated Monte Carlo "weather-only" simulations of the historical record using the SSWMs for each location. Simulations are run a single day at a time to make use of the short-term memory of the chain models — for each day, the probability of occurrence and the probability distribution for intensity are determined by the current model state, a function of day of year and the last 0–5 days of simulated data. For each station, 1000 simulations of the 80–110 year record are produced. These simulations then form a distribution of precipitation records which match the observed record in length, and — for each day of the year — match the probability of precipitation occurrence, the likelihood of precipitation given the past few days of occurrence, and the intensity distribution. The important way in which they do not match the observed record is that — despite being stochastically-generated and therefore displaying interannual variability — the SSWM data-generating mechanism is stationary from year-to-year, and therefore fails to incorporate low-frequency processes that might affect the occurrence/intensity distributions like those driven by SSTs, atmospheric modes, trends in greenhouse gas concentrations, and land-cover changes. Any long-term aggregation of the SSWM simulations, then, should generally display lower variability than the observed records upon which they are based.

More importantly for our purposes, variability in the observed record that extends beyond the window of likely variabilities in the SSWM simulations suggests the existence of low-frequency (climate) variability. While not all low-frequency variability is deterministically predictable, this difference between the variability of the observed record and the variability of the distribution of simulated SSWM records is the basis of our estimate for potential predictability. For example, if in a given location the interannual variance of the observed autumn precipitation was equal to the interannual variance of autumn precipitation in the SSWM, then the observed variance can be explained entirely by a high-frequency noise process (our SSWM), and there is no separable low-frequency signal, hence no PP.

To account for missing data, we simulate missing days using the SSWM for that location. This is more conservative than either imputing missing values with re-sampled observations or omitting the data entirely (the SSWMs have no persistence beyond five days, and thus any precipitation sequences [longer than five days] that incorporate SSWM data are under-dispersed relative to the actual conditions, and will therefore tend to moderately underestimate the PP signal).

Since the observations and simulations are at daily resolution and we recreate the seasonal cycle by using 365 occurrence/intensity models to comprise the SSWM for each station, we can look at interannual variability for any period greater than or equal to a single day. For this analysis, however, we will focus on the variance and PP of seasonal precipitation totals in an 89-day window, as well as the portion of that variance due to occurrence versus intensity. The variance in total seasonal precipitation in the observed station data ($\sigma^2_{obs,tot}$) is shown for 89-day windows centered on January 15th, April 15th, July 15th, and October 15th in Figure 3.1a–d, respectively. Variability is generally highest in winter on the leeward sides of the Southern Appalachians and the Pacific Coastal Ranges; in spring along the Gulf Coast and reaching into the Ozarks; in summer in the Central Plains, Gulf Coast, and the Atlantic Plains; and in fall the broader Southeast and Pacific Northwest Coast.

3.2.2 Variance Decomposition: Occurrence and Intensity

Previous research in the tropics has suggested that precipitation occurrence may be more predictable than intensity (Moron et al., 2006; Robertson et al., 2009). To separate the portion of the total variance that is due to occurrence versus intensity, we can use the law of conditional variance:

$$\mathrm{Var}[S_{89}] = \mathrm{E_N}\left[\mathrm{Var}\left[S_{89}|N_{89}\right]\right] + \mathrm{Var_N}\left[E\left[S_{89}|N_{89}\right]\right], \tag{3.2}$$

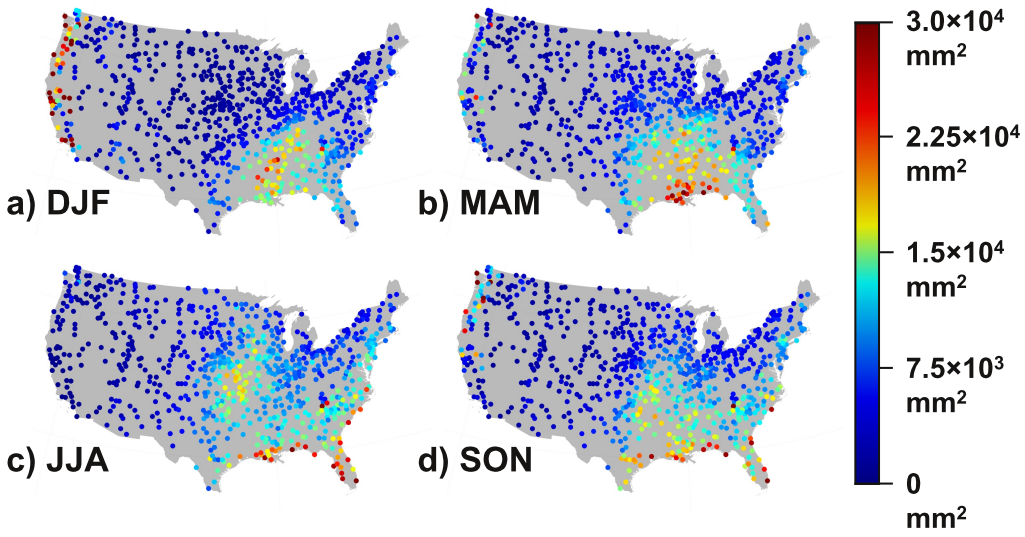

Figure 3.1: The variance of seasonal precipitation totals in the U.S. from at least 1930–2009 (when available, data goes back as far as 1900) for a) December–February, b) March–May, c) June–August, and d) September–November. Daily data taken from 774 United States Historical Climate Network stations.

where S_{89} is the total precipitation at a given location in an 89-day window and N_{89} is the number of days with measured precipitation (≥ 0.254 mm) in that window. The first term on the right-hand-side is referred to as the "expectation of conditional variances," and the second as the "variance of conditional means;" for simplicity's sake we will denote them as $\sigma^2_{obs,int}$ and $\sigma^2_{obs,occ}$, respectively so that Equation 3.2 can be re-written as

$$\sigma^2_{obs,tot} = \sigma^2_{obs,int} + \sigma^2_{obs,occ} \tag{3.3}$$

It is important to note that unless occurrence and intensity are independent, then $\sigma^2_{obs,occ}$ is different (by more than just a linear scaling factor) from the variance in the number of wet days (Var$[N_{89}]$, as used in Gianotti et al. (2013)), and $\sigma^2_{obs,int}$ is not the same as the variance of the intensity on wet days. They are the portions of total precipitation due to occurrence or intensity, rather than the variances of occurrence or intensity. For further information on the partitioning of occurrence and intensity see Appendix B.

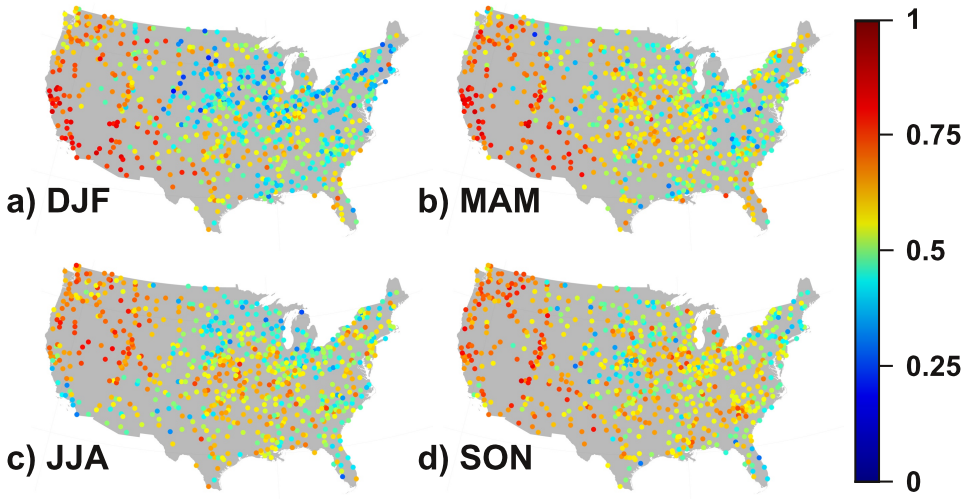

Figure 3.2: The fraction of variance in the observed seasonal total precipitation record attributable to occurrence, $\sigma^2_{obs,occ}/\sigma^2_{obs,tot}$ for a) December–February, b) March–May, c) June–August, and d) September–November. Subtract plotted values from one to obtain the fraction of variance attributable to intensity.

Figure 3.2, the fraction of the total variability that can be attributed to occurrence ($\sigma^2_{obs,occ}/\sigma^2_{obs,tot}$), shows that occurrence contributes substantially to the total observed interannual variance — 56% when averaged across all stations and seasons. The portion of variance due to occurrence is generally higher in the West than the East and is generally highest in the fall.

3.3 Simulation Results

The mean variance (mean of all 1000 simulations) in total precipitation from the SSWM simulations, $\sigma^2_{sim,tot}$, is shown in Figure 3.3, broken down again by 89-day season. This is the simulated analogue of $\sigma^2_{obs,tot}$, shown in Figure 3.1, but represents only the amount of variability that can be well-modeled by stationary processes. Unsurprisingly, the spatial patterns are similar, with the major difference being that the simulated (i.e., due to high-frequency processes) variance of aggregated totals is lower in magnitude than the observed

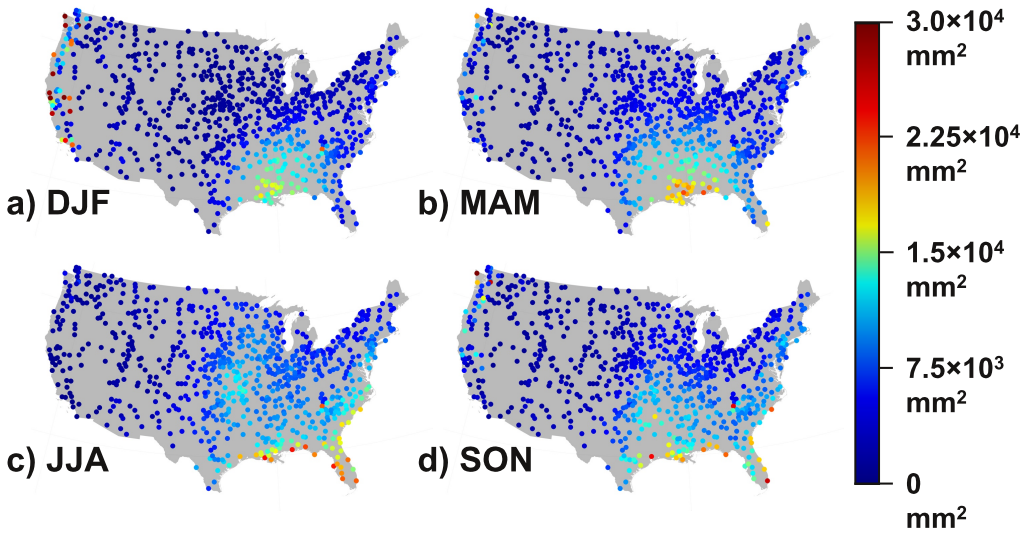

Figure 3.3: The variance of seasonal precipitation totals in the U.S., as simulated by interannually-stationary models fit to the historical record for a) December–February, b) March–May, c) June–August, and d) September–November. This variance serves as a best estimate for the degree of "weather noise," i.e. variance due to only short-memory processes with time-scales on the order of days. While the models used to generate the simulated data are designed to recreate the observed day-to-day variance exactly, they lack low-frequency, annually-varying climate signals found in the observed system, and hence have lower overall variance for seasonal and longer accumulations (cf. Fig. 3.1).

(high and low-frequency) variance upon which it is based. To determine the contributions to $\sigma^2_{sim,tot}$ variability from intensity ($\sigma^2_{sim,int}$) and occurrence ($\sigma^2_{sim,occ}$), we decompose the variability of each simulated data set in the same manner as for the observational data using Equation 3.2. The fraction of the mean total simulated variability due to occurrence, $\overline{\sigma^2}_{sim,occ}/\overline{\sigma^2}_{sim,tot}$, is shown in Figure 3.4 as an analogue to Figure 3.2 (with over-bars used to represent the mean of the distribution of simulations).

Focusing first on winter (DJF), we notice a difference between the fraction of total variance due to occurrence in the observations (Fig. 3.2a) and simulations (Fig. 3.4a), specifically that the simulated data has a less pronounced West/East split in the fraction of variance due to occurrence (or intensity), as well as a smaller range than the observed data. This suggests that much of the low-frequency variability in the observations (absent

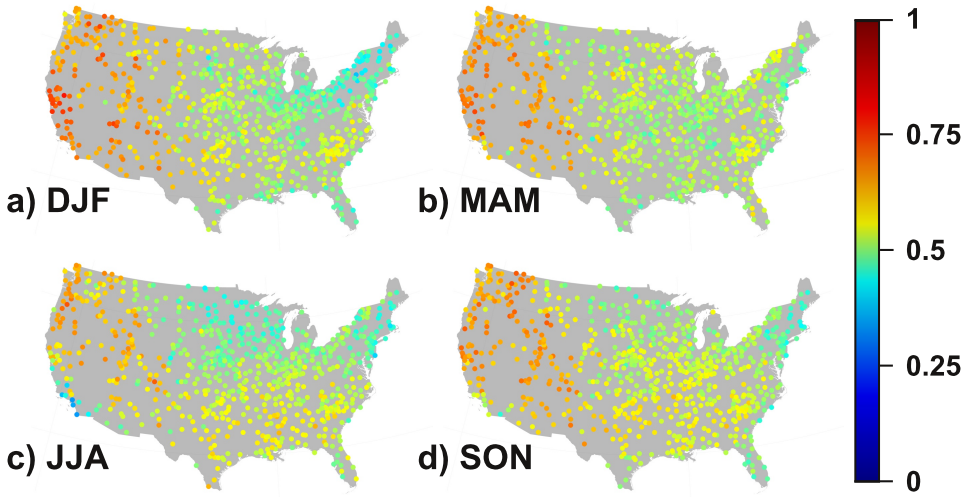

Figure 3.4: The fraction of variance in the simulated seasonal total precipitation records attributable to occurrence, $\overline{\sigma^2}_{sim,occ}/\overline{\sigma^2}_{sim,tot}$, serving as an estimate of the fraction of high-frequency "weather noise" caused by occurrence. Subtract plotted values from one to obtain the fraction of variance attributable to intensity. a) December–February, b) March–May, c) June–August, and d) September–November.

in the simulations) is due to processes that effect occurrence or intensity preferentially, with low-frequency occurrence patterns predominant in the Western U.S. (and along the southern ENSO track), and low-frequency intensity variability stronger in the Eastern U.S. For the other seasons, the same pattern of stronger observed occurrence/intensity separation holds, although the spatial patterns change (particularly in the Northwest in the summer and fall, the Mississippi/Ohio/Tennessee River Valleys in the summer and fall, the Central Great Plains in the spring through fall, and Northern Appalachia in the winter through summer).

3.4 Results of Potential Predictability Analysis

Using this variability data, we can now determine the seasonal potential predictability qualitatively suggested by the variance maps in the previous sections. The potential predictability is defined as the fraction of the observed variance that can be ascribed to climatic (low-frequency) variability:

$$\mathrm{PP}_{89,tot} = \frac{\sigma^2_{obs,tot} - \overline{\sigma^2}_{sim,tot}}{\sigma^2_{obs,tot}} \tag{3.4}$$

The subscript 89 in $\mathrm{PP}_{89,tot}$ is a reminder that this is the potential predictability of 89-day seasonal precipitation, and the subscript "tot" is to clarify that this is defined for the variability of the total accumulated precipitation, rather than the portions of variability due to occurrence or intensity. The values of $\mathrm{PP}_{89,tot}$ fall in the range from zero to one, although in some cases when there is little non-stationary signal and $\sigma^2_{obs,tot}$ lies in the middle of the $\sigma^2_{sim,tot}$ distribution, $\overline{\sigma^2}_{sim,tot}$ may exceed $\sigma^2_{obs,tot}$, leading to negative $\mathrm{PP}_{89,tot}$ values, which imply no ability to extract low-frequency variability from the envelop of high-frequency noise (and hence no PP). As a reminder, if this same analysis was performed using two-day rather than 89-day accumulations (for a location with a chain order of, say, two or three days), the model's variance would exactly match the observed variance, the average of the simulated variances across all 1000 simulations would asymptotically equal

the observed variance, and the potential predictability would be zero. It is only when accumulating beyond the chain order that the persistence of climatic boundary conditions (not represented by our models) can lead to precipitation totals that are unusual when compared to the stationary assumptions of an SSWM.

Using the decomposed simulated and observed variances, we can also establish the potential predictability of climate variability as driven by occurrence and intensity. $\mathrm{PP}_{89,int}$ is the potential predictability of total 89-day precipitation due to intensity, defined as $\mathrm{PP}_{89,int} = (\sigma^2_{obs,int} - \overline{\sigma^2}_{sim,int})/\sigma^2_{obs,int}$. Similarly, $\mathrm{PP}_{89,occ} = (\sigma^2_{obs,occ} - \overline{\sigma^2}_{sim,occ})/\sigma^2_{obs,occ}$ is the potential predictability of 89-day precipitation totals due to occurrence.

To determine whether a PP value is statistically significantly different from zero or indistinguishable from zero, we calculate $\mathrm{PP}_{89,tot}$, $\mathrm{PP}_{89,int}$, and $\mathrm{PP}_{89,occ}$ values using each of the simulations in place of the observed record. If the PP value (using the observed record) exceeds $1 - \alpha$ of the simulated values, then we can reject the null hypothesis that our SSWM adequately reproduces the variability of the observed record (and therefore we can confidently say that $\mathrm{PP} > 0$). If the PP value does not exceed $1 - \alpha$ of the simulated values, then we can not reject the null hypothesis (equivalently our SSWM statistically reproduces the full variability of the observed record).

Seasonal maps of $\mathrm{PP}_{89,tot}$, $\mathrm{PP}_{89,int}$, and $\mathrm{PP}_{89,occ}$ are shown in Figures 3.5–3.8 (a–c, respectively). Markers show station locations, colored by their PP values. Stations shown as circles are significantly different from $\mathrm{PP} = 0$ at the $\alpha = 0.1$ level, and stations shown as triangles are not significantly different from $\mathrm{PP} = 0$ at the $\alpha = 0.1$ level. Higher PP values are (perhaps not surprisingly) more likely to be statistically significantly different from zero, and thus represent conditions that cannot be represented by SSWMs with 0–5 days of memory. Lower PP values are typically not significantly different from zero, and thus the observed precipitation statistics at these locations can easily be recreated by our weather models.

Gridded values shown in the background of Figures 3.5–3.8 are determined using a spatial analysis of the station-based PP values. As is evident in both the station val-

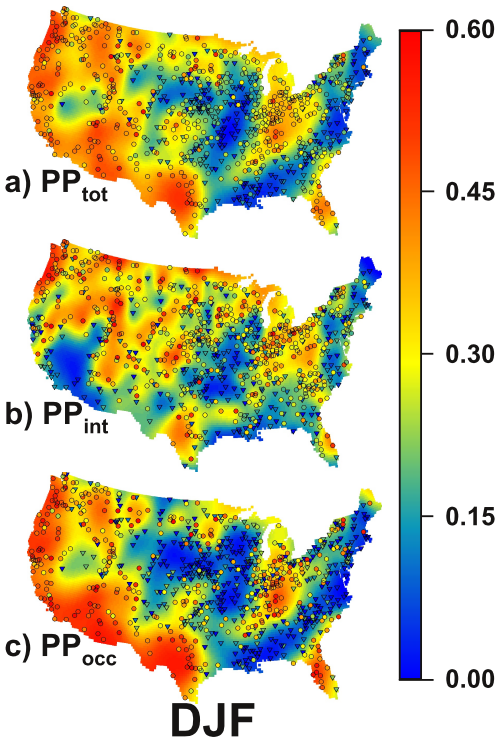

Figure 3.5: Potential Predictability of precipitation for a seasonal (89-day) window centered on January 15, shown as the fraction of seasonal precipitation variance attributable to low-frequency (climate) processes. a) Potential Predictability of total precipitation ($PP_{89,tot}$, unitless). b) Potential Predictability of intensity processes ($PP_{89,int}$, unitless). c) Potential Predictability of occurrence processes ($PP_{89,occ}$, unitless). Circles denote locations with PP significantly different from 0, and triangles denote locations where PP is not significantly different from 0 ($\alpha = 0.1$). Gridded underlying values are modeled based on spatial coherence. DJF displays the highest $PP_{89,tot}$, $PP_{89,int}$, and $PP_{89,int}$ values of all seasons when averaged across stations. The DJF station-averaged PP values for all three metrics are nearly equal at roughly 0.3. $PP_{89,tot}$ and $PP_{89,occ}$ have a very similar spatial pattern, with large potential climate signals along the West Coast, Rockies, Southwest, Florida, and the Michigan/Ohio River Valley area. Intensity signals are notably lower in the Southwest and are slightly higher along the Piedmont/Atlantic Plains.

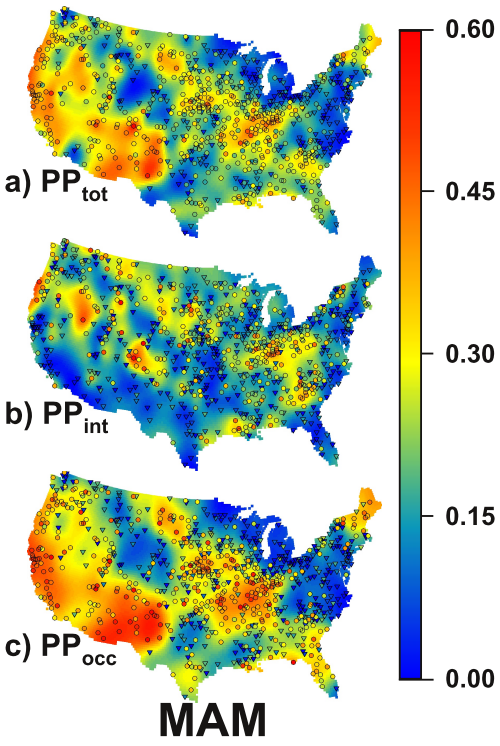

Figure 3.6: Same as Fig. 3.5, but for a seasonal (89-day) window centered on April 15. All of the PP metrics are generally lower in MAM than in DJF throughout the contiguous 48 states, with $PP_{89,int}$ particularly low and incoherent (station-based average of roughly 0.2). Springtime occurrence processes show higher Potential Predictability in the Central Great Plains and New England as compared to winter, but lower in the Northwest and Great Lakes regions. $PP_{89,tot}$ and $PP_{89,occ}$ have a similar pattern, although on average $PP_{89,occ}$ exceeds $PP_{89,tot}$ by 25%.

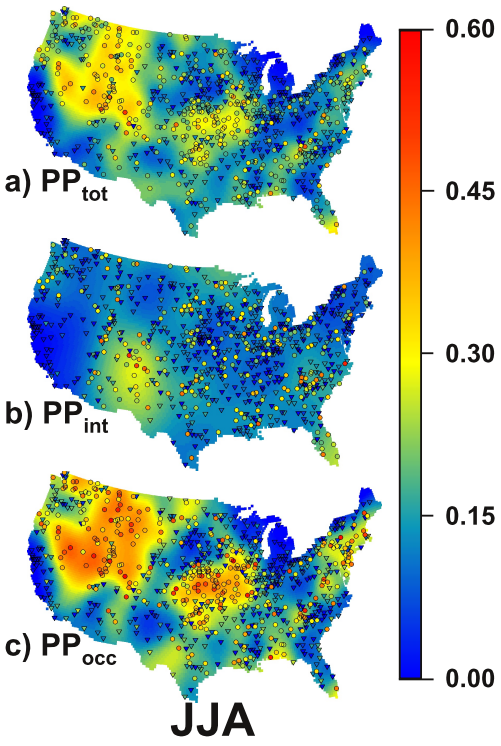

Figure 3.7: Same as Fig. 3.5, but for a seasonal (89-day) window centered on July 15. Summer is the Potential Predictability minimum for each of $PP_{89,tot}$, $PP_{89,int}$, and $PP_{89,int}$, with station-averaged total and occurrence values in a similar range of roughly 0.15 and intensity PP at half that. The most predictable areas are the Rockies and Central Great Plains for total precipitation and occurrence, and the North American Monsoon region of Arizona/New Mexico for intensity.

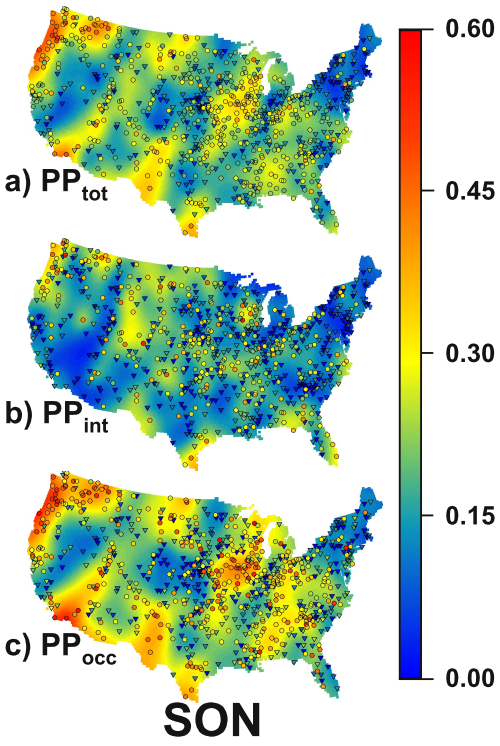

Figure 3.8: Same as Fig. 3.5, but for a seasonal (89-day) window centered on October 15. Autumn PP is highest in the Northwest and the "Cornbelt" region, although intensity processes do not seem to demonstrate much coherent potential predictability. The Cornbelt area demonstrates the "destructive interference" of convolving intensity and occurrence signals — $PP_{89,tot}$ is lower than $PP_{89,occ}$ because high-frequency noise in the intensity processes mask climate signals in the occurrence processes when combined as total precipitation.

ues (symbols) and the interpolated background, there is a large degree of geographical coherence in the estimated PP values. This coherence is not a simple result of overall precipitation climatology (Pearson correlation between all 365×774 $PP_{89,tot}$ values and daily mean precipitation is -0.03), and so any spatial coherence that emerges from the analysis lends further confidence to our estimates. Gridded estimates are determined using a generalized cross validation (GCV) scheme for fitting thin plate splines (ANUSPLIN – Wahba, 1979; Hutchinson, 1998). 90% confidence intervals from the GCV analysis around the gridded values are fairly spatially homogeneous, suggesting a spatial uncertainty in our PP estimates of roughly ± 0.05 (maps of 90% confidence intervals are shown in the Supplementary Materials).

Looking at $PP_{89,tot}$ (Figs. 3.5a–3.8a), winter displays the most potential predictability, including broadly across the West, West Texas, Florida, and in the Western Appalachian/ Great Lakes regions. The West and Florida are strongly associated with ENSO precipitation (Ropelewski and Halpert, 1986), and the area south of the Great Lakes may have connections to variability in lake effect snow, as driven by variability in winter storm tracks (Niziol et al., 1995). The Pacific Northwest shows high $PP_{89,tot}$ in the fall, also likely connected to ENSO patterns (Ropelewski and Halpert, 1986). Also significantly, the Gulf and southern Atlantic Coast, which displayed some of the highest observed precipitation variability during the JJA/SON Atlantic hurricane season, shows little $PP_{89,tot}$ at these times, suggesting that the low-frequency total precipitation signal in this area is small compared with the weather noise, a discouraging sign for seasonal forecasting of tropical storm-induced precipitation locally.

$PP_{89,int}$ is shown in Figures 3.5b–3.8b. Like $PP_{89,tot}$, $PP_{89,int}$ has a station-averaged maximum in the winter, with a similar PP "hot spot" in the Western Appalachia/Great Lakes region and high values throughout much of the broader West, but without the low values across the Southwest. PP due to intensity is low and spotty in the summer and fall (with the highest summertime values in the Southwest monsoon region), and is much more spatially correlated with the intensity climate signal ($\sigma^2_{obs,int}$- $\overline{\sigma^2}_{sim,int}$, Pearson correlation

$r = 0.573$) than it is anti-correlated with the weather noise ($\overline{\sigma^2}_{sim,int}$, $r = -0.064$). The relatively low values of $\mathrm{PP}_{89,int}$ (DJF: 0.257, MAM: 0.181, JJA: 0.123, SON: 0.156, when averaged across all stations) indicate that our SSWMs capture intensity variability well, suggesting that the low-frequency processes that drive the intensity of precipitation variance on seasonal scales are typically small relative to the high-frequency noise.

$\mathrm{PP}_{89,occ}$ is shown in Figures 3.5c–3.8c. $\mathrm{PP}_{89,occ}$ values (DJF: 0.237, MAM: 0.243, JJA: 0.203, SON: 0.237) are generally higher than $\mathrm{PP}_{89,tot}$ values; and, like $\mathrm{PP}_{89,int}$, $\mathrm{PP}_{89,occ}$ is much more highly correlated to its climate signal ($\sigma^2_{obs,occ}$- $\overline{\sigma^2}_{sim,occ}$, $r = 0.551$) than anti-correlated with its weather noise ($\overline{\sigma^2}_{sim,occ}$, $r = 0.023$). $\mathrm{PP}_{89,occ}$ is also much more spatially correlated with $\mathrm{PP}_{89,tot}$ than is $\mathrm{PP}_{89,int}$ ($\mathrm{PP}_{89,occ} \propto \mathrm{PP}_{89,tot}$, $r = 0.821$; $\mathrm{PP}_{89,int} \propto \mathrm{PP}_{89,tot}$, $r = 0.448$). The highest values of $\mathrm{PP}_{89,occ}$ (greater than 0.7) occur in the West, through Texas, and in the Ohio River Valley region in winter (Fig. 3.5c) ; in California, the Southwest, and the Cornbelt in the spring (Fig. 3.6c); in the Rockies and Cornbelt in summer (Fig. 3.7c); and in the Pacific Northwest and along a swath between Milwaukee and Kansas City in the fall (Fig. 3.8c). From September through May (Figs. 3.5, 3.6, 3.8), the patterns are similar to slight amplifications of the total PP, but the summer shows $\mathrm{PP}_{89,occ}$ hot spots in the Central Plains and the Rockies that are much higher than in the $\mathrm{PP}_{89,tot}$ map (Fig. 3.7a,c).

This highlights an important point, namely that $\mathrm{PP}_{89,tot} \neq \mathrm{PP}_{89,occ} + \mathrm{PP}_{89,int}$, and in fact the separated occurrence and intensity PP values are often higher than the total PP signal. We can think of this as a destructive interference of sorts, where the weather noise in the intensity process masks an otherwise persistent occurrence signal (or vice versa) when the two are examined together as $\mathrm{PP}_{89,tot}$. The summer Central Plains $\mathrm{PP}_{89,occ}$ hot spot is a good example of this; most of the variability in the region is tied to occurrence processes (Figure 3.2c), but roughly half of the weather noise is due to intensity processes (Figure 3.4c), and so the climate signal stands out against the weather noise for occurrence (high values in Figure 3.7c) much more so than for intensity (low values in Figure 3.7b). When combined, however, the low signal-to-noise ratio from intensity weakens the overall

signal-to-noise ratio for total precipitation, and so $PP_{89,tot}$ (Figure 3.7a) is lower than $PP_{89,occ}$. This area is also noteworthy as the site of high modeled potential predictability due to land-atmosphere coupling (Koster et al., 2004). We might imagine, then, that this strong signal from occurrence is due to a precipitation recycling amplification in the number of rainy days in years with sufficient moisture in the area (as has been previously found by Findell et al., 2011). The difference between years in which sufficient moisture is present to trigger recycling (and thereby extra rainy days) and those with insufficient moisture for this feedback leads to interannual variability not present in our annually-stationary models, yielding potential predictability (i.e., $\sigma^2_{obs,occ} > \overline{\sigma^2}_{sim,occ}$, meaning that $PP_{89,occ} > 0$). In general, processes driving PP in these "hot spot" areas will need a more thorough investigation to establish regional modes of predictability.

3.5 Summary and Discussion

We have used observational data directly to obtain estimates of both the "weather noise" and non-stationary climate signal contributions to precipitation variability, and have additionally decomposed these signals into contributions from occurrence and intensity processes. This is significant in that precipitation occurrence displays more potential predictability than intensity, suggesting that those processes responsible for triggering occurrence may be particularly important when attributing variability. This finding is in agreement with previous work in the tropics and Southern Hemisphere (Moron et al., 2007; Robertson et al., 2009; Verbist et al., 2010), suggesting that precipitation occurrence may generally be a more fruitful avenue for improving seasonal forecasts than intensity-related processes. We find that occurrence also displays higher PP in most cases than total precipitation, suggesting that modeling and forecasting efforts might be well-served by focusing on occurrence first, rather than seasonal totals. Intensity is also occasionally more potentially predictable than total precipitation, and the two components should be investigated separately to avoid the loss of information that comes with convolving the two

signals. Some care will be necessary in approaching occurrence using either global climate models (GCMs) or gridded datasets, as the spatially-amalgamated discrete data will almost definitely yield different results at finite spatial scales rather than point-measurements (Hughes et al., 1999).

Generally, the potential predictability results shown in Figures 3.5–3.8 should help guide those attempting to attribute precipitation variability to processes. PP "hot spot" regions with high potential climate signals should guide researchers to both investigate processes and to work towards improved local forecasts if possible. Equally importantly, PP "cold spots" highlight regions in which climate signals may be difficult to disentangle from chaotic, high-frequency weather noise. Where PP values are low — and particularly where PP is not significantly different from zero — our SSWMs represent the interannual variability of 89-day precipitation totals well, despite having only at most five days of persistence and no representation of interannual processes. In some cases, the relative role of weather noise may not matter so much as the absolute magnitude of the potential climate signal, $\sigma^2_{obs} - \sigma^2_{sim}$. Maps of this absolute potential signal strength for total precipitation, occurrence, and intensity are available as Figures C.2–C.5.

In the introduction, we suggested applying PP to 1) estimate the limits of our ability to confidently forecast longer-term climate behaviors in analyses of precipitation time series, 2) determine where and when our present forecasts are limited by an understanding of the system versus an absence of useful climate signals, and 3) test whether or not our current forecast models display signals and noise in the same manner as the real climate system. This study provides information on the first of these applications. Precipitation variability at locations with low PP is difficult to distinguish from the noise of a simple stationary weather model with only a few days of persistence, and so climate signals in those regions cannot be distinguished from what we think of as weather noise. Locations with higher PP values show significant divergences from stationary processes, and so may be predictable; but the ability to predict precipitation from climate indices in these locations is limited both by our knowledge of the proper climate indices, and the coherence of their relationship

with precipitation. Determining the degree to which each of these limits our ability to make optimal forecasts of precipitation (i.e., the second proposed application for PP) is a substantial area of research, one which weather and climate forecast centers address each day. The third suggested application of PP information is to see whether global climate models follow patterns of potential predictability similar to the observational data upon which they are based (see previous work by Moron et al., 2006). These latter applications are important areas of future research.

Additional open questions remain regarding the drivers of potential predictability, specifically, how much of PP is attributable to climate oscillations versus trends in the historic precipitation record? Also, the prediction of extremes is important both for planning and because increases in extreme precipitation events are one of the most consistent findings from past decades. Assessment of PP values for extreme events can be performed using our dataset and methods, but will require different metrics than those applied in this study.

Chapter 4

A kernel-auto-regressive weather generator for improved subseasonal-to-seasonal precipitation statistics

4.1 Abstract

The kernel auto-regressive (KA) method can be used to represent the daily to multi-day auto-correlation structure of daily precipitation time series, using information both in the occurrence and intensity of measured rainfall events. The method is able to capture a larger fraction of the memory in multiple time series than commonly-used occurrence-based Markov chain models, even when the intensity distribution is allowed to be conditioned on the Markov state. The kernel auto-regressive method is less sensitive to the spatial scale at which the data is reported, as it is not strictly reliant on patterns of wet and dry days for providing correlation. Output from the KA model can be used as weather generator model simulations, as empirical representations of process structure, or as climatological null models against which observations can be compared for statistical significance. The KA method demonstrates improvements in each of these over occurrence Markov chain models and daily independent climatology.

4.2 Introduction

There is no shortage of existing weather generator models (WGs), for any number of weather variables, and for any number of specialty purposes (Wilks and Wilby, 1999;

Ailliot et al., 2015). Weather generators are used by researchers, practitioners, businesses, and agencies for purposes ranging from estimating natural resource availability, forecasting hazard risk, understanding fundamental meteorologic processes, and driving other complex natural systems models. The basic motivations that these models share in common is the desire to represent some probabilistic structure of weather variables and a need for simulations of weather that meet some basic statistical criteria.

Among the classes of WGs are those that represent single versus multiple variables (e.g., precipitation, temperature, radiation, etc.), those that represent some level of physical process detail versus purely statistical methods, and those that assume some level of climate process stationarity versus those that represent process variability at climate time-scales. The difference in model form is dependent on the use the WG will play: the classic "Richardson-type" WG for precipitation represents daily rainfall, typically fit as twelve distinct parameterizations to represent the seasonal cycle, with a single-lagged Markov-chain representation of occurrence and a parameterized univariate distribution for intensity (classically exponential, but more typically gamma) (Richardson, 1981; Wilks and Wilby, 1999). This is useful for representing the scale of seasonal variability and for driving other physical models that may not require any sort of long-term change analysis. A model focused on sub-seasonal-scale extreme events will likely carefully fit more complex distributions to the tails of the distribution and require attention to daily-to-monthly-scale auto-correlation of these extremes (Koutsoyiannis, 2004; Min et al., 2011), while a WG used in downscaling output from a global climate model (GCM) may focus mainly on a spatial covariance structure, conditioned on the state of multiple climate indices or a given mean value (Wood et al., 2004).

In this study, we propose a method for the stochastic simulation of precipitation to fit a specific set of criteria. First, we are interested in a WG for use as a *climatological null* model — that is, an entirely probabilistic data model that represents processes on weather time-scales as well as possible, using only lagged local precipitation as a predictor, while explicitly not representing variability due to processes on climate time-scales.

Following on the work in Gianotti et al. (2013) and Short Gianotti et al. (2014), this method will provide a means of investigating potential predictability (PP), but allowing for a stronger representation of the interdependence of occurrence and intensity processes in precipitation. Secondly, we are interested in a model that can explicitly be used at *multiple spatial scales*. Due to the dependence of occurrence probability on the spatial scale of observations, the methods used in Gianotti et al. (2013) and Short Gianotti et al. (2014) are best suited for scales at which occurrence probability is far from either zero or unity and are not applicable for inter-comparison between different spatial scales. The method proposed in this chapter will provide the foundation through which climate and weather variability can be compared among global climate models and gridded observational datasets in Chapter 5.

4.2.1 Climatological null models

Forecast skill is often measured relative to climatology (e.g., Heidke and Brier Skill Scores), and that climatology is typically enumerated as the probability distribution of a single variable for a given time period, marginalized over all states of the Earth system (including climate states, atmospheric states, secular trends, representations of spatial teleconnections, land surface conditions, etc.). For this model, we wish to explicitly acknowledge the daily-scale temporal correlation structure inherent in precipitation data by modeling it rather than marginalizing over it. This representation of a climatology with serial correlation serves two major purposes.

First, by representing precipitation as a data generating process which can include auto-correlation we create a more stringent baseline for quantifying weather forecast skills than daily-independent climatologies. When we test to see if a variable serves as a skillful predictor for precipitation, we compare the forecasts to climatology because we want to determine if that variable contains any useful information not already hidden in the precipitation data itself. If lagged precipitation values are more skillful than another predictor, those lagged precipitation values should be used in place of (or in conjunction with) that

predictor. Thus, our weather generator model is a stronger reference climatology (null model) for skill score calculations and assessment of predictor utility.

Second, by representing the auto-correlation of daily-scale precipitation, we explicitly start separating stochasticity from processes on weather time-scales and climate time-scales. Climate and weather are often difficult to extricate from one another, partially due to conflicting definitions. Climate is sometimes defined as "average," "expected," or "marginal" weather; sometimes as boundary conditions acting upon the atmosphere; and sometimes as low-frequency processes (as compared to high-frequency weather). Weather, similarly, can refer to the atmospheric state, that atmospheric state with some low-frequency climate signal removed (i.e., as anomalies from a slowly varying climate signal), or broadly anything with persistence shorter than the atmosphere's chaotic time-scale on the order of weeks. By explicitly representing auto-correlation in precipitation data, we characterize atmospheric persistence as partially deterministic, in the same sense that modelers represent the climate state as partially deterministic by calculating the annual seasonal cycle explicitly in weather generator models. Thus, our model is not only a more strict climatology for weather forecasts, but also a null model for climate variability in that it represents some climate-scale variability via weather time-scale processes.

4.3 Methods

4.3.1 Overview and Data

To capture the correlation structure of daily precipitation without decoupling occurrence and intensity processes, we combine an inverse-CDF transformation of each day's data with a generalized non-parametric auto-correlation model using Gaussian kernels. The transformation is known as a "rank-based inverse normal" transformation (Akritas, 1990; Cai et al., 2016), and it serves to allow us to work in an unbounded domain, reducing some of the common complications inherent in both bounded and zero-inflated data. It also allows us to provide correlation structure between wet and dry days in the same manner

that we represent the correlation structure between serial wet days. The kernel model is used to represent the joint probability density of m-day series of precipitation values without relying on the assumption that the covariance structure is multivariate normal (as in the typical AR time-series paradigm), or even that it follows any specific family of parametric distributions. By using Gaussian kernels, the kernel model is a specific instance of the broad class of Gaussian process models, common in machine learning applications due to their flexibility and somewhat analytically tractable nature (Rasmussen and Williams, 2004).

We use precipitation data from three sources: station data from the Global Historical Climatology Network (Menne et al., 2012), 1/4 degree gridded data from the Climate Prediction Center's (CPC) Unified Gauge-based Analysis of Daily Precipitation over the Continental United States (Chen and Xie, 2008), and a 1 degree gridding of the same CPC data (U.S. Climate Prediction Center, 2015). In each case the data is from the years 1948–2004, inclusive.

4.3.2 Fitting

The model fitting procedure is shown in Figure 4.1. To preserve the seasonal patterns of precipitation occurrence, intensity, interannual variability, and short-term correlation structure, we fit a model to each day of the year for a given location. For any given day's model, we use data drawn from a width-p window around that day (the "pooling window" — $p = 31$ days in the subsequent analysis) to improve our estimates of the serial correlation (see Figure 4.1a). For each of the days in the pooling window, using all N years of observations, we use that day's observation and the $m - 1$ previous observations to form a length m vector of serial precipitation. Each of these vectors represents a single point in an m-dimensional space, to which we will fit a joint probability density function (PDF). The marginal distributions are simply empirical histograms of daily observed precipitation, and the conditional distribution of the m^{th} dimension given the other $m - 1$ is the probability density for a single day given that you have just observed a specific $m - 1$ days of

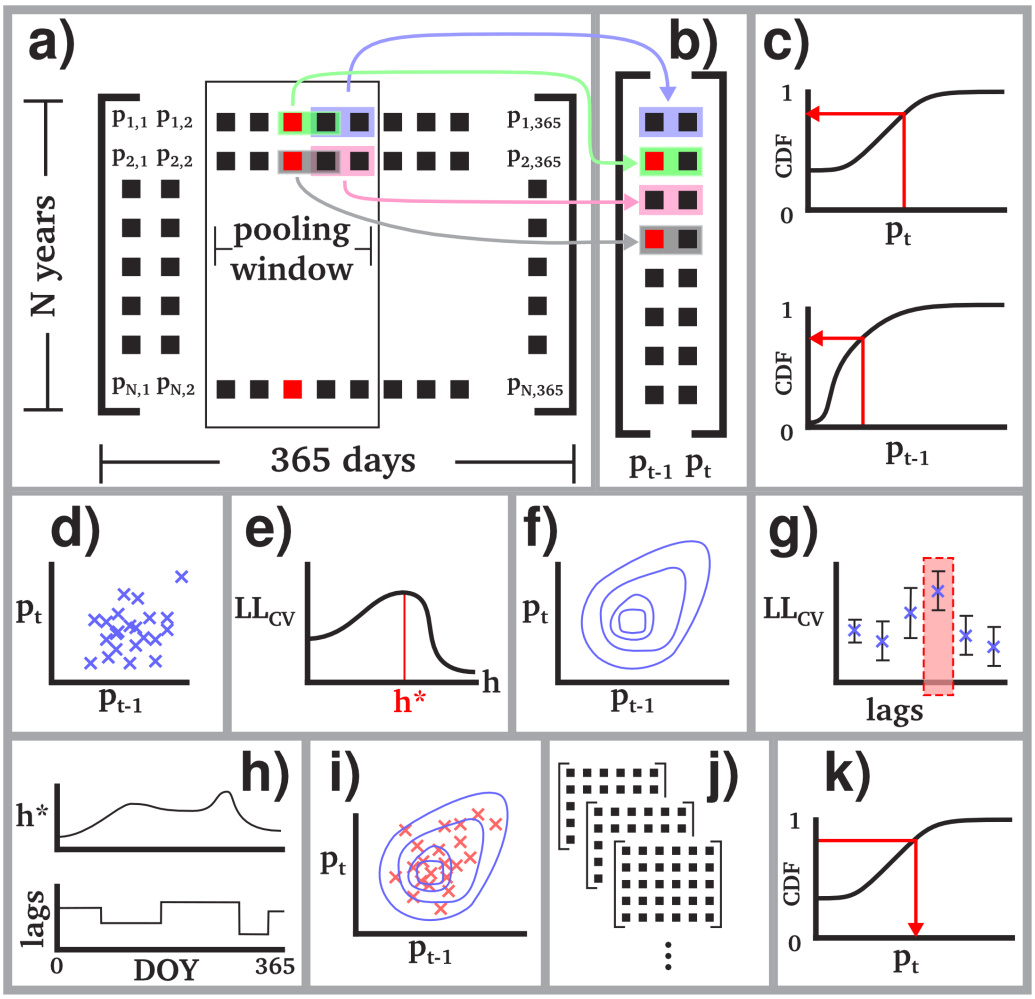

Figure 4.1: A schematic showing the steps for fitting the kernel-auto-regressive weather generator and simulating precipitation data. a) N years of daily data form the basis for 365 daily models. For each day (e.g, the column in red), a pooling window is used to optimize covariance estimation. b) All m-day serial vectors of observed precipitation from the pooling window are used to form an m-dimensional (2D shown) empirical distribution of precipitation. (c) Each column (day) is rank-transformed so that all marginal distributions of the joint distribution (d) are exactly Gaussian. (e) A bandwidth, h, is selected using cross-validation to create a kernel density (f) from the observations. (g) Selecting models with zero lags (1D) to five lags (6D) using cross-validation for each day yields 365 selected model dimensions and 365 optimal bandwidths (h), which together comprise the model for the location. (i) Simulating one day at a time using the corresponding kernel model and conditioning on the previous $m-1$ days produces ensembles (j) of stochastic precipitation data in the CDF-transformed domain, which are then back-transformed (k) for analysis.

precipitation.

The most common means of quantifying the correlation structure of serial data is the auto-regressive (AR) model, the simplest member of the ARCH/ARFIMA/ARMAX families. The AR(1) model fits a bivariate normal distribution to 2-dimensional vectors of observations, usually either maximizing the likelihood of the joint distribution or the likelihood of the conditional distribution. Since daily precipitation clearly does not fit the assumptions of normality, a typical AR-type model is inappropriate. The multivariate normal (MVN) distribution of the typical AR model can of course be replaced with other multivariate parametric models, or can be represented more empirically using a multivariate binned histogram (or probability mass function) to capture exotic distributions (examples given in Wilks and Wilby, 1999). However, for zero-inflated data (such as precipitation from weather stations), the size of the bin has a strong impact on the correlation structure of the model. Smaller bins will assign more likelihood weight to occurrence processes, and larger bins will assign relatively larger weight to intensity, and any finite bin width is effectively an arbitrary trade-off in the role of occurrence in the model. Even more problematic than selecting a bin size, is that for any bin size the climatological occurrence frequency has a very large impact on the joint distribution (and thus model parameter likelihoods), making comparisons of parameters or simulations between different locations or the same location at different spatial scales (any gridded scale or point measurements) impossible. The same issue arises for other parametric distributions (such as a multivariate gamma): datasets with more dry days will lead to huge inflation of likelihood weight towards those identically-valued dry days, essentially forcing a continuous model to emulate a Bernoulli model as best as possible to maximize the zero-inflated likelihood.

To circumvent this inherent dependency of the simplest probability models on occurrence frequency, we transform our data into an unbounded domain and "un-inflate" our zero-inflated data. For our $(N \cdot p)$ by m matrix of observations for a given day (see Figure 4.1a–b, in which $m = 2$, or a 1-lag model), we transform each of the m columns through a rank-based univariate inverse normal CDF, assigning the smallest observed value

$\Phi^{-1}(1/(Np+1))$ and the largest observed value $\Phi^{-1}(1-1/(Np+1))$ so that each column of the transformed matrix is exactly normally-distributed (Figure 4.1c). Duplicate values (notably zeros) are assigned random relative ranks (and so are asymptotically uncorrelated with each other), and are handled as special cases when calculating likelihoods.

In the CDF-transformed domain (Figure 4.1d), each dimension of the data is marginally normal, but the joint distribution is not necessarily MVN. To allow for as flexible a representation of the covariance structure as possible, we represent the joint distribution between the m days of serial observations using a kernel density. Since all dimensions of our data are scaled identically, we use a simple spherical Gaussian kernel, which has one scalar parameter — the bandwidth, h. We select the optimal kernel bandwidth for that day of year and for each number of lags ($1 \leq m \leq 6$ in this analysis) using cross-validation (Figure 4.1e). We perform a nested grid search of possible bandwidths and use a leave-out 20% repeated-random-subsampling cross-validation scheme. The likelihood to be optimized is that of the validation data using the full joint PDF of the training data kernel model. By selecting a bandwidth, we have selected a probability model for our data (Figure 4.1f).

Once we select an optimal bandwidth for each potential number of lags, we then pick the optimal number of lags using a second cross-validation step (Figure 4.1g). The entropy of the joint distribution scales with the dimension m, and so the comparison between models of differing dimension is scaled by the dimension of the model. Alternatively, one could compare the univariate conditional likelihood of the last day's precipitation given $m-1$ previous days for a more prediction-focused approach to model selection. The model with the highest mean likelihood across all repeated subsampling cross-validations is selected, and the dimension of that model becomes the dimensionality of the kernel model for that day of year. The dimension and bandwidth are the two critical parameters for each daily model, and the full model for a dataset at any location is specified by 365 dimension values and 365 corresponding bandwidth values (Figure 4.1h).

When determining bandwidth, likelihoods are calculated as typical for a Gaussian kernel model. Given N d-dimensional kernel means in the $N \times d$ matrix $\mathbf{T}$, a bandwidth h, and

a d-dimensional vector x at which to calculate the density or likelihood,

$$f(x) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\sqrt{h(2\pi)^d}}\exp\left[-\frac{1}{2h}(\mathbf{T}_i - x)(\mathbf{T}_i - x)'\right] \tag{4.1}$$

where the summation is over all N kernel means, or equivalently all N rows of $\mathbf{T}$. The log-likelihood for a set of M observations is then

$$LL = \sum_{i=1}^{M}\ln f(x_i) \tag{4.2}$$

In contrast, when calculating likelihoods used to determine the appropriate number of lags, since all dry day zero-values are equivalent, we force the distance $\mathbf{T}_i - x)$ to be zero in any dimension where both $\mathbf{T}_i$ and x correspond to dry days.

Figure 4.2 shows the 1-lag (2-dimensional) joint distribution for precipitation at Fairhope, AL at three different spatial scales for day of year 180 (June 29). In Figure 4.2a (station data), the majority of observations are dry days (left and/or below the threshold lines), which are uncorrelated with each other, but still provide the appropriate covariance between occurrence processes and intensity processes. At the 1/4°-scale (Figure 4.2b), occurrence frequency is higher than 50%, and the positive correlation structure of wet-day/wet-day pairs is more evident. Additionally, since vertical cross-sections give the conditional distribution of p_t given p_{t-1}, we can see that the heaviest wet events (the upper-most points in Figure 4.2b) tend to occur after other heavy wet days. At the 1°-scale (Figure 4.2c), dry days are rare and two day dry spells are non-existent in the observational data (and highly unlikely in the kernel model, though not impossible). The relatively symmetric PDF shows that the light-then-heavy pattern is essentially as probable as the heavy-then-light pattern, and that dry days are likely to be followed by light precipitation days.

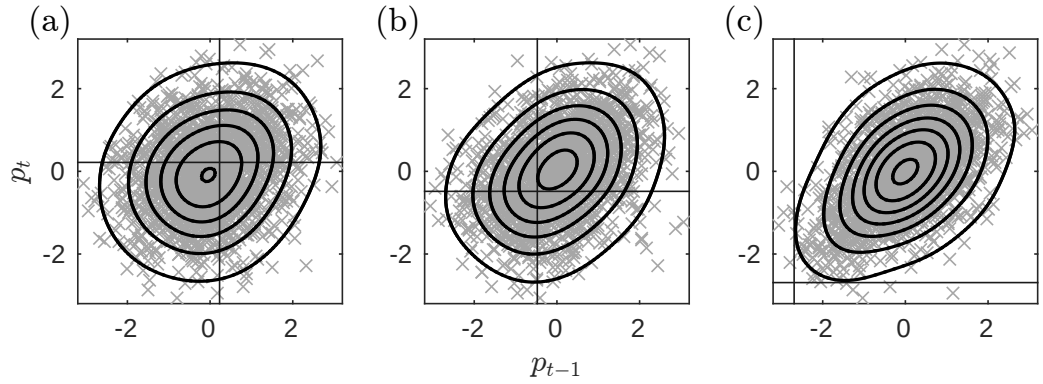

Figure 4.2: Joint probability densities for daily precipitation, p_t, and the previous day's precipitation, p_{t-1}, for day of year 180 (June 29) at Fairhope, AL using a 31-day pooling window. The spatial scale of the data increases from left to right: (a) shows the joint density for a single GHCN station, (b) for the co-located CPC Unified 1/4° gridded data, and (c) for the 1° CPC Unified data. Individual two-day observations are represented as "x-es" in Gaussian z-score units, contours show equal density levels, and vertical/horizontal lines show the cut-off threshold for precipitation occurrence: points above and right of the lines are wet days, below and left are dry days, and the dry-dry two-day pairs have no covariance structure. The marginal distributions are, by design, identically $\mathcal{N}(\mu = 0, \sigma = 1)$. As the occurrence probability increases (from left-to-right), the "wet quadrant" covariance structure becomes the more dominant feature of the joint density as a whole.

4.3.3 Simulation

Simulation of precipitation is performed in the CDF-transformed domain, where the correlation structure is more simply represented, then transformed back through an inverse CDF transformation to the domain of the actual observations. The key step in the simulation process is conditioning the model for the given day of the year on the simulated values for the previous $m - 1$ days so that the daily correlation structure is maintained.

The dimensionality of the model changes from day to day, but the maximum number of days used in conditioning is one less than the maximum dimensionality of the model over all days of the year. To keep track of this, a vector of length $\max(m) - 1$ is used as a buffer to store the relevant conditioning data. Since we initially have no data to condition with, the buffer is set to a random draw of climatological values for the appropriate days of the year, and then a one-year burn-in period is used (and later discarded) to represent proper correlation statistics.

For each day of the year, the marginal probability of producing the $m-1$ values in the buffer (marginalizing over the single dimension representing precipitation on the current day) is determined for all of the data-point/kernels in in the joint PDF. If the model is one-dimensional for that day (no lags/memory), the marginal probability is uniformly 1/N, where N is the number of data points/kernels. A single data point/kernel is selected based on the marginal probabilities, and the conditional probability is determined for the previous $m-1$ days' rain. Since the kernels are Gaussian and spherically symmetric, the conditional PDF is a univariate normal distribution, the conditional mean is simply the m^{th} (last/current) value of the data point used as the multivariate mean, x_{ij}, and the (scalar) conditional variance is just the bandwidth, h_i, squared. Thus, the simulated precipitation (in the CDF-transformed domain) is just a random draw from $\mathcal{N}(\mu = x_{ij}, \sigma^2 = h^2)$, where i corresponds to day of the year, and j corresponds to the j^{th} (last) entry in an m-dimensional vector, x, representing the selected observed m-day precipitation data point/kernel.

Following simulation, the data are transformed back to the observational domain by interpolation using the original data and its CDF-transformed values. Before transformation, the simulated data are re-standardized to ensure proper variance representation (see Appendix D for further details). Any values below the no-rain cut-off in the original data are converted to zeros, and any values larger than the largest value in the observational data set need to be extrapolated. We use the tail of a gamma distribution to fit the extrapolated values. The gamma distribution is fit to the wet days for that day of year, we align the z-score of the largest observed value in the CDF-transformed observational data with the corresponding quantile of the gamma distribution; the extrapolated values are mapped to the appropriate part of the upper tail by normal-to-gamma quantile matching.

Simulations can be run for as many years as necessary to calculate asymptotic statistics, or can be run in independent ensemble modes (likely in multiples of the observational record length) for statistical assessment of climatological phenomena.

4.4 Results

The kernel-auto-regressive model (KA) was fit to data at three locations: Fairhope, Alabama; Blue Hill Massachusetts; and Forks, Washington. In each location, separate annual models (each comprised of 365 daily models) were fit for each of the three data sources (GHCN, CPC-1/4°, and CPC-1°). In addition, a chain-based model — referred to as the "Occurrence Markov Chain" or OMC model (Short Gianotti, 2016), also comprised of 365 daily submodels — and a no-memory, "zero-lag" (ZL) occurrence/intensity model were fit to the same datasets for model comparison. The OMC model uses a variable order Markov chain to represent the auto-correlation in the occurrence process and a flexible five-parameter gamma-gamma mixture model to represent intensity, also with a 31-day pooling window for parameter estimation. The chain order (number of lags) is determined for each day of the year using the corrected Akaike Information Criterion (Hurvich and Tsai, 1989), and the parameters for the intensity model are selected for each day of the year by maximum likelihood estimation (see Short Gianotti et al., 2014, for further details). The zero-lag model uses the same distribution family for intensity as the OMC model, but daily occurrence does not depend on the previous days' precipitation and simply follows the climatological probability of occurrence for that day of year (within the 31-day pooling window).

Each of the three models was used to simulate 1000 57-year ensembles of stochastic precipitation at each of the three spatial scales at each location. By design, all of the models asymptotically reproduce the probability of occurrence, mean intensity, and variance of intensity for every day of the year. Each of the models create interannual variability stochastically, but none of them represent climate variability processes, and thus are likely to be "under-dispersed" in their representation of accumulated totals relative to the observations (Katz and Parlange, 1998). Additionally, the kernel-auto-regressive model and the OMC model each represent serial correlation (although the OMC only represents correlation in occurrence), and so precipitation totals accumulated over multi-day-to-multi-

year periods will likely be more under-dispersed for the zero-lag model than for the KA or OMC models. The mean accumulated totals match the observations asymptotically for each model. The KA model is able to represent any processes captured in the OMC model, but with more flexibility, and the ZL model is explicitly a restricted version of the OMC model with no memory, so we would expect the KA model to be most able to represent complex variability structure, followed by the OMC model, and then the ZL model.

Figure 4.3 shows empirical distributions of December–February seasonal precipitation for the observed data and simulations from the three models for each of the nine datasets (three locations times three spatial scales). At the station level (top row), all models similarly miss some of characteristics of the observations, presumably because none of them represent interannual variability other than through zero-to-six day correlation structure. Any variability caused by slower processes (such as climate modes) will not be well represented. Notably, for Forks, WA (the wettest location), the OMC model outperforms the zero-lag model, and the kernel-auto-regression model outperforms both of the simpler models. At the 1/4-degree scale (CPC-1/4°, second row), the same pattern holds, but with more notable performance improvements for the kernel-auto-regressive model. This is not surprising, as the increased frequency of occurrence makes the proper modeling of the intensity auto-correlation more important for characterizing the patterns of synoptic scale precipitation events. At the largest spatial scale (CPC-1°, third row), the KA model's performance is enhanced further, while the OMC model is effectively no better than the zero-lag model.

To investigate the role of the temporal accumulation scale as well as the spatial scale, we can compare the interannual variance of accumulated totals over a range of accumulation period lengths (effectively comparing the variance of the PDFs in Figure 4.3 for different sub-seasonal to annual windows). Figure 4.4 shows the interannual variance for each model/location/spatial-scale as a function of accumulation period, scaled by the period length, and averaged over the annual cycle. In each of the nine plots, the zero-lag model (a basic climatological null) shows essentially no response to accumulation period; this is because

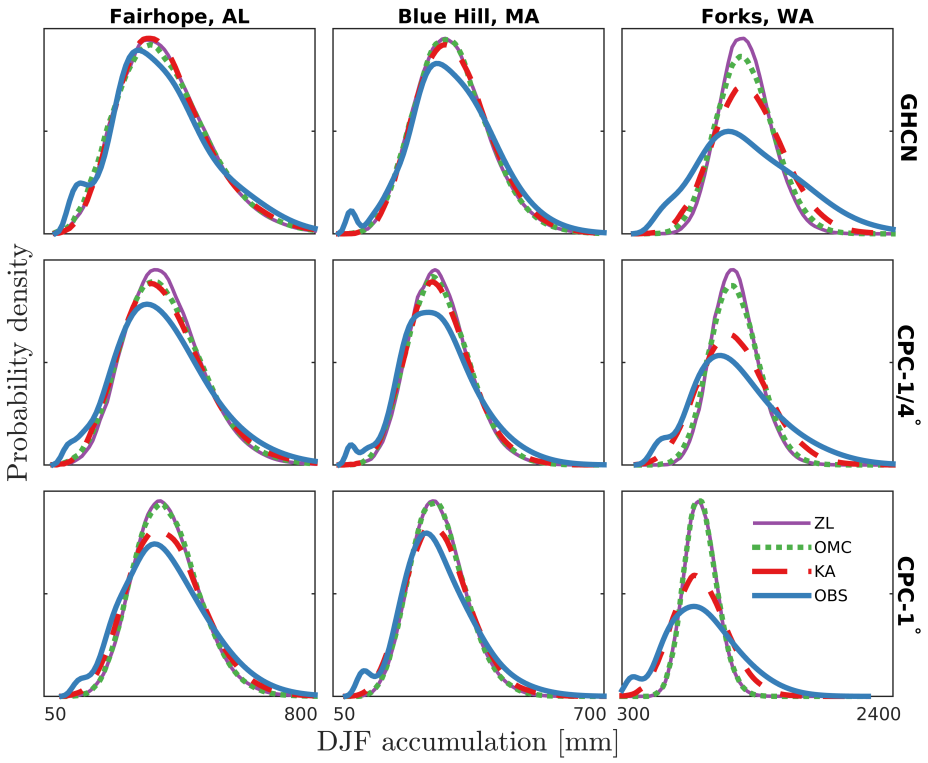

Figure 4.3: Probability density functions (PDFs) of December–February precipitation totals at Farihope, Alabama; Blue Hill, Massachusetts; and Forks, Washington: observations and three models: a zero-lag model with no daily-scale correlation byond climatology (ZL), a Markov chain based model which represents memory in occurrence processes (OMC), and the kernel auto-regressive model which represents memory in occurrence and intensity (KA). Blue lines correspond to observations (OBS) from each of the three data sets (GHCN, CPC gridded at 1/4°, and CPC gridded at 1°). All three models underestimate the variability of accumulated precipitation although they each are fit to optimally represent precipitation at the daily scale. At the wettest location (Forks, WA) and at larger spatial scales (lower two rows) the KA model's ability to represent the serial correlation in both intensity and occurrence enhances its ability to represent the 57-year distribution of winter precipitation totals.

with no serial correlation, the observations are independent, and so the variance of the sum of the precipitation is equal to the sum of the (averaged) daily variances, which is constant. These lines lie at the same value as the annual average of the 365 daily variances from the observations. In the upper two rows (GHCN and CPC-1/4°), the OMC model represents more interannual variability than the ZL model for periods longer than a single day, but at the 1°-scale (third row), the daily occurrence probability is effectively 1, and so there is no useful memory structure in occurrence for improving the multi-day variability representation. The KA model consistently outperforms the OMC and ZL models, but seems to asymptote around 30 days, while the observational data continues to increase in variability. This is not surprising, as the KA model does not represent any explicit drivers of variability at those time scales, and we know that there are earth system processes that would lead to variability at those scales (ranging in time scale from the Madden Julian Oscillation to the El Nino Southern Oscillation, multidecadal oscillations, and secular trends).

Comparing different rows, we see that there is less variability in the observations at larger spatial accumulation scales (roughly a factor of two difference in Var(OBS) between GHCN and CPC-1° at all three locations), and that the KA model represents more of the observed variability at larger spatial scales than at smaller spatial scales (by comparing the distance between the KA and OBS lines relative to the ZL null).

4.4.1 Potential Predictability

Weather generator models have previously been used as climatological nulls to estimate bounds on predictability or precipitation (the OMC model specifically, see Short Gianotti et al., 2014; Anderson et al., 2015a,b, etc.), with the potential predictability scaling with the magnitude of the interannual variability in the model:

$$\mathrm{PP}_{2w+1}(\tau) = 1 - \frac{\mathrm{E}\{\mathrm{Var}\{\Sigma_{i=\tau-w}^{\tau+w} x_{i,sim}\}\}}{\mathrm{Var}\{\Sigma_{i=\tau-w}^{\tau+w} x_{i,obs}\}} \tag{4.3}$$

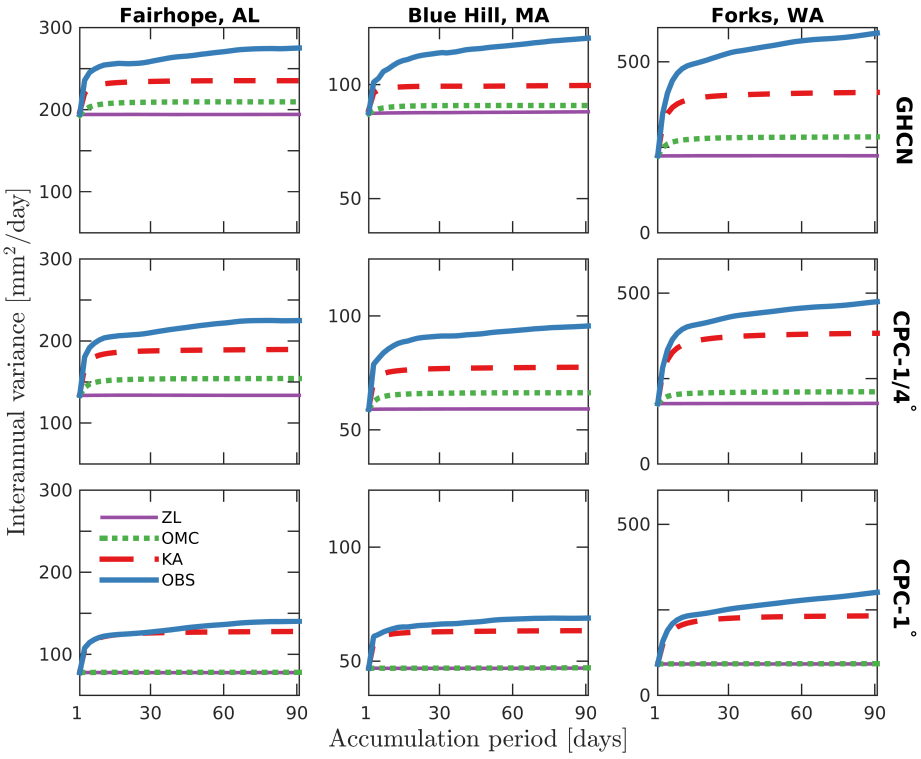

Figure 4.4: Comparison of different models' abilities to represent the variability of precipitation as a function of both spatial scale and temporal scale. As in Figure 4.3, ZL is a zero-lag model with no daily-scale correlation byond climatology, OMC is a Markov chain based model which represents memory in occurrence processes, and KA is the kernel auto-regressive model which represents memory in occurrence and intensity. Blue lines correspond to observations (OBS) from each of the three data sets (GHCN, CPC gridded at 1/4°, and CPC gridded at 1°). At larger spatial scales, the OMC model's occurrence-based memory structure is no better than the climatological null (ZL model). The KA model, alternatively, seems to represent more of the observed varaibility at larger spatial scales, suggesting that either short-term "weather-scale" variability is more dominant at larger spatial scales relative to longer-term "climate-scale" variability, or that the model fit is more effective at larger spatial scales for a fixed-length data record. Variance values are scaled by the accumulation period, and averaged across the annual cycle.

where Var$\{\cdot\}$ is the interannual variance (i.e., across years of observations or simulations), E$\{\cdot\}$ is the expectation across an ensemble of simulations, τ is the day of year, x_i are individual days of observed or simulated precipitation, and the number of sequential x_i days — $2w + 1$ — to be accumulated depends on the precipitation metric under study (typically seasonal or annual precipitation totals). Since the magnitude of the interannual variability in both the observations and the model output changes with the spatial scale at which precipitation is represented, the scale of the data will likely impact predictability estimates. Specifically, judging from Figure 4.4, both the numerator and denominator decrease at larger spatial scales, but with the denominator decreasing more rapidly, leading to smaller PP values.

More explicitly, for each day of the year we define a seasonal potential predictability index, PP_{91}:

$$PP_{91}(\tau) = 1 - \frac{\mathrm{E}\{\mathrm{Var}\{\Sigma_{i=\tau-45}^{\tau+45} x_{i,sim}\}\}}{\mathrm{Var}\{\Sigma_{i=\tau-45}^{\tau+45} x_{i,obs}\}} \tag{4.4}$$

where the 91-day precipitation totals are centered on day of year τ. This can be thought of as the fraction of the interannual variance of 91-day precipitation which is not explained by knowledge of the preceding few days' rainfall and the mean daily climatology. The periodic cycle of PP_{91} over the year are shown in Figure 4.5. The values for the OMC model using the GHCN data are effectively identical to the values shown for the same locations in Chapter 3, aside from a slightly different length of data record. The specific seasonal patterns of PP_{91} are relevant to those who are attempting to estimate seasonal precipitation (for example, the low PP in the drier summer months in Forks, WA could create high uncertainty for forecasts of seasonal rainfall which drives forest growth rates, of interest to the timber industry and the ecological modeling community), and these patterns are explored in depth for the contiguous 48 United States in Chapter 3. For the sake of this study, the most noteworthy points are 1) that PP_{91} values are similar but consistently lower using the KA method than using the OMC or ZL methods at the point scale (top row of Figure 4.5), meaning that the KA method is able to capture as much or more memory in

short-term precipitation; 2) the seasonal patterns are similar in shape (but not magnitude) across spatial scales (columns in Figure 4.5); and 3) the KA method is far more consistent in magnitude across spatial scales than the OMC or ZL methods. Importantly, as the spatial scale increases (and the frequency of precipitation occurrence increases), the OMC method becomes identical with the memoryless ZL method, meaning that it fails to represent any memory processes beyond daily climatological occurrence, rendering it effectively useless for PP studies. The KA method, on the other hand, continues to represent PP at roughly the same magnitude at all spatial scales, suggesting that it is a more appropriate technique for PP studies at spatial scales with high probability of occurrence. The seasonal pattern is fairly similar for Fairhope, AL and Blue Hill, MA for all spatial scales, suggesting some degree of spatial homogeneity at the 1° scale. The same is not true for Forks, WA, which is on the Olympic Peninsula, at low elevation on the windward side of the Olympic Mountains, and likely has a very different precipitation regime than other locations in the same 1° CPC grid box.

While seasonal accumulations are a common long-lead forecast target, the signal to noise ratio of different climate processes may peak at different accumulation time scales (Anderson et al., 2016a), and potential predictability indices can be selected to match. At annual time scales, we define the potential predictability as

$$\mathrm{PP}_{365}(\tau) = 1 - \frac{\mathrm{E}\{\mathrm{Var}\{\Sigma_{i=\tau-182}^{\tau+182} x_{i,sim}\}\}}{\mathrm{Var}\{\Sigma_{i=\tau-182}^{\tau+182} x_{i,obs}\}} \tag{4.5}$$

The PP_{365} for each τ is shown in Figure 4.6, where selection of τ (the central day of the 365-day window) allows for different definitions of the "water year." Appropriate water year selection is particularly relevant for locations with strongly seasonal climatology, where starting an accumulation period in the middle of the wet season could easily inflate interannual variability (for example, if a storm comes slightly "early" or "late" in Forks, WA using a $\tau = 183$ Julian calendar-based year); in these climates, centering the water year around the wet season is advised, both to reduce weather-based noise and to more

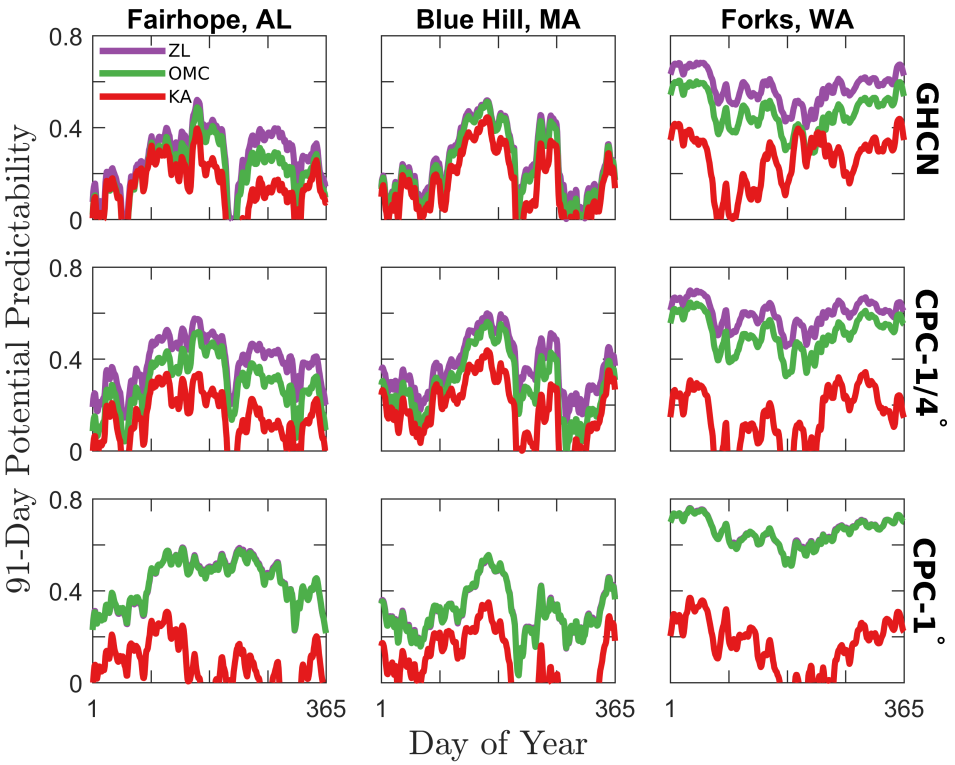

Figure 4.5: The potential predictability (PP) of 91-day precipitation totals, as defined as in Equation 4.4 at Fairhope, Alabama; Blue Hill, Massachusetts; and Forks, Washington. Lower PP values suggest a larger fraction of the interannual variability in 91-day (seasonal) precipitation totals is due to short-term ("weather time-scale") process; higher PP values suggest more variabilty due to slower variability-generating processes and hence more potential for prediction on climate time-scales. Red lines show PP over the course of the annual cycle using the kernel auto-regressive (KA) model as the no-climate-variability null, and are consistantly lower than when using the occurrence Markov chain (OMC — green) or zero-lag (ZL — purple) models at all spatial scales.

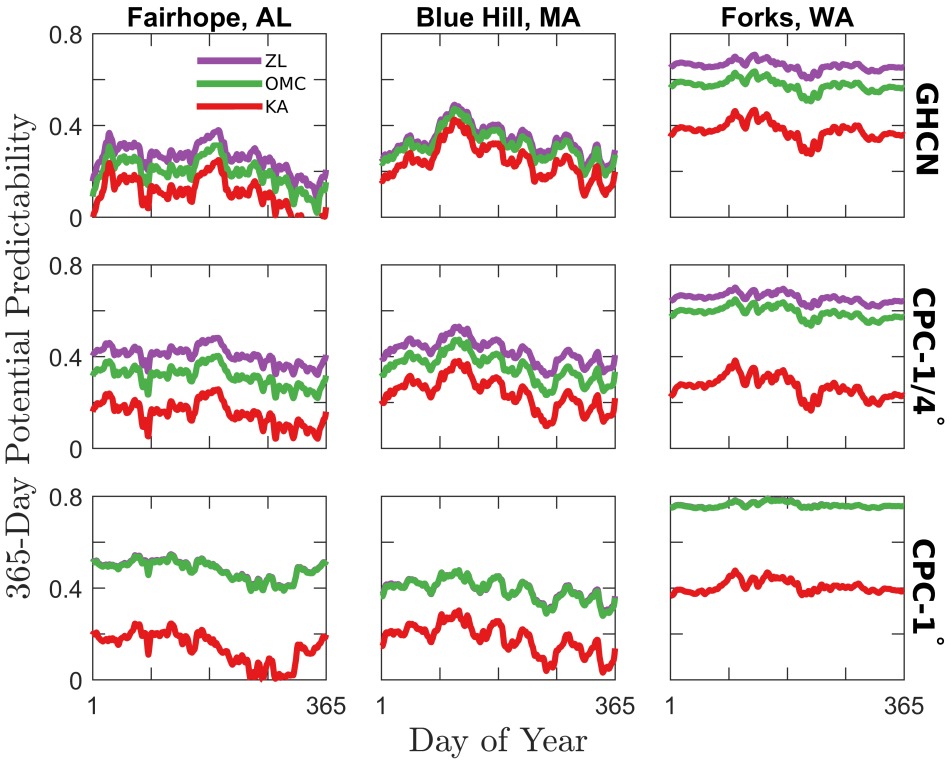

Figure 4.6: The potential predictability (PP) of 365-day precipitation totals, as defined in 4.5 at Fairhope, Alabama; Blue Hill, Massachusetts; and Forks, Washington. Similarly to Figure 4.5, the kernel auto-regressive model shows lower PP at all spatial scales than the occurrence Markov chain and zero-lag models.

accurately detect long-term changes in seasonality (e.g., Pal et al., 2013).

As in Figure 4.5, Figure 4.6 shows 1) a generally higher PP for Forks, WA than for Blue Hill, MA or Fairhope, AL; 2) a tendency for the ZL and OMC models to converge to higher PP at large spatial scales (demonstrating their inability to represent daily occurrence-based memory processes when the occurrence probability approaches 1; 3) lower PP for the KA model, suggesting better representation of memory using combined occurrence and intensity; and 4) a fairly consistent PP magnitude and seasonal pattern at all spatial scales using the KA method.

4.5 Discussion

The kernel-auto-regressive model is better able to capture the variability of accumulated precipitation than an advanced occurrence-chain-based model (OMC). Even when the OMC model was able to condition intensity on previous occurrence patterns to provide additional memory structure, the added benefits were almost never justifiable from an information criterion perspective for any day at hundreds of U.S. locations (see Short Gianotti et al., 2014). And while most of the memory in station data is in the occurrence signal (Short Gianotti et al., 2014), once we look at larger scales local occurrence information is lost, variability is reduced, and methods that represent occurrence and intensity separately will under-represent the daily correlation structure of precipitation.

The KA model's improved ability to represent daily scale process memory — particularly when the probability of precipitation is near unity — demonstrates that it can provide better estimates of potential predictability when using spatially-aggregated precipitation datasets, or when looking at regions with very high probability of daily precipitation (e.g., the tropics). This can prove useful for looking at potential predictability or the relative magnitudes of weather-scale and climate-scale processes from remotely-sensed precipitation products or the output of global climate models.

Although this research does not provide an exhaustive analysis of the relationship between potentially predictable precipitation and spatial scale, there are reasons to believe that such a relationship might exist. Climate-scale processes can act in a manner that might be predictable at one spatial scale, but not another. If, for example, heavy rainfall was tied to a decadal-scale process, but the precise location of that heavy rainfall varied from year-to-year, the process might be potentially predictable at coarse spatial scale, but not at the scale of a single weather station, increasing potential predictability with scale. Alternatively, depending on the spatial correlation scale of unpredictable atmospheric processes, integrating precipitation over large spatial extents could serve to either increase or decrease high-frequency variability, leading to lower or higher potential predictability with

increasing spatial scale.

To fully investigate these scaling relationships, the KA model could be applied to datasets at a wide range of spatial scales. There are presumably spatial scales at which interannual variability of accumulated seasonal total precipitation from the KA model reaches a maximum, at which the optimal number of lagged days changes, and at which the potential predictability peaks. The correlation structure between lagged days of precipitation is seen to change with scale in Figure 4.2, with a stronger positive correlation between adjoining days' precipitation at coarser spatial scale. It remains to be seen if this relationship is more broadly true, and what determines the exact relationship between correlations structure and spatial extent across a wide range of scales. If such a relationship could be quantified, the KA model could prove to be a useful tool for down-scaling or up-scaling of earth-system data.

Beyond its uses in establishing variances and potential predictability estimates for precipitation at varying scales, or as a stochastic weather generator model, the kernel auto-regressive model can be used to represent conditional probabilities and empirical probability distributions for any regularly-sampled variable — particularly those with some degree of auto-correlation or non-stationarity that may be poorly represented by a multivariate Gaussian correlation structure. Examples include weather generators for meteorologic variables (temperature, wind speed, pressure levels, etc.); conditional distributions of one earth system variable on another and/or their lags (evaporation given wind speed, convective available potential energy given net surface radiation, etc.); and representations of auto-correlation in non-linear biological and ecological processes (vegetation transpiration rates, population dynamics, cellular metabolic processes, etc.). The kernel auto-regressive method also provides a means of representing complex auto-correlation in advanced modeling situations, e.g., empirical distributions for use in mixed process/data models in machine learning (the KA method with Gaussian kernels is a class of Gaussian Process Models); as a means of assessing the degree of (non-)stationarity in time series analysis; and as a means of reducing assumptions inherent in common auto-regressive models (for

example, providing a means of properly dealing with all members of “Anscombe’s Quartet” and similarly devious statistical relationships).

Chapter 5

Weather noise correctly dominates climate signals in global climate model precipitation

5.1 Abstract

In the physical climate system, interannual variability in local precipitation is driven by transient atmospheric processes with meteorologic time-scales on the order of days to weeks as well as more slowly-changing boundary conditions on time-scales typically considered climatic rather than meteorologic (annual, decadal, etc.). The magnitude of interannual variability driven by daily-to-weekly scale processes is found to dominate that from lower-frequency processes in precipitation data over the contiguous 48 states from 1948–2004 in both observational data and the output from five global climate models (GCMs) in historical runs over the same period. The GCMs appropriately represent the general magnitudes of the two components over the region, and accurately represent the spatial patterns of the meteorologic-scale variability.

5.2 Introduction

Among the Frequently Asked Questions in the most recent IPCC report is, "If you cannot predict the weather next month, how can you predict climate for the coming decade?" (Kirtman et al., 2013). While the response properly highlights the difference between climate state and atmospheric state, boundary condition and butterfly effect, a reasonable follow-up question might be, "Can you predict the weather well *enough* to be confident in your

climate predictions?"

As part of the Coupled Model Intercomparison Project (most recently phase 5 — CMIP5), modeling teams ran a global climate model (GCM) experiment over the historical period from 1850–2005, tracking the observed atmospheric composition and land-cover changes for the purpose of intercomparison and validation with observational data sets (Taylor et al., 2012). The models were broadly found to represent the proper spatial patterns of mean temperature and precipitation (Sheffield et al., 2013a), with more mixed results for drivers of interannual, decadal, and secular variability (Sheffield et al., 2013b). Previous studies have established the relative magnitudes of variability in precipitation due to secular changes versus internal variability (Hawkins and Sutton, 2011), as well as the variability driven by individual annual-to-multidecadal processes (Sheffield et al., 2013b), but there has been less investigation into separating variability into what we might term "weather" scale versus "climate" scale processes.

Additionally, despite frequent success in representing climatological process, precipitation typically has lower forecast/hindcast skill than surface temperature. This has been attributed to precipitation's low signal-to-noise ratio, often taken as given due to our difficulty in modeling (although signal to noise ratios have been established for anthropogenic signals versus all internal variability (Hawkins and Sutton, 2011)). This suggests that either precipitation is inherently difficult to predict on climate time scales, or that GCMs may be recreating climatology correctly, but perhaps for the wrong reasons. To probe this question, this study asks whether the interannual variability in GCM precipitation is driven by an appropriate mix of weather-scale versus climate-scale processes, and on the path to this answer determines what this "appropriate" mix is from observational data.

In particular, for the sake of predictions and forecasts, we would like to know that GCMs are representing low-frequency "climate-scale" variability (here defined as annual and longer) properly. It is well-known that some of this variability is effectively stochastic — that is driven by high-frequency "weather-scale" processes (transient atmospheric conditions) — and that some is driven by more persistent low-frequency boundary condi-

tions such as sea-surface temperatures, greenhouse gas concentrations, and decadal-scale climate modes (Mann and Lees, 1996; Feldstein, 2000; Short Gianotti et al., 2014). This is important for GCM forecast use, as improper calibration of weather and climate in climate models can lead to false attribution of extreme events (Bellprat and Doblas-Reyes, 2016), and extremes are the primary drivers of global changes in precipitation (Trenberth, 2011).

Using a method developed for estimating the stochastic component of coarse-scaled precipitation data, the interannual variability of accumulated precipitation totals (over seasons or years) can be separated into a stochastic weather-scale component and a climate-scale component that can not be explained by daily- or seasonal-scale processes (see Chapter 4). This allows us determine the interannual variability of precipitation in GCM-modeled datasets due to climate-scale processes and that due to weather-scale processes. These climate signal and noise components are then compared to observational data for validation.

5.3 Data

Our goals in this study are 1) to determine the degree to which interannual precipitation variability is driven by weather-scale versus climate-scale processes, and 2) to determine whether GCMs display the same balance between these time scales as observational data. Our observational dataset is the NOAA Climate Prediction Center's Unified Daily Precipitation data (Chen and Xie, 2008), and our GCM data sets are the CMIP5 baseline historical runs. The Climate Prediction Center's (CPC) Unified Precipitation dataset provides a gridded precipitation product using roughly 8,000 weather stations daily for the period 1948–1992 and roughly 13,000 weather stations daily after 1992 (Xie, 2010). Station data is quality controlled for duplicates and overlapping stations, and stations are checked for anomalous standard deviations and compared against nearby "buddy stations" prior to gridding. Station data is gridded to a $1^\circ \times 1^\circ$ from 140°W–60°W, 10°N–60°N using a modified Cressman scheme.

Model Abbreviation in Paper	**Modeling Center/Group**	**Institute ID**	**Model Name**
CCSM	National Center for Atmospheric Research	NCAR	CCSM4
CSIR	Commonwealth Scientific and Industrial Research Organization in collaboration with Queensland Climate Change Centre of Excellence	CSIRO-QCCCE	CSIRO-Mk3-6-0
GE2R	NASA Goddard Institute for Space Studies	NASA-GISS	GISS-E2-R
GFDL	NOAA Geophysical Fluid Dynamics Laboratory	NOAA GFDL	GFDL-CM3
MIR5	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	MIROC	MIROC5

Table 5.1: Global Climate Models used in this study. Data from each model is from the CMIP5 baseline historical run from 1948–2004.

GCM data comes from the baseline historical runs of five models from the CMIP5 archive as shown in Table 5.3 (Taylor et al., 2012). Simulations are performed in each model's native resolution, as are accumulated total and interannual variability calculations. Comparisons between individual models and the CPC dataset are performed in the GCM's native resolution by calculating all quantities in the native resolution of the parent dataset, then regridding the quantity from the CPC dataset to the GCM grid using two-dimensional linear interpolation. Comparisons of errors (differences between GCM and CPC data) *between* GCMs are performed on the CPC grid by performing all possible calculations in the native resolution, then regridding each GCM quantity to the CPC resolution to calculate differences (again using two-dimensional linear interpolation).

5.4 Methods

Previous studies have shown that there is substantial memory in daily-scale precipitation at spatial scales ranging from weather stations to coarsely-gridded products (Mandelbrot and Wallis, 1968; Gianotti et al., 2013; Short Gianotti et al., 2014; Anderson et al., 2015a, and Chapter 4 of this document). While the memory structure can be highly non-linear (particularly at small scale or in dry regions), sufficiently flexible (empirical) models of auto-correlation processes can represent this memory in both intensity and occurrence across spatial scales (see Chapter 4). We apply the method of Gianotti et al. (Chapter 4) to estimate the stochastic component of interannual variability from accumulated precipitation totals in a null-climate probability framework.

By fitting empirical probability models for each day of the year and marginalizing over years to effectively ignore any interannual drivers of climate variability, we can create a stochastic model of daily-scale precipitation that exactly matches the daily probability density (PDF) of precipitation for every day of the year, and thereby empirically matches the daily mean, variance, and further moments. By further representing precipitation as an n-dimensional probabilistic process, using $n-1$ days of lagged values, we can represent the empirical PDFs of memory, including the probability of precipitation on a given day conditioned on $n-1$ days of previously-observed precipitation. We fit these empirical probability models for each location in a dataset, selecting the appropriate number of lags and a kernel bandwidth for each day using cross-validation, and then use the model as a weather-generator to simulate 1,000 representations of the historical precipitation record. As in Gianotti et al. (Chapter 4), the lag and bandwidth selection occurs in a rank-based inverse normal transformed space (Akritas, 1990; Cai et al., 2016) determined separately for each day of the year, using all of the observed values within a 25-day pooling window for determining correlation structure. The probability model is represented in the transformed space as a Gaussian process using symmetric Gaussian kernels (Rasmussen and Williams, 2004). Simulations are back-transformed through the reverse CDF transformation. Indi-

vidual simulations within the 1,000-member ensemble may vary slightly from the historical mean (and further moments) at the daily, seasonal, and annual scales due to sampling variability.

Each ensemble member at each location has some degree of interannual variability due entirely to stochastic processes with 0–10 days of memory. The simulations as an ensemble match the historical mean precipitation at daily and longer time-scales, and the interannual variability matches the observed interannual variability at daily time-scales. The interannual variability for longer (e.g., seasonal or annual) accumulation periods, however, is generally smaller (under-dispersed) in the simulations relative to the observed data because the simulations have no drivers of interannual variability other than 0–10 day autocorrelation processes. The "historical" data (both based on actual observations and based on the GCM runs directly), in contrast, has both a stochastic high-frequency component due to processes at weather time-scales and low-frequency variability due to changes in the atmospheric boundary conditions.

For each of our six historical datasets (one observation-based: CPC; and five GCM-based: CCSM, CSIR, GE2R, GFDL, MIR5), we simulate 1,000 high-frequency-only recreations of the historical data and calculate summary statistics (seasonal mean precipitation, interannual variability of seasonal totals) using 91-day accumulation periods centered on January 15 (DJF), April 15 (MAM), July 15 (JJA), and October 15 (SON). This yields a single set of summary statistics for each location and season of the historical datasets, and a distribution of 1,000 summary statistics for each location and season of the simulated datasets. By subtracting the simulated seasonal interannual variances from the historical interannual variance, we can calculate a residual variance that is not well-explained by our high-frequency-only simulations:

$$\sigma^2_{\text{residual}} = \sigma^2_{\text{historical}} - \sigma^2_{\text{simulated}} \tag{5.1}$$

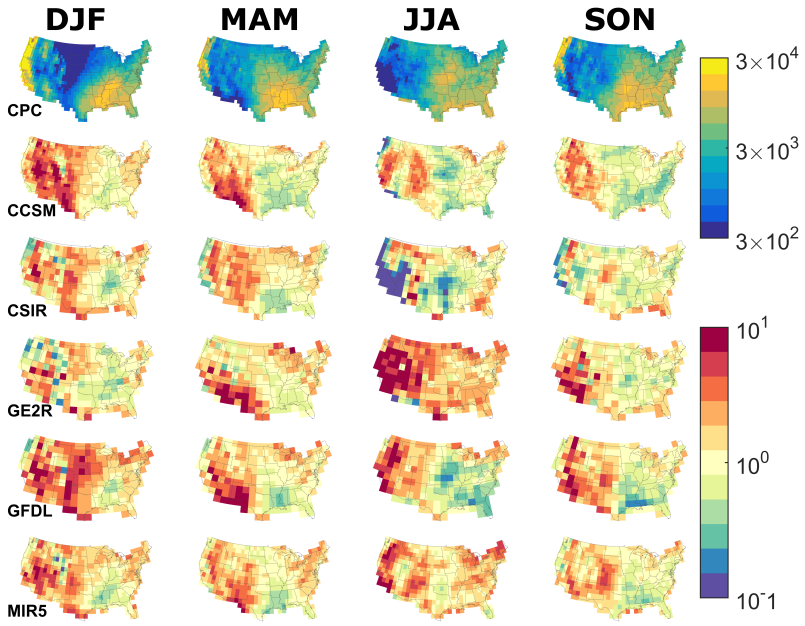

Figure 5.1: Interannual variance of 91-day (seasonal) accumulated precipitation from one observational dataset and historical runs from five GCMs from 1948–2004 [mm^2]. GCM variances (lower color bar) are shown as log-scale anomalies from the CPC seasonal values (values greater than 1 imply more interannual variability in the historical GCM dataset than the historical CPC dataset).

or, equivalently,

$$\sigma^2_{\substack{\text{low}\\\text{frequency}}} = \sigma^2_{\text{historical}} - \sigma^2_{\substack{\text{high}\\\text{frequency}}} \tag{5.2}$$

which provides our means for separating climate-scale variability from weather-scale variability. The "historical" variance, the distribution of "simulated" (high-frequency, weather-scale) variances, and the distribution of "residual" (low-frequency, climate-scale) variances are the primary analysis variables of this study.

5.5 Results

The mean seasonal precipitation values for each data set are shown in Supplementary Figures E.1 and E.2, with the broad spatial patterns well represented by the GCMs (mean across all GCMs and seasons of correlations between GCM and CPC: $r = 0.786$, see

Supplementary Table E.1), albeit with consistent over-estimates of precipitation in the arid West and fairly consistent under-estimates in the Southeast relative to the CPC dataset. The interannual variability of the seasonal precipitation totals for each historical dataset is shown in Figure 5.1, with the GCM data shown as anomalies relative to the CPC dataset. The spatial patterns of interannual variability in the historical GCM data show most of the same coarse structure as the observational CPC data (see Supplementary Figure E.3), but with inflated levels of variability in the West clearly dominating the error anomaly structure for most models in most seasons. Despite the clear biases regionally, the spatial patterns remain well-correlated with the CPC data (mean Pearson correlation coefficient of $r = 0.651$ across all seasons, all GCMs), but less so than the mean correlation, suggesting a tendency towards climatology in the GCMs.

Figure 5.2 shows the median interannual variability from 1,000 stochastic simulations of precipitation in which all variability comes from processes with memory on time-scales of days, which we refer to as "weather-scale." The daily memory structure in the weather-scale simulations follows the empirical conditional distribution of precipitation for each day in each location, but has no other interannual variance generation mechanism (no "climate-scale" processes). As in Figure 5.1, the weather-scale variability for the GCM datasets is shown as anomalies from the CPC data (see Supplementary Figure E.4 for non-anomaly data), and black dots represent locations in which the 1,000-member distribution of weather-scale variances from the CPC simulations is statistically indistinguishable from the distribution of weather-scale variances from the GCM simulations using the Wilcoxson rank-sum test ($\alpha = 0.05$). Statistical indistinguishability is rare for most models in all seasons due to anomalies of the same order as in Figure 5.1. The spatial patterns of both the CPC weather-scale variability and the GCM anomalies are also very similar to the variability in the historical data (Fig. 5.1), demonstrating that weather-scale processes contribute a substantial fraction of the historical variability. The spatial correlations between GCM and CPC weather-scale variability are higher for every model in every season than the for the historical variability (with a mean over all models and all seasons of $r = 0.731$,

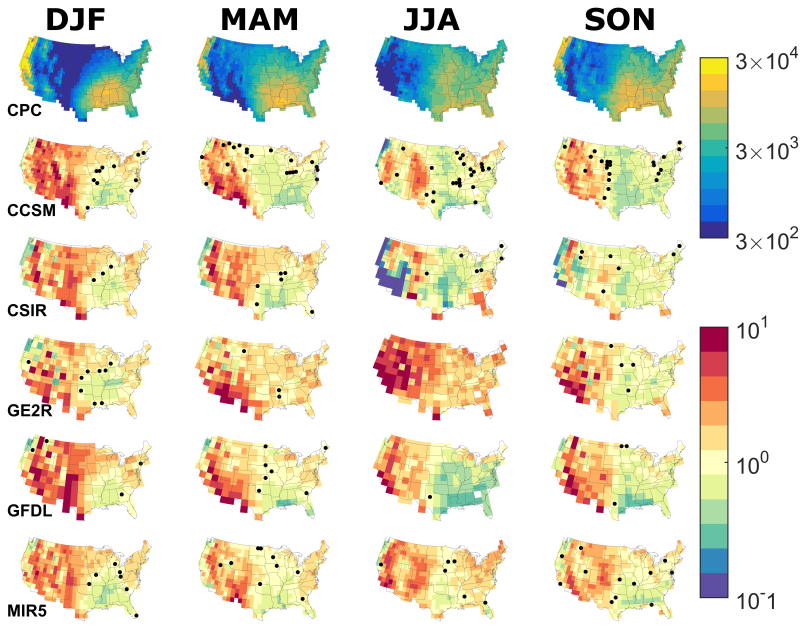

Figure 5.2: Interannual 91-day (seasonal) precipitation variance attributable to weather-scale phenomena (1948–2004) [mm^2]. Variances shown for the CPC dataset are the median values of 1,000 stochastic weather-generator type datasets as fit to the historical data. Interannual variability is generated only by stochastic processes with memory on the order of days. Values shown for the five GCM datasets (lower color bar) make use of variances generated in the same manner for each weather-scale-only GCM dataset as for the weather-scale-only CPC dataset, but are shown as log-scale anomalies from the median CPC variance (values greater than 1 imply more interannual variability due to weather-scale processes in the GCM dataset than the CPC dataset). Black dots show grid cells in which the distribution of 1,000 variances from the GCM dataset is statistically indistinguishable from the distribution of variances from the CPC dataset (Wilcoxson rank-sum test, $\alpha = 0.05$).

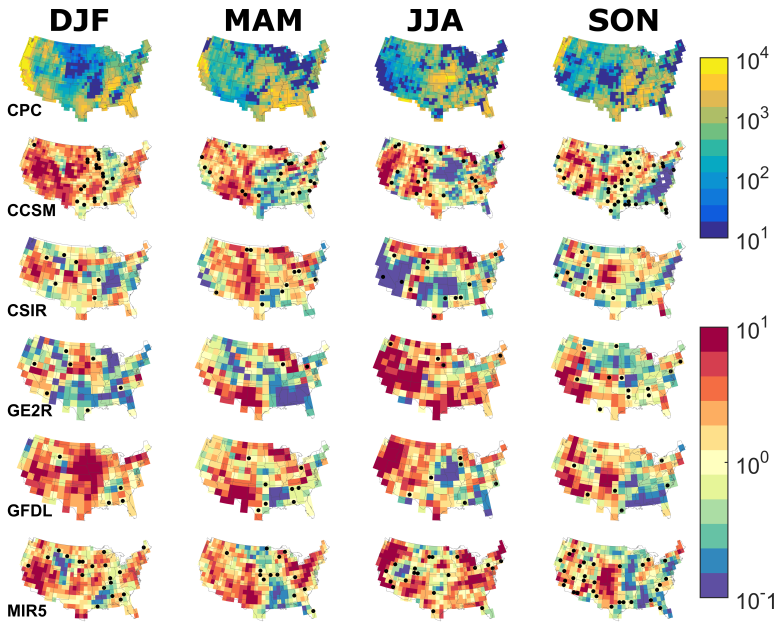

Figure 5.3: Interannual 91-day (seasonal) precipitation variance attributable to climate-scale phenomena (1948–2004) [mm^2]. Variances shown for the CPC dataset are the median values of 1,000 residual variance terms (Median$\{\sigma^2_{\text{historical}} - \sigma^2_{\text{simulated}}\}$), and represent variability that cannot be captured by a stochastic model with memory on the order of days. Values shown for the five GCM datasets use the corresponding median residual variances for each GCM dataset, but are shown as log-scale anomalies from the median CPC climate-scale variance (values greater than 1 imply more interannual variability due to climate-scale processes in the GCM dataset than the CPC dataset). Black dots show grid cells in which the distribution of variances from the GCM dataset is statistically indistinguishable from the distribution of variances from the CPC dataset (Wilcoxson rank-sum test, $\alpha = 0.05$).

see Supplementary Table E.1), implying that weather-scale processes may in fact be more accurately represented in the GCM runs than lower-frequency climate-scale processes.

This is demonstrated in Figure 5.3, which shows the climate-scale variability, determined as residuals from the historical and weather-scale variability. The magnitude of climate variability is substantially lower than that of weather variability, both in the observational CPC data set and the GCM datasets (note the difference in colorbar scale between Figs. 5.2 and 5.3), and less correlated to mean climatological precipitation as well (mean across-dataset, across-season spatial Pearson correlation between mean precipitation and historical variability: 0.740; between mean precipitation and weather-scale

variability: 0.819; between mean precipitation and climate-scale variability: 0.295), as would be expected in an environment dominated by weather-scale processes. Beyond this, the magnitudes of anomalies between the GCM and CPC datasets are more pronounced for climate-scale variability than for either weather-scale or historical variability, and across-model, across-season averaged spatial correlation is accordingly lower ($r = 0.564$). The lower correlation between climate-scale GCM/observational variability than weather-scale GCM/observational variability is not of specific concern with regard to GCM fidelity, as the climate-scale processes are not prescribed in the historical runs (e.g., a GCM may have a different number of ENSO events than observations).

Together these results point to a precipitation regime in the contiguous U.S. that is dominated by weather-scale processes; interannual variability attributable to weather-scale processes is more correlated to mean precipitation than total interannual variability in the historical datasets for both the observational data and global climate models. The median portion of interannual variability explained by weather-scale processes across the region is between 72% (winter) and 84% (summer) in the observational CPC data (with a range from 62%–95% for all seasons and models in the GCM datasets, see Supplementary Figure E.8 for maps of climate-scale variance fractions). Figure 5.4 shows three Taylor diagrams for GCM performance relative to the CPC (reference) dataset. Across all seasons, the GCM weather-scale variability (Figure 5.4b) more closely matches the observations than does either the climate-scale variability (Figure 5.4c) or the total historical variability (Figure 5.4a), in terms of spatial patterns (correlations) and magnitude (RMSE).

5.6 Discussion

The historical runs from global climate models are imperfect in their representations of both mean precipitation and interannual variability (particularly in the arid West), but broadly capture the spatial patterns of both. We would expect spatial patterns of bias in interannual variability to be similar to those in mean precipitation due to the relationships

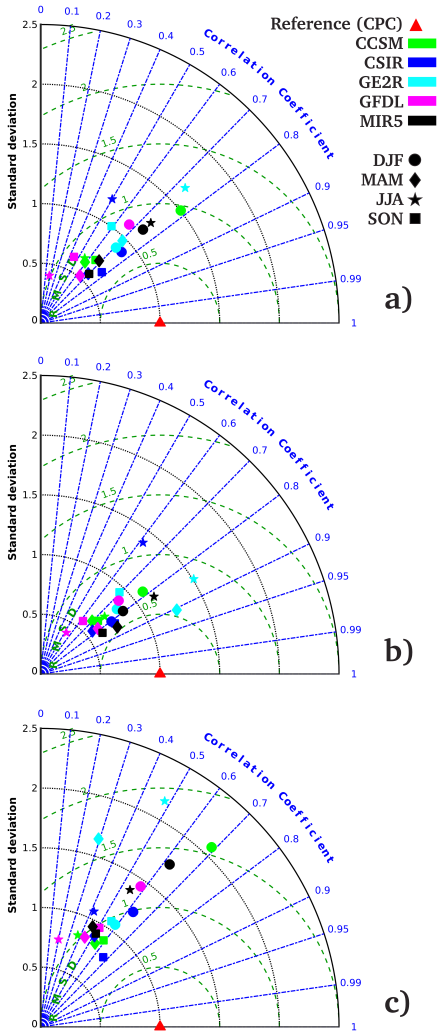

Figure 5.4: Taylor Diagram of grid-cell-wise a) historical, b) weather-scale, and c) climate-scale variability for each season, using CPC data as reference. GCM standard deviations are normalized to the grid-cell-wise standard deviation (of interannual variances) of the CPC dataset, and are shown as radial coordinates. Correlation coefficients (angle coordinate) are determined grid-cell-wise between the GCM dataset and the CPC dataset, as are RMSE values (green dashed circles). Points closer to the CPC reference (red triangle) are closer to matching the reference data. Notably, the GCMs demonstrate higher correlations and lower RMSE for the weather-scale variability than for either climate-scale variability or the total historical variability, suggesting that the spatial climatology of hydro-meteorological processes may be better represented than the spatial climatology of hydro-climatogical variability in the GCMs. Beyond that, the spatial patterns in weather-scale variability are somewhat under-dispersed (more standard deviations values below 1) — perhaps not surprising given the inflated variability (and mean precipitation) in the arid West in GCMs — and the spatial patterns in climate-scale variability are somewhat over-dispersed (more standard deviations values above 1).

between means and variances in typical distributions used for probabilistic representations of precipitation, which is the case for variability attributable to weather-scale processes, and — to lesser degrees — total historical variability and climate-scale variability.

Stochastic weather-scale variability is found to be the major component of interannual variability in the contiguous U.S., accounting for roughly 80% of the total interannual variability, which means that on annual and longer time-scales the noise dominates. The magnitude of weather-scale variability was generally larger than that of climate-scale variability by a factor of 4:1 (inverse of the climate signal-to-noise ratio). This corresponds to potential predictability values of 20% for both the observational data and the GCMs (Supplementary Figure E.8). It is worth noting that this is true for a period in which statistically significant (increasing) precipitation trends have been reported (Karl and Knight, 1998) and in which precipitation intensity has been found to be shifting towards heavy events (Karl et al., 1995). This highlights the importance of sufficiently long data records for detection and attribution studies as well as providing a reminder of the importance of accounting for high-frequency stochastic drivers of low-frequency fluctuations and trends (e.g., Hamed and Rao, 1998), particularly in scenario evaluations and projections on interannual-to-decadal horizons.

We find that GCMs are reasonably well-correlated with observations in their representations of interannual variability, but are better-correlated in their representations of interannual variability driven by purely weather-scale processes. Despite this, there are relatively few locations in which the weather variability is a statistically ideal match with the observational data, primarily occurring in regions of mid-level variability, demonstrating a tendency for GCMs to under-represent the spatial variability of temporal-variability generating processes. This may be either a driver of or a symptom of the biases in mean precipitation (towards too much precipitation in arid regions and too little in wet regions).

We also find that global climate models' representation of weather-scale variability is more closely correlated with that of the observational data than is climate-scale variability. To some degree, this can be explained by sampling variability in quasi-periodic climate

fluctuations (e.g., ENSO) in the GCM datasets, but it may also lead to areas in which improvements can be made in GCM hydroclimate fidelity.

5.7 Acknowledgements

We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table 5.3 of this paper) for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

We would also like to thank Dara Entekhabi for his support during the writing of this manuscript.

Chapter 6

Conclusions and future directions

This dissertation provides methods to characterize variability in precipitation data on weather time-scales, and applies these methods to determine signal and noise characteristics of precipitation on climate time-scales.

Chapter 2 expanded the Markov chain model for precipitation occurrence to provide a method by which to represent short-term memory processes in precipitation occurrence data. The results of this chapter determined the spatial and seasonal patterns of potential predictability in precipitation occurrence, with notably higher PP in the winter months, and low PP in the western U.S. in the summer months.

Chapter 3 combined the method in Chapter 2 with a precipitation intensity model for a more complete occurrence/intensity stochastic representation of precipitation and then further decomposed the potential predictability of precipitation into a total value, and a contribution from occurrence and from intensity.

Chapter 4 developed a new method for a stochastic representation of auto-correlated precipitation using a conditional kernel density method. This provides a means to treat precipitation occurrence and intensity as a single process, which is necessary for inter-comparison of weather-scale and climate-scale variability on multiple spatial scales. The method shows substantial improvement over earlier methods in ability to capture variability from high-frequency (weather-scale) processes when using gridded precipitation products. This is particularly necessary when analyzing coarse-scale data, such as the output from global climate models (GCMs).

Chapter 5 applies the method of Chapter 4 to an observational data set as well as

the historical runs from five GCMs. The weather-scale variability is found to be the source of roughly 80% of the total interannual variability in precipitation data, both in the observational data and the GCM data. Broadly, the GCMs represent the weather-scale variability well, including spatial patterns, although there is some definite bias towards mean values of precipitation.

The introduction of this dissertation proposed three research questions:

1. **How can we best model daily precipitation as a null model for climate variability studies?**

2. **Can we quantify the magnitude of the stochastic component of precipitation on climate time-scales?**

3. **How predictable is precipitation beyond the window of atmospherically-driven numerical weather predication?**

Chapters 2–4 propose methods to answer the first question, for station based precipitation occurrence, station based total precipitation intensity, and coarse-scaled precipitation from multiple data sources.

Chapters 2–4 also serve to answer the second question, by using the null models as a representation of the weather-scale component of interannual variability. This variability is effectively stochastic at climate time-scales, and represents an envelop of noise which partially masks climate-signal detection.

Finally, each of the chapters answer the third question by providing estimates of the potential predictability of precipitation on annual and longer time-scales. As a rough summary, perhaps a quarter of the interannual variability in precipitation can be potentially tied to climate-scale processes over the last 60–120 years in the contiguous 48 United States, but this value has a quite large range with season and location, from essentially 0 to 70%.

The methods used in this paper have also been used beyond the estimation of the potential predictability of seasonal precipitation totals. Pal et al. (2013) uses a weather-scale null model to establish significance of changes in phase and seasonality of the annual

precipitation occurrence cycle. Anderson et al. (2015a) uses the simulations from Chapter 3 to establish the potential predictability of heavy precipitation and dry spells, as does Anderson et al. (2015b) to establish a null model for trend detection. Anderson et al. (2016b) follows on these to calculate the spectral signature of potential predictability, and Anderson et al. (2016a) uses weather noise simulations to establish the significance of a new decadal pattern of atmospheric pressure.

Since the modeling frameworks used in each chapter are inherently probabilistic, they can be used either as weather simulators (as in Chapters 2–5), or as likelihood models for establishing anomalous events. As an example, the recent drought in California was by many metrics record breaking (Salvucci et al., 2014), but to answer whether or not it was statistically unlikely, one could determine its likelihood from the weather-generator frameworks of Chapters 2–4 and compare that likelihood against the distribution of observed likelihoods (or *entropy*). A particularly unlikely event might be considered anomalous from within a climate-stationary modeling perspective, suggesting that it was likely driven by climate-scale processes, which could be determined through physical or statistical attribution analysis.

The method of Chapter 4 can be extended to include predictors of precipitation beyond lagged precipitation itself. One future application is to use this method, but with climate indices as possible predictors to attribute specific short- or long-term precipitation patterns to certain states of the climate system. Beyond climate indices, local system states such as soil moisture (for land-atmosphere feedbacks), meteorology (e.g., pressure, temperature, convective available potential energy), or land surface cover could serve as predictors which would allow researchers to determine distributions of likely precipitation conditioned on current conditions, both for forecast uncertainty and for determining drivers of variability in precipitation processes.

The kernel auto-regressive method is also not explicitly tied to precipitation data. It can be used generally as a machine learning technique, but is particularly well-suited for representation of bounded and/or inflated variables. Work in progress suggests that the

method can be used for estimating land surface fluxes (specifically evapotranspiration) from surface and atmospheric conditions; and with sufficient data, the cross-validation method of dimension selection can be used for variable selection (in the same manner as likelihood tests, information criteria, and stepwise regression methods). Although the work in this dissertation has focused on variances under null assumptions of climate stationarity, the method is more likely to see use in a hidden variable estimation setting, or as an update step in a data assimilation or Bayesian forecast system, since observed predictors yield conditional distributions of the forecast variables.

As well as time-series analysis, the kernel auto-regressive framework can be used broadly as a non-linear empirical description of probabilistic relationships between related covariates. As an example, co-measured values from a multi-spectral imager could be used as predictors of a vegetation index (based on some training data), and the model could be used for spatial/temporal extrapolation; or the same lagged indices along with ground-based nutrient measurements could be used as predictors to determine the effects of nutrient "innovations" in a Granger Causality sense.

For hydroclimate in particular, potential predictability could be expanded to include spatial representations of processes, which may suggest avenues of prediction. The results of Chapter 5 should be expanded to global scale, with temperature as well as precipitation (perhaps with direct representation of their covariance using the kernel auto-regressive model as well). And, finally, by making use of water and energy conservation, multiple variables may be able to be more tightly constrained in surface process modeling.

Appendix A

Supplementary material for JCLI1

A.1 Chain Orders

The order, m, of the models is the number of days of precipitation data used in determining the probability of rain on a given calendar day. A first order model has two states for each day of the year; one state in which it rained on the previous day, and one in which it did not, each with a different associated probability of rain for the present day. Let $X(t, d)$ denote a discrete daily time series of precipitation occurrence, where t is the day of the series and $d \in \{1, \ldots, 365\}$ is the calendar day ($X(t, d) = 1$ represents a rainfall event and $X(t, d) = 0$ represents no rain); and let $x_1, x_2, x_3, \ldots$ represent the events of days $t - 1$, $t - 2$, and $t - 3$ respectively. If we allow our models statistical behavior to depend only on the day of year d and the events of the previous m days, then we can express the probability of rain on any given day as

$$\Pr\{X(t, d) = 1\} = \Pr\{X(t, d) = 1 | x_1, \ldots, x_m\}, \tag{A.1}$$

known as the Markov property. For each day of the year, d, the probability of rain can be written in shorthand, then, as $p_{x_m \ldots x_1}(d)$, so that $p_{11}(5)$ is the probability of rain on Jan 5 given that it rained on both of the previous days. This second order model has four states for each day of the year, corresponding to the four possible two-day histories, and four independent transition probabilities $\{p_{00}, p_{01}, p_{10}, p_{11}\}$ to be calculated (probabilities of dry days can be calculated from the probabilities of wet days). An m^{th}-order model has,

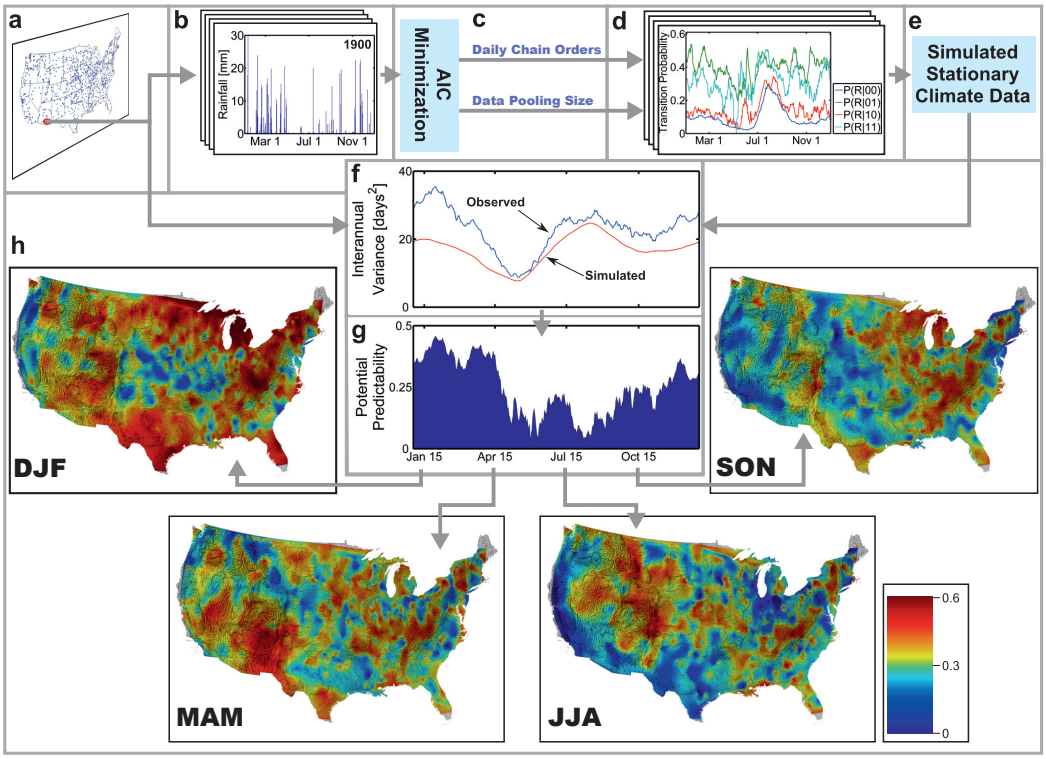

Figure A.1: The method for determining a daily time-series of potential predictability (PP) for an example station (Tuscon, AZ). Precipitation data from each of 774 weather stations (a) are used to create annual time series of precipitation occurrence (b). The Akaike Information Criterion (AIC) is used to select the model pooling size and daily chain orders that optimize the likelihood/parsimony trade-off (c). Transition probabilities (d) are established using results from c and are used to create a stationary chain model with no low-frequency drivers of variability. This model is in turn used to create 1,000 stochastic simulations of $\sim$ 100 years of precipitation data (e). The simulated data typically has lower interannual variability than the observed data due to the models ignorance of all longer-than-weekly scale processes (f). The difference, $\sigma^2_{89,obs}(t) - \left\langle \sigma^2_{89,sim}(t) \right\rangle$, as a fraction of $\sigma^2_{89,obs}(t)$ is the potential predictability (g) — the variance unexplained by stochastic processes alone (shown on maps in h).

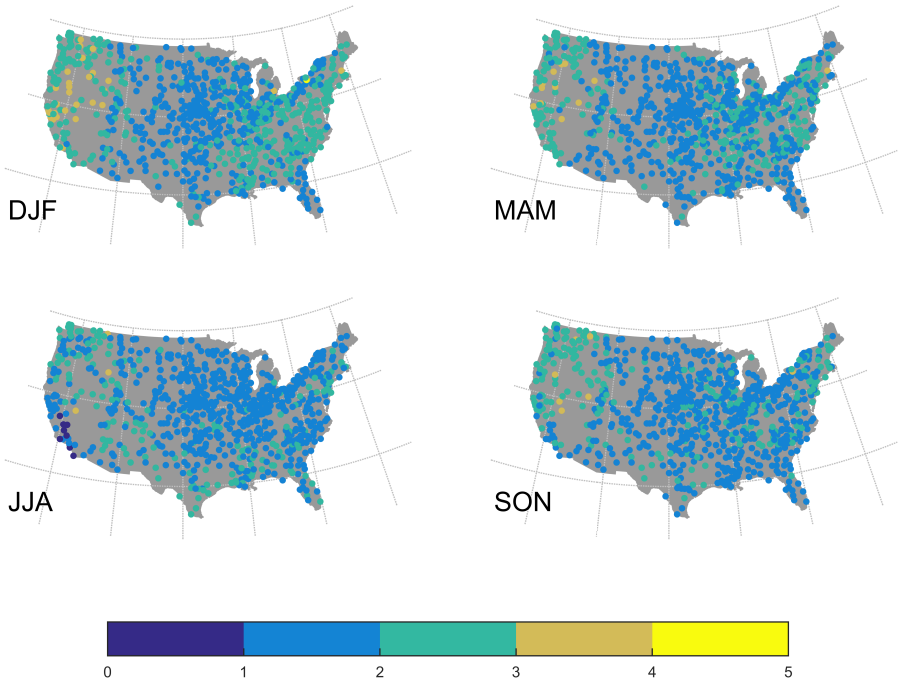

Figure A.2: Mean chain orders for each station for December–February, March–May, June–August, and September–November. Values shown are the mean of the roughly 90-daily values in each season. All model fitting and simulation uses the daily values explicitly. The highest mean values are found in the mountainous West, particularly in the winter, and just south of the Great Lakes.

similarly, 2^m states and independent transition probabilities for each day of the year.

A.2 Data Pooling

To determine the transition probabilities to be used in our chain models (Katz, 1977; Wilks and Wilby, 1999), we allow pooling of our training data to improve the estimation of extreme (rare) events, again using the AIC for parameter selection. A pooling size of 1 day is equivalent to using the historic data with no alteration, while a pooling size of 5 days uses all observations within a 5-day window (the central day ± 2 days on either side) as historical instances of the central day when looking for precipitation patterns. If the historical record spans T years and the pooling size is b days, then the effective pooled

record spans bT years, more appropriately representing the frequency of rare events in the training data. We again use the AIC to optimize the trade-off between complexity (a pooling size of 365 days removes all seasonality) and goodness-of-fit (larger pooling sizes increase our confidence in parameter estimation).

A.3 Akaike Information Criterion (AIC)

To select the appropriate model for each day of the year, models are ranked using the corrected Akaike Information Criterion (AIC_c) for each day (Hurvich and Tsai, 1989):

$$\mathrm{AIC}_c = \frac{2kN}{N-k-1} - 2\ln l \tag{A.2}$$

where k is the number of parameters (2^m), N is the number of (pooled) observations for that day (bT if no missing data), and l is the likelihood of witnessing the observed distribution of occurrences given each of their m-day histories (the product of N corresponding transition probabilities). By selecting the chain order (using $0 \leq m \leq 5$) with the minimum AIC value (Fig. A.5a), we create a variable order chain model for the year for a given pooling size (Katz and Parlange, 1993). Letting the pooling size b be an odd integer $1 \leq b \leq 119$, we create 60 models with different pooling sizes each with different progressions of chain orders throughout the year, and different corresponding transition probabilities for each day (Fig. A.5b). To compare between the pooling sizes, we again use the AIC, in a slightly rewritten form to account for the comparison in sample size (Fig. A.5c–e):

$$\mathrm{AIC}^* = \frac{2}{b}\left[K_b^* - L_b^*\right] \tag{A.3}$$

where

$$K_b^* = \sum_{i=1}^{365} 2^{m_{i,b}} \tag{A.4}$$

is the sum of the number of parameters (independent transition probabilities) necessary for the annual variable-order model with pooling size b, and

$$L_b^* = \sum_{i=1}^{365} \ln l_{i,b} \tag{A.5}$$

is the sum of the log-likelihoods associated with observing the (pooled) historical data for each day based on the appropriate transition probabilities. The pooling size with the minimum AIC* value is selected (Fig. A.5e), and the associated variable-order chain model is used as the basis of our simulation (Fig. A.5f).

Since the first and last years of the observed data are needed for edge effects with the pooling and m-day histories, they are discarded from the analysis and simulation. 1,000 stochastic precipitation occurrence simulations, each of length $T-2$ years, are created one day at a time using the daily transition probabilities and the previous m days of simulated data for each of the 774 stations. To allow a more direct comparison between the simulated and observed data, any missing values in the observed datasets are imputed using the simulation models.

For each of the simulated and observed time series, the number of rainy days within a seasonal 89-day window (± 44 days) is computed for each day. These seasonal occurrence counts form a new $365(T-2)$ day time series, and interannual variability of these counts is calculated for each day of the year for each time series. The interannual variance of the observed time series,

$$\sigma^2_{89,obs}(t) = \mathrm{Var}\left[\sum_{i=-44}^{44} X_{obs}(t+i)\right], \tag{A.6}$$

is compared with the mean of the interannual variances of the simulated time series (Fig. A.1f),

$$\left\langle \sigma^2_{89,sim}(t) \right\rangle = \left\langle \mathrm{Var}\left[\sum_{i=-44}^{44} X_{sim}(t+i)\right] \right\rangle, \tag{A.7}$$

and the potential predictability (Fig. A.1g) is calculated as

$$\mathrm{PP}_{89}(t) = \frac{\sigma^2_{89,obs}(t) - \left\langle \sigma^2_{89,sim}(t) \right\rangle}{\sigma^2_{89,obs}(t)} \tag{A.8}$$

Individual components of potential predictability are shown in Figure A.3, and the timings of maximum and minimum potential predictability are shown in Figure A.4.

A.4 Statistical significance

To assess the statistical significance of the PP_{89} signal as shown in Figure 2.1 (Chapter 2), we created distributions of PP_{89} values using the individual simulations in place of the observed time series. A single $\mathrm{PP}_{sim,i}$ value was determined as

$$\mathrm{PP}_{sim,i} = \frac{\sigma^2_{sim,i} - \left\langle \sigma^2_{sim} \right\rangle}{\sigma^2_{sim,i}}, \tag{A.9}$$

where $\left\langle \sigma^2_{sim} \right\rangle$ is the ensemble mean of the 1000 simulations for a given station. Using all 1000 values for i, we create a distribution of simulated PP values from which we can calculate the statistical significance of our observationally-based PP.

Our null assumption is that the observed variability is stationary and not significantly different from the distribution of simulated variabilities, i.e. that $\mathrm{PP} = 0$. This means that areas with low p-values (statistically different from the null) have potential predictability. Those areas with *high* p-values (not statistically significant) are not statistically different from the null, and thus we can not reject the hypothesis that $\mathrm{PP} = 0$.

PP and its statistical significance are strongly correlated however — for those values where σ^2_{obs} lies in the top half of the σ^2_{sim} distribution, the correlation between PP and the (standard, normal) inverse CDF of the confidence level is $r = 0.90$. This means that for significant (high PP) results, we most often conclude that there is potential predictability (as estimated), and for non-significant results (low PP) we are typically unable to assume that there is predictability. It follows then, that our most justifiable PP estimate for areas

with near-zero PP and high p-values is, in fact, zero.

PP is shown station-by-station in Figure A.6, with values significant at the $p = 0.05$ level shown as filled circles and non-significant values shown as X's. The percentages of stations with PP values significantly different from zero by season are: DJF — 89%, MAM — 86%, JJA — 76%, SON — 81%. There are additionally a number of stations with seasonal σ^2_{obs} values greater than any value in the σ^2_{sim} distribution, and the corresponding PP values would thus be significant at any confidence interval. By season the fraction of stations for which this is the case is: DJF — 58%, MAM — 46%, JJA — 34%, SON — 39%.

A.5 PP-metrics other than seasonal

Although this paper focuses on PP_{89}, windows other than an 89-day seasonal window can be used to evaluate time periods of interest. Agricultural growing seasons, snow-pack periods, dry seasons, and monsoons might all be of interest in a particular region, and the analysis window can be selected to fit the objective.

Equation A.8 can be rewritten for any window, τ, as

$$PP_\tau(t) = \frac{\sigma^2_{\tau,obs}(t) - \left\langle \sigma^2_{\tau,sim}(t) \right\rangle}{\sigma^2_{\tau,obs}(t)}, \tag{A.10}$$

where σ^2_τ is the variance (between years) in the total number of occurrence events in a window τ for either the observed or simulated data series.

As an example, Figure A.7 shows PP_{365} — the potential predictability of the number of occurrence events annually. As compared to the seasonal PP_{89} maps (Chapter 2 Fig. 2.1e–h) or even their mean (Fig. A.3 bottom-right), PP_{365} is generally higher, suggesting that some years may have precipitation occurrence correlations between seasons (as we might expect, say, from ENSO) which are not well-represented by stochastic processes with only a few days' memory.

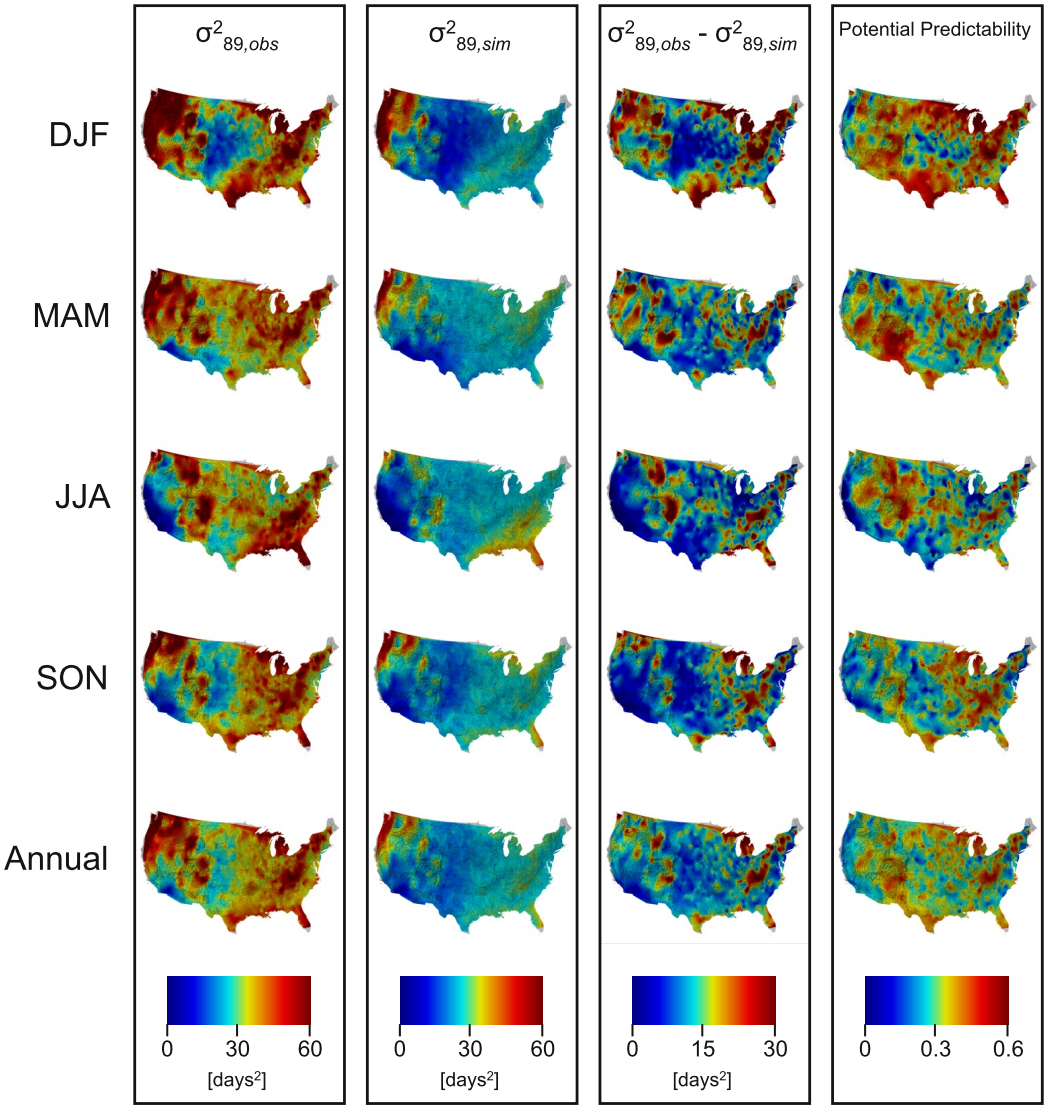

Figure A.3: Seasonal values of the components of potential predictability. Rows 1–4 show values for 89-day windows centered on Jan 15, Apr 15, Jul 15, and Oct 15; row 5 shows the mean of 365 89-day windows, centered on each day of the year. The first column shows the interannual variance in number of wet days in an 89-day for each season from the historical data. The second column shows the mean value of the interannual variance for 1,000 simulated time series with only short-memory process knowledge. The third column is the difference between the first two columns (the numerator in the potential predictability). The last column is the potential predictability (as shown in Fig. 2.1 and Fig. A.1).

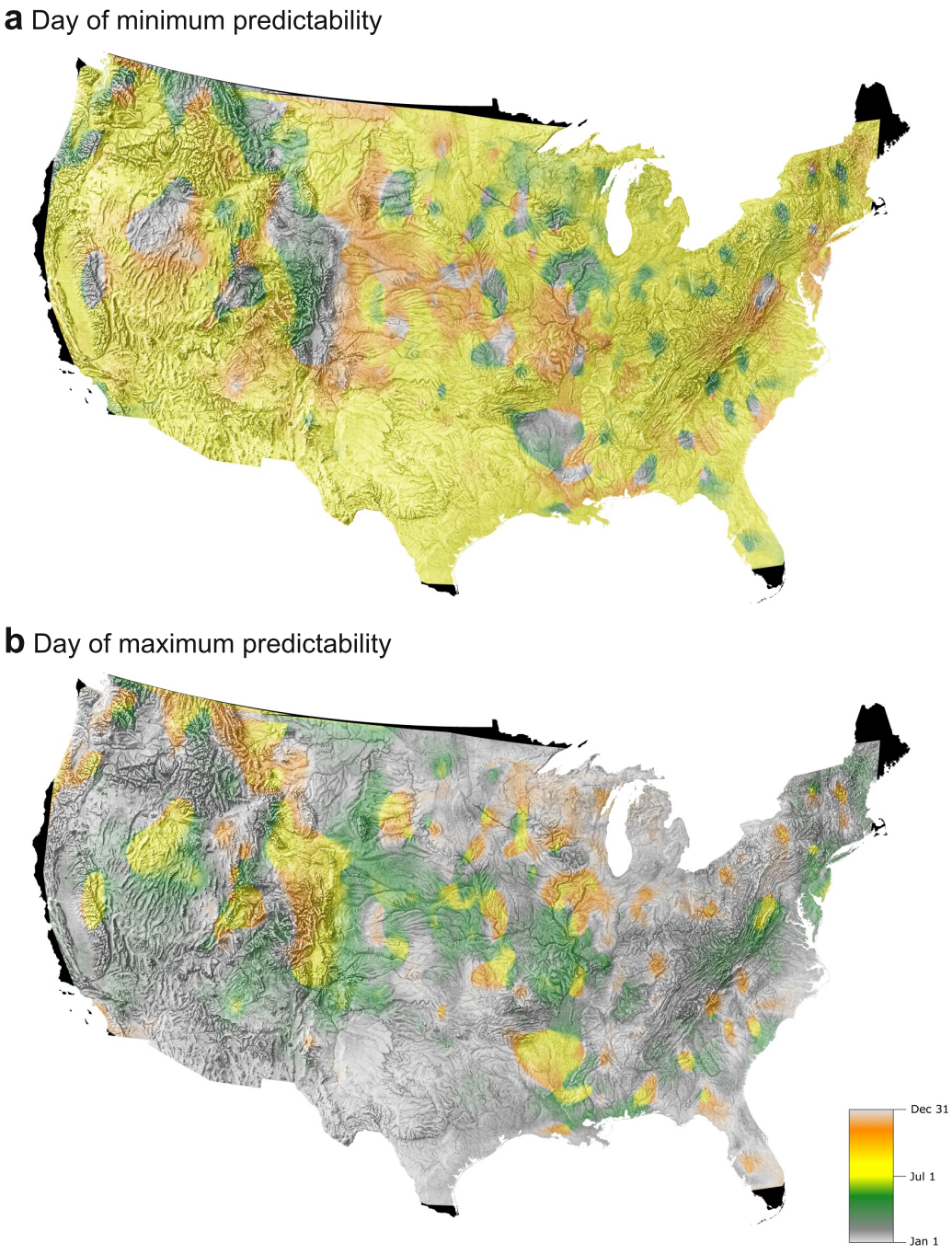

Figure A.4: Timings of (a) minimum and (b) maximum potential predictability.

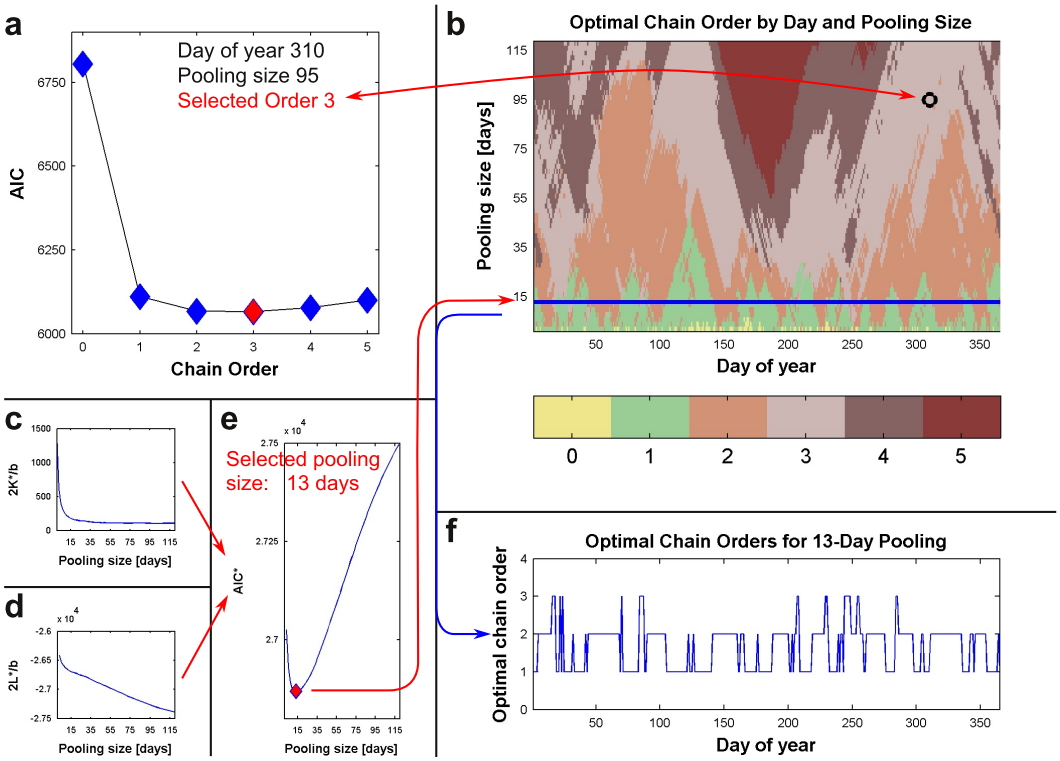

Figure A.5: Chain-order and pooling-size determination process using the Akaike Information Criterion (AIC) for Tuscon, AZ. For each day of the year and possible pooling size, the AIC is calculated for chain orders 0–5 (a). The order which gives the lowest AIC value is selected as optimal. For November 6^{th} (day 310) in Tuscon, using a pooling size of 95 days, the 3^{rd}-order model has the lowest AIC (a) and is selected as optimal (shown as the selected chain-order in b). This is performed for every day and possible pooling size (b). To select the optimal pooling size for the station, K^* and L^* are calculated from Equations A.4 and A.5 (c and d, respectively), and are combined to give the AIC* (e, Eqn. A.3). For Tuscon, the selected pooling size is 13 days. The model transition probabilities are determined from the 13-day pooled data, and the chain orders corresponding to 13-day pooling (as determined in b, shown in f) are used in our stochastic simulations.

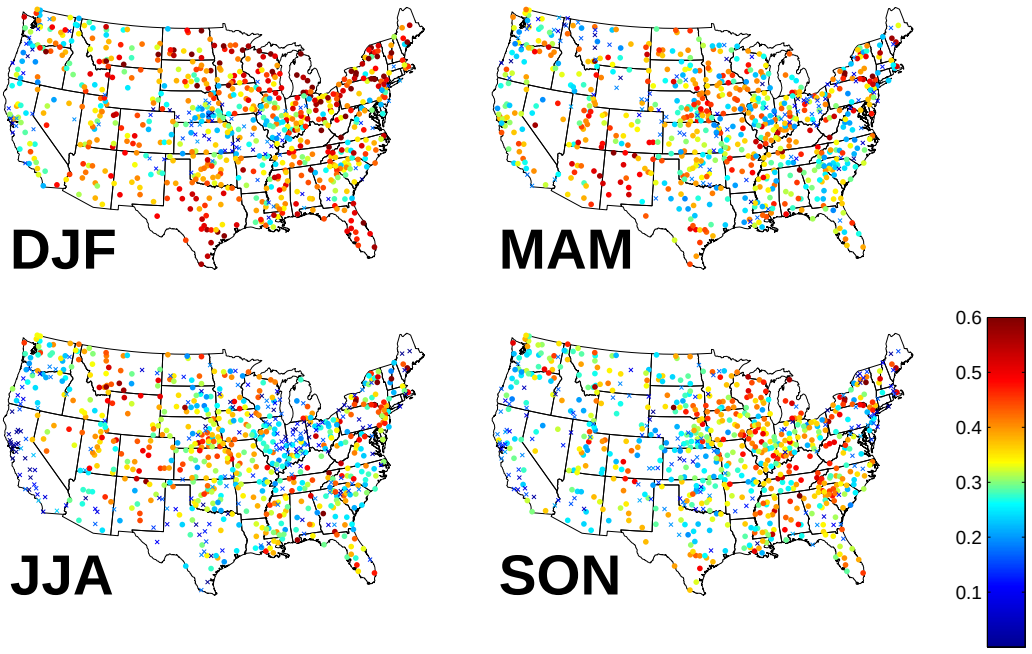

Figure A.6: PP and corresponding statistical significance station-by-station. Colors denote the PP value. PP values which are significantly different from zero at the 0.05 level are shown as filled circles, while those not significantly different from 0 are shown as X's. Data in this figure match identically with those in Figure 2.1 (Chapter 2).

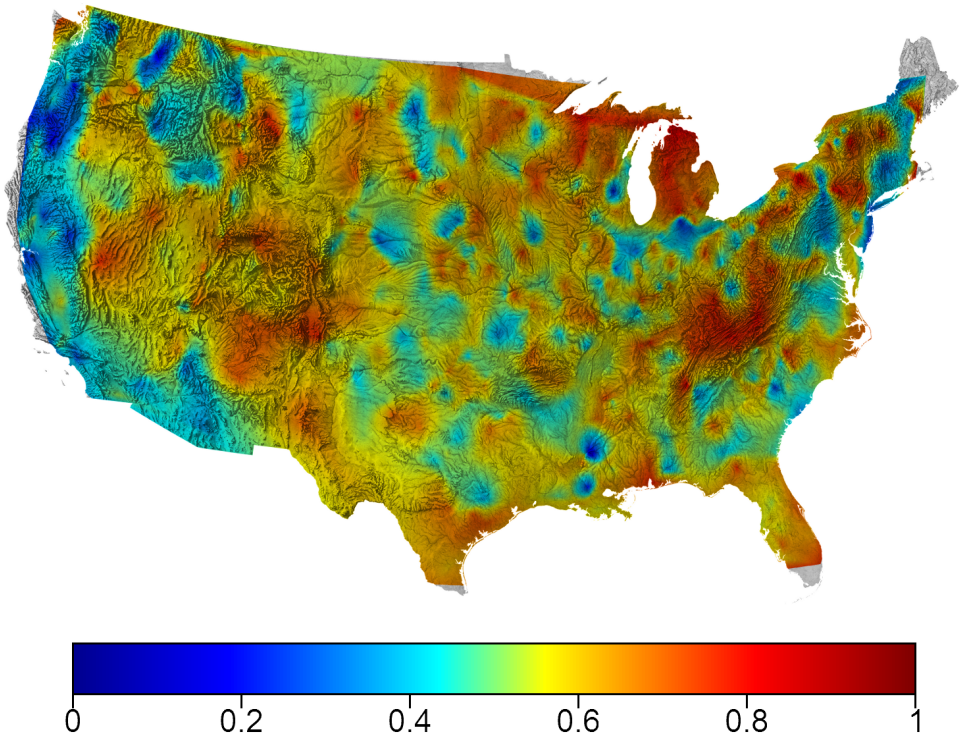

Figure A.7: PP_{365}: annual-scale potential predictability. An annual window is used to calculate the total number of occurrence events, and variances of these annual totals (observed and simulated) are used to calculate PP_{365}. Values are generally higher (note the difference in scale) than for any of the seasonal PP_{89} maps (Fig. 2.1) or the mean of seasonal PP_{89} (Fig. A.3, bottom-right).

Appendix B

Variance Decomposition into Occurrence and Intensity

The law of conditional variance states,

$$\mathrm{Var}[\Pi] = \mathrm{E}_\Theta\left[\mathrm{Var}\left[\Pi|\Theta\right]\right] + \mathrm{Var}_\Theta\left[E\left[\Pi|\Theta\right]\right], \tag{B.1}$$

We will use this to decompose total precipitation in a seasonal window into components due to the number of days of precipitation occurrence in that window, and the intensity on those days. Let π be the total precipitation observed in an 89-day window with θ days of precipitation occurrence in those 89 days. Π is a vector $[\pi_1, \pi_2, \ldots, \pi_m]$ and Θ is similarly $[\theta_1, \theta_2, \ldots, \theta_m]$, where each element of the vectors is the seasonal observation for one year (i.e. in the first year there were θ_1 rainy days which sum to a total of π_1 mm of rain.

The green diamonds in Figure B.1a each represent a single winter's (θ_i, π_i) pair for a rain gauge at McConnelsville Lock, Ohio, and Figure B.1b is a histogram of the number of precipitation events over all observed years. The second term in the right-hand side of Equation B.1, $\mathrm{Var}_\Theta\left[E\left[\Pi|\Theta\right]\right]$, is referred to as the variance of conditional means and is displayed visually in Figure B.1c. The blue circles in Figure B.1a are the mean seasonal precipitation conditioned on the number of wet days, as are the green lines in Figure B.1c. The red dot in Figure B.1c is the mean of the conditional means, and the error bars show one standard deviation — the square root of σ^2_{occ}, as shown in Figure B.1e.

The first term on the right-hand side of Equation B.1 is referred to as the expectation

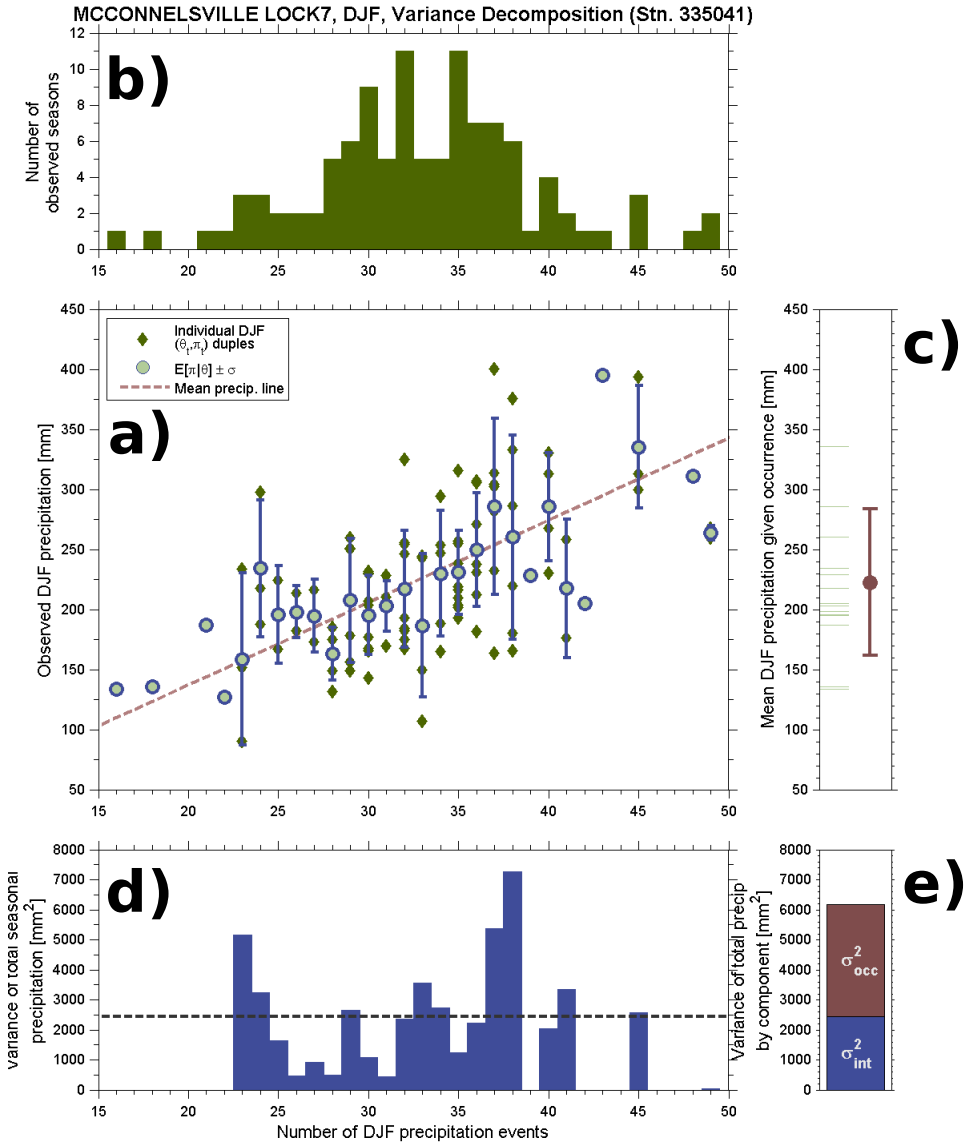

Figure B.1: An example of the decomposition of total precipitation into occurrence and intensity contributions following the law of conditional variance for USHCN weather station 335041 in McConnelsville, OH. a) Observed 89-day precipitation vs. number of precipitation events. b) Histogram of seasons with a given number of precipitation events. c) Mean precipitation values conditioned on number of events (green lines) and the standard deviation of the conditional means (red bar), the square root of the variance of conditional means, σ^2_{occ}. d) Variance of precipitation conditioned on number of events (blue bars) and the mean of conditional variances, σ^2_{int} (dashed line). e) Partitioning of variance, as from c) and d), as per the law of conditional variance.

of conditional variances and is calculated by first taking the variances of the seasonal total precipitation as conditioned on the number of wet days (blue error bars in Figure B.1a show one standard deviation, blue bars below in Figure B.1d show the variance). The mean across all observed θ values is shown with the dashed blue bare in Figure B.1d and as the blue region in Figure B.1e, also called σ^2_{int}.

In this example, σ^2_{occ} is greater than σ^2_{int} due to the comparatively large spread between the blue dots in Figure B.1a as compared to the within-θ deviations around those means. This partitioning is equivalent to the notion of "variance explained" by a model (in this case the model is the mean precipitation per event dashed line in Figure B.1a) in a typical regression setting. If it always rained precisely the same amount on each wet day, our occurrence-based model would fit the data perfectly (no residual sum of squares), all of the blue dots would fall on the dashed red line in Figure B.1a, and there would be no σ^2_{int}. Similarly, if it always rained the same amount in each season no matter how many wet days that season had, occurrence would explain none of the variance in the data, there would be no dispersion about the mean in Figure B.1c, and hence no σ^2_{occ}.

The decomposition could potentially show less stability for seasons with very few events, as the estimated precipitation per event would be calculated using fewer observed data points (see Supplementary Materials for discussion). In this analysis, this potential source of uncertainty is negligible for two reasons: 1) the coefficients of variation for the $\sigma^2_{occ}/\sigma^2_{tot}$ ratio are generally low and not a function of the number of wet days except for in one autumn and nine summertime locations (see Supplementary Figure C.11), and 2) those locations with higher coefficients of variation (> 0.25) are in the driest regions which display consistently low (and not significantly different from zero) PP values for all of $\mathrm{PP}_{89,tot}$, $\mathrm{PP}_{89,int}$, and $\mathrm{PP}_{89,occ}$.

Appendix C

Supplementary Material for Chapter 3

C.1 Methods schematic and overview

To further clarify the methods used in this study, Figure C.1 shows a schematic representation of the methods used to calculate the potential predictability (PP), with data shown taken from the station in Tuscon, AZ. Daily precipitation data is taken from the 774 stations in the United States Historical Climate Network (USHCN) (Williams et al., 2006) which have 80–110 years of data ending in 2009 and less than 5% missing — this data forms our observational ("obs") record. Each day of the year is represented by its own model, with data from all observed years used to determine parameters for each model. The models for precipitation occurrence on each day are variable-order Markov chains conditioned on the last m days of occurrence data, where $0 \leq m \leq 5$. Probabilities of occurrence given an m-day history are determined from the 80–110 years of observed data for each day, and data pooling (using data from neighboring days) is allowed to improve parameter estimation. The width of the pooling window, W, and the chain order, m, are determined by comparing models using all combinations of W and m, and selecting the parsimony/goodness-of-fit trade-off using the Corrected Akaike Information Criterion (Hurvich and Tsai, 1989). Transition probabilities for the selected 365 Markov models are stored for each station.

The precipitation intensities for each day are modeled as parameterized probability

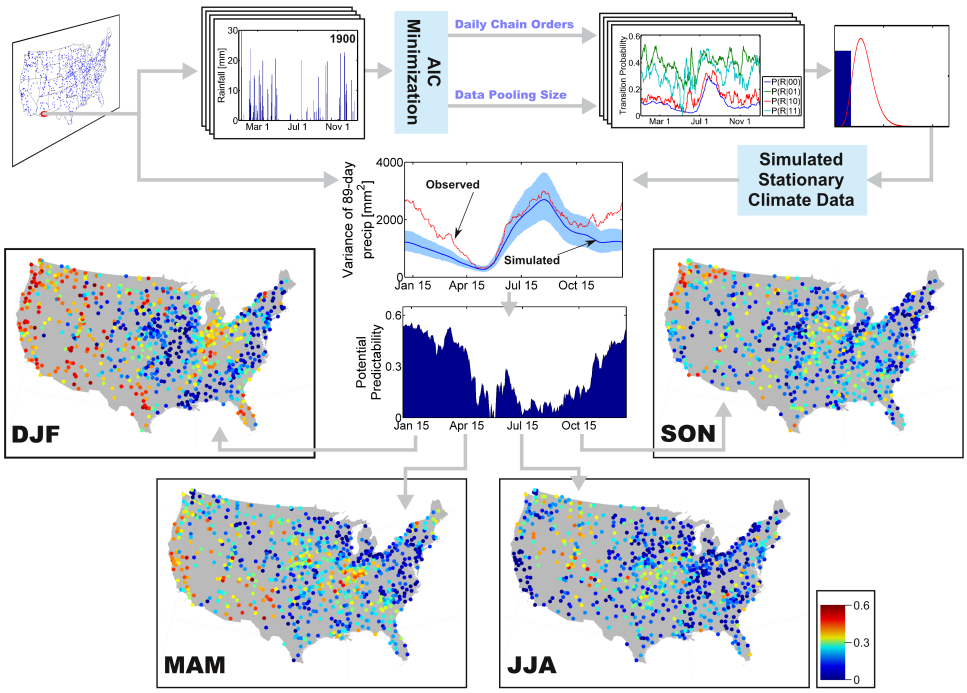

Figure C.1: Schematic representation of the methods used to calculate potential predictability. Data shown in two central plots (Variance of 89-day precipitation and Potential Predictability) are from station in Tuscon, AZ. Potential predictabiliy maps shown are for $PP_{89,tot}$ (same as Figures 3.5–3.8a).

distributions which take the following form:

$$f(x) = \lambda \frac{\beta_1^{\alpha_1} x^{\alpha_1 - 1} e^{-\alpha_1 \beta_1}}{\Gamma(\alpha_1)} + (1 - \lambda) \frac{\beta_2^{\alpha_2} x^{\alpha_2 - 1} e^{-\alpha_2 \beta_2}}{\Gamma(\alpha_2)}, \tag{C.1}$$

where $0 \leq \lambda \leq 1$, α_1 and α_2 are shape parameters, β_1 and β_2 are rate parameters, and $\Gamma(\cdot)$ is the gamma function. As with the occurrence process, a model is constructed for each combination of station, day of year, pooling size, and chain order (conditioning on the previous m days' occurrence, rather than intensity), the data pooling size is selected using AIC_c, and then the chain order is selected also using AIC_c.

Using the combined occurrence and intensity models, simulations of daily precipitation are created which match the observed record in terms of day-of-year mean precipitation as well as day-of-year interannual variance due to short-memory (m-day) processes. 1000 simulations, each the length of the "obs" record, are created for each station, which form a ("sim") distribution of "weather noise" against which to compare the "obs" record. The metric for comparison in this analysis is 89-day (seasonal) precipitation totals. In Figure C.1, the variance of the seasonal total precipitation in the observed record ($\sigma^2_{obs,tot}$, shown in red) is much higher than the mean variance of the seasonal total precipitaion from the simulations ($\overline{\sigma^2}_{sim,tot}$, shown in blue) in the winter, and so the corresponding potential predictabilty,

$$\text{PP}_{89,tot} = \frac{\sigma^2_{obs,tot} - \overline{\sigma^2}_{sim,tot}}{\sigma^2_{obs,tot}} \tag{C.2}$$

is high in the winter for Tuscon, AZ. The light blue shading in the variance plot is a 95% confidence interval from the simulations.

After calculating the potential predictability for each day of the year at each station, maps of the seasonal potential predictability can be created using the $\text{PP}_{89,tot}$ values for January 15, April 15, July 15, and October 15.

C.2 "Climate signal" estimates

Some readers may be interested in the direct estimate of the variance due to potentially predictable processes, rather than the fraction of total variance that the these processes represent. The variance due to potentially predictable processes (or potential "climate signals") is simply the numerator of the potential predictability as defined in Equation C.2, $\sigma^2_{obs,tot} - \overline{\sigma^2}_{sim,tot}$. The values of the numerators corresponding to the potential predictability shown in the main manuscript (Figures 3.5–3.8) are shown in Figures C.2–C.5, respectively.

In each figure, part a) shows the numerator of $PP_{89,tot}$, b) shows the numerator of $PP_{89,int}$, and c) shows the numerator of $PP_{89,occ}$ for an 89-day window centered on January 15 (Figure C.2), April 15 (Figure C.3), July 15 (Figure C.4), or October 15 (Figure C.5). σ^2_{obs} is the variance of the total precipitation in the historical record for each station. This is designated as total (tot) when taken as a simple variance, occurrence (occ) when depicting the portion of tot explained by conditioning on the number of days with non-zero precipitation, and intensity (int) when depicting the portion *not* explained by conditioning on occurrence. Similarly, $\overline{\sigma^2}_{sim}$ is the mean variance of total precipitation from each of 1000 simulations of the historical record representing only interannually-stationary, short-memory ("weather noise") processes.

It is important to note that while the PP values shown in Figures 3.5–3.8 are normalized by the observed variance, the un-normalized numerators have units of precipitation variance (mm^2), and thus are likely to be higher in areas with more precipitation.

C.3 Stability of variance decomposition

The variance decomposition (Appendix A) is determined using the law of conditional variance, which is exact in that the left-hand side of Equation A1 is exactly equal to the right-hand side, but is also an estimate of the "true" variance decomposition using the data as a sample. As the number of wet days increases, that estimate is determined using

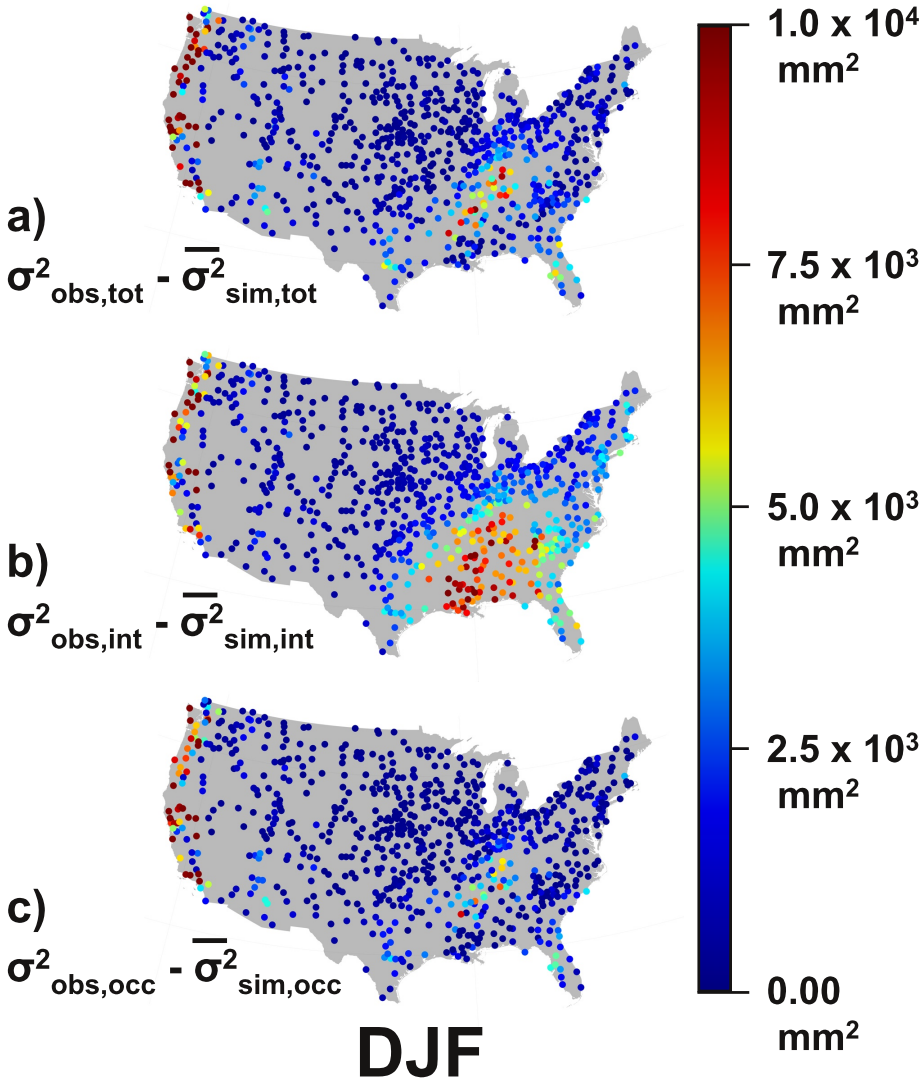

Figure C.2: The numerators of the a) $PP_{89,tot}$, b) $PP_{89,int}$, and c) $PP_{89,occ}$ fractions (respective PP values shown in the Figure 3.5) for an 89-day window centered on January 15. σ^2_{obs} is the variance of the total precipitation in the historical record for each station. This is designated as total (tot) when taken as a simple variance, occurrence (occ) when depicting the portion of tot explained by conditioning on the number of days with non-zero precipitation, and intensity (int) when depicting the portion *not* explained by conditioning on occurrence. Similarly, $\overline{\sigma^2}_{sim}$ is the mean variance of total precipitation from each of 1000 simulations of the historical record representing only interannually-stationary, short-memory ("weather noise") processes.

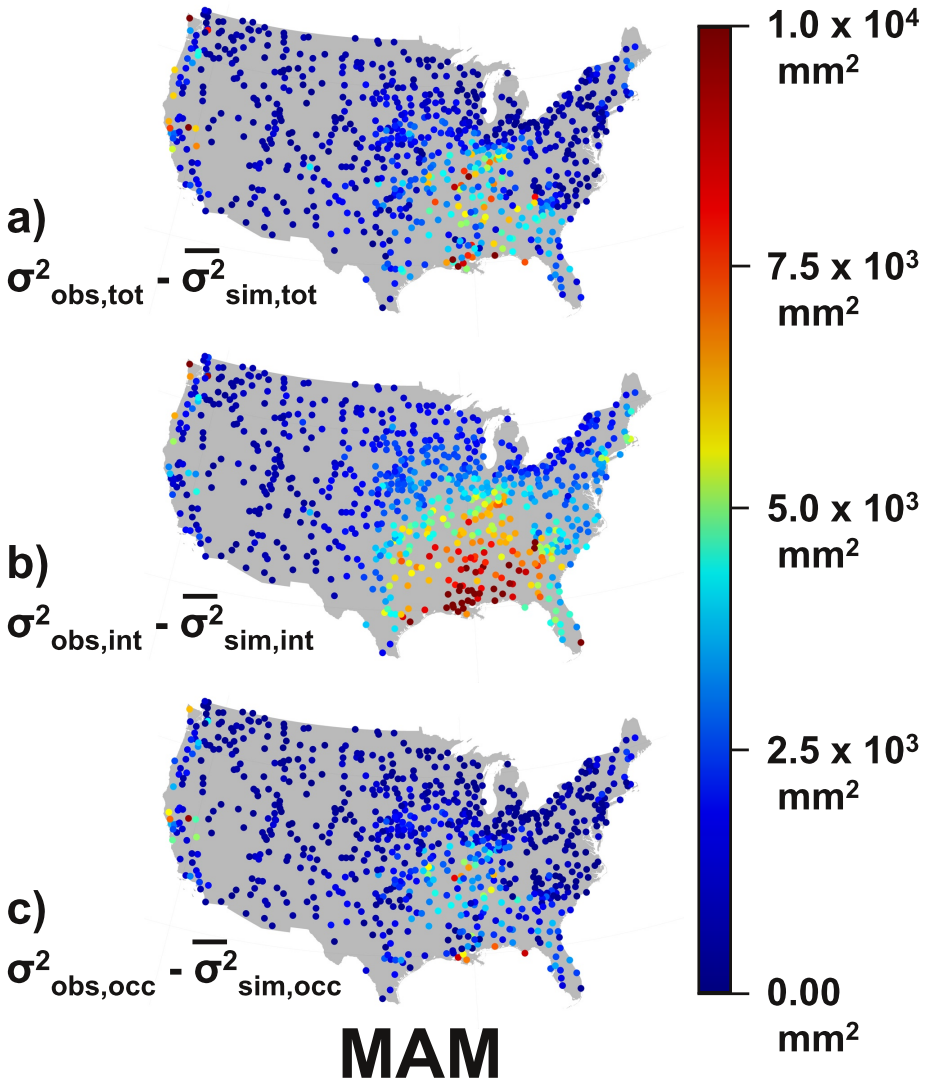

Figure C.3: Same as Figure C.2, but for an 89-day window centered on April 15.

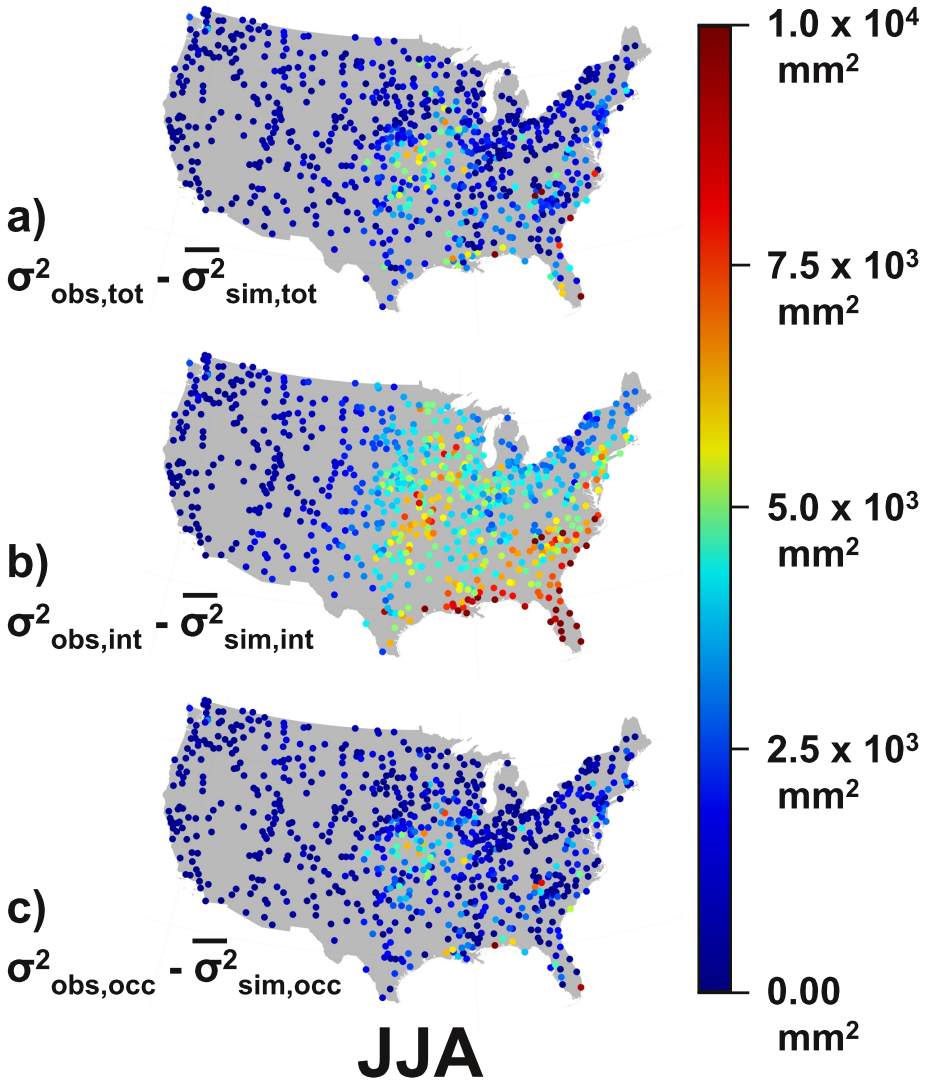

Figure C.4: Same as Figure C.2, but for an 89-day window centered on July 15.

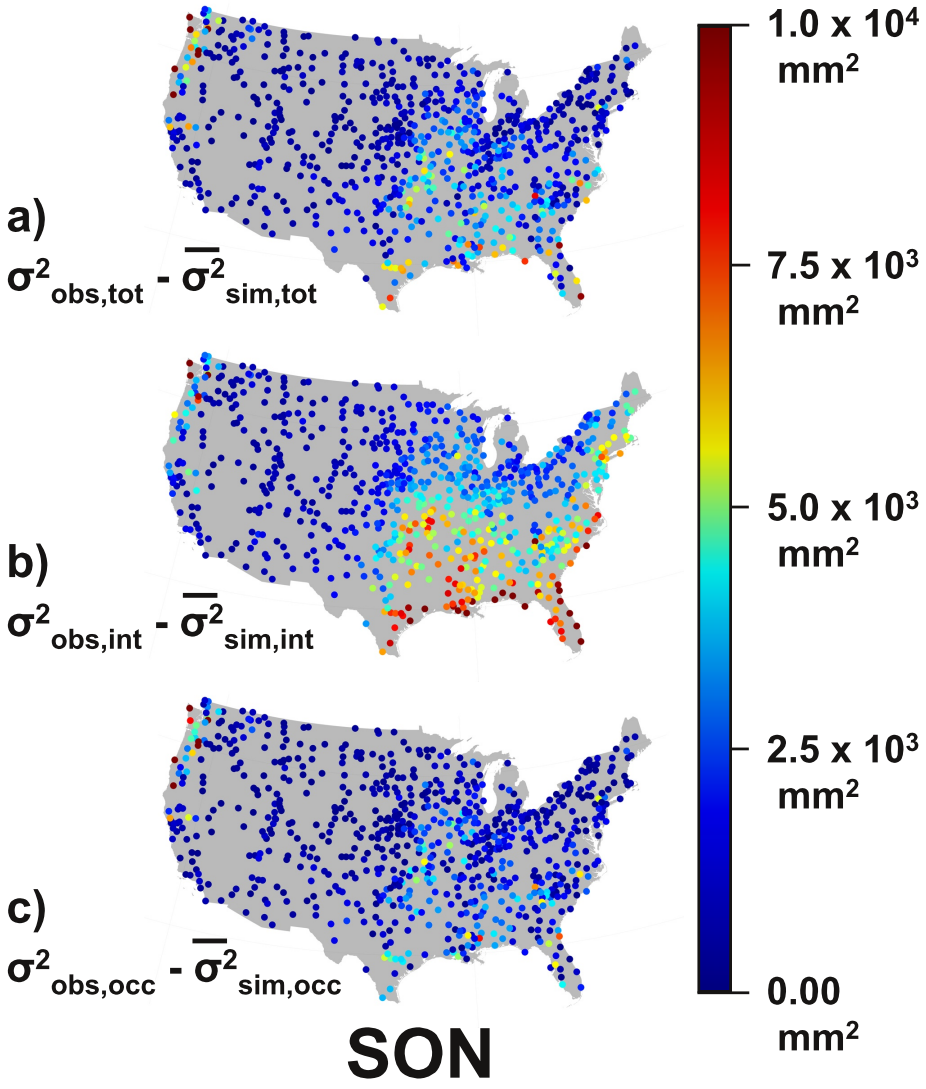

Figure C.5: Same as Figure C.2, but for an 89-day window centered on October 15.

more data points, and so the estimate is likely more robust for seasons with more wet days. To determine if the estimates of the variance decomposition (as seen in Figures 3.2 and 3.4) are robust for seasons with small numbers of wet days, Figure C.11 shows seasonal plots of the coefficient of variation (CV) of the $\sigma^2_{occ}/\sigma^2_{tot}$ ratio versus the number of wet days in the simulations. Each y-coordinate of a point in the plots is the standard deviation of the $\sigma^2_{occ}/\sigma^2_{tot}$ ratio across all 1000 simulations for a given station in a given season divided by the mean $\sigma^2_{occ}/\sigma^2_{tot}$ ratio across all 1000 simulations for that station and season. The x-coordinate is the mean number of wet days in that 89-day season across all 1000 simulations. As seen in Figure C.11, the CV is not clearly a function of the number of wet days; however, for very dry seasons (with 5 wet days or fewer, see JJA plot) many CV values are notably higher than for those seasons with more than 5 wet days. This suggests less confidence in the decomposition into σ^2_{occ} and σ^2_{int}. However, for these stations in the Southwest (9 locations in JJA, and one in SON with CV > 0.25), the potentially predictable variance components are low to begin with, and so all three of $\mathrm{PP}_{89,tot}$, $\mathrm{PP}_{89,occ}$, and $\mathrm{PP}_{89,int}$ are very small.

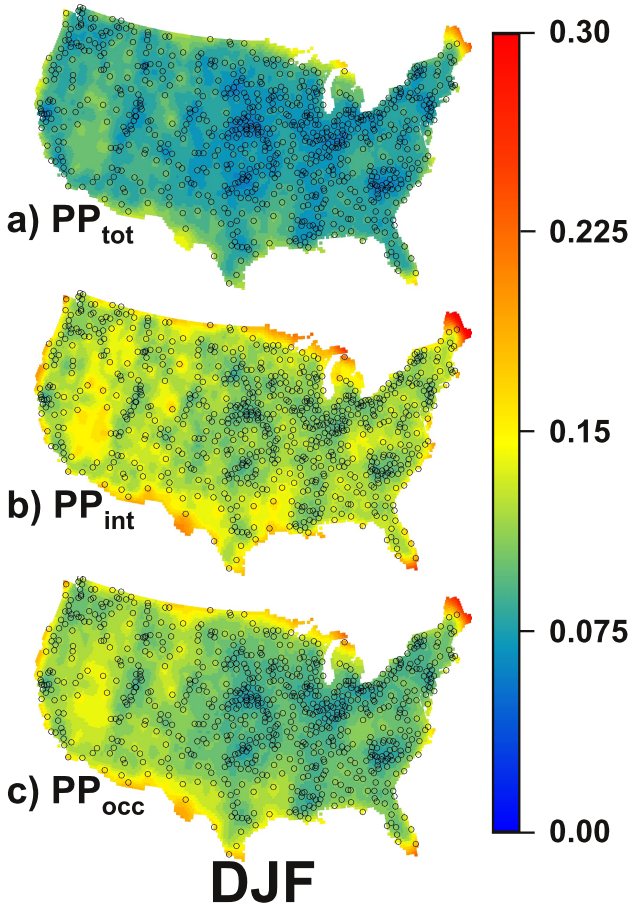

Figure C.6: 90% confidence interval range for spatial model of December–February potential predictability using generalized cross validation fitting of thin plate splines for a) PP_{tot}, b) PP_{int}, and c) PP_{occ}. 0.5 times the values from this figure added to the gridded (background) values in Figure 3.5 give the 95% confidence bound, and 0.5 times the values from this figure subtracted from the gridded values in Figure 3.5 give the 5% confidence bound for PP.

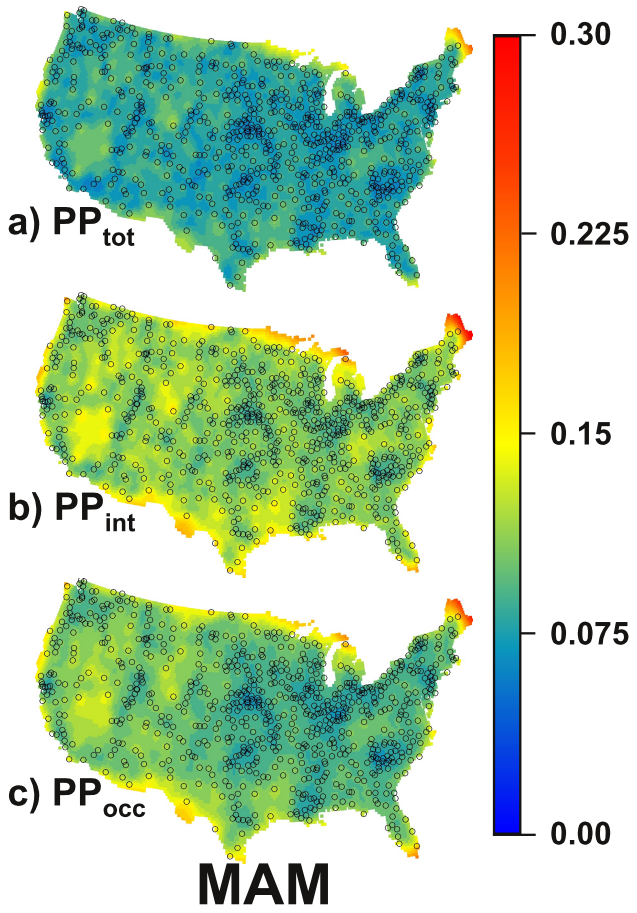

Figure C.7: 90% confidence interval range for spatial model of March–May potential predictability using generalized cross validation fitting of thin plate splines for a) PP_{tot}, b) PP_{int}, and c) PP_{occ}. 0.5 times the values from this figure added to the gridded (background) values in Figure 3.6 of the main text give the 95% confidence bound, and 0.5 times the values from this figure subtracted from the gridded values in Figure 3.6 give the 5% confidence bound for PP.

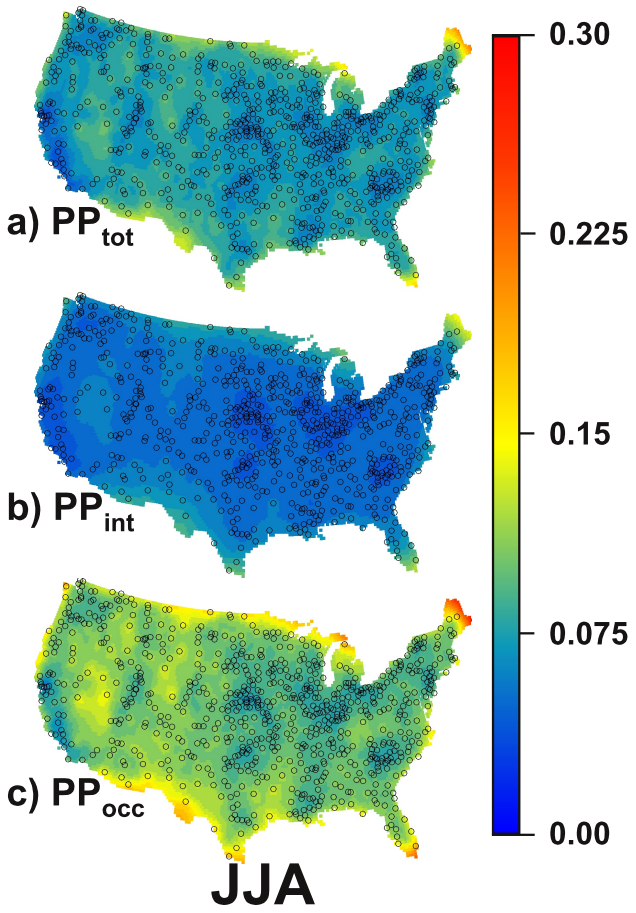

Figure C.8: 90% confidence interval range for spatial model of June–August potential predictability using generalized cross validation fitting of thin plate splines for a) PP_{tot}, b) PP_{int}, and c) PP_{occ}. 0.5 times the values from this figure added to the gridded (background) values in Figure 3.7 of the main text give the 95% confidence bound, and 0.5 times the values from this figure subtracted from the gridded values in Figure 3.7 give the 5% confidence bound for PP.

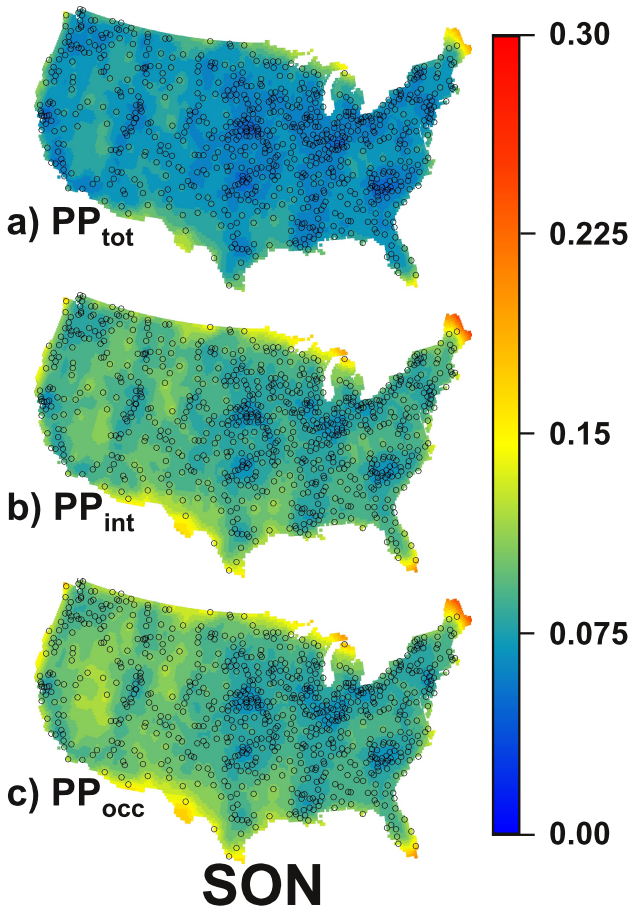

Figure C.9: 90% confidence interval range for spatial model of September–November potential predictability using generalized cross validation fitting of thin plate splines for a) PP_{tot}, b) PP_{int}, and c) PP_{occ}. 0.5 times the values from this figure added to the gridded (background) values in Figure 3.8 of the main text give the 95% confidence bound, and 0.5 times the values from this figure subtracted from the gridded values in Figure 3.8 give the 5% confidence bound for PP.

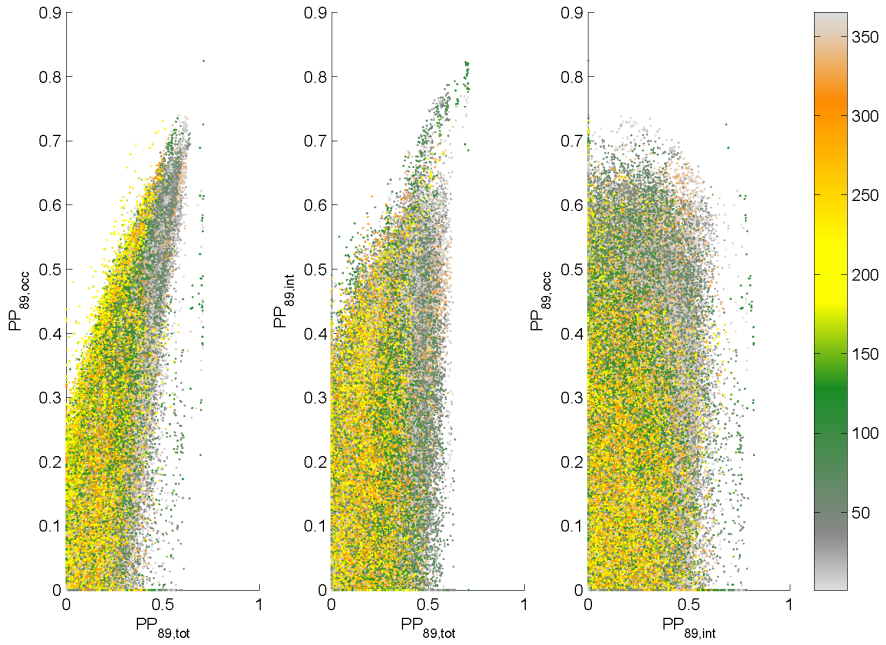

Figure C.10: Potential predictability station by station for each day of the year. a) $PP_{89,occ}$ versus $PP_{89,tot}$, b) $PP_{89,int}$ versus $PP_{89,tot}$, and c) $PP_{89,occ}$ versus $PP_{89,int}$. Colors correspond to the day of the year (January 1 = 1, December 31 = 365).

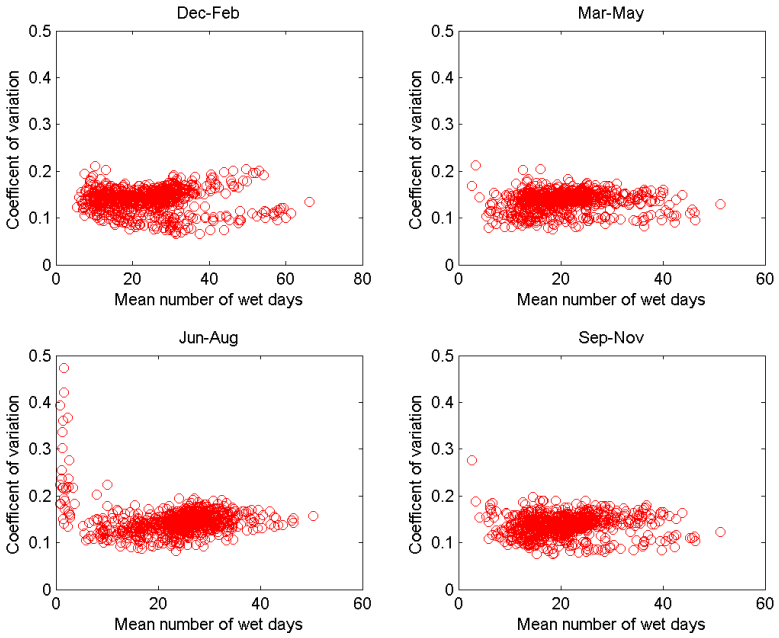

Figure C.11: Coefficent of variation (CV) versus the mean number of wet days in the simulations by season.

Appendix D

Variance of a kernel density estimate

Although using kernel density methods to estimate empirical probability densities is fairly common practice, there are important caveats for their use, specifically regarding variance in this context. Kernel-based probability density functions are essentially mixture model distributions, using one mixture for each observed data point, and are not appropriate for either maximum-likelihood or method-of-moments fitting techniques (both of which will be optimized with a bandwidth of zero, equivalent to bootstrapping). Because of this, cross-validation (or some approximation thereof) is typically employed, but this does not preserve the variance of the sample to which the kernel density was fit, nor is it tied to asymptotic estimators for the variance of the population from which the data sample was drawn. While the bias in the variance is typically small, when variance is a key feature of your model, this needs to be addressed (and is the reason why simulated samples in this paper are re-standardized prior to transformation back to the data domain). It is worth noting that when using an axially-symmetric kernel (as is typical), the sample mean is always preserved.

D.0.1 Kernel density

For a set of N observed points, $\{y_i\}$, the univariate probability density using Gaussian kernels with bandwidth (standard deviation in this case), h, is

$$f_K(x;\{y_i\},h) = \frac{1}{N}\sum_{i=1}^{N}\phi(x;y_i,h), \tag{D.1}$$

where $\phi(x; y_i, h)$ is the normal density function with mean y_i and standard deviation h.

D.0.2 Mean of kernel density

Using $E\{\cdot\}$ to denote expectation and $\mathcal{N}$ to denote the normal (Gaussian) distribution, the mean of the Gaussian kernel density is

$$\begin{aligned} E\{x\} &= \int x f_K(x)\,\mathrm{d}x && \text{(D.2)} \\ &= \int \frac{x}{N}\sum_{i=1}^{N}\phi(x; y_i, h)\,\mathrm{d}x && \text{(D.3)} \\ &= \frac{1}{N}\sum_{i=1}^{N}\int x\phi(x; y_i, h)\,\mathrm{d}x && \text{(D.4)} \\ &= \frac{1}{N}\sum_{i=1}^{N} E\{\mathcal{N}(y_i, h)\} && \text{(D.5)} \\ &= \frac{1}{N}\sum_{i=1}^{N} y_i && \text{(D.6)} \\ &= \overline{y_i}, && \text{(D.7)} \end{aligned}$$

which is just the sample mean of the points $\{y_i\}$ used to define the kernel densities.

D.0.3 Variance of kernel density

The variance of the kernel density function is

$$\begin{aligned} \mathrm{Var}\{x\} &= E\{x^2\} - (E\{x\})^2 && \text{(D.8)} \\ &= \int x^2 f_K(x)\,\mathrm{d}x - (\overline{y_i})^2 && \text{(D.9)} \\ &= \int \frac{x^2}{N}\sum_{i=1}^{N}\phi(x; y_i, h)\,\mathrm{d}x - (\overline{y_i})^2 && \text{(D.10)} \\ &= \frac{1}{N}\sum_{i=1}^{N}\int x^2\phi(x; y_i, h)\,\mathrm{d}x - (\overline{y_i})^2 && \text{(D.11)} \end{aligned}$$

But for any individual normal distribution with mean y_i and standard deviation h,

$$E\{(x-y_i)^2\} = E\{x^2\} - (E\{x\})^2 \tag{D.12}$$

Rearranging gives

$$\begin{aligned} E\{x^2\} &= E\{(x-y_i)^2\} + (E\{x\})^2 && \text{(D.13)} \\ \int x^2 \phi(x; y_i, h)\,\mathrm{d}x &= E\{(x-y_i)^2\} + (E\{x\})^2 && \text{(D.14)} \\ &= h^2 + y_i^2, && \text{(D.15)} \end{aligned}$$

since for kernel density centered at y_i the variance is just h^2 and the mean is y_i. Substituting back into D.11,

$$\begin{aligned} \mathrm{Var}\{x\} &= \frac{1}{N}\sum_{i=1}^{N}\left(h^2 + y_i^2\right) - (\overline{y_i})^2 && \text{(D.16)} \\ &= \left[\overline{y_i^2} - (\overline{y_i})^2\right] + h^2 && \text{(D.17)} \end{aligned}$$

The term on the left in D.17 is just the sample variance, and so the variance of the kernel density is effectively additively inflated by the squared bandwidth, h^2. In this paper, where N for any given daily model is on the order of (31 days) $\cdot$ (57 years) = 1767 data points, and h^2 is on the order of 0.1 (in squared z-score units), the bandwidth variance inflation is larger than the effect of the typical multiplicative "Bessel correction" (i.e., $n/(n-1)$) used for unbiased population variance estimates. If our data did not go through an explicitly relative rank-driven CDF transformation (but rather some absolute mapping from the data domain to the Gaussian CDF-transformed domain and back), very large bandwidth values could lead to biases in the correlation structure of the simulated data. The rank-based transformation, however, eliminates this potential problem, but requires that any simulated data be re-standardized before back-transformation to explicitly preserve the proper variance.

This re-standardization can lead to statistical problems of its own if an insufficient number of simulated data points are used. To prevent the daily simulated variances from matching the observed sample variance exactly, a large number of simulations can be performed, re-standardized, and back-transformed. Then a subsample of the simulated data can be used for analysis. As an example, in this research, 1000 stochastic recreations of the historic record were simulated, and for each day of the year the 57,000 simulated daily values match the observed mean and sample variance, but the means and variances of each individual 57-year simulation can vary stochastically.

Appendix E

Global Climate Model Supplementary Material

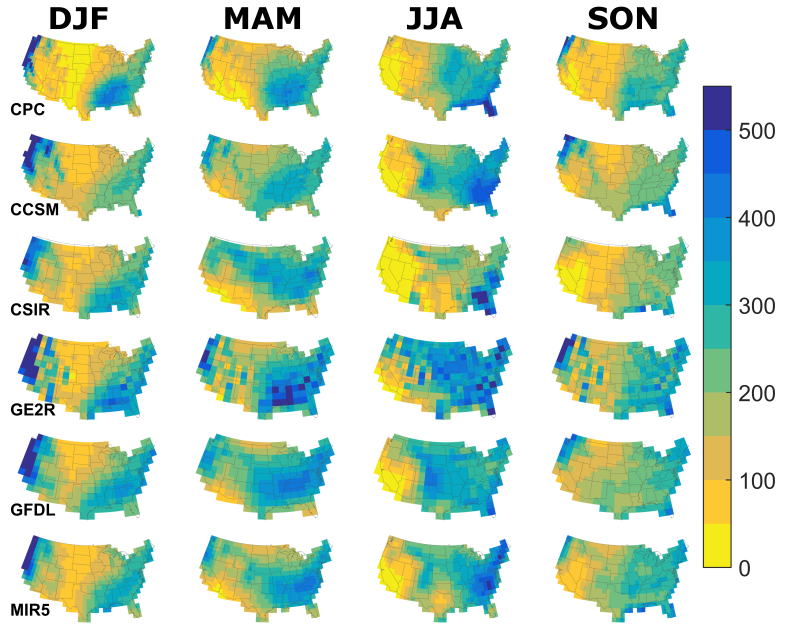

Figure E.1: Mean 91-day (seasonal) accumulated precipitation from one observational dataset and historical runs from five GCMs from 1948–2004 [mm].

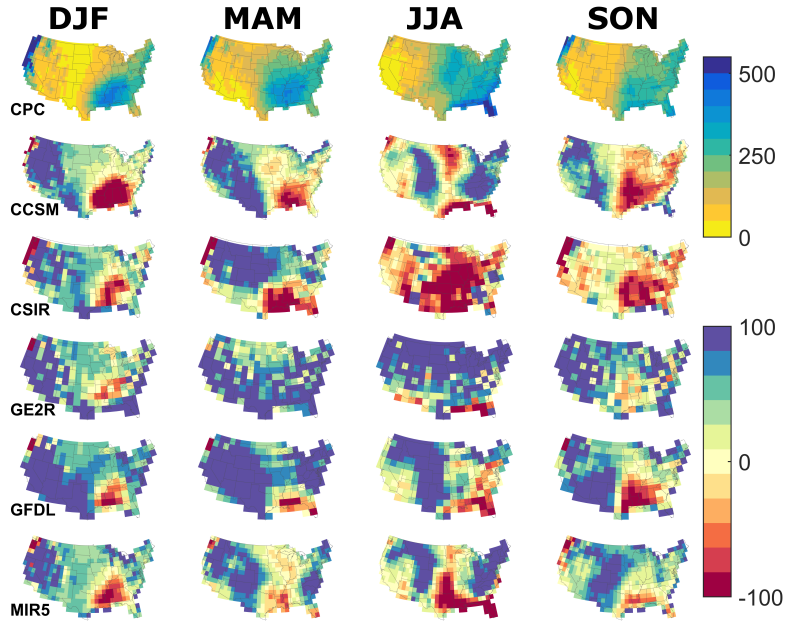

Figure E.2: Mean 91-day (seasonal) accumulated precipitation from one observational dataset and historical runs from five GCMs from 1948–2004 [mm]. GCM totals (lower color bar) are shown as anomalies from the CPC seasonal values [in mm] (values greater than 0 imply more average precipitation in the historical GCM dataset than the historical CPC dataset).

	Model	DJF	MAM	JJA	SON
Mean	CCSM	0.800	0.816	0.826	0.762
	CSIR	0.847	0.584	0.793	0.763
	GE2R	0.829	0.859	0.697	0.755
	GFDL	0.816	0.772	0.742	0.718
	MIR5	0.908	0.847	0.724	0.860
$\sigma^2_{\text{historical}}$	CCSM	0.830	0.625	0.545	0.726
	CSIR	0.729	0.661	0.506	0.588
	GE2R	0.731	0.714	0.609	0.639
	GFDL	0.694	0.677	0.141	0.590
	MIR5	0.815	0.721	0.713	0.763
$\sigma^2_{\text{simulated}}$	CCSM	0.823	0.688	0.704	0.747
	CSIR	0.787	0.720	0.581	0.626
	GE2R	0.778	0.858	0.733	0.695
	GFDL	0.775	0.755	0.389	0.628
	MIR5	0.878	0.857	0.777	0.828
$\sigma^2_{\text{residual}}$	CCSM	0.827	0.555	0.391	0.686
	CSIR	0.696	0.580	0.405	0.532
	GE2R	0.667	0.502	0.500	0.579
	GFDL	0.647	0.548	0.021	0.539
	MIR5	0.771	0.576	0.645	0.623

Table E.1: Grid-cell-wise Pearson correlation coefficients (unweighted) between GCM and CPC datasets. All correlation values are determined after regridding the summary statistics from the CPC dataset to the native grid of the GCM.

	Model	DJF	MAM	JJA	SON
$\sigma^2_{\text{historical}}$	CCSM	2532	-293	-498	-757
	CSIR	547	487	-8	-707
	GE2R	388	1398	5002*	913
	GFDL	2684	415	-643	8
	MIR5	1601	637	2297	-104
$\sigma^2_{\text{simulated}}$	CCSM	773	-311	-372	-723
	CSIR	469	270	210	-502
	GE2R	704	2002	3797*	747
	GFDL	1171	247	-854	-93
	MIR5	267	265	1382	109
$\sigma^2_{\text{residual}}$	CCSM	1759	19	-124	-34
	CSIR	78	217	-219	-205
	GE2R	316	-604	1201	166
	GFDL	1514	167	218	101
	MIR5	1333	372	916	-213

Table E.2: Spatial mean of the absolute errors of the interannual variability for each model by season in $(\text{mm/day})^2$. Positive (negative) values denote over-estimation (under-estimation) of interannual variability by the GCM, using the CPC interannual variability as reference. These values correspond to the mean values of the PDFs in the first row of Figure E.6. The only statistically significant difference (two-tailed, $\alpha = 0.05$, denoted by a superscript $*$) from the observed CPC variability is the GE2R model in the summer (JJA) months — the model consistently *over*-estimates the interannual variability across the contiguous 48 United States. Other models/seasons display over-dispersion or under-dispersion, but are not consistently over- or under-dispersed when considering all locations as a single data set.

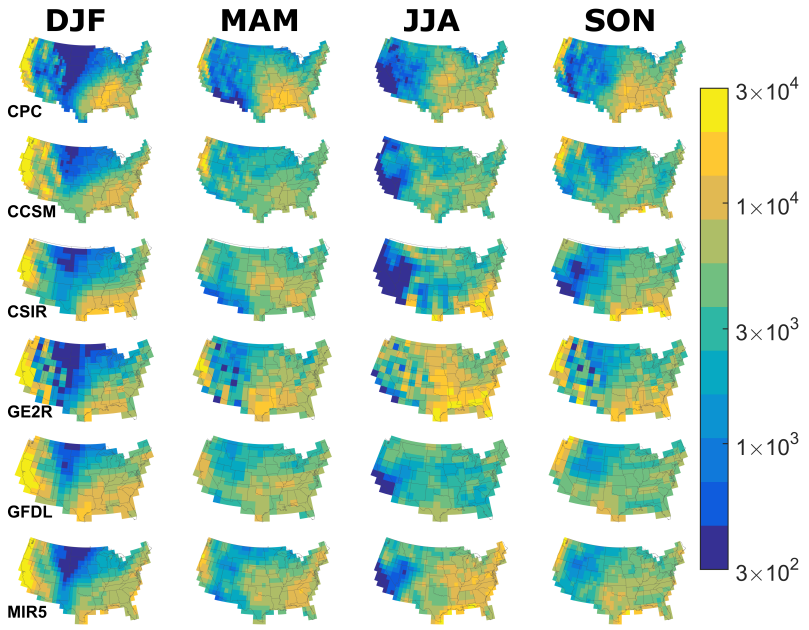

Figure E.3: Interannual variance of 91-day (seasonal) accumulated precipitation from one observational dataset and historical runs from five GCMs from 1948–2004 [mm^2].

	Model	DJF	MAM	JJA	SON
$\sigma^2_{\text{historical}}$	CCSM	1.66	0.7	0.39	0.32
	CSIR	0.97	0.81	0.17	0.09
	GE2R	0.74	1.11	5.85*	1.07
	GFDL	2.19	1.03	1.06	0.94
	MIR5	1.23	0.75	1.44	0.54
$\sigma^2_{\text{simulated}}$	CCSM	1.52	0.77	0.24	0.31
	CSIR	1.25	0.89	0.20	0.10
	GE2R	1.16	1.38	3.71*	1.01
	GFDL	2.02	0.94	0.55	0.88
	MIR5	1.09	0.60	1.02	0.57
$\sigma^2_{\text{residual}}$	CCSM	2.26	0.53	-1.16	0.98
	CSIR	1.08	1.55	-1.17	0.83
	GE2R	-0.03	-1.77	-0.83	0.34
	GFDL	3.50	2.32	0.18	0.33
	MIR5	1.91	1.19	-2.00	0.10

Table E.3: Same, but relative errors

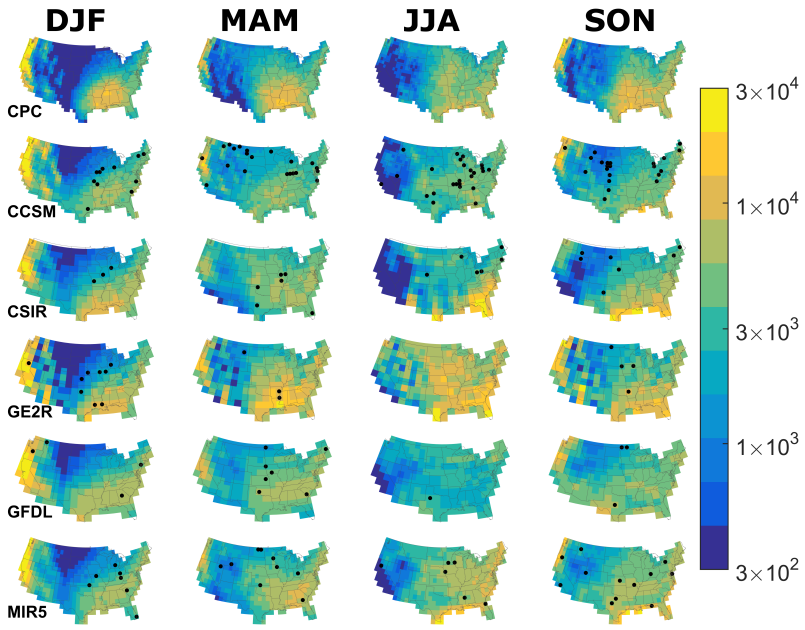

Figure E.4: Interannual 91-day (seasonal) precipitation variance attributable to weather-scale phenomena (1948–2004) [mm^2]. Variances shown are the median values of 1,000 stochastic weather-generator type datasets as fit to the historical data. Interannual variability is generated only by stochastic processes with memory on the order of days. Black dots show grid cells in which the distribution of variances from the GCM dataset is statistically indistinguishable from the distribution of variances from the CPC dataset (Wilcoxson rank-sum test, $\alpha = 0.05$).

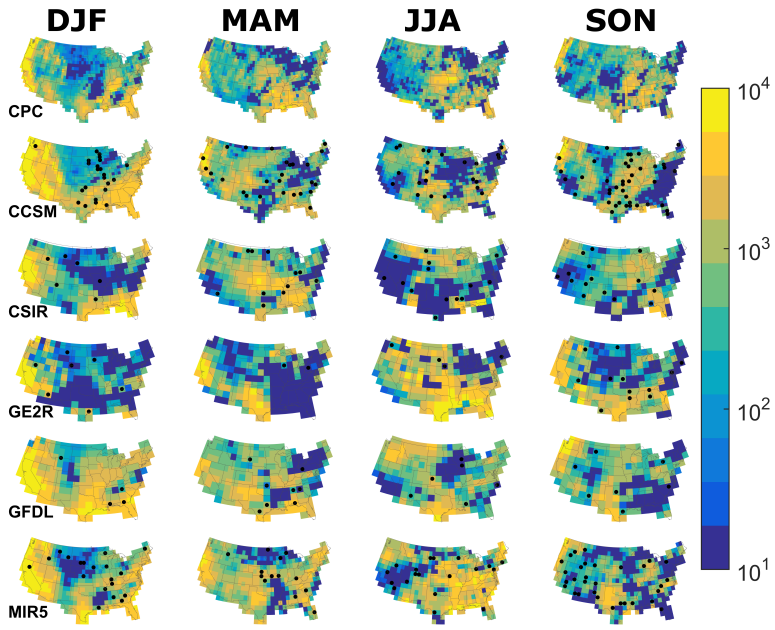

Figure E.5: Interannual 91-day (seasonal) precipitation variance attributable to climate-scale phenomena (1948–2004) [mm^2]. Variances shown are the median values of 1,000 residual variance terms (subtracting the weather-scale variance from the historical variance), and represent variability that cannot be captured by a stochastic model with memory on the order of days. Black dots show grid cells in which the distribution of variances from the GCM dataset is statistically indistinguishable from the distribution of variances from the CPC dataset (Wilcoxson rank-sum test, $\alpha = 0.05$).

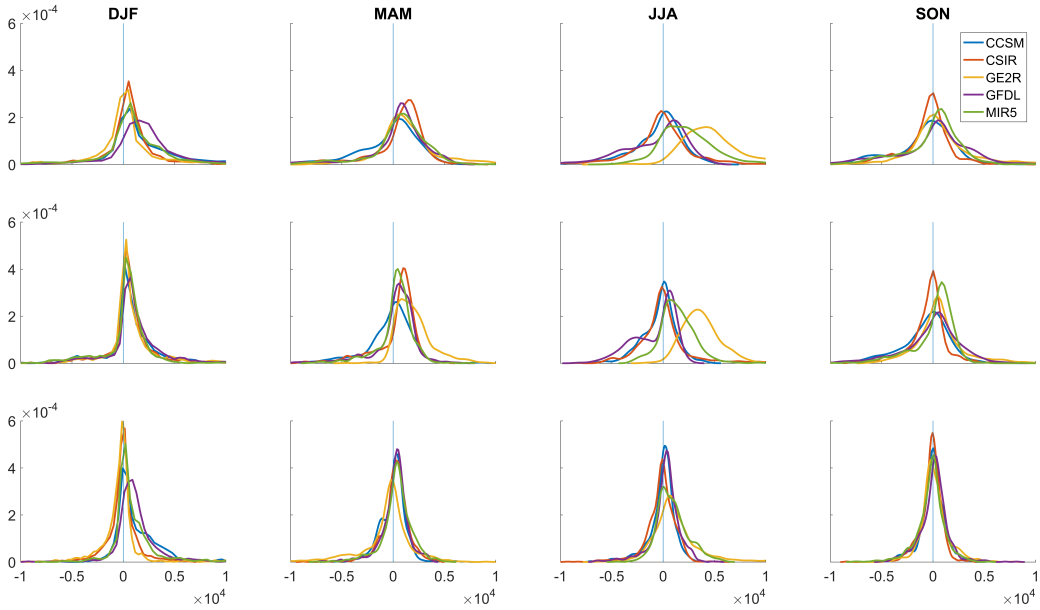

Figure E.6:

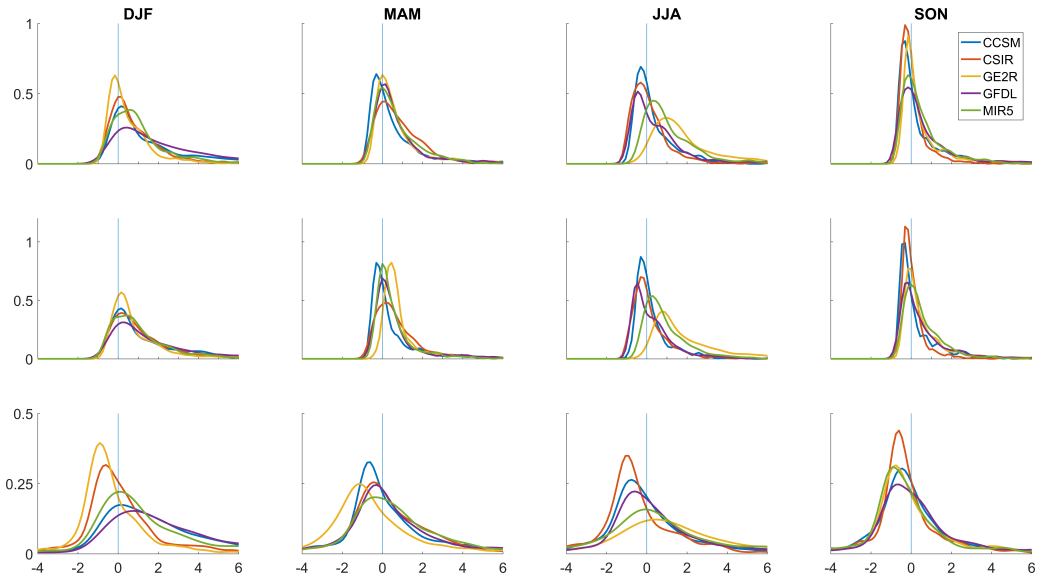

Figure E.7:

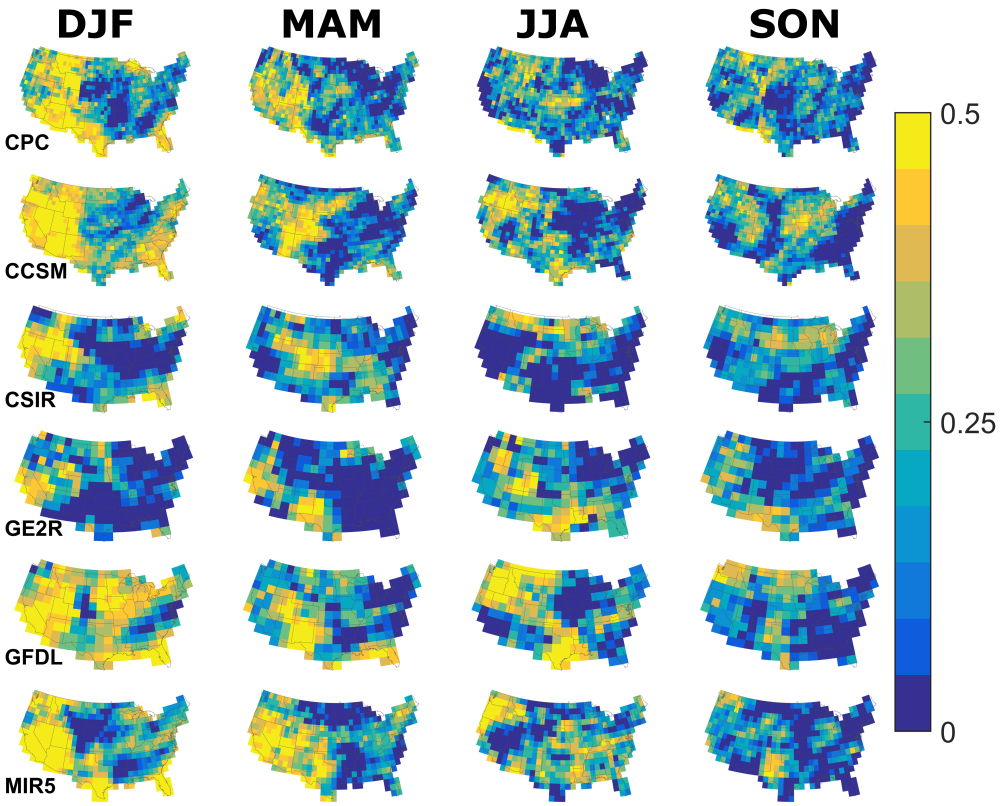

Figure E.8: The fraction of historical interannual variability attributable to climate-scale processes (commonly refered to as *Potential Predictability*.

Bibliography

Pierre Ailliot, Denis Allard, Valérie Monbet, and Philippe Naveau. Stochastic weather generators: an overview of weather type models. *Journal de la Société Française de Statistique*, 156(1):101–113, 2015.

Michael G. Akritas. The Rank Transform Method in Some Two-Factor Designs. *Journal of the American Statistical Association*, 85(409):73–78, 1990.

Myles R Allen and William J Ingram. Constraints on future changes in climate and the hydrologic cycle. *Nature*, 419(6903):224–32, 2002.

Bruce T. Anderson, Jingyun Wang, S. Gopal, and Guido Daniel Salvucci. Influence of Daily Rainfall Characteristics on Regional Summertime Precipitation over the Southwestern United States. *Journal of Hydrometeorology*, 10(5):1218–1230, oct 2009.

Bruce T. Anderson, Daniel J. Short Gianotti, and Guido D. Salvucci. Characterizing the Potential Predictability of Seasonal, Station-Based Heavy Precipitation Accumulations and Extreme Dry Spell Durations. *Journal of Hydrometeorology*, 16(2):843–856, 2015a.

Bruce T. Anderson, Daniel J. Short Gianotti, and Guido D. Salvucci. Detectability of historical trends in station-based precipitation characteristics over the continental United States. *Journal of Geophysical Research*, 120(10):4842—-4859, 2015b.

Bruce T. Anderson, Dan Gianotti, Guido Daniel Salvucci, and Jason Furtado. Dominant timescales of potentially predictable precipitation variations across the continental United States. *Journal of Climate1*, In press, 2016a.

Bruce T. Anderson, Daniel J S Gianotti, Jason C. Furtado, and Emanuele Di Lorenzo. A decadal precession of atmospheric pressures over the North Pacific. *Geophysical Research Letters*, 1:3921–3927, 2016b.

Shigeru (Japan) Aoki, Edmo (Brazil) Campos, Don (USA) Chambers, Richard A. (USA) Feely, Sergey (Russian Federation) Gulev, Gregory C. (USA) Johnson, Simon A. (UK) Josey, Andrey (Russian Federation) Kostianoy, Cecilie (Norway) Mauritzen, Dean (USA) Roemmich, and Fan (China) Wang. Observations: Ocean. Technical report, Intergovernmental Pannel on Climate Change, 2013.

Aristotle and E. W. (trans.) Webster. *Meteorology*. Digireads, Overland Park, KS., 2006.

Spencer F. Baird. *Annual record of science and industry for 1878*. Harper and Brothers, New York, 1879.

M Barlow, S Nigam, and Ernesto Hugo Berbery. ENSO, Pacific decadal variability, and US summertime precipitation, drought, and stream flow. *Journal of Climate*, (1997): 2105–2128, 2001.

Emily J. Becker, Ernesto Hugo Berbery, and R. Wayne Higgins. Modulation of Cold-Season U.S. Daily Precipitation by the MaddenJulian Oscillation. *Journal of Climate*, 24(19): 5157–5166, oct 2011.

Omar Bellprat and Francisco Doblas-Reyes. Unreliable climate simulations overestimate attributable risk of extreme weather and climate events. *Geophysical Research Letters*, pages n/a–n/a, 2016.

Asit K . Biswas. The Automatic Rain-Gauge of Sir Christopher Wren, F.R.S. *Notes and Records of the Royal Society of London*, 22(1):94–104, 1967.

Maggie Black and Janet King. *The Atlas of Water*. University of California Press, Oakland, CA, 2nd edition, 2009.

John C. Bloxam. Atmospheric pressure and rainfall. *Quarterly Journal of the Royal Meteorological Society*, 2(13):277–280, 1875.

George Blumenstock. Drought in the United States Analyzed by Means of the Theory of Probabihty. *Technical Bulletin*, (819):1–64, 1942.

G.J. Boer. Long time-scale potential predictability in an ensemble of coupled climate models. *Climate dynamics*, 23(1):29–44, jun 2004.

GJ Boer. Changes in interannual variability and decadal potential predictability under global warming. *Journal of Climate*, 22(11):3098–3109, jun 2009.

GS Bohr and Edward Aguado. Use of April 1 SWE measurements as estimates of peak seasonal snowpack and total coldseason precipitation. *Water resources research*, 37(1): 51–60, 2001.

T. A. Buishand. Some remarks on the use of daily rainfall models. *Journal of Hydrology*, 36(3-4):295–308, 1978.

Xiaoyu Cai, Huiyun Li, and Aiyi Liu. A marginal rank-based inverse normal transformation approach to comparing multiple clinical trial endpoints. *Statistics in Medicine*, 35:3259–3271, 2016.

Daniel R. Cayan, Michael D. Dettinger, Henry F. Diaz, and Nicholas E. Graham. Decadal variability of precipitation over western North America. *Journal of Climate*, pages 3148–3166, 1998.

Daniel R. Cayan, KT Redmond, and LG Riddle. ENSO and Hydrologic Extremes in the Western United States. *Journal of Climate*, pages 2881–2893, 1999.

Stanley A. Changnon. Thunderstorm rainfall in the conterminous United States. *Bulletin of the American Meteorological Society*, 82(9):1925–1940, 2001.

M. Chen and P. Xie. CPC Unified Gauge-based Analysis of Global Daily Precipiation. In *Western Pacific Geophysics Meeting (9 July - 1 August)*, Cairns, Australia, 2008.

Ha-man Cho, Sang-Won Kim, Young-sin Chun, Hye-Yeong Park, and Woo-Jeong Kang. A Historical Review on the Introduction of Chugugi and the Rainfall Observation Network during the Joseon Dynasty. *Atmosphere*, 25(4):719–734, 2015.

Philippe Ciais, Christopher Sabine, G Bala, Laurent Bopp, Victor Brovkin, J. Canadell, A Chhabra, R DeFries, J. Galloway, Martin Heimann, C Jones, C. Le Quéré, R.B. Myneni, S Piao, and P Thornton. Carbon and Other Biogeochemical Cycles. Technical report, Intergovernmental Pannel on Climate Change, 2013.

W. G. Cochran. An extension of Gold's method of examining the apparent persistence of one type of weather. *Quarterly Journal of the Royal Meteorological Society*, 64:631–634, 1938.

M. Collins. Climate predictability on interannual to decadal time scales: the initial value problem. *Climate Dynamics*, 19(8):671–692, oct 2002.

Daniel Cooley, Douglas Nychka, and Philippe Naveau. Bayesian Spatial Modeling of Extreme Precipitation Return Levels. *Journal of the American Statistical Association*, 102 (479):824–840, 2007.

Mark J. Costello, Simon Wilson, and Brett Houlding. Predicting total global species richness using rates of species description and estimates of taxonomic effort. *Systematic Biology*, 61(5):871–883, 2012.

Griffith N. Cox. On the rainfall of Jamaica during the seven years 18701876. *Quarterly Journal of the Royal Meteorological Society*, 4(25):15–25, 1878.

Aiguo Dai. Precipitation characteristics in eighteen coupled climate models. *Journal of Climate*, 19(18):4605–4630, sep 2006.

Timothy DelSole and Michael K. Tippett. Predictability: Recent insights from information theory. *Reviews of Geophysics*, 45(4):1–22, 2007.

P D'Odorico and L Ridolfi. Preferential states of seasonal soil moisture: The impact of climate fluctuations. *Water Resources Research*, 36(8):2209–2219, 2000.

David R. Easterling, JL Evans, Pavel Ya. Groisman, Thomas R. Karl, Kenneth E. Kunkel, and P. Ambenje. Observed Variability and Trends in Extreme Climate Events: A Brief Review. *Bulletin of the American Meteorological Society*, 81(3):417–425, 2000.

Nicole Feldl and GH Roe. Climate variability and the shape of daily precipitation: A case study of ENSO and the American West. *Journal of Climate*, 24(10):2483–2499, may 2011.

S. B. Feldstein. The timescale, power spectra, and climate noise properties of teleconnection patterns. *Journal of Climate*, 13(24):4430–4440, 2000.

X Feng, Timothy DelSole, and P Houser. Bootstrap estimated seasonal potential predictability of global temperature and precipitation. *Geophysical Research Letters*, 38(7): L07702, 2011.

Ian M. Ferguson, Philip B. Duffy, Thomas J. Phillips, Xu Liang, John A. Dracup, Siegfried D. Schubert, and Philip Pegion. Non-stationarity of the signal and noise characteristics of seasonal precipitation anomalies. *Climate dynamics*, 36(3-4):739–752, jun 2011.

KL Findell, Pierre Gentine, BR Lintner, and Christopher Kerr. Probability of afternoon precipitation in eastern United States and Mexico enhanced by high evaporation. *Nature Geoscience*, 4(7):434–439, jun 2011.

P Frich, Lisa V. Alexander, and P Della-Marta. Observed coherent changes in climatic extremes during the second half of the twentieth century. *Climate . . .*, 19:193–212, 2002.

K R Gabriel and J Neumann. A Markov chain model for daily rainfall occurrence at Tel Aviv. *Quarterly Journal of the Royal Meteorological Society*, 88(375):90–95, 1962.

Daniel J. Gianotti, Bruce T. Anderson, and Guido D. Salvucci. What Do Rain Gauges Tell Us about the Limits of Precipitation Predictability? *Journal of Climate*, 26(15): 5682–5688, aug 2013.

Harry R. Glahn. Statistical Weather Forecasting. In Allan H. Murphy and Richard W. Katz, editors, *Probability, Statistics, and Decision Making in the Atmospheric Sciences*, pages 263–310. Westview Press, Boulder, CO, 1982.

E. Gold. Note on the frequency of occurrence of sequences in a series of events in two types. *Quarterly Journal of the Royal Meteorological SocietyQuarterly Journal of the Royal Meteorological Society*, 55:307–309, 1929.

Xiaofeng Gong, Anthony G. Barnston, and M. Niel Ward. The Effect of Spatial Aggregation on the Skill of Seasonal Precipitation Forecasts. *Journal of Climate*, 16(18):3059–3071, 2003.

William Gore and D. Halton Thomson. Note on Rainfall and Stream-Flow Statistics. *The Geographical Journal*, 34(6):645–650, 1909.

A. W. Greely. Charts showing the probability of rainy days, prepared from observations for eighteen years, 1891.

JM Gregory, TML Wigley, and PD Jones. Application of Markov models to area-averaged daily precipitation series and interannual variability in seasonal totals. *Climate Dynamics*, 8(6):299–310, 1993.

Pavel Ya. Groisman and David R. Easterling. Variability and trends of total precipitation and snowfall over the United States and Canada. *Journal of Climate*, 1994.

Pavel Ya. Groisman and Richard W. Knight. Contemporary changes of the hydrological cycle over the contiguous United States: Trends derived from in situ observations. *Journal of Hydrometeorology*, pages 64–85, 2004.

Pavel Ya. Groisman and Richard W. Knight. Trends in intense precipitation in the climate record. *Journal of Climate*, 18(9):1326–1350, 2005.

Pavel Ya. Groisman and Richard W. Knight. Prolonged dry episodes over the conterminous United States: new tendencies emerging during the last 40 years. *Journal of Climate*, 21(9):1850–1862, may 2008.

Pavel Ya. Groisman, Thomas R. Karl, David R. Easterling, Richard W. Knight, Paul F. Jamason, Kevin J. Hennessy, Samasamy Suppiah, Cher M. Page, Joanna Wibig, Krzysztof Fortuniak, Vyacheslav N. Razuvaev, Arthur Douglas, Eirik Forland, and Pan-Mao Zhai. Changes in the probability of heavy precipitation: Important indicators of climatic change. *Climatic Change*, 42(1):243–283, 1999.

Brian Groombridge and Martin Jenkins. *World Atlas of Biodiversity: Earth's Living Resources in the 21st Century*. University of California Press, Berkeley, CA, 2002.

T Hall, He Brooks, and Ca Doswell Iii. Precipitation forecasting using a neural network. *Weather and Forecasting*, 14(3):338–345, 1999.

Khaled H Hamed and A Ramachandra Rao. A modified Mann-Kendall trend test for autocorrelated data. *Journal of Hydrology*, 204(1-4):182–196, 1998.

Alan F. Hamlet and Dennis P. Lettenmaier. Production of Temporally Consistent Gridded Precipitation and Temperature Fields for the Continental United States. *Journal of Hydrometeorology*, 6(3):330–336, 2005.

Lars S. Hanson and Richard Vogel. The probability distribution of daily rainfall in the United States. In *World Environmental and Water Resources Congress*, 2008.

Dennis J Hartmann, Albert M G Klein Tank, Matilde Rusticucci, Lisa V Alexander, Stefan Brönnimann, Yassine Abdul-Rahman Charabi, Frank J Dentener, Edward J Dlugokencky, David R Easterling, Alexey Kaplan, Brian J Soden, Peter W Thorne, Martin Wild, and Panmao Zhai. Observations: Atmosphere and Surface. Technical report, Intergovernmental Pannel on Climate Change, 2013.

Ed Hawkins and Rowan Sutton. The potential to narrow uncertainty in projections of regional precipitation change. *Climate Dynamics*, 37(1):407–418, 2011.

R. Wayne Higgins and VE Kousky. Changes in Observed Daily Precipitation over the United States between 1950-79 and 1980-2009. *Journal of Hydrometeorology*, 14(1):105–121, feb 2013.

R. Wayne Higgins, VBS B. S. Silva, W. Shi, and J. Larson. Relationships between Climate Variability and Fluctuations in Daily Precipitation over the United States. *Journal of Climate*, 20(14):3561–3579, jul 2007.

James P. Hughes, Peter Guttorp, and Stephen P. Charles. A non-homogeneous hidden Markov model for precipitation occurrence. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 48(1):15–30, feb 1999.

J W Hurrell. Decadal trends in the North Atlantic Oscillation: regional temperatures and precipitation. *Science*, 269:676–679, aug 1995.

H.E. Hurst. Long-term storage capacity of reservoirs. *Trans. Amer. Soc. Civil Eng.*, pages 770–808, 1951.

Clifford M. Hurvich and Chih-Ling Tsai. Regression and time series model selection in small samples. *Biometrika*, 76(2):297–307, 1989.

Michael F Hutchinson. Interpolation of Rainfall Data with Thin Plate Smoothing Splines - Part I: Two Dimensional Smoothing of Data with Short Range Correlation. *Journal of Geographic Information and Decision Analysis*, 2(2):139–151, 1998.

Liwei Jia and Timothy DelSole. Diagnosis of Multiyear Predictability on Continental Scales. *Journal of Climate*, 24(19):5108–5124, oct 2011.

J. Kallmeyer, R. Pockalny, R. R. Adhikari, D. C. Smith, and S. D'Hondt. From the Cover: Global distribution of microbial abundance and biomass in subseafloor sediment. *Proceedings of the National Academy of Sciences*, 109(40):16213–16216, 2012.

Thomas R. Karl and Richard W. Knight. Secular trends of precipitation amount, frequency, and intensity in the United States. *Bulletin of the American Meteorological Society*, 79 (2):231–241, 1998.

Thomas R. Karl, Richard W. Knight, and Neil Plummer. Trends in high-frequency climate variability in the twentieth century. *Nature*, 337:217–220, 1995.

Richard W. Katz. Precipitation as a chain-dependent process. *Journal of Applied Meteorology*, 16(7):671–676, 1977.

Richard W. Katz. Probabalistic Models. In A. H. Murphy and Richard W. Katz, editors, *Probability, Statistics, and Decision Making in the Atmospheric Sciences*, pages 261–288. Westview Press, Boulder, CO, 1985.

Richard W. Katz and MB Parlange. Overdispersion phenomenon in stochastic modeling of precipitation. *Journal of Climate*, 11(4):591–602, 1998.

Richard W. Katz and MB Marc B. Parlange. Effects of an index of atmospheric circulation on stochastic properties of precipitation. *Water Resources Research*, 29(7):2335–2344, jul 1993.

Ben Kirtman, Scott B. Power, John Adedoyin Adedoyin, George J. Boer, Roxana Bojariu, Ines Camilloni, Francisco J. Doblas-Reyes, Arlene M. Fiore, Masahide Kimoto, Gerald A. Meehl, Michael Prather, Abdoulaye Sarr, Christoph Schär, Rowan Sutton, Geert Jan van Oldenborgh, Gabriel Vecchi, and Hui-Jun Wang. Near-term Climate

Change: Projections and Predictability. Technical report, Intergovernmental Pannel on Climate Change, New York, 2013.

BP Kirtman and PS Schopf. Decadal variability in ENSO predictability and prediction. *Journal of Climate*, pages 2804–2822, 1998.

William Kleiber, Richard W. Katz, and Balaji Rajagopalan. Daily spatiotemporal precipitation simulation using latent and transformed Gaussian processes. *Water Resources Research*, 48(1):1–17, 2012.

Alan K Knapp, Philip a Fay, John M Blair, Scott L Collins, Melinda D Smith, Jonathan D Carlisle, Christopher W Harper, Brett T Danner, Michelle S Lett, and James K McCarron. Rainfall variability, carbon cycling, and plant species diversity in a mesic grassland. *Science*, 298(5601):2202–2205, dec 2002.

David B. Knight and Robert E. Davis. Contribution of tropical cyclones to extreme rainfall events in the southeastern United States. *Journal of Geophysical Research: Atmospheres*, 114(D23), dec 2009.

Randal D. Koster, Max J. Suarez, and Mark Heiser. Variance and predictability of precipitation at seasonal-to-interannual timescales. *Journal of Hydrometeorology*, 1(1):26–46, 2000.

Randal D. Koster, Max J. Suarez, R. Wayne Higgins, and Hung M. Van den Dool. Observational evidence that soil moisture variations affect precipitation. *Geophysical Research Letters*, 30(5), 2003.

Randal D. Koster, Paul A. Dirmeyer, Zhichang Guo, Gordon B. Bonan, Edmond Chan, Peter Cox, C.T. Gordon, Shinjiro Kanae, Eva Kowalczyk, David Lawrence, Ping Liu, Cheng-Hsuan Lu, Sergey Malyshev, Bryant McAvaney, Ken Mitchell, David Mocko, Taikan Oki, Keith Oleson, Andrew Pitman, Y.C. Sud, Christopher M. Taylor, Diana Verseghy, Ratko Vasic, Yongkang Xue, and Tomohito Yamada. Regions of strong coupling between soil moisture and precipitation. *Science*, 305(5687):1138–1140, aug 2004.

Demetris Koutsoyiannis. The Hurst phenomenon and fractional Gaussian noise made easy. *Hydrological Sciences Journal*, 47(August 2002):573–596, 2002.

Demetris Koutsoyiannis. Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation. *Hydrological Sciences Journal*, 49(4):575–590, 2004.

Kenneth E. Kunkel, Karen Andsager, and David R. Easterling. Long-term trends in extreme precipitation events over the conterminous United States and Canada. *Journal of Climate*, (1998):2515–2527, 1999.

Kenneth E. Kunkel, David R. Easterling, Kelly Redmond, and Kenneth Hubbard. Temporal variations of extreme precipitation events in the United States: 18952000. *Geophysical Research Letters*, 30(17), 2003.

Kenneth E. Kunkel, David R. Easterling, David a. R. Kristovich, Byron Gleason, Leslie Stoecker, and Rebecca Smith. Meteorological Causes of the Secular Variations in Observed Extreme Precipitation Events for the Conterminous United States. *Journal of Hydrometeorology*, 13(3):1131–1141, jun 2012.

Kenneth E. Kunkel, Thomas R. Karl, David R. Easterling, Kelly Redmond, John Young, Xungang Yin, and Paula Hennon. Probable maximum precipitation and climate change. *Geophysical Research Letters*, 40(7):1402–1408, apr 2013.

CE Leith. Predictability of climate. *Nature*, 276(23):352–355, 1978.

Edward Norton Lorenz. A study of the predictability of a 28-variable atmospheric model. *Tellus*, 17(3):321–333, 1965.

Edward Norton Lorenz. Predictability and periodicity: A review and extension. *Proc. Third Conf. on Predictability and Statistics in the ...*, 1973.

S Lovejoy and Benoit B. Mandelbrot. Fractal properties of rain, and a fractal model. *Tellus A*, pages 209–232, 1985.

Wp Lowry and D Guthrie. Markov Chains of Order Greater than one. *Monthly Weather Review*, 96(11):798–801, 1968.

Peter Lynch. The origins of computer weather prediction and climate modeling. *Journal of Computational Physics*, 227(7):3431–3444, mar 2008.

J MacQueen. Some methods for classification and analysis of multivariate observations. *Proceedings of the fifth Berkeley symposium on ...*, 233(233):281–297, 1967.

RA Madden and DJ Shea. Estimates of the natural variability of time-averaged temperatures over the United States. *Monthly Weather Review*, 1978.

Roland A. Madden, Dennis J. Shea, Richard W. Katz, and John W. Kidson. The potential long-range predictability of precipitation over New Zealand. *International Journal of Climatology*, 19(4):405–421, mar 1999.

BB Mandelbrot. A fast fractional Gaussian noise generator. *Water Resources Research*, 7, 1971.

Benoit B. Mandelbrot and JR Wallis. Noah, Joseph, and operational hydrology. *Water Resources Research*, 4(5):909–918, 1968.

Michael E. Mann and Jonathan M. Lees. Robust estimation of background noise and signal detection in climatic time series. *Climatic Change*, 33(3):409–445, 1996.

Nathan J. Mantua and Steven R. Hare. The Pacific Decadal Oscillation. *Journal of Oceanography*, 58(1):35–44, 2002.

Andrey Andreyevich Markov. Extension of the law of large numbers to dependent quantities. *Izv. Fiz.-Matem. Obsch. Kazan Univ. (2nd Ser)*, 15:135–156, 1906.

GJ McCabe and MD Dettinger. Decadal variations in the strength of ENSO teleconnections with precipitation in the western United States. *International Journal of Climatology*, 19(13):1399–1410, nov 1999.

Bernard N. Meisner and Phillip A. Arkin. Spatial and annual variations in the diurnal cycle of large-scale tropical convective cloudiness and precipitation. *Monthly Weather Review*, 115(9):2009–2032, 1987.

Merab Menabde and Murugesu Sivapalan. Modeling of rainfall time series and extremes using bounded random cascades and levy-stable distributions. *Water Resources Research*, 36(11):3293, 2000.

M.J. Menne, I. Durre, B. Korzeniewski, S. McNeal, K. Thomas, X. Yin, S. Anthony, R. Ray, R.S. Vose, B.E. Gleason, and T.G. Houston. Global Historical Climatology Network - Daily (GHCN-Daily), Version 3.22, 2012.

Varaha Mihira and V. (trans.) Panditbhushan. *Brihat Samhita (with English translation)*. V.B. Soobbiah and Sons, M.B.D. Electronic Printing Works, Bangalore, 1946.

P C D Milly, Julio Betancourt, Malin Falkenmark, Robert M Hirsch, W Zbigniew, Dennis P Lettenmaier, and Ronald J Stouffer. Stationarity Is Dead: Whither Water Management? *Science*, 319:573–574, 2008.

Seung-Ki Min, Xuebin Zhang, Francis W Zwiers, and Gabriele C Hegerl. Human contribution to more-intense precipitation extremes. *Nature*, 470(7334):378–381, 2011.

Annalisa Molini, Gabriel G. Katul, and Amilcare Porporato. Causality across rainfall time scales revealed by continuous wavelet transforms. *Journal of Geophysical Research Atmospheres*, 115(14):1–16, 2010.

Vincent Moron, Andrew W. Robertson, and M. Neil Ward. Seasonal Predictability and Spatial Coherence of Rainfall Characteristics in the Tropical Setting of Senegal. *Monthly Weather Review*, 134(11):3248–3262, 2006.

Vincent Moron, Andrew W. Robertson, M. Neil Ward, and Pierre Camberlin. Spatial Coherence of Tropical Rainfall at the Regional Scale. *Journal of Climate*, 20(21):5244–5263, nov 2007.

Gunnar Myhre, Drew Shindell, Francois-Marie Bréon, William Collins, Jan Fuglestvedt, Jianping Huang, Dorothy Koch, Jean-Francois Lamarque, David Lee, Blanca Mendoza, Teruyuki Nakajima, Alan Robock, Graeme Stephens, Toshihiko Takemura, and Hua Zhang. Anthropogenic and Natural Radiative Forcing. Technical report, Intergovernmental Pannel on Climate Change, 2013.

National Climate Assessment and Development Advisory Committee. Third National Climate Assessment Report. Technical report, U.S. Global Change Research Program, 2013.

National Oceanic and Atmospheric Administration. Land-Based Datasets and Products, a.

National Oceanic and Atmospheric Administration. WPC Verification: Quantitative Precipitation Forecasts, b.

National Oceanic and Atmospheric Administration. CPC Verification Summary, c.

National Oceanic and Atmospheric Administration Climate Prediction Center. CPC Verification Summary, 2013.

National Research Council. Assessment of intraseasonal to interannual climate prediction and predictability. Technical report, National Research Council, 2010.

National Weather Service. History of the National Weather Service.

Joseph Needham. *Science and Civilization in China: Volume 3, Mathematics and the Sciences of the Heavens and the Earth*. Caves Books Ltd., Taipei, 1986.

TA Niziol, WR Snyder, and JS Waldstreicher. Winter weather forecasting throughout the eastern United States. IV: Lake effect snow. *Weather and forecasting*, 10(1):61–77, 1995.

Jonas Olsson and Janusz Niemczynowicz. Multifractal analysis of daily spatial rainfall distributions. *Journal of Hydrology*, 187(1-2):29–43, dec 1996.

Indrani Pal, Bruce T. Anderson, Guido Daniel Salvucci, and Daniel J. Gianotti. Shifting seasonality and increasing frequency of precipitation in wet and dry seasons across the U.S. *Geophysical Research Letters*, 40(15):4030–4035, 2013.

T. N. Palmer. Predictability of Weather and Climate –from theory to practice. In T. N. Palmer, editor, *Predictability of Weather and Climate*, pages 1–29. Cambridge University Press, 2006.

Hans R. Pruppacher and James D. Klett. *Microphysics of Clouds and Precipitation*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1st. edition, 1997.

S. C. Pryor and J. T. Schoof. Changes in the seasonality of precipitation over the contiguous USA. *Journal of Geophysical Research: Atmospheres*, 113(D21), nov 2008.

Adolphe Quetelet. *Sur le climat de la Belgique: des pluies, des greles, et des neiges*. L'Academie Royale, Bruxelles, 1852.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. University Press Group Limited, 2004.

V. Raulin. On the rainfall at Athens. *Quarterly Journal of the Royal Meteorological Society*, 2(16):448–450, 1875.

CW Richardson. Stochastic simulation of daily precipitation, temperature, and solar radiation. *Water Resources Research*, 17(1):182–190, 1981.

Andrew W. Robertson, Sergey Kirshner, Padhraic Smyth, Stephen P. Charles, and Bryson C. Bates. Subseasonal-to-interdecadal variability of the Australian monsoon over North Queensland. *Quarterly Journal of the Royal Meteorological Society*, 132 (615):519–542, jan 2006.

Andrew W. Robertson, Vincent Moron, and Yunus Swarinoto. Seasonal predictability of daily rainfall statistics over Indramayu district, Indonesia. *International Journal of Climatology*, 29(10):1449–1462, 2009.

C. F. Ropelewski and Michael S. Halpert. North American precipitation and temperature patterns associated with the El Niño/Southern Oscillation (ENSO). *Monthly Weather Review*, 114(12):2352–2362, 1986.

DP Rowell. Assessing potential seasonal predictability with an ensemble of multidecadal GCM simulations. *Journal of Climate*, pages 109–120, 1998.

Guido Salvucci, Daniel Gianotti, and Bruce T. Anderson. Stochastic analysis of California's recent precipitation drought in the context of the last one hundred years. In *AGU Fall Meeting Abstracts*, San Francisco, CA, 2014. American Geophysical Union.

Guido Daniel Salvucci and Conghe Song. Derived Distributions of Storm Depth and Frequency Conditioned on Monthly Total Precipitation: Adding Value to Historical and Satellite-Derived Estimates of Monthly Precipitation. *Journal of Hydrometeorology*, 1 (2):113–120, 2000.

J. T. Schoof and S. C. Pryor. On the Proper Order of Markov Chain Model for Daily Precipitation Occurrence in the Contiguous United States. *Journal of Applied Meteorology and Climatology*, 47(9):2477–2486, sep 2008.

Gideon Schwarz. Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2): 461–464, 1978.

Justin Sheffield, Andrew P. Barrett, Brian Colle, D. Nelun Fernando, Rong Fu, Kerrie L. Geil, Qi Hu, Jim Kinter, Sanjiv Kumar, Baird Langenbrunner, Kelly Lombardo, Lindsey N. Long, Eric Maloney, Annarita Mariotti, Joyce E. Meyerson, Kingtse C. Mo, J. David Neelin, Sumant Nigam, Zaitao Pan, Tong Ren, Alfredo Ruiz-Barradas, Yolande L. Serra, Anji Seth, Jeanne M. Thibeault, Julienne C. Stroeve, Ze Yang, and Lei Yin. North American Climate in CMIP5 experiments. Part I: Evaluation of historical simulations of continental and regional climatology. *Journal of Climate*, 26(23): 9209–9245, 2013a.

Justin Sheffield, Suzana J. Camargo, Rong Fu, Qi Hu, Xianan Jiang, Nathaniel Johnson, Kristopher B. Karnauskas, Seon Taekim, Jim Kinter, Sanjiv Kumar, Baird Langenbrunner, Eric Maloney, Annarita Mariotti, Joyce E. Meyerson, J. David Neelin, Sumant Nigam, Zaito Pan, Alfredo Ruiz-Barradas, RICHARD SEAGER, YOLANDE L. SERRA, DE-ZHENG SUN, CHUNZAI WANG, SHANG-PING XIE, JIN-YI YU, TAO ZHANG, and MING ZHAO. North American Climate in CMIP5 Experiments . Part II : Evaluation of Historical Simulations of Intraseasonal to Decadal Variability. *Journal of Climate*, 26:9247–9290, 2013b.

Daniel J. Short Gianotti. Occurrence Markov Chain daily precipitation model, 2016.

Daniel J. Short Gianotti, Bruce T. Anderson, and Guido D. Salvucci. The Potential Predictability of Precipitation Occurrence, Intensity, and Seasonal Totals over the Continental United States. *Journal of Climate*, 27(18):6904–6918, 2014.

R. Sibson. A brief description of natural neighbor interpolation. In Ed. V. Barnett, editor, *Interpreting Multivariate Data*, pages 21–36. Wiley, 1981.

Amy Solomon, Lisa Goddard, Arun Kumar, James Carton, Clara Deser, Ichiro Fukumori, Arthur M. Greene, Gabriele Hegerl, Ben Kirtman, Yochanan Kushnir, Matthew Newman, Doug Smith, Dan Vimont, Tom Delworth, Gerald a. Meehl, Timothy Stockdale, B Kirtman10, Y Kushnir11, and Others. Distinguishing the roles of natural and anthropogenically forced decadal climate variability: Implications for prediction. *Bulletin of the American Meteorological Society*, 92(2):141–156, feb 2011.

RCJ Somerville. The predictability of weather and climate. *Forecasting in the Social and Natural Sciences*, 11:239–246, 1987.

RD Stern and R Coe. A model fitting analysis of daily rainfall data. *Journal of the Royal Statistical Society. Series A*, 147(1):1–34, 1984.

R. Strachan. On the weather of thirteen autumns. *Quarterly Journal of the Royal Meteorological Society*, 2(11):129–139, 1875.

Y Sun, S Solomon, Aiguo Dai, and RW Portmann. How often does it rain? *Journal of Climate*, 19(6):916–934, 2006.

Y Sun, S Solomon, Aiguo Dai, and RW Portmann. How often will it rain? *Journal of Climate*, 20(19):4801–4818, 2007.

Karl E. Taylor, Ronald J. Stouffer, and Gerald A. Meehl. An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, 93(4):485–498, 2012.

R. Trachan. On the rainfall at Calcutta. *Quarterly Journal of the Royal Meteorological Society*, 3(17):21–26, 1877.

Kevin E. Trenberth. Changes in precipitation with climate change. *Climate Research*, 47 (1):123, 2011.

Kevin E. Trenberth, Aiguo Dai, Roy M. Rasmussen, and David B. Parsons. The Changing Character of Precipitation. *Bulletin of the American Meteorological Society*, 84(9):1205–1217, sep 2003.

Hristos Tyralis and Demetris Koutsoyiannis. Simultaneous estimation of the parameters of the HurstKolmogorov stochastic process. *Stochastic Environmental Research and Risk ...*, 25(1):21–33, jul 2011.

U.S. Climate Prediction Center. NOAA NCEP CPC REGIONAL US Mexico daily gridded archive, 2015.

Koen Verbist, Andrew W. Robertson, Wim M. Cornelis, and Donald Gabriels. Seasonal Predictability of Daily Rainfall Characteristics in Central Northern Chile for Dry-Land Management. *Journal of Applied Meteorology and Climatology*, 49(9):1938–1955, sep 2010.

Grace Wahba. How to smooth curves and surfaces with splines and cross-validation. Technical report, University of Wisonsin, Madison, 1979.

Hailan Wang, Siegfried Schubert, Max Suarez, and Randal D. Koster. The Physical Mechanisms by Which the Leading Patterns of SST Variability Impact U.S. Precipitation. *Journal of Climate*, 23(7):1815–1836, apr 2010.

J Wang, Bruce T. Anderson, and Guido D. Salvucci. Stochastic modeling of daily summertime rainfall over the southwestern United States. Part I: Interannual variability. *Journal of Hydrometeorology*, 7(4):739–754, 2006.

Robert de C. Ward. Rainy Days and Rain Probability in the United States. *Geographical Review*, 7(1):44–48, 1919.

Colin N Waters, Jan Zalasiewicz, Colin Summerhayes, Anthony D Barnosky, Clément Poirier, Agnieszka Ga, Alejandro Cearreta, Matt Edgeworth, Erle C Ellis, Michael Ellis, Catherine Jeandel, Reinhold Leinfelder, J R Mcneill, Daniel Richter, Will Steffen, James Syvitski, Davor Vidas, Michael Wagreich, Mark Williams, An Zhisheng, Jacques Grinevald, Eric Odada, Naomi Oreskes, and Alexander P Wolfe. The Anthropocene is functionally and stratigraphically distinct from the Holocene. *Science*, 351(6269): aad2622–1–aad2622–10, 2016.

Seth Westra and Ashish Sharma. An Upper Limit to Seasonal Rainfall Predictability? *Journal of Climate*, 23(12):3332–3351, jun 2010.

Seth Westra, Lisa V. Alexander, and Francis W. Zwiers. Global Increasing Trends in Annual Maximum Daily Precipitation. *Journal of Climate*, 26(11):3904–3918, jun 2013.

D. S. Wilks. Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. *Agricultural and Forest Meteorology*, 93(3):153–169, mar 1999.

D.S. S. Wilks and R.L. L. Wilby. The weather generation game: a review of stochastic weather models. *Progress in Physical Geography*, 23(3):329–357, jul 1999.

C. N. Williams, Russell S. Vose, David R. Easterling, and M. J. Menne. United States historical climatology network daily temperature, precipitation, and snow data, 2006.

Susan Wills and Steven R. Wills. *Meteorology: Predicting the Weather*. Oliver Press, Minneapolis, 2003.

A. W. Wood, L. R. Leung, V. Sridhar, and D. P. Lettenmaier. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change*, 62(1-3):189–216, 2004.

P. Xie. CPC unified gauge-based analysis of global daily precipitation. In *24th Conference on Hydrology*, 2010.

M. Tugrul Yilmaz and Timothy DelSole. Predictability of Seasonal Precipitation Using Joint Probabilities. *Journal of Hydrometeorology*, 11(2):533–541, apr 2010.

Xuebin Zhang, FW Zwiers, GC Hegerl, and FH Lambert. Detection of human influence on twentieth-century precipitation trends. *Nature*, 448(7152):461–5, jul 2007.

Curriculum Vitae

Daniel J. Short Gianotti

Department of Earth & Environment
Boston University
685 Commonwealth Avenue
Boston, MA 02215 U.S.A.

email: gianotti@bu.edu
URL: http://people.bu.edu/gianotti

Current position

PhD Student, Boston University

Areas of specialization

Hydroclimatology · Ecohydrology · Precipitation Predictability
Stochastic Modeling · Statistical Representations of Variability

Education

2011-2016 PhD in Geography and Environment, Boston University
Dissertation Title: Potential Predictability of Precipitation over the Continental United States
Defense Date: August 9, 2016
Committee: Bruce T. Anderson, Guido D. Salvucci, Michael C. Dietze, Dara Entekhabi, & Anthony C. Janetos (chair)

1999-2003 BS in Mathematics, Harvey Mudd College

Publications & talks

Journal articles

2016d Short Gianotti, DJ, BT Anderson, & GD Salvucci (2016), "Weather noise correctly dominates climate signals in global climate model precipitation," (*in preparation*).

2016c Short Gianotti, DJ, BT Anderson, & GD Salvucci (2016), "A kernel-auto-regressive weather generator for improved subseasonal-to-seasonal precipitation statistics," (*in preparation*).

2016b Anderson, BT, **DJ Short Gianotti**, GD Salvucci, & J Furtado (2016), "Dominant timescales of potentially predictable precipitation variations across the continental United States," Journal of Climate (*in early release*).

2016a Anderson, BT, **DJ Short Gianotti**, J Furtado, & E Di Lorenzo (2016), "A decadal precession of atmospheric pressures over the North Pacific," Geophysical Research Letters 43 (8) 3921–3927.

2015c Anderson, BT, **DJ Short Gianotti**, & GD Salvucci (2015), "Detectability of historical trends in station-based precipitation characteristics over the continental United States," Journal of Geophysical Research 120 (10) 4842–4859.

2015b Gill, AL, AS Gallinat, R Sanders-DeMott, AJ Rigden, **DJ Short Gianotti**, JA Mantooth, & PH Templer (2015), "Changes in Autumn Senescence in Northern Hemisphere Deciduous Trees: a Meta-Analysis of Autumn Phenology Studies," Annals of Botany, (Special Issue on Plants and Climate Change) 116, 875–888.

2015a Anderson, BT, **D Gianotti**, & G Salvucci (2015), "Characterizing the potential predictability of seasonal, station-based heavy precipitation accumulations and extreme dry-spell durations," Journal of Hydrometeorology 16 (2) 843–856.

2014 Short Gianotti, DJ, BT Anderson, & GD Salvucci (2014), "The Potential Predictability of Precipitation Occurrence, Intensity, and Seasonal Totals over the Continental United States," Journal of Climate 27 (18), 6904–6918.

2013b Pal, I, BT Anderson, GD Salvucci, & **DJ Gianotti** (2013), "Shifting seasonality and increasing frequency of precipitation in wet and dry seasons across the US," Geophysical Research Letters 40 (15), 4030–4035.

2013a Gianotti, D, BT Anderson, & GD Salvucci (2013), "What Do Rain Gauges Tell Us about the Limits of Precipitation Predictability?" Journal of Climate 26 (15), 5682–5688.

Conference presentations

2015 Short Gianotti, DJ*, GD Salvucci, & BT Anderson (2015) "California Drought, Weather Variability, and Climate Variability," AGU Chapman Conference on California Drought: Causes, Impacts, and Policy, Irvine CA.

2014 Short Gianotti, DJ*, BT Anderson, & GD Salvucci (2014) "Characterizing weather and climate variability for precipitation: A data-based stochastic modeling framework," American Geophysical Union Fall Meeting, San Francisco CA.

2014 Short Gianotti, DJ*, BT Anderson, & GD Salvucci (2014) "Stochastic analysis of California's recent precipitation drought in the context of the last one hundred years," American Geophysical Union Fall Meeting, San Francisco CA.

2014 Dietze, M*, HE Emery, D Gergel, **D Gianotti**, JA Mantooth, AJ Rigden (2014), "Integrating satellite and tower phenology: a case-study in real-time ecological forecasting" American Geophysical Union Fall Meeting, San Francisco CA.

2014 Dietze, M*, HE Emery, D Gergel, **D Gianotti**, JA Mantooth, AJ Rigden (2014), "Predicting phenology: A case-study in real-time ecological forecasting," Ecological Society of America Annual Meeting, Sacramento CA.

2013 Gianotti, DJ*, BT Anderson, & GD Salvucci (2013), "Potential Predictability of Precipitation: Occurrence or Intensity?" 38th Climate Diagnostic and Prediction Workshop, College Park MD.

2012 Gianotti, DJ*, BT Anderson, & GD Salvucci (2012), "Establishing Potential Predictability of U.S. Precipitation Using Rain Gauge Data," 37th Climate Diagnostic and Prediction Workshop, Fort Collins CO.

2012 Pal, I*, BT Anderson, G Salvucci, & **D Gianotti** (2012), "Magnitude and significance of observed trends in precipitation frequency over the U.S.," 37th Climate Diagnostic and Prediction Workshop, Fort Collins CO.

2012 Anderson, BT*, **D Gianotti**, & GD Salvucci (2012), "Historical expansion of the summertime monsoon over the southwestern United States: What can regional models tell us about its causes?" Regional Spectral Modeling Workshop, Scripps Institution of Oceanography, San Diego CA.

2012 Pal, I*, BT Anderson, G Salvucci, & **D Gianotti** (2012), "Magnitude and significance of observed trends in precipitation frequency over the US," American Geophysical Union Fall Meeting, San Francisco CA.

2011 **Gianotti, D***, BT Anderson, & G Salvucci (2011), "Stochastic and deterministic aspects of observed seasonal-mean precipitation variations and extreme event occurrences over the United States," American Geophysical Union Fall Meeting, San Francisco CA.

2011 Anderson, BT*, **D Gianotti**, & GD Salvucci (2011), "Detection of historical summertime monsoon precipitation variations and trends over the southwestern United States," WCRP Open Science Conference, Denver CO.

2011 Anderson, BT*, D Gianotti, & GD Salvucci (2011), "Detection of historical precipitation variations and trends over the continental United States," Department of Energy Principal Investigators Meeting, Washington DC.

2007 Schubert, DH*, **DJ Gianotti**, & K Sauers (2007), "Upgrades to a wastewater lagoon treatment system in a rural sub-Arctic community in Alaska," International Symposium on Cold Region Development, Tampere Finland.

2007 Schubert, DH*, **DJ Gianotti**, & G Jones (2007), "Application of a Thermal-hydraulic Model to Analyze and Design a Circulating Water System in Alaska," International Symposium on Cold Region Development, Tampere Finland.

2005 **Gianotti, DJ***, C Woolard, & D White (2005),"Wastewater treatment lagoon design in rural Alaska," 45th Alaska Water and Wastewater Management Association Annual Statewide Conference, Juneau AK.

** denotes presenting author*

Non-refereed research documents

2007 Schubert, DH, **DJ Gianotti**, & K Sauers (2007), "Upgrades to a wastewater lagoon treatment system in a rural sub-Arctic community in Alaska," Proceedings of the 8th International Symposium on Cold Region Development.

2007 Schubert, DH, **DJ Gianotti**, & G Jones (2007), "Application of a Thermal-hydraulic Model to Analyze and Design a Circulating Water System in Alaska," Proceedings of the 8th International Symposium on Cold Region Development.

2005 Woolard, C, **D Gianotti**, K Hardie, D White, & A Pinto (2005), "Waste Stabilization Pond Design and Performance Study," Prepared for the Alaska Department of Environmental Conservation.

2003 Gianotti, DJ (2003), "Fluid drop coalescence in a Hele-Shaw cell," Undergraduate Mathematics Thesis, Advised by A Nadim, *Harvey Mudd College*.

2002 Lampe, K, K Hultman, K Hedstrom, **D Gianotti**, E Deyo, & R Seat (2002), "Internal metrology for the Space Interferometry Mission," Undergraduate Physics Clinic Report, Advised by R Haskell, D MacDonald, & B Nemati, *Harvey Mudd College & NASA-JPL*.

Non-conference presentations

2015 Gianotti, DJ, (2015) "Weather models for climate variability," Dept. of Earth & Env. Graduate Student Presentations, Boston University.

2014 Gianotti, DJ, (2014) "Real weather, fake weather, and the California Drought," Dept. of Earth & Env. Graduate Student Presentations, Boston University.

2012 Gianotti, DJ, (2012) "How predictable is rain?" Dept. of Geography & Env. Graduate Student Presentations, Boston University.

2012 Gianotti, D, BT Anderson, & G Salvucci (2012), "Stochastic and deterministic aspects of observed seasonal-mean precipitation variations and extreme event occurrences over the United States," Science and Engineering Research Symposium, Boston University.

Published software packages

2016 Short Gianotti, DJ (2016) "Occurrence Markov Chain daily precipitation model," http://github.com/dgianotti/OMC-precip, DOI:10.5281/zenodo.45435.

Appointments held

2016-Present Research Programmer, Massachusetts Institute of Technology

2011-2015 Research Assistant, Boston University

2011 Math Teacher, Boston Public Schools

2004-2010 Private Tutor, Anchorage & Los Angeles

2007-2008 Lab Technician, California Institute of Technology

2005-2006 Environmental Engineering Associate, GV Jones & Associates

2004-2005 Research Assistant, University of Alaska, Anchorage

2003-2005 Substitute Teacher, Anchorage School District

2004 Staff, National Youth Science Camp

2001-2003 Writing Consultant, Harvey Mudd College

2002 Research Assistant, Lawrence Berkeley National Lab

Professional development

2015 *ComSciCon 2015* Communicating Science Workshop, Harvard University.

Professional service

Journal reviews

Hydrology and Earth System Sciences
Journal of Hydrometeorology
International Journal of Climatology

Memberships

American Geophysical Union
Boston Water Group