Spring August 2014

Essays on the Quantification and Propagation of Uncertainty in Climate Change Impact Assessments for Water Resource Systems

Scott Steinschneider

Essays on the Quantification and Propagation of Uncertainty in Climate Change Impact Assessments for Water Resource Systems

A Dissertation Presented

by

SCOTT STEINSCHNEIDER

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2014

Department of Civil and Environmental Engineering,
Environmental and Water Resources Engineering

Essays on the Quantification and Propagation of Uncertainty in Climate Change Impact Assessments for Water Resource Systems

A Dissertation Presented

by

SCOTT STEINSCHNEIDER

Approved as to style and content by:

_____
Casey Brown, Chair

_____
Richard Palmer, Member

_____
David Ahlfeld, Member

_____
Bernard Morzuch, Member

_____
Paul Moody, Member

_____

Richard N. Palmer, Department Head

Civil and Environmental Engineering Department

DEDICATION

To my parents, for decades of hard work that gave me the opportunity to pursue the career of my dreams.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Casey Brown, for all of his guidance and encouragement throughout my years here at UMass Amherst. His persistent dedication to my educational and professional development has been invaluable. I have only gotten to where I am today because of him. I would also like to thank Richard Palmer for his endless resolve to help guide me through my graduate education; his efforts have been greatly appreciated. I would like to extend further gratitude to the other members of my committee, David Ahlfeld, Bernard Morzuch, and Paul Moody, for all of their helpful comments and insights. Finally, I would like to acknowledge the entire faculty in the Civil and Environmental Engineering Department, whose combined support and dedication has made my experience at the University of Massachusetts Amherst truly great.

This dissertation presents work that was part of a larger effort supported by the Department of Defense through the Strategic Environmental and Research Development Program (SERDP). I would like to thank the Department of Defense for their support of this research.

Finally, I would like to thank my friends and family (and *mushmushti*) for all their encouragement and support in my time here at UMass.

ABSTRACT
ESSAYS ON THE QUANTIFICATION AND PROPAGATION OF UNCERTAINTY
IN CLIMATE CHANGE IMPACT ASSESSMENTS FOR WATER RESOURCE
SYSTEMS
MAY 2014
SCOTT STEINSCHNEIDER, B.A., TUFTS UNIVERSITY
M.S., UNIVERSITY OF MASSACHUSETTS AMHERST
Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST
Directed by: Professor Casey Brown

Sustainable water resources planning and management under climate change requires a proper treatment of uncertainties that emerge in an impacts analysis. A primary source of this uncertainty originates from the difficulties in projecting how anthropogenic greenhouse gas emissions will evolve over time and influence the climate system at regional and local scales. However, other sources of uncertainty, such as errors in modeling hydrologic response to climate and the influences of internal climate variability, compound the effects of climate change uncertainty and further obscure our understanding of water resources performance under future climate conditions. This work presents an approach to quantify the interactions, propagation, and relative contributions of different sources of uncertainty in a water resources impacts assessment under climate change. Hydrologic modeling uncertainty is addressed using Bayesian methods that can quantify both parametric and structural errors. Hydrologic uncertainties are propagated

through an ensemble of climate projections to explore their joint uncertainty. A new stochastic weather generator is presented to develop a wide ensemble of climate projections that can extend beyond the limited range of change often afforded by global climate models and better explore climate risks. The weather generator also enables the development of multiple realizations of the same mean climate conditions, allowing an exploration of the effects of internal climate variability. The uncertainties from mean climate changes, internal climate variability, and hydrologic modeling errors are then integrated in two climate change analyses of a flood control facility and a multi-purpose surface reservoir system, respectively, to explore their separate and combined effect on future system performance. The primary goal of this work is to present methods that can better estimate the precision associated with future projections of water resource system performance under climate change, and through this provide information that can guide the development of adaptation strategies that are robust to these uncertainties.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

Water resources planning and management, as traditionally practiced, has long relied on the assumption that hydroclimatic variables of interest follow a time-invariant probability distribution (i.e. they are stationary). This assumption has enabled engineers to plan for the future by testing the expected benefits and costs associated with different projects and management plans under historic hydroclimatic conditions and choosing those that best achieve a set of objectives. Over the last two decades, however, the validity of the stationary assumption has been strongly challenged [Solomon et al., 2007], with some arguing that enough evidence has been presented to preclude the use of stationarity as a justifiable, default assumption for water resource planning [Milly, 2008].

The water resources community has largely accepted that nonstationarity needs to be considered in planning considerations moving forward, and many have also recognized that significant uncertainty in future projections may hinder a clear understanding of how climate change will impact local hydrology and water resources. Over the past decade there has been an increasing emphasis in the literature on better accounting of climate change uncertainty in long-term planning efforts, with recent work emphasizing the need for risk-based approaches. Risk-based planning methods attempt to provide probabilistic information about potential impacts using scenario ensembles and relative scenario probabilities [Brekke et al., 2009], allowing decision-makers to choose a level of acceptable risk and discount impacts that do not exceed that threshold. The goal of these

planning efforts is often to identify robust decisions - those that provide an adequate level of performance across a range of climate change uncertainty - provided that some of that uncertainty can be characterized with probabilistic information drawn from climate information sources (e.g. climate projections).

There are several sources of uncertainty that need to be accounted for in risk-based approaches to water resources planning under climate change. The primary source of uncertainty stems from the difficulties in projecting how anthropogenic greenhouse gas emissions will evolve over time and influence the climate system at regional and local scales. Most studies adopting risk-based approaches have relied on downscaled future climate projections from global circulation models (GCMs) to provide an ensemble of climate scenarios and then use the relative frequencies of those downscaled projections to inform the probability analysis of future change [Dessai and Hulme, 2007; Brekke et al., 2009; Lempert and Groves, 2010]. One potential issue with this approach is that the scenarios generated from climate models may not adequately characterize the true uncertainty surrounding future climate [Stainforth et al. 2007a, Stainforth et al. 2007b], especially in hydrologic applications that depend on variables like precipitation that GCMs reproduce poorly. To circumvent this issue, *Brown et al.* [2012] introduced the methodology of Decision-Scaling, a risk-based planning approach that employs climate scenarios that are independent of and extend beyond the range of GCM projections to identify system vulnerabilities. Future climate projections produced by GCMs, as well as other climate information sources, can then be used to develop probabilistic estimates of future change in order to estimate risk. By separating the identification of system

vulnerabilities from the assessment of likelihoods of future change, this method is arguably less sensitive to climate model uncertainties because it can identify system vulnerabilities potentially unrealized under downscaled GCM projections that may be inadequate for exploring certain types of climate change. The Decision-Scaling approach has recently been successfully tested in a comprehensive study of climate risk and adaptation planning for the Great Lakes water system [Moody and Brown, 2013]. However, the literature on this topic is relatively young, and limited tools have been investigated for the production of altered climate time series over which to conduct the vulnerability assessment. This presents an opportunity to develop improved climate generation tools that can be used to extend the approach of Decision-Scaling to systems that are sensitive to a wide range of nuanced changes in climate variability at multiple temporal scales, an opportunity taken up in the second chapter of this dissertation.

In addition to the uncertainty surrounding future climate, there is also significant uncertainty surrounding the ability to estimate the hydrologic response of a local watershed. This uncertainty should not be ignored when considering the adequacy of water resource systems under future climate change. Over the past decade, there have been significant efforts to explore the integrated uncertainty of future river flows stemming from both climate and hydrologic model uncertainties. *Arnell* [1999] presented the first such study, separately exploring how a small subset of different GCMs, climate sensitivities to greenhouse gas concentrations, hydroclimatic model structures, and hydrologic model parameters influenced the projections of future continental runoff across Europe. This study, and a few others that have followed [Prudhomme and Davies,

2009; Kay et al. 2009], explored these uncertainties in isolation without investigating their integrated effects. This shortcoming has been partly resolved in other work. For instance, *Wilby and Harris* [2006] presented a comprehensive evaluation of the separate and integrated effects of various uncertainties of low river flows in England, including: 1) four different GCMs, 2) two downscaling techniques, 3) two emission scenarios, 4) two hydrologic model structures, and 5) two hydrologic parameter sets. *Chen et al.* [2011] extended this work by including more samples of several of these uncertainty sources and also included the effect of internal climate variability by considering multiple GCM initial conditions.

The results of the studies above generally conclude that GCM structure, followed by downscaling technique and internal climate variability, can have a substantial influence on the outcome of a water resources impact assessment. The influence of hydrologic uncertainties, on the other hand, is generally much smaller, particularly for hydrologic parameterization error. The insignificant effects often associated with hydrologic modeling uncertainty can in part be attributed to the limited sampling schemes used to explore this uncertainty source. While a small set of hydrologic model structures and parameter sets can provide some insight regarding this source of error, it is difficult to determine whether such a small sample size can adequately quantify the full uncertainty inherent to the hydrologic modeling process. In fact, the efforts mentioned above have largely ignored the recent advancements in methods used to formally quantify hydrologic model uncertainty. These include Pseudo-Bayesian and formal Bayesian techniques that account for both predictive and parameter uncertainties [Beven and Binley, 1992; Beven

and Freer, 2001; Bates and Campbell, 2001; Marshall et al., 2004; Stedinger et al., 2008; Schoups and Vrugt, 2010], as well as other methods that separate out input and response data errors from the analysis [Kavetski et al., 2006a, 2006b; Thyer et al., 2009; Renard et al., 2010] and address structural model uncertainty through Bayesian model averaging or other such techniques [Duan et al., 2007; Marshall et al., 2007].

There have been a handful of studies that have extended some of these more advanced methods to the integrated hydrologic and climate uncertainty problem. *Cameron et al.* [2000] was the first, using a Psuedo-Bayesian technique (the GLUE methodology [Beven and Binley, 1992]) to develop a more complete understanding of the influence hydrologic parameter uncertainty on flood response under climate change. This study was extended in *Cameron et al.* [2006] to include more GCM scenarios. A more formal Bayesian technique was used in *Kwon et al.* [2011] and was again applied in a flood frequency analysis under climate change. All of these studies found a more substantial influence from hydrologic uncertainty on impact results than those studies that only explore a handful of hydrologic model parameter sets. Yet even these more advanced studies failed to fully account for structural hydrologic uncertainties, as quantified by the model's predictive error, despite the fact that prediction error can often dominate total model uncertainty [Stedinger et al. 2008].

To date, the author has only been able to identify one study that estimated the joint effects of parametric and predictive hydrologic uncertainties in a climate change analysis using a formalized statistical approach [Khan and Coulibaly, 2010]. This study quantified

hydrologic uncertainty using a formal Bayesian analysis and compared it to the mean hydrologic model results under an ensemble of climate projections composed of two GCMs, two emissions scenarios, and two downscaling techniques. The results showed that the integrated predictive and parametric error associated with the hydrologic model encompassed the spread of mean hydrologic projections under all the individual climate ensemble members, suggesting that the fully integrated uncertainty stemming from the hydrologic modeling process can be highly significant in impacts assessments. However, this work made highly simplifying assumptions to facilitate the quantification of predictive hydrologic model error (i.e. an independent, homoscedastic, Gaussian error model). Furthermore, the fully integrated hydrologic model uncertainty was only propagated through the mean of an ensemble of climate change projections, rather than separately through each ensemble member. This approach artificially deflates the true uncertainty in future hydrologic model projections because hydrologic model error should be integrated with the range of uncertainties stemming from GCMs and downscaling techniques. The third chapter of this dissertation seeks to build upon this study to develop a more comprehensive and accurate assessment of the interactions between hydrologic and climate change uncertainties for use in water resources studies.

To this point the discussion has focused on the progress and gaps of previous work exploring integrated uncertainty assessments in hydrologic impacts studies under climate change. Yet, there is an even more disparate gap in the literature exploring the propagation of these integrated climatic and hydrologic uncertainties through a water resource systems analysis in order to delimit how they influence planning decisions. Most

studies that have attempted to account for hydrologic modeling uncertainty in water resources planning mainly focus on short-term (daily-seasonal) decision-making timescales [Georgakakos et al., 1998; Faber and Stedinger, 2001; Yao and Georgakakos, 2001; Alemu et al., 2011] and do not utilize the most recent advances in hydrologic modeling uncertainty methods referenced above. One notable exception is the work presented in Ajami et al. [2008]. In this study, hydrologic model uncertainty was quantified using the IBUNE method [Ajami et al., 2007] and propagated through a long-term planning study to assess the range of reliability, resilience, and vulnerability realized by a water supply system under different management rules. This study did not, however, consider planning uncertainties related to climate change. To the author's knowledge, there have been no attempts to formally quantify hydrologic model uncertainty, couple it with an analysis of climate change uncertainty, and assess their integrated impact on long-term water resource planning decisions. The fourth and fifth chapters of this dissertation address this research question, with the fourth chapter focused on flood control operations and the fifth chapter exploring the same issue in a long-term planning effort of a multi-objective water system.

The primary contribution of this dissertation is to present a series of tools and methods for revealing whether a water resource system or adaptations thereof are robust under integrated uncertainties from long-term climate change, internal climate variability, and hydrologic modeling capabilities. What follows are four separate chapters that present the methods explored to improve the current state of the science of integrated climate change uncertainty analyses for water resource systems.

CHAPTER 2

A SEMIPARAMETRIC MULTIVARIATE, MULTI-SITE WEATHER GENERATOR WITH LOW-FREQUENCY VARIAIBLITY FOR USE IN CLIMATE RISK ASSESSMENTS

2.1. Abstract

A multivariate, multi-site daily weather generator is presented for use in decision-centric vulnerability assessments under climate change. The tool is envisioned useful for a wide range of socioeconomic and biophysical systems sensitive to different aspects of climate variability and change. The proposed stochastic model has several components, including 1) a wavelet decomposition coupled to an autoregressive model to account for structured, low-frequency climate oscillations, 2) a Markov Chain and k-nearest-neighbor (KNN) resampling scheme to simulate spatially-distributed, multivariate weather variables over a region, and 3) a quantile mapping procedure to enforce long-term distributional shifts in weather variables that result from prescribed climate changes. The Markov Chain is used to better represent wet and dry spell statistics while the KNN bootstrap resampler preserves the covariance structure between the weather variables and across space. The wavelet-based autoregressive model is applied to annual climate over the region and used to modulate the Markov Chain and KNN resampling, embedding appropriate low-frequency structure within the daily weather generation process. Parameters can be altered in any of the components of the proposed model to enable the generation of realistic time series of climate variables that exhibit changes to both lower-order and higher-order statistics at long-term (inter-annual), mid-term (seasonal), and short-term (daily) timescales. The tool can be coupled with impact models in a bottom-up risk

assessment to efficiently and exhaustively explore the potential climate changes under which a system is most vulnerable. An application of the weather generator is presented for the Connecticut River basin to demonstrate the tool's ability to generate a wide range of possible climate sequences over an extensive spatial domain.

2.2. Introduction

The reluctance of the global community to mitigate greenhouse gas emissions and the legacy of past emissions already produced spurs the need for climate change adaptation. Recently, bottom-up or "decision-centric" approaches to identifying robust climate change adaptations have become more popular in the literature [Jones, 2001; Johnson and Weaver, 2009; Lempert and Groves, 2010; Prudhomme et al., 2010; Wilby and Dessai, 2010; Brown et al., 2011; Brown and Wilby, 2012]. These approaches focus on a system of interest (e.g. agricultural lands, an ecosystem, a reservoir, etc.) and systematically identify its vulnerabilities to climate; this contrasts "scenario-led" methods that limit the analysis to a set of climate model projections that may or may not reveal a system's climate sensitivities. A critical step in decision-centric methods involves testing the performance of a system over a range of plausible climate changes to identify harmful climate states that could cause the system to fail. As the literature on this topic is relatively young, limited tools have been investigated for the production of altered climate time series over which to conduct the vulnerability assessment. This study presents a new stochastic weather generator specifically designed to aid in these assessments. The model can be used to generate time series of weather expressing various changes in the climate at multiple temporal scales. Such time series may be especially

useful for exploring changes that are expected to occur, such as increasing intensity and decreasing frequency of precipitation consistent with the acceleration of the hydrologic cycle, or changes to low-frequency climate variability, that are not well simulated in current global climate model projections.

Bottom-up or vulnerability-based approaches to climate change adaptation form a relatively new area of research that attempts to appraise possible adaptations of a system to climate stressors by first identifying the climate vulnerabilities of that system over a wide range of potential climate changes. After system vulnerabilities are identified, different adaptation strategies can be evaluated over threatening climate states in order to identify robust adaptation measures. The likelihood of harmful climate conditions can also be assessed using available climate information, including the most up-to-date climate modeling results (e.g. global circulation model (GCM) projections). By detaching the identification of system vulnerabilities from climate projections produced by GCMs, bottom-up approaches differ from more traditional top-down approaches that depend on a limited number of internally consistent climate scenarios to explore the range of potential climate change impacts [Christensen et al., 2004; Wiley and Palmer, 2008]. It has been argued that bottom-up methods are better equipped to provide more decision-relevant information useful in identifying robust adaptation measures under deep future uncertainty [Lempert et al., 1996]. In part, this is because bottom-up approaches can better explore a full range of plausible climate changes, whereas GCM projections provide only a limited view and do not delimit the possible range (although they are often interpreted to do so) [see Stainforth et al., 2007; Deser et al., 2012].

Despite the growing interest in decision-centric approaches, technical methods for actually conducting the vulnerability assessment (i.e. generating perturbed climate sequences over which to test system vulnerability) are relatively underdeveloped. To date, only a handful of methods have been utilized. The most popular approach has been to apply simple change factors to the historic record of precipitation and temperature, effectively testing system sensitivities to mean climate shifts [Johnson and Weaver, 2009; Gober et al., 2010; Lempert and Groves, 2010; Brown et al., 2012]. Other studies have explored more detailed changes, including shifts in intra-annual climate [Prudhomme et al., 2010] and high-order statistics (e.g. variance, serial correlation) of annual hydroclimate data [Moody and Brown, 2013]. While all of these approaches were appropriate for their specific application, these methods exhibit limited ability to perturb the entire distribution of climate variables or alter their behavior at multiple temporal scales. For instance, none of the methods mentioned are equipped to simulate climates exhibiting shifts in both long-term (decadal) precipitation persistence and extreme daily precipitation amounts. Yet both of these changes are possible under climate change [Timmermann et al., 1999; Collins, 2000; IPCC, 2007] and may be important in a climate sensitivity analysis for a particular system (e.g. a reservoir jointly managed for flood risk reduction and water supply). Thus, there is a need for more generalized and comprehensive tools to conduct climate vulnerability assessments for systems sensitive to different climate variables across multiple temporal scales.

We propose stochastic weather generators as one possible tool that can fulfill this need. Stochastic weather generators are computer algorithms that produce long series of synthetic daily weather data. The parameters of the model are conditioned on existing meteorological records to ensure the characteristics of historic weather emerge in the daily stochastic process. Weather generators are a popular tool for extending meteorological records [Richardson, 1985], supplementing weather data in a region of data sparsity [Hutchinson, 1995], disaggregating seasonal hydroclimatic forecasts [Wilks, 2002], and downscaling coarse, long-term climate projections to fine-resolution, daily weather for impact studies [Wilks, 1992; Kilsby et al., 2007; Groves et al., 2008; Fatichi et al. 2011; Fatichi et al. 2013]. Their use for climate sensitivity analysis of impact models has also been tested, particularly in the agricultural sector [Semenov and Porter, 1995; Mearns et al., 1996; Riha et al. 1996; Dubrovsky et al. 2000; Confalonieri, 2012]. These sensitivity studies systematically change parameters in the model to produce new sequences of weather variables (e.g. precipitation) that exhibit a wide range of change in their characteristics (e.g. average amount, frequency, intensity, duration, etc.). By incrementally manipulating one or more parameters in the model, many climate scenarios can be simulated that exhaustively explore potential futures that exhibit slight differences in nuanced climate characteristics, such as the intensity and frequency of daily precipitation, the serial correlation of extreme heat days, or the recurrence of long-term droughts. Previous bottom-up climate impact assessments, which have relied heavily on simple change factors to generate new climate sequences, have not been able to test system vulnerabilities over such a wide range of plausible climate changes. To the authors' knowledge, only one study has used a weather generator to investigate a

system's climate sensitivity in the context of a decision-centric climate change analysis [Jones, 2000], and this study only examined changes in mean temperature and precipitation. The potential of weather generators for driving vulnerability assessments in bottom-up climate change studies has not yet been adequately explored, particular with respects to nuanced aspects of climate variability.

While the use of stochastic weather generators for bottom-up risk assessments is very attractive in theory, there are many challenges that arise in practical application. As mentioned earlier, socioeconomic and biophysical systems are often vulnerable not only to changes in mean climate, but also to changes in nuanced climate variability. Therefore, the chosen weather generator should be able to easily perturb any of these climate characteristics, which not all models in the literature can easily accomplish [Wilks and Wilby, 1999]. Additionally, impact models often require sequences of several weather variables at multiple locations that exhibit a realistic covariance structure between variables and across sites. The production of spatially distributed, correlated weather variables continues to challenge certain approaches to stochastic weather generation [Beersma and Buishand, 2003]. Weather variables can also exhibit long-term persistence [Hurst, 1951; Koutsoyiannis, 2003] on timescales up to decades that can significantly impact system performance, requiring that the chosen weather generator be capable of replicating (and possibly altering in a bottom-up analysis) structured low-frequency climate variability.

The literature is rich with examples of stochastic weather generators that can address some subset of the challenges listed above. Both parametric and non-parametric models have been proposed to maintain correlation structures between variables and across sites [Wilks, 1998; Wilks, 1999; Rajagopalan and Lall, 1999; Buishand and Brandsma, 2001; Wilby et al., 2003, Apipattanavis et al., 2007]. Some have argued that non-parametric models may be more capable than their parametric counterparts to reproduce the spatial covariance structure of multivariate weather variables [Buishand and Brandsma, 2001], but the ability to specify distributional shifts in weather variables is often more straightforward using parametric approaches [Wilks and Wilby, 1999]. Several models have also been proposed to preserve low-frequency variability observed in the historic record [Hansen and Mavromatis, 2001; Dubrovsky et al., 2004; Wang and Nathan, 2007; Chen et al., 2010; Fatichi et al. 2011; Kim et al., 2011], but these approaches have not been generalized to multi-site applications. After a substantial literature review, the authors were only able to identify one stochastic weather generator in the literature with the ability to specify distributional shifts in weather variables while simultaneously maintaining low frequency climate variability and inter-variable and inter-site correlations [Srikanthan and Pegram, 2009], and the simulation of multi-decadal climate persistence may still be difficult with this model formulation. In the context of vulnerability-based climate change assessments, a new model is required that can simultaneously simulate weather variables exhibiting accurate correlations between variables and across sites, appropriate long-term persistence at inter-annual and inter-decadal time scales, and shifted distributional characteristics hypothesized under climate change.

This study presents a stochastic weather generator with greater ability to support bottom-up vulnerability assessments under climate change for a wide range of socioeconomic and biophysical systems sensitive to different aspects of climate variability and change. The proposed stochastic model addresses all of the challenges mentioned above with several components, including 1) a wavelet decomposition coupled to an autoregressive model to account for structured, low-frequency climate oscillations, 2) a Markov Chain and k-nearest-neighbor (KNN) resampling scheme to simulate spatially-distributed, multivariate weather variables over a region, and 3) a quantile mapping procedure to enforce long-term distributional shifts in weather variables under climate change. Parameters that govern each model component can be altered to perturb various statistics of the climate system at different temporal scales. The tool can be coupled with impact models in a decision-centric risk assessment to determine the potential climate changes under which a system is most vulnerable. This allows the analyst to evaluate system performance over a wide range of possible climate changes to identify risk or to investigate specific climate change effects that are of concern (e.g. less frequent but more intense rainfall). An application of the weather generator is presented for the Connecticut River basin to demonstrate the tool's ability to generate a wide range of possible climate sequences over an extensive spatial domain. The remainder of the paper proceeds as follows. The proposed weather generator is presented in section 2.3. The model is evaluated in section 2.4, and section 2.5 demonstrates the ability of the model to produce various climate sequences for use in a bottom-up climate change analysis. The article then concludes with a discussion in section 2.6.

2.3. The Weather Generator

A flexible weather generator is desired that can accurately reproduce various characteristics of the historic climate regime while introducing the capacity to alter many of these characteristics in a decision-centric climate change analysis. The model considered in this work couples an autoregressive wavelet decomposition [Kwon et al., 2007] for extracting and simulating low-frequency structure in annual climate with a multivariate weather generator [Apipattanavis et al., 2007] that effectively captures daily weather characteristics, including dry and wet spell statistics, cross-correlations between weather variables, and spatial correlations across multiple sites. The two models are linked by conditioning the daily weather generator on simulations of annual climate produced by the autoregressive wavelet decomposition. Time series of weather variables produced by the coupled modeling approach are then altered in a third step used to enforce distributional shifts in the climate. For precipitation, a quantile mapping procedure is utilized to implement this change. Long-term shifts in other variables are enforced using simpler additive and scaling methods. A flow diagram of the overall modeling framework is given in Figure 2.1. The various sub-models and algorithms used are described in detail below.

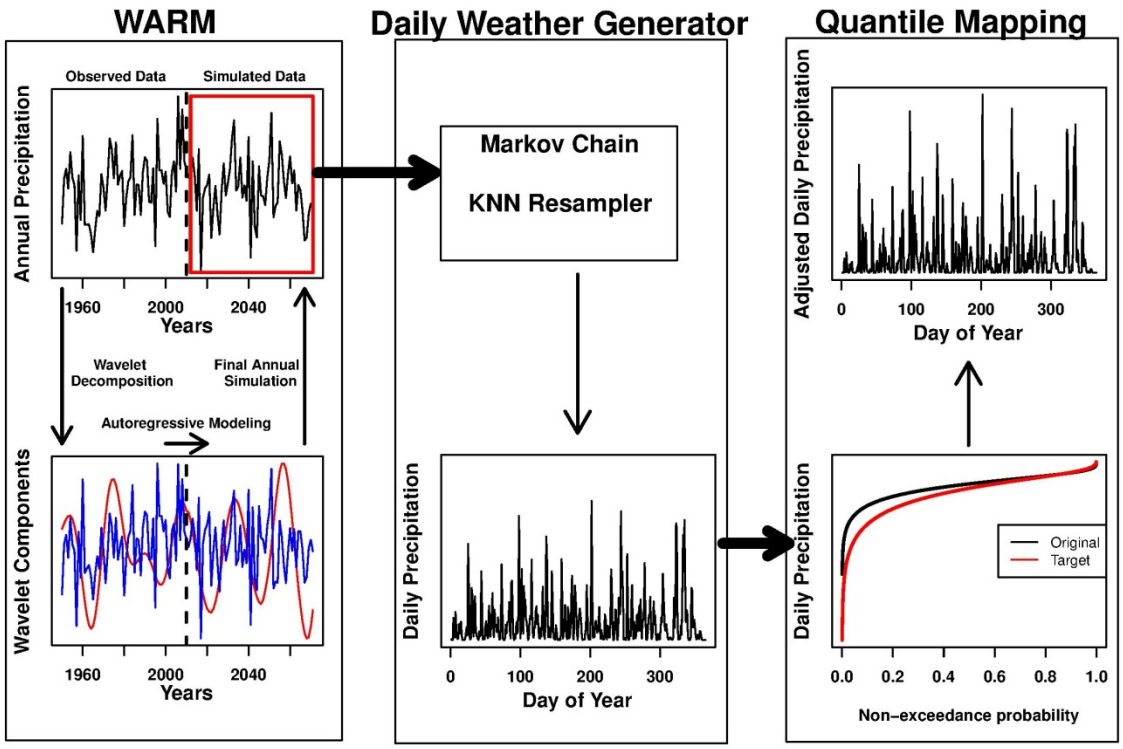

Figure 2.1. Schematic flowchart of the daily weather generation process conditional on annual simulations of climate and subject to post-process distributional adjustments.

2.3.1. Wavelet Auto-Regressive Model for the Preservation of Low-Frequency Structure

Most daily weather generators produce weather simulations that tend to be over-dispersed at inter-annual timescales and fail to reproduce observed low-frequency persistence. Several studies have proposed methods to correct for over-dispersion in weather simulations [Hansen and Mavromatis, 2001; Dubrovsky et al., 2004; Wang and Nathan, 2007; Chen et al., 2010; Fatichi et al. 2011; Kim et al., 2011]. This study utilizes a relatively new approach put forth in *Kwon et al.* [2007] that extracts low-frequency signals in climate data using wavelet decomposition and then stochastically simulates each signal using autoregressive time series models. By simulating each signal

separately, the Wavelet Autoregressive Model (WARM) can better reproduce a time series of climate exhibiting a similar spectral signature to the observed data. In our methodology, the WARM approach is applied to annual, area-averaged precipitation over the region of interest. Each year of generated annual precipitation is then used to inform a single-year simulation of the daily weather generator (described below), embedding appropriate low-frequency structure within the daily weather generation process.

Let $\tilde{\mathbf{x}}$ represent a time series of annual, area-averaged precipitation for a region. The WARM approach decomposes this series into *H* orthogonal component series, $\mathbf{z}_h$, that represent different low-frequency signals, as well as a residual noise component $\boldsymbol{\varepsilon}$.

$$\tilde{\mathbf{x}} = \sum_{h=1}^{H} \mathbf{z}_h + \boldsymbol{\varepsilon} \tag{2.1}$$

A simulation of $\tilde{\mathbf{x}}$ is generated with time series models of each low-frequency component and the residual noise. Following *Kwon et al.* [2007], we consider linear autoregressive (AR) models for each term:

$$\tilde{\mathbf{x}}_t = \sum_{h=1}^{H} \left(\sum_{v=1}^{\rho_h} \alpha_{h,v} \times \mathbf{z}_{h,t-v} + e_{h,t} \right) + \sum_{u=1}^{\rho} \beta_u \times \varepsilon_{t-u} + \xi_t \tag{2.2}$$

Here, ρ_h is the order of the AR model for the h^{th} low-frequency component, ρ is the model order for the residual noise term, e and ξ are independently and identically distributed white noise processes, and $\alpha_{h,v}$ and β_u are the AR model coefficients. Wavelet decomposition is used to generate the low-frequency components and residual noise term in equation 2.2. The wavelet transform is an analysis tool that enables the decomposition of a signal into orthogonal components in both the time and frequency domain [Torrence and Compo, 1998]. In-depth details on the implementation of the wavelet transform and its use in the WARM approach can be found in the Appendix. Time series models can be fit to each low-frequency component and the residual noise term using well-documented model fitting procedures [Box and Jenkins, 1970]. A simulated time series of annual precipitation, $\tilde{\tilde{\mathbf{x}}}$, can then be generated by summing the simulations of each component.

The daily weather generator (presented in section 2.3.2) must be conditioned on the annual climate simulations produced using WARM to embed appropriate low-frequency structure within the daily weather generation process. To achieve this, the WARM simulation is used to generate a new climate dataset for each simulation year that is composed of a weighted resampling of historic years. The daily weather generator is then iteratively fit to each new dataset for a given simulation year and run for 365 days. The methodology for conditioning the daily weather generator on WARM simulations proceeds as follows:

1. Generate a simulation of annual precipitation of length T_a using the WARM procedure.

2. For simulation year t_a, calculate the Euclidean distances $\mathbf{d} = \sqrt{\left(\tilde{\tilde{x}}_{t_a} - \tilde{\mathbf{x}}\right)^2}$ between the WARM simulated area-averaged precipitation value, $\tilde{\tilde{x}}_{t_a}$, and the vector of annual, historic, area-averaged precipitation, $\tilde{\mathbf{x}}$.

3. Order the distances from smallest to largest and assign weights to the k smallest distances using a discrete kernel function given as:

$$K[d^j] = \frac{1/j}{\sum_{j=1}^{k} 1/j} \tag{2.3}$$

Here, j indexes the first k ordered distances d^j. These weights, which are greatest for the nearest neighbor and smallest for the k^{th} neighbor, sum to 1 and thus form a discrete probability mass function. We follow the heuristic approach suggested by *Lall and Sharma* [1996] and set k equal to the square root of the number of years of historic data.

4. Sample with replacement 100 of the k-nearest neighbors based on the kernel weights from step 3. Determine the associated years of the 100 selected neighbors. Gather all of the daily data from the 100 selected years into a new dataset to be associated with simulation year t_a. We note that data may be repeated in this new dataset because years can be sampled more than once.

5. Build the daily weather generator using this conditional dataset and run it over the length of one year.

6. Repeat steps 1-5 for all T_a years of the annual WARM simulation.

2.3.2. Semiparametric Multivariate and Multisite Weather Generating Algorithm

The daily weather generation process utilized in this study is based on the methods proposed in *Apipattanavis et al.* [2007]. That study coupled a Markov Chain and KNN resampling scheme to simulate spatially-distributed, correlated, multivariate weather variables over a region. The Markov Chain is used to better represent wet and dry spell statistics while the KNN bootstrap resampler preserves the covariance structure between the weather variables and across space. Since the details of the method can be found in *Apipattanavis et al.* [2007], only a brief overview will be provided here.

Assume a simulated, daily time series of R weather variables $\mathbf{X}^l = \left\{x_{1,t}^l, x_{2,t}^l, ..., x_{R,t}^l \mid t = 1,2,...,T\right\}$ is desired at L different locations, where $x_{i,t}^l$ represents the i^{th} weather variable (e.g. precipitation) at time t and location l, and T is the length of the simulation. A weather generation scheme is designed to simulate area-averaged weather variables, $\overline{\mathbf{X}}$, that can then be immediately disaggregated to individual locations. The weather generation approach is based on the common practice of first simulating precipitation occurrence, S_t, as a chain-dependent process. A three-state (extremely wet (S_t=2), wet (S_t=1) or dry (S_t=0)) Markov Chain of order 1 is used to simulate the occurrence of area-averaged precipitation across the L locations. The number of states

and chain order can be chosen to maximize performance while maintaining model parsimony using quantitative criteria such as Akaike's information criterion [Akaike, 1974], though this study simply follows the chain structure suggested in *Apipattanavis et al.* [2007]. Nine transition probabilities ($p_{00}, p_{01}, p_{02}, p_{10}, p_{11}, p_{12}, p_{20}, p_{21}, p_{22}$) for the three-state Markov Chain are fit to the area-averaged precipitation occurrence time series by month using the method of maximum likelihood. Here, p_{ab} denotes the probability of precipitation state *b* occurring given the occurrence of state *a* on the previous day. A threshold of 0.3 mm is chosen to distinguish between wet and dry days at the area-averaged scale, while the 80th percentile of area-averaged precipitation (by month) is used as the threshold for extremely wet conditions. Again, these values are taken directly from *Apipattanavis et al.* [2007].

Area-averaged precipitation occurrence can be simulated from the fitted Markov Chain using standard procedures well documented in the previous weather generation literature. After simulating the occurrence of area-averaged precipitation states, a vector of weather variables $\overline{\mathbf{X}}$ must be simulated and then disaggregated to each of the *L* locations. A KNN resampling algorithm of lag-1 is used to generate the values for all the weather variables. This algorithm follows a six-step process:

1. Let $\overline{\mathbf{X}}_{t-1}$ be a vector of area-averaged weather variables already simulated for day *t*-1. Also assume, without loss of generality, that the Markov Chain had simulated day *t*-1 and day *t* as wet days.

2. Partition the historic record to find all pairs of days in a 7-day window centered on day t (if day t is January 15th, then the window includes all historic days from January 12th through January 18th) that have the same sequence of area-averaged precipitation states simulated by the Markov Chain for day t-1 and day t (in this case, two wet days in a row). Assume there are Q such pairs, each containing two days of area-averaged weather, $\overline{\mathbf{X}}_q^1$ and $\overline{\mathbf{X}}_q^2$.

3. Calculate the weighted Euclidean distance, d_q, between the simulated, area-averaged vector of weather variables, $\overline{\mathbf{X}}_{t\text{-}1}$, and each of the Q vectors of historic, area-averaged variables:

$$d_q = \sqrt{\sum_{i=1}^{R} w_i \times ((\overline{x}_{i,t-1} - \overline{x}_i) - (\overline{x}_{i,q}^1 - \overline{x}_i))^2} \qquad (2.4)$$

Here, $\overline{x}_{i,t-1}$ denotes the i^{th} area-averaged weather variable already simulated for time t-1, $\overline{x}_{i,q}^1$ denotes the same area-averaged weather variable on the first day of the q^{th} historic pair sampled in step 2, $\overline{x}_i$ is the mean of the ith area-averaged weather variable across all time steps, and w_i denotes the weight. In this study each weight w_i is set equal to the inverse of the standard deviation of the i^{th} weather variable, though there are methods in the literature for selecting weights in KNN resampling procedures to produce optimal forecasts [Karlsson and Yakowitz, 1987]. By centering each variable in the distance

equation about its mean and dividing by its standard deviation, we standardize values and give near-equal importance to each variable in the nearest-neighbor calculation. Prior to normalization, transformations may be required for non-Gaussian weather variables.

4. Order the distances d_q from smallest to largest. The k smallest distances are assigned weights using the same discrete kernel function presented in equation 2.3. Again, we follow the heuristic approach suggested by *Lall and Sharma* [1996] and set $k = \sqrt{Q}$.

5. Sample one of the k-nearest neighbors based on the weights developed in step 4 and record the historic date associated with that selected neighbor. Then, use vectors of weather variables $\mathbf{X}^l$ on the successive day to the recorded date for each of the L locations to simulate the multivariate, multisite weather for day t.

6. Repeat steps 1-5 for all T days of the simulation.

To begin the algorithm and generate initial values for all weather variables, data for a random day from the simulation starting month is selected from the historic record that is consistent with the first precipitation state simulated by the Markov Chain.

2.3.3. Quantile Mapping Technique to Enforce Long-Term Climate Changes

By just using the coupled models of sections 2.3.1 and 2.3.2, it is not feasible to generate weather outside of the range of historic variability, nor is it possible to change the distribution of those variables. In the context of a vulnerability assessment, this capability is critically important, particularly for precipitation, which often dominates system performance. The approach developed here incorporates a quantile mapping method to

alter the distribution of daily precipitation. Alterations to other weather variables are treated more simply using standard additive or multiplicative factors.

Let $\hat{X}_p^{m,l}$ be daily, non-zero precipitation values for month *m* and location *l* simulated from the daily weather generator. Assume the simulated precipitation amounts can be modeled by a theoretical cumulative distribution function $F_{\hat{X}_p^{m,l}}(x_0 \mid \boldsymbol{\eta})$ with parameters $\boldsymbol{\eta}$. A "target" cumulative distribution function, $F^*_{\hat{X}_p^{m,l}}(x_0 \mid \boldsymbol{\eta}^*)\, \mathrm{F}^*_{\mathrm{Pm}}(\mathrm{P_m} \le \mathrm{p} \mid \eta^*)$, is introduced that represents the projected distribution of future precipitation under a climate change. For simplicity we assume that $F_{\hat{X}_p^{m,l}}$ and $F^*_{\hat{X}_p^{m,l}}$ arise from the same distribution but differ between their parameter sets, $\boldsymbol{\eta}$ and $\boldsymbol{\eta}^*$. The parameter set $\boldsymbol{\eta}^*$ can be altered to control how the distribution of future precipitation differs from the historic observations. Many possible changes in precipitation characteristics are possible through adjustments to $\boldsymbol{\eta}^*$, including shifts in the mean, standard deviation, or extremes. For example, assume historic and projected precipitation for month *m* follow two-parameter Gamma distributions with shape and scale parameters $\boldsymbol{\eta} = \{\kappa, \theta\}$ and $\boldsymbol{\eta}^* = \{\kappa^*, \theta^*\}$. The parameter set $\boldsymbol{\eta}$ can be estimated by fitting a Gamma distribution to $\hat{X}_p^{m,l}$. Then, a new mean μ^* and variance σ^{2^*} can be specified for the target Gamma distribution, and the parameter set $\boldsymbol{\eta}^*$ can be inferred using the relationships between the parameters and the first two moments, $\mu^* = \kappa^* \times \theta^*$ and $\sigma^{2^*} = \kappa^* \times \theta^{2^*}$. If changes in the first two moments do not sufficiently account for particular shifts in higher order statistics that are of interest, the target parameter set $\boldsymbol{\eta}^*$ can be further tailored to better impose this change. Once the parameter

set $\boldsymbol{\eta}^*$ of the target distribution is specified, a quantile mapping procedure can be used to alter the distribution $F_{\hat{X}_p^{m,l}}$ of simulated non-zero precipitation to match that specified by $F^*_{\hat{X}_p^{m,l}}$ (Figure 2.2). To do this, we first determine the exceedance probability of the t^{th} value of synthesized precipitation for month *m*, $\hat{x}_{p,t}^{m,l}$, from the cdf $F_{\hat{X}_p^{m,l}}$. Then, the target cdf $F^*_{\hat{X}_p^{m,l}}$ is used to map this exceedance probability to a new precipitation amount, $(\hat{x}_{p,t}^{m,l})^*$, that is consistent with the specified distribution for climate-altered monthly precipitation:

$$(\hat{x}_{p,t}^{m,l})^* = F^*_{\hat{X}_p^{m,l}}{}^{-1}(F_{\hat{X}_p^{m,l}}(\hat{x}_{p,t}^{m,l})) \tag{2.5}$$

This procedure is repeated for each non-zero precipitation amount synthesized by the weather generator.

2.4. Model Evaluation

To evaluate the performance of the proposed weather generator, we apply it to daily weather data distributed across the Connecticut River basin in the New England region of the United States. Daily precipitation and maximum and minimum temperature are the variables included in the analysis. The data are available between January 1, 1949 and December 31, 2010 as gridded observations with a spatial resolution of approximately 144 km^2 [Maurer et al., 2002]. The Connecticut River basin drains over 31,000 square kilometers and contains a large number (260) of grid cells, enabling an evaluation of the

multi-site performance of the approach. The spatial extent of the proposed model application is quite large, and so adequate performance of the model at this spatial scale greatly supports its use for vulnerability assessments of large, spatially expansive systems. For evaluation, the model is run 50 separate times, each 62 years long (the length of the historic record). We examine the reproduction of multiple characteristics of each weather variable at several different time scales.

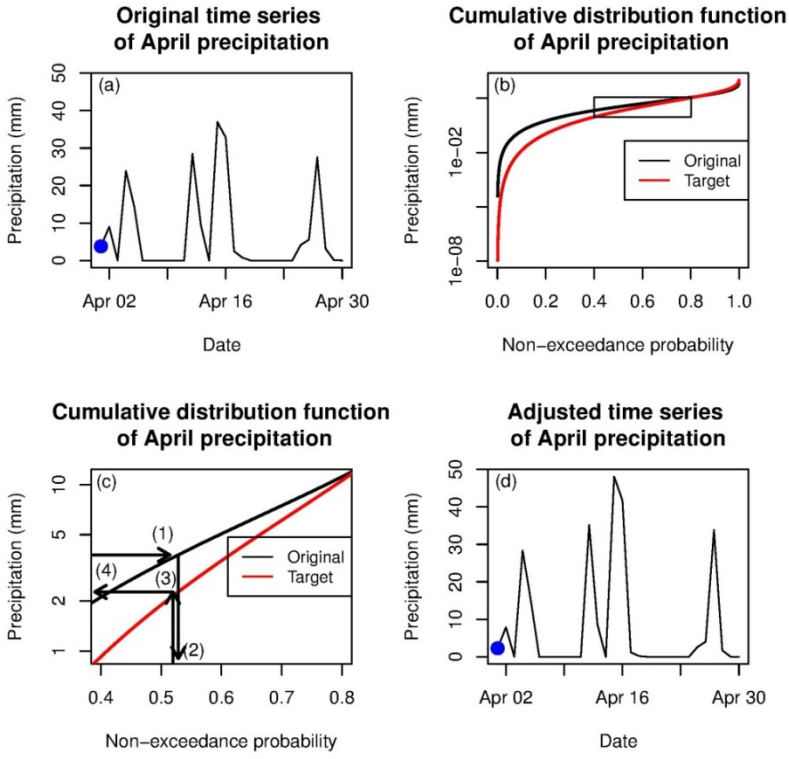

Figure 2.2. The quantile mapping procedure to adjust daily, non-zero precipitation values. a) A sample of an original time series of April precipitation simulated by the weather generator. The blue point represents a sample precipitation value to be adjusted. b) The cdf for the fitted gamma distribution to the original simulation of April precipitation (black), as well as the target cdf used to make the adjustments (red). The rectangle delimits an inset, shown in detail in (c). Here, the precipitation value represented by the blue point in (a) is mapped to a new precipitation value via four steps. The new, adjusted precipitation time series, including the adjusted point (blue), is shown in (d).

Figure 2.3 shows the mean, standard deviation, and skew of non-zero daily precipitation amounts, daily maximum temperature, and daily minimum temperature for all combinations of months and grid cells. The median values of these statistics are taken over the 50 different simulations for comparison against the historic statistics. The results suggest good performance for all variables and statistics except for the skew of daily precipitation, which tends to be underestimated in the simulations for some grid cells.

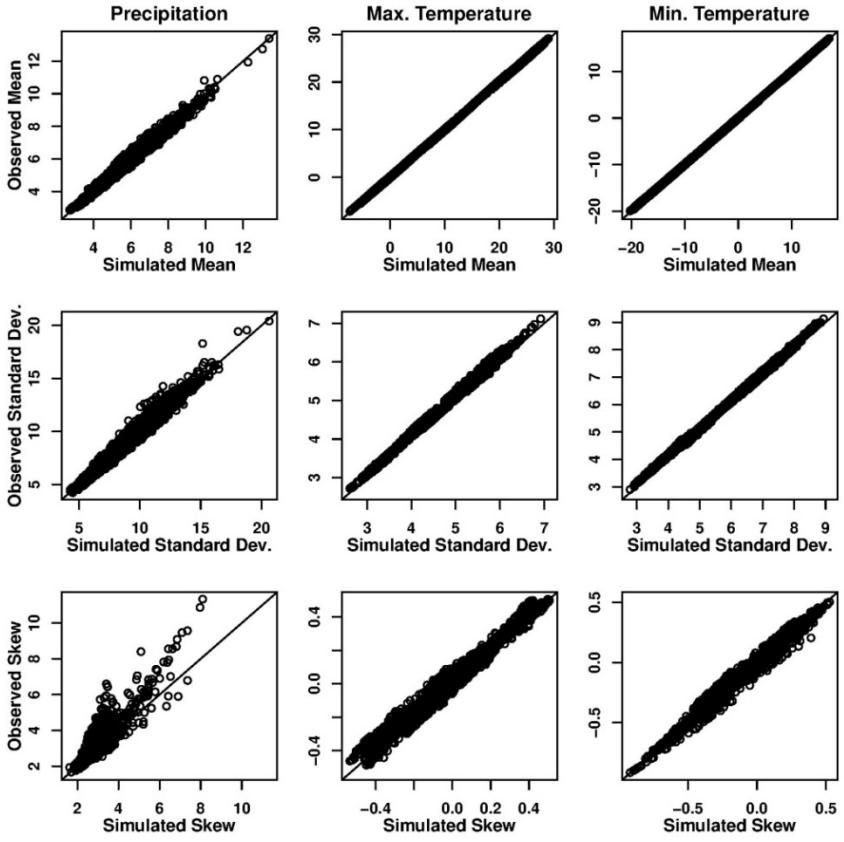

Figure 2.3. Daily performance statistics for all grid cells and months, including the mean, standard deviation, and skew of precipitation, maximum temperature, and minimum temperature. Median values across the 50 different simulations are shown against the observed values.

Correlations of a given variable across sites and cross-correlations between different variables for a given site are shown in Figure 2.4. Again, median values across the 50 simulations are shown. Both types of correlation are very well preserved, as is expected given the resampling techniques used to generate the daily weather sequences. The simulations also capture the average number of dry and wet days across all sites and months rather well (Figure 2.5). There is a slight underestimation of the average lengths of wet and dry spells, particularly for those grid cells with larger spell lengths, but this underestimation is slight (less than a day).

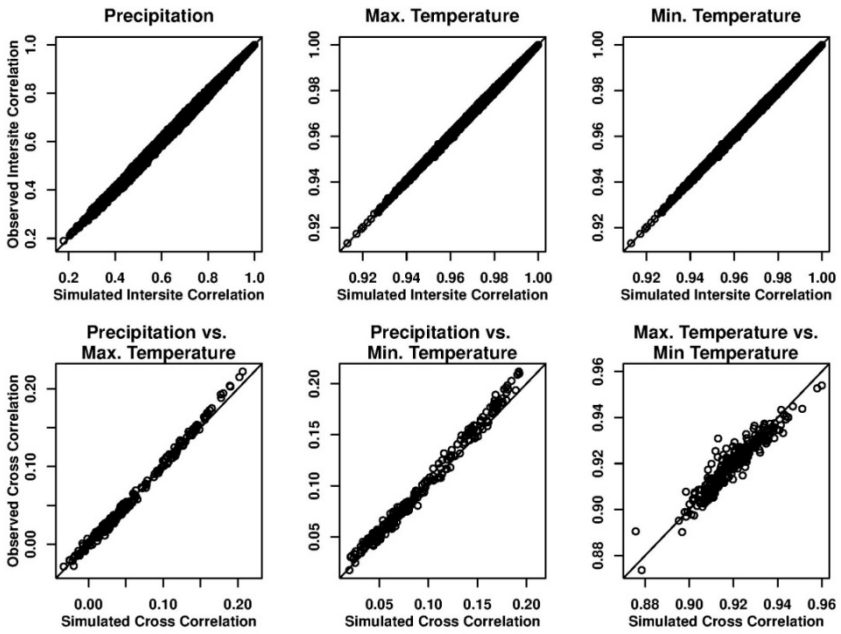

Figure 2.4. Inter-site correlations for daily precipitation, maximum temperature, and minimum temperature, as well as cross correlations between each pair of variables. Median values across the 50 different simulations are shown against the observed values for all grid cells. Correlations are taken across the entire simulation/observed record.

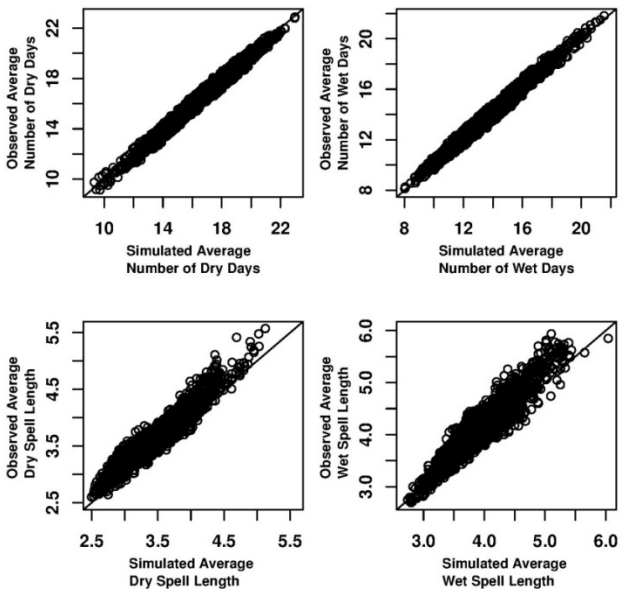

Figure 2.5. Average number of dry and wet days per month, as well as the average dry and wet spell length per month, across all grid cells. Median values across the 50 different simulations are shown against the observed values.

The spread of lag-1 autocorrelations across the 50 different simulations are shown in Figure 2.6. For each variable, the distribution of this statistic is shown for the average autocorrelation across all sites. There is a negative bias in the lag-1 autocorrelations for daily precipitation, although this bias is slight. Similarly, the simulations tend to consistently underestimate the autocorrelation in the temperature fields, but again this bias is actually rather small in magnitude. The slight underestimation of serial correlation for all variables could likely be improved by increasing the order of the Markov Chain, but no such correction was made here.

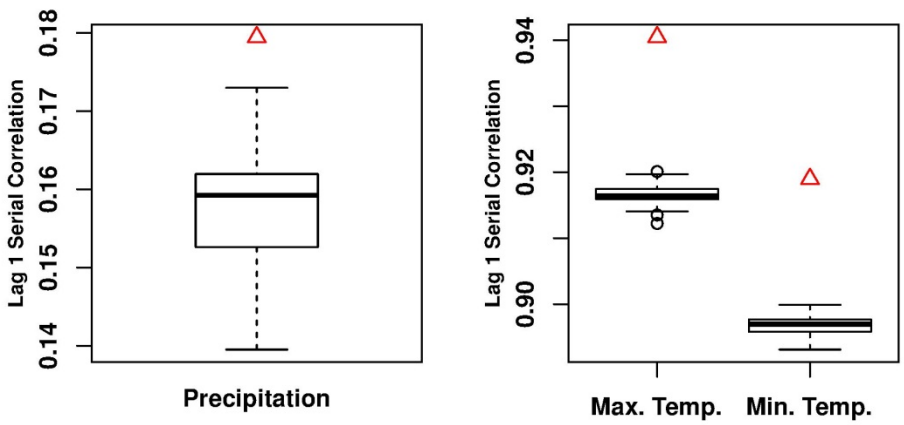

Figure 2.6. Distributions of lag 1 serial correlation values for precipitation and maximum and minimum temperature across the 50 model simulations. The average serial correlation across all grid cells is shown. Observed values are shown by the red triangles. All serial correlations are taken across the entire simulation/observed record.

To explore the reproduction of extremes, Figure 2.7 shows the distribution of 10-year and 20-year maximum annual precipitation events, as well as the average number of extreme heat days, across the 50 simulations. The precipitation extreme value estimates were developed for each grid cell by fitting a Generalized Extreme Value (GEV) distribution to the time series of annual maximum precipitation at that location. The temperature extremes were taken as the average number of days per year above 32°C. The distributions for the average of these statistics across all locations are shown for the ensemble of 50 simulations. The model tends to underestimate the magnitude of extreme rainfall events, although the spread of model simulations contains the observed value for the 10-year event and nearly reaches the observed value for the 20-year event. For temperature extremes, the model again shows a slight negative bias, although the range of simulations does contain the observed value. Overall, there is a moderate negative bias in

the extremes, an effect that can often emerge in weather generators that rely on data resampling [Lee et al. 2012].

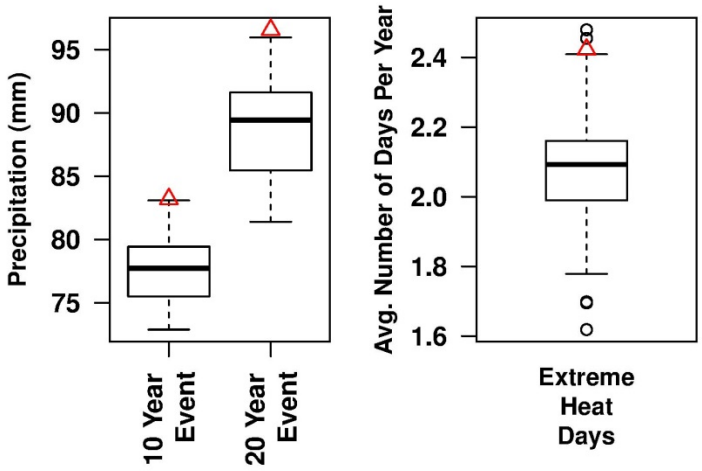

Figure 2.7. Distributions of the 10-year and 20-year precipitation event, as well as the average number of extreme heat days per year (>32°C), across the 50 model simulations. The average of each extreme event across all grid cells is shown. Estimated from the observed data are shown by the red triangles. All precipitation extreme value estimates are derived from a fitted GEV distribution.

Statistical comparisons for annual precipitation totals and temperature averages are shown in Figure 2.8. The mean precipitation and temperature fields are well preserved at the annual timescale. The standard deviation of precipitation is adequately captured for all but a few grid cells. The standard deviation of both temperature fields tends to be under-simulated, particularly for those grid cells exhibiting greater annual temperature variability. The skew for all three variables is not well captured by the model, although we note that there is significant uncertainty in the observed skew values due to the small number of annual observations available for its calculation. For precipitation and maximum temperature, the skew is overestimated for those grid cells with small skew

values and underestimated for those grid cells with larger skew values. This particular model discrepancy may be due to the fact that basin-averaged climate fields are being used to drive the model over a large and somewhat heterogeneous region.

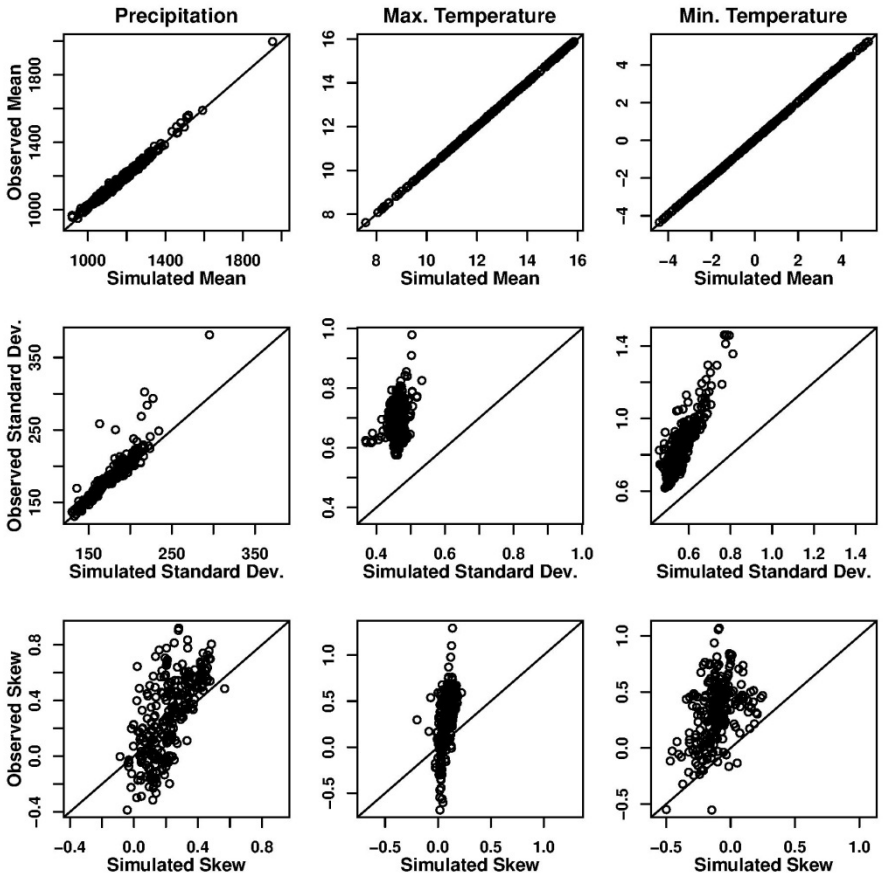

Figure 2.8. Annual performance statistics for all grid cells, including the mean, standard deviation, and skew of cumulative precipitation, maximum temperature, and minimum temperature. Median values across the 50 different simulations are shown against the observed values.

Finally, the power spectra of annual precipitation values are examined in Figure 2.9. One low-frequency component (H=1) with significant periods between 1-4 years was modeled in the WARM approach. The mean simulated power spectrum across the 50 simulations

matches that seen for the observations reasonably well. Most importantly, the mean simulated spectra become statistically insignificant at around the same period length (~4 years) as in the observations. Furthermore, the observed spectra are completely within the 95% uncertainty bounds.

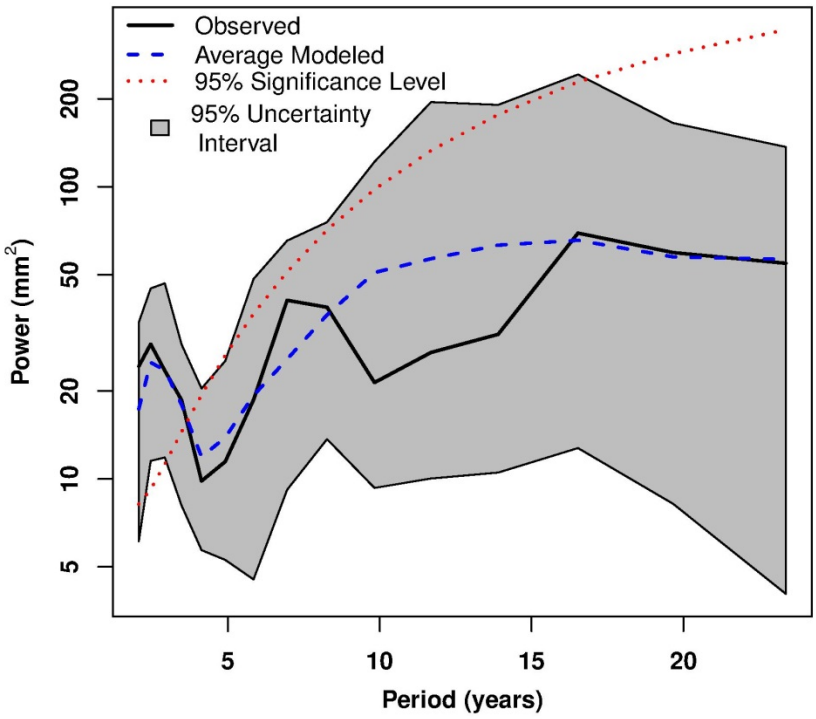

Figure 2.9. Power spectra for annual precipitation. The observed spectra (black solid) are compared against the mean power spectra (dashed blue) of the 50 simulations, along with range bounded by the 2.5th and 97.5th percentiles of the power spectra for the ensemble (grey). Also shown is the 95% significance level (red dotted) developed from a red noise background process. The power spectra of the observations and simulations become statistically significant if they rise above the red dotted line.

Overall, the performance of the model for most statistics is either good or adequate, with only some moderate discrepancies in the higher-order statistics. This is promising given that the model is being applied to a very large region subject to various changes in

topography, which can often be quite challenging for weather generation procedures. Furthermore, we note that these performance statistics are comparable to those seen in the weather generator presented in *Srikanthan and Pegram* [2009], which is the only other weather generator in the literature with the ability to specify distributional shifts in weather variables while simultaneously maintaining low-frequency climate variability and inter-variable and inter-site correlations.

2.5. Model Demonstration for a Climate Stress Test

The daily weather generator was specifically designed to facilitate a decision-centric climate risk assessment of systems sensitive to several components of the climate at various temporal scales. In the modeling framework presented here an emphasis was placed on altering precipitation patterns in the climate system because this variable often dominates the performance of biophysical and socioeconomic systems. Several parameters can be adjusted in the model to vary different components of precipitation (see Table 2.1). These include the parameters for the target distribution in the quantile mapping scheme, the transition probabilities of the Markov Chain, the coefficients of the AR model for low-frequency components, and the standard deviation of white noise for those AR models. By changing these parameters, shifts in daily precipitation amounts, daily persistence, inter-annual persistence, and inter-annual variability can be implemented in a bottom-up climate change assessment. The exact outcome of some of these perturbations will be known *a priori*, such as with the quantile mapping procedure, while outcomes from other perturbations can only be approximated prior to the simulation due to the stochastic formulation of the model. This is the case for changes in

annual persistence forced by alterations to the parameters of the WARM model. Furthermore, scaling factors and delta shifts can be applied to other climate fields (e.g. daily temperatures, wind speeds, etc.) to explore other system sensitivities to potential climate changes. Many of these changes, including those related to the quantile mapping, delta shifts, and transition probabilities, can be implemented differently by month, allowing for seasonal climate changes to be explored.

Table 2.1. Model parameters that can be altered to perturb the climate system at various temporal scales.

Climate Field	Model Component	Parameter	Effect	Timing		
				Daily	Seasonal	Inter-Annual
Precipitation	Quantile Mapping	Target Distribution Parameters (η*)	Change distribution of daily precipitation by month	X	X	
	Daily Weather Generator	Transition Probabilities (pab)	Alter daily persistence of daily precipitation by month	X	X	
	WARM	Coefficients of the AR model (α_h)	Adjust persistence of low-frequency signals			X
	WARM	Standard deviation of AR white noise (σe)	Adjust magnitude of low-frequency signals			X
Temperature	Daily Weather Generator	Delta Shifts (δt)	Shift daily temperature by month	X	X	

To demonstrate how this model could be used in a decision-centric climate risk assessment, the weather generator is used to generate several sequences of weather representing various types of climate change for the Connecticut River basin. Five types of climate change are examined here, including alterations to the mean of daily precipitation, the coefficient of variation of daily precipitation, the daily persistence of precipitation, the magnitude of low-frequency variability, and the level of persistence in that low-frequency variability. All adjustments are applied as step changes in the model rather than trended changes. The model parameters being changed and the magnitude of their perturbation are given in Table 2.2. Various combinations of these changes are presented below in order to illustrate the types of climate change that can be explored with the tool, as well as the potential, unintended consequences that may arise in other variables from the imposed parameter changes.

Figure 2.10a,b shows the changes to the distribution of non-zero daily precipitation at one grid cell in April caused by increasing the mean and coefficient of variation, respectively, for that month by 30% in the quantile mapping procedure. All other components of the climate system were kept unchanged from their historic, fitted values. Comparisons are made against a baseline model run with no changes imposed. When the mean value is increased in the quantile mapping approach, the entire distribution of daily precipitation values is shifted upwards (Figure 2.10a). These values are shifted in such a way to ensure that the variability of precipitation (i.e. the coefficient of variation) does not change. Correlations between precipitation and maximum temperature are examined to determine whether mean changes under the quantile mapping procedure degrade relationships

between precipitation and other variables (Figure 2.10d). For mean changes, these relationships appear well preserved. The distribution of daily April precipitation looks quite different when the mean is kept constant but the coefficient of variation is increased (Figure 2.10b). Here, the distribution is stretched to increase the highest events (>0.85 non-exceedance level) while lowering all of the remaining, smaller precipitation values in order to maintain the same mean value. This stretching of the distribution causes distortions in the correlations between precipitation and temperature, producing a negative bias in the correlation values across most grid cells (Figure 2.10e).

Figure 2.10c shows the average number of dry days per month across all grid cells for a model run under baseline transition probabilities in the Markov Chain and a run with increased persistence in dry days. As expected, the run with a greater persistence in dry days exhibits an increased number of these events. Unlike the results from the quantile mapping procedure, however, the change in this statistic for each grid cell can only be determined after imposing the alternative model parameterization and exploring the resulting climate sequence, because daily precipitation persistence is being modeled (and altered) at the basin-average scale. We also note that alterations to daily precipitation persistence can change the distribution of certain temperature statistics that depend on the occurrence of precipitation. For instance, increases in dry day persistence also lead to more extreme heat days (>32°C) across most grid cells (Figure 2.10f).

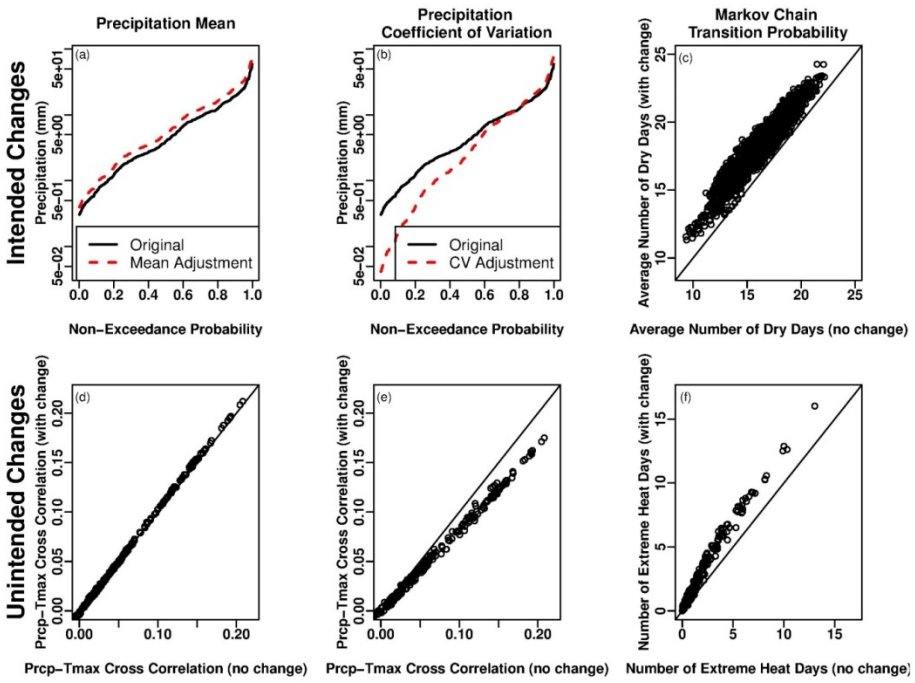

Figure 2.10. Intended (first row) and unintended (second row) changes to various weather characteristics due to forced changes in model parameters, including the mean of daily precipitation (first column), the coefficient of variation (CV) (second column), and transition probabilities in the Markov Chain (third column). Comparisons are made between a model run with the change imposed and a baseline run without any parameter changes. a,b) Baseline (black solid) and adjusted (red dashed) empirical distributions of non-zero April precipitation for a single grid cell. c) The average number of dry days per month across all grid cells. d,e) The cross correlation between non-zero precipitation and maximum temperature at each grid cell. f) The average number of extreme heat days (>32°C) per year across all grid cells.

Table 2.2. Climate changes included in the stress test. All adjustments are applied as step changes in the model rather than trended changes.

Climate Change	Model Parameter Adjusted	Size of Adjustment (All values show the size of the change above baseline values)
Mean precipitation	Mean of daily precipitation $\left(\mu^*\right)$	+/-30%
Precipitation variability	Coefficient of variation of daily precipitation $\left(\frac{\sigma^*}{\mu^*}\right)$	+30%
Daily precipitation persistence	Transition probabilities p0,1 and p0,0	-0.2 (p0,1) +0.2 (p0,0)
Magnitude of low-frequency variability	Standard deviation of white noise for all AR models (σe, $\sigma\xi$)	+/-30%
Persistence of low-frequency variability	Lag-1 coefficient for low-frequency component ($\alpha 1$)	-0.2

Finally, we present a sample of model runs exhibiting changes to the magnitude, variability, and frequency of annual precipitation. The model runs are compared against an ensemble of GCM projections to demonstrate how the weather generator can produce a much wider range of potential climate changes than the limited view afforded by the GCMs. Figure 2.11 shows the mean, coefficient of variation, and lag-1 autocorrelation coefficient for annual precipitation averaged over the entire Connecticut River basin. The statistics from several climate scenarios are presented, including those from the observed record, 234 downscaled GCM projections for the 2050-2099 period, and many different weather generator runs. The GCM projections were gathered from the World Climate Research Programm's (WCRP's) Coupled Model Intercomparison Project phase 5 (CMIP5) multi-model dataset and were downscaled using the bias-correction spatial

disaggregation technique [Wood et al., 2004; Reclamation, 2013]. Three, 20-member ensembles of weather generator runs, each 62-years long, are presented. The first set is run under baseline conditions, while the second set is run with a 30% reduction in mean precipitation and a 30% increase in the standard deviation of annual precipitation. The final ensemble is run with a 30% increase in mean precipitation, a 30% reduction in the standard deviation of annual precipitation, and a significant decrease in the lag-1 autocorrelation of annual precipitation.

Several conclusions emerge from the results in Figure 2.11. First, the ensemble of 2050-2099 GCM runs shows an increase in mean precipitation over the historic average, with a mean increase of 110% and a range of 100% and 122%. These projections show a slight decline in the average coefficient of variation, but this change is largely driven by an increase in the mean with little change in the standard deviation. Also, the projections exhibit much lower serial correlation values than that seen in the observed record, with only a handful of scenarios showing comparable levels of persistence. The historic (1950-2000) time period from these projections (not shown) exhibit the same low level of persistence as the future scenarios, suggesting that the downscaled GCM projections may not exhibit realistic, higher order climate characteristics over an aggregate region. Importantly, the magnitude, variability, and persistence of annual precipitation under these future GCM projections only exhibit a limited range of possible outcomes. This narrow view of possible future climate outcomes limits the utility of these projections in a climate change risk analysis, in which all climate possibilities, particularly high-impact, low-probability events, are important to the discovery and quantification of risk.

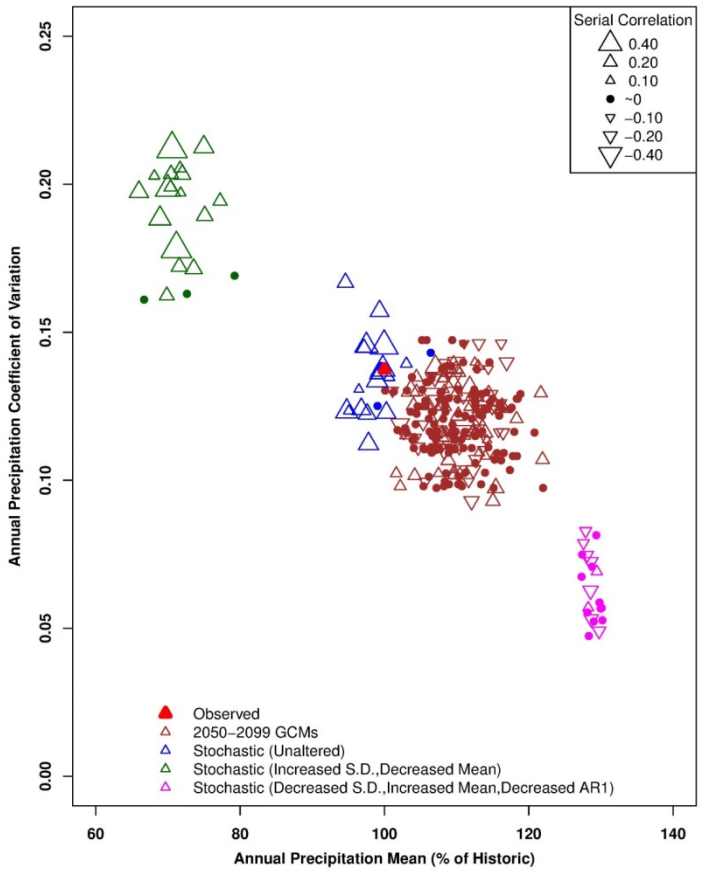

Figure 2.11. The mean, coefficient of variation, and lag-1 serial correlation coefficient of annual precipitation. Statistics for several climate scenarios are show, including 1) the observed record (red), 2) future (brown) BCSD downscaled GCM projections from the CMIP5 archive, 3) 20 baseline weather generator simulations (blue), 4) 20 simulations with a decreased mean and increased standard deviation (green), and 5) 20 simulations with an increased mean, decreased standard deviation, and decreased autocorrelation (magenta). The observed lag-1 serial correlation is 0.19.

In contrast, the 20-member ensemble of weather generator runs under baseline conditions exhibit climate characteristics that are directly comparable to the observed record. The magnitude, variability, and lag-1 autocorrelation of annual precipitation are all relatively unbiased. Furthermore, the ensemble of runs presents a range of plausible climates that

could occur even without climate change, providing an analyst with climate sequences that could be used to test the robustness of a system to internal climate variability.

A much wider range of possible future outcomes can be explored using the proposed weather generator. Figure 2.11 exhibits two possible combinations of change simulated by the model, including a set of climate sequences with significantly less but more variable annual precipitation, as well as a set of climate sequences with more annual precipitation, but with depressed variability and persistence. These two sets of changes are just a sample of what could be simulated by the weather generator, but their expansive range across climate change space demonstrates how the model could be used to explore a wide range of possible climate outcomes under climate change. This affords analysts more flexibility in how they examine the weaknesses of a system of interest and enables a more thorough exploration of climate risk. Given the tendency of planners and managers to underestimate the possibility of potential hazards, we feel that there are significant advantages to exploring system weaknesses over a wide range of possible climate outcomes, an analysis made possible by the proposed weather generator.

2.6. Discussion

2.6.1. Model Limitations

It is important to recognize the limitations of any tool when trying to infer insight from model results. While the weather generator presented in this study was designed to simulation multiple forms of climate variability at several different time scales, there are certain components of climate variability that are still challenging for the model to

account for or modulate. For one, a resampling algorithm drives the model, so at the daily time scale the tool implicitly assumes that the spatial correlation structure of the weather variables is stationary. This may not be the case under future climate changes, yet such a change cannot be simulated with this model. At inter-annual timescales, the tool currently simulates low-frequency variability based on an annual precipitation time series and ignores any signal in the annual temperature data. Also, it may be difficult to estimate robust parameters for certain low-frequency signals in the WARM model if the length of the annual precipitation time series is not sufficiently long. One approach to circumvent both of these issues would be to replace the annual precipitation time series with an alternative climate proxy that relates to both precipitation and temperature (such as an ENSO index) for which there is more data available through climate reconstructions [Kwon et al., 2009]. This requires, however, that a significant climate proxy with a long record can be found for the region of interest. Additionally, if monotonic trends, as opposed to quasi-oscillatory variability, are present in the annual data, then the WARM approach may identify spurious low-frequency components [Kwon et al. 2007]. Such trends, if identified, should be removed from the data before building the WARM model, but distinguishing trends from low-frequency oscillations is not straightforward. Finally, this model is data intensive, and therefore may be difficult to use in data-sparse regions. Despite these limitations, however, this tool does provide a step forward in the simulation of climate across multiple temporal and spatial scales for use in vulnerability assessments of human and ecological systems.

2.6.2. Determining Scenario Plausibility, Selecting the Scenario Range, and Linking to Climate Science

The model presented here was designed to support decision-centric climate change studies by enabling an analyst to test a system under a wide range of plausible climate scenarios and identify potential climate hazards. However, the analyst faces two immediate questions when trying to conduct this "climate stress test": 1) what constitutes a plausible climate change? and 2) how large should the range of climate changes be? Finding limitations on how far the climate can be perturbed before the scenario should be considered implausible is a difficult task. Expert opinion may be useful in defining these bounds, as may very large simulation ensembles of simpler (computationally faster) climate models [Piani et al. 2005]. However, the plausibility of each climate change scenario may not be critical when identifying system hazards as long as implausible changes are discounted or disregarded later in the analysis when developing estimates of climate risk [Brown et al., 2012]. The important factor is to determine how far the climate must change before the system no longer functions properly so that the analyst is aware of the potential climate hazards. Therefore, a promising strategy in bottom-up approaches may be to identify those climate variables and time scales that influence the performance of the system and then extend the range of climate changes for those variables wide enough to stress the system to failure. When those failures emerge judgments can be made regarding the plausibility of the conditions causing them; they need not be made earlier. In practice, there may be computational challenges for exploring so many scenarios, but with parallel computing capabilities the cost of an additional simulation run is often rather small. Also, adaptive sampling techniques may be utilized to reduce

the number of simulations needed to discover performance thresholds in climate change space.

Once performance thresholds in climate change space are identified, information on the likelihood of harmful climate states can be used to estimate climate risks facing the system. If certain scenarios used in the stress test are truly implausible, then the likelihood assessment should reveal this and discount these scenarios when estimating climate risk. Downscaled GCM projections are a logical starting place to garner this likelihood information, and recently there have been significant efforts in the climate science community to develop formal probability distributions of global and regional climate variables from these projections. These approaches utilize initial condition ensembles [Stainforth et al., 2005], perturbed physics ensembles [Rougier et al., 2009], multi-model ensembles [Tebaldi et al., 2005], or combinations thereof [Sexton et al., 2012] to develop pdfs of response variables. Expert opinion can also be very valuable in forming these likelihood estimates, as can data from the paleo-record. In addition, imprecise probabilities could be utilized to express uncertainty regarding the estimated values (Rinderknecht et al., 2012). Potentially, more reliable probability estimates may be developed for discrete thresholds (i.e. the likelihood of climate change beyond a threshold associated with system failure), rather than continuous probabilities across the entire climate space. In all of these cases, the probabilities of change should likely be considered subjective, but they can still be coupled with the results of the vulnerability assessment to quantitatively appraise the robustness of different adaptation measures across the range of climate change space [Moody and Brown, 2013]. More research is

needed to explore approaches for gathering this probabilistic information and coupling it with the results of an extensive vulnerability assessment.

2.7. Conclusion

The most recent scientific knowledge suggests that the impacts of climate change on socioeconomic and biophysical systems could be very significant, yet they remain highly uncertain. Recently, decision-analytic approaches have been proposed to better handle this uncertainty and frame adaptation studies under climate change in terms more relevant for decision-makers. These approaches, often bottom-up by design, require an understanding of system sensitivities to various changes in the climate system to better identify vulnerabilities and develop an understanding of potential risks to the system. However, technical methods for conducting these vulnerability assessments are relatively underdeveloped in the literature. This study presented a stochastic weather generator that can help facilitate the discovery of system vulnerabilities to several components of the climate system. When coupled with impact models, the weather generator enables a more complete identification of system vulnerabilities that can help inform risk management strategies and the selection of robust adaptation measures.

The tool is designed to work not only for specific sites but also for systems that cover large spatial extents, such as trans-state river basins or ecosystems. However, future work is needed to explore how spatially expansive the model can be made before its skill degrades. Future studies will also utilize the weather generator tool to conduct stress tests

on various socioeconomic and biophysical systems in order to appraise potential improvements from available adaptation measures.

As climatic records continue to show increasing nonstationary in their probabilistic behavior, decision makers across a range of fields will seek actionable information that directly informs a choice between measures they can take to safeguard their system from further shifts in the climate. The high degree of uncertainty that surrounds these changes hinders the utility of a traditional predict-then-act framework for adaptation decision making. A shift in philosophy may be needed to provide the information truly needed to adapt our society to potential environmental changes that we cannot foresee. This study hopefully adds to a developing body of literature exploring new methods to analyze and present climate change adaptation information that can help better inform decision makers as they navigate an uncertain future.

CHAPTER 3

TOWARDS A STATISTICAL FRAMEWORK TO QUANTIFY THE UNCERTAINTIES OF HYDROLOGIC RESPONSE UNDER CLIMATE CHANGE

3.1. Abstract

The cascade of uncertainty that underscores climate impact assessments of regional hydrology undermines their value for long-term water resources planning and management. This study presents a statistical framework that quantifies and propagates the uncertainties of hydrologic model response through projections of future streamflow under climate change. Different sources of hydrologic model uncertainty are accounted for using Bayesian modeling. The distribution of model residuals is formally characterized to quantify predictive skill, and Markov chain Monte Carlo sampling is used to infer the posterior distributions of both hydrologic and error model parameters. Parameter and residual error uncertainties are integrated to develop reliable prediction intervals for streamflow estimates. The Bayesian hydrologic modeling framework is then extended to a climate change impacts assessment. Ensembles of baseline and future climate are downscaled from global circulation models and used to drive simulations of streamflow over parameters drawn from the posterior space. Time series of streamflow statistics are calculated from baseline and future ensembles of simulated flows. Uncertainties in hydrologic model response, sampling error, and the range of future climate projections are integrated to help determine the level of confidence associated with hydrologic alteration between baseline and future climate regimes. A case study is conducted on the White River in Vermont, USA. Results indicate that the framework can

be used to present a reliable depiction of the range of hydrologic alterations that may occur in the future.

3.2. Introduction

The threat of nonstationary hydrology has motivated significant research efforts investigating the potential impacts of climate change on regional hydrology and implications for local water resource systems. Despite these efforts, uncertainty in both future climate conditions and regional hydrologic response confounds the interpretation of results and diminishes their utility in water resources planning [Lopez et al., 2009]. A systematic approach is required to account for the uncertainty in hydrologic impact assessments so that decision-makers can consider adaptation strategies contextualized by the uncertainty in design statistics critical to the decision-making process. In this paper we propose a statistical framework that quantifies several sources of uncertainty in long-range projections of hydrologic alteration, including uncertainties in future climate, hydrologic model predictive skill, model parameterization, and sampling error of estimated hydrologic statistics. These uncertainties are integrated to develop a probabilistic description of potential alterations to regional hydrology useful for water resources planning.

In the vast majority of studies, hydrologic alteration under climate change is assessed using future climate scenarios, as simulated by global circulation models (GCMs), that are downscaled to a location of interest and used to force a regional hydrologic model.

The simulated hydrologic response is then compared to a baseline response based on historic climate data, and measures of hydrologic alteration are computed [Gleick, 1986]. There are multiple sources of uncertainty that degrade this process, including those associated with the GCMs (i.e. inaccuracy at sub-continental scales, inconsistencies across models, parameterization, uncertain boundary conditions, difficulty in assessing predictive skill), the ambiguity between different downscaling techniques, and the hydrologic model (i.e. model structure, input and output data used for calibration, parameterization) [Wood, 1997]. GCM accuracy and consistency, along with the choice of downscaling methodology, are considered to be the primary sources of uncertainty and have garnered significant research attention [Raisanen and Palmer, 2001; Palmer and Raisanen, 2002; Piani et al., 2005; Stainforth et al., 2005, Fowler et al., 2007; Stainforth et al., 2007a; Lopez et al., 2009]. Errors associated with the hydrologic model, however, have received less emphasis in studies considering hydrologic alteration under climate change. In the majority of climate change impact assessments, hydrologic simulations of future climate are treated largely as deterministic output that can be used to directly identify hydrologic alterations [Chao, 1999; Hamlet and Lettenmaier, 1999; Lettenmaier et al., 1999; Nijssen et al., 2001]. Some studies have explored the impacts of hydrologic model uncertainty on climate impact assessment results, but they often only investigate uncertainties in parameterization [Arnell, 1999; Cameron et al., 2001; Wilby, 2005], model structure [Boorman and Sefton, 1997; Jiang et al., 2007], or a combination of both [Wilby and Harris, 2006; Kay et al., 2009; Prudhomme and Davies, 2009a; Prudhomme and Davies, 2009b], and almost never formally account for prediction error, which can often dominate total model uncertainty [Stedinger et al. 2008].

While parameter and structural errors are important components of the total uncertainty in hydrologic model results, accounting for these uncertainties alone may not guarantee reliable predictive bounds for streamflow estimates. For a watershed exhibiting significant heterogeneity or unexplainable behavior, many types of hydrologic response may be challenging to simulate even with an ensemble of model structures or parameterizations. The assumption that a set of hydrologic models with multiple parameterizations is complete enough to reliably bound true hydrologic response is difficult to verify [Renard et al., 2010]. This is especially true if the models struggle to reproduce certain aspects of the observed streamflow and exhibit errors that vary across the magnitude and timing of hydrologic responses. To generate reliable predictive bounds, a formal quantification of residual error is needed. If predictive uncertainty associated with the hydrologic model is not formally addressed and propagated through climate change impact analyses, claims of hydrologic alteration from such studies can be overstated and misguide water resources decision makers.

In a related line of research, predictive uncertainty in hydrologic modeling has been extensively explored and mature methods for quantifying error have been developed. Early efforts focused on pseudo-Bayesian methods [Beven and Binley, 1992; Beven and Freer, 2001], and later more formal Bayesian techniques emerged to properly account for both residual and parameter uncertainties [Bates and Campbell, 2001; Marshall et al., 2004; Stedinger et al., 2008; Schoups and Vrugt, 2010]. Further studies have dissected

model error into its component parts, investigating the impacts of uncertain input and response data on model predictions [Kavetski et al., 2006a, Kavetski et al., 2006b; Thyer et al., 2009; Renard et al., 2010]. Other innovative approaches for assessing hydrologic model uncertainty include Bayesian Recursive Estimation [Thiemann et al., 2001], Bayesian Hierarchical Mixture of Experts [Marhall et al., 2007], and Simultaneous Parameter Optimization and Data Assimilation [Vrugt et al., 2005; Clark and Vrugt, 2006], among others. These techniques can be extended to climate impact studies to quantify the total uncertainty in hydrologic models and demonstrate the extent to which it obscures the differences between future and baseline hydrologic conditions.

To the authors' knowledge, only one study has attempted to simultaneously quantify hydrologic model prediction and parameterization error and then propagate that uncertainty through climate impact assessments of hydrologic alteration [Khan and Coulibaly, 2010]. This study employed a Bayesian neural network rainfall-runoff model to explore climate-impacted hydrology. In this study, the posterior distribution of model parameters and the final distribution of model predictions were assumed Gaussian to improve the tractability of Bayesian integrals, despite the availability of Markov chain Monte Carlo (MCMC) sampling procedures that allow for more complex and accurate distributional assumptions. More importantly, uncertainty bounds were only generated for the streamflow trace generated using the mean of ensemble climate members, rather than for each climate member individually. This approach artificially deflates the true uncertainty in future hydrologic model projections because hydrologic model error

should be integrated with the range of uncertainties stemming from GCMs and downscaling techniques.

The study presented here will contribute to the science of hydrologic uncertainty analysis under climate change by developing a framework in which hydrologic model error is formally characterized and appropriately integrated with other sources of future climate uncertainty to better quantify the total uncertainty of hydrologic alterations under future climates. This allows a comparison of the range of projected changes in streamflow due to climate change to be compared with the uncertainty due to hydrologic model error. Hydrologic model prediction error is formally characterized with an appropriate likelihood function and combined with prior distributions of model parameters using Bayes' Theorem. MCMC sampling is used to evaluate the posterior distributions of hydrologic and error model parameters. Reliable uncertainty bounds for streamflow estimates are constructed from the integration of parameter and residual uncertainties and evaluated over the historic record. The Bayesian hydrologic modeling framework is then extended to a climate change impacts assessment. Ensembles of baseline and future climate data are downscaled from GCMs and used to drive simulations of streamflow over parameter samples drawn from the posterior space. While GCM projections do not fully capture climate change uncertainty, the range of climate projections can be described as an estimate of the irreducible range of climate uncertainty, a minimum bound [Stainforth et al., 2007a; Wilby and Dessai, 2010]. Time series of streamflow statistics are generated from baseline and future ensembles of simulated flows. Appropriate probability distributions are then fit to these statistics, enabling the

estimation of streamflow quantiles and their sampling error for the ensemble of baseline and future conditions. Quantile estimates are directly compared between baseline and future scenarios in the context of their cumulative uncertainties. The framework can be used to highlight the complex interactions between different sources of uncertainty and their effects on future estimates of design flow statistics used in decision-making. An application of this framework is presented for the White River Basin in Vermont using a version of the monthly ABCD hydrology model [Thomas, 1981] with a snow component.

The paper will proceed as follows. Section 3.3 provides background on Bayesian inference techniques in rainfall-runoff modeling and their potential use for error propagation in future hydrologic simulations. Section 3.4 delineates the methodology used to quantify the total uncertainty of hydrologic alteration under future climate change scenarios. The methodology is applied and results presented in Section 3.5, and the study concludes in Section 3.6 with a discussion of future research needs.

3.3. Bayesian methods in hydrologic modeling and their use in climate change studies

Bayesian methods provide a formal mechanism to characterize the error in hydrologic model predictions, along with uncertainties surrounding parameterization. In a Bayesian framework, previous knowledge about parameter values can be incorporated into model calibration through a probability density function (pdf) known as the prior distribution. A joint pdf is then used to summarize the distribution of model residuals, and MCMC sampling procedures can be used to characterize the posterior distributions of hydrologic

and error model parameters. If the error model correctly represents the distribution of model residuals, parameter and residual uncertainties can be integrated to develop predictive bounds for streamflow estimates. A relatively simple Bayesian formulation for rainfall-runoff modeling is described below that can be employed to help propagate uncertainties in a climate change impacts analysis. The Bayesian formulation presented below can be used to emphasize the importance of prediction error in uncertainty analyses under climate change and highlight the complex interactions between different sources of modeling uncertainty. Later on (Section 3.6), we discuss other challenges (e.g. source separation of uncertainties, choice of error model, and model structural errors) facing a complete quantification of hydrologic modeling uncertainty and their implications for the framework presented in this work.

3.3.1. Bayesian hydrologic modeling

Let a conceptual rainfall runoff model be formulated as follows,

$$\boldsymbol{Q} = M\ (\boldsymbol{\theta}_{\mathrm{M}}, \boldsymbol{X}\ ) + \boldsymbol{\varepsilon} \tag{3.1}$$

where $\boldsymbol{Q}$ equals the vector of observed streamflows of length n, $\boldsymbol{\theta}_M$ equals the set of hydrologic model parameters, $\boldsymbol{X}$ equals the matrix of inputs, $\hat{\boldsymbol{Q}}$ =M($\boldsymbol{\theta}_M$, $\boldsymbol{X}$) represents the streamflow model predictions, and $\boldsymbol{\varepsilon}$ equals residual model errors. Model residuals are assumed to follow a probability distribution described by a hypothesized joint pdf with a set of residual error model parameters $\boldsymbol{\theta}_{\square}$. Initially, no assumptions are made regarding the functional form of the error model $\boldsymbol{\varepsilon}(\boldsymbol{\theta}_{\varepsilon})$. That is, model residuals may be autocorrelated, non-Gaussian, or heteroskedastic. However, we assume that errors

associated with input data measurements, response data measurements, and model structure are aggregated into the error term ε. The implications of this simplifying assumption are discussed in Section 3.6.1.

Before proceeding with calibration, all previous knowledge about the set of hydrologic and error model parameters, $\boldsymbol{\theta} = \{\theta_M, \theta_\square\}$, is summarized in a prior distribution, denoted $P(\boldsymbol{\theta})$. If no prior information is available, vague priors can be used so that calibration is driven by observed data only. The likelihood function, $L(\boldsymbol{Q} | \boldsymbol{\theta}, \boldsymbol{X})$, is based on the error model and is essentially a measure of hydrologic model skill. For certain hydrologic models applied at coarse temporal resolutions the choice of error model may be relatively simple, while many other applications may require more care in the identification of an appropriate error model [Kuczera, 1983]. These issues are discussed further in Section 3.6.2. With an error model and associated likelihood function chosen, Bayes' Theorem can then provide the joint posterior distribution of all model parameters,

$$P(\boldsymbol{\theta} | \boldsymbol{Q}, \boldsymbol{X}) = \frac{L(\boldsymbol{Q} | \boldsymbol{\theta}, \boldsymbol{X}) \times P(\boldsymbol{\theta})}{\int L(\boldsymbol{Q} | \boldsymbol{\theta}, \boldsymbol{X}) \times P(\boldsymbol{\theta}) \times d\boldsymbol{\theta}} \tag{3.2}$$

The integral in the denominator is a constant of proportionality required to ensure that the right hand side term is a well-defined probability density function. MCMC methods can be used to evaluate the joint posterior distribution by sampling parameter values that are consistent with the combined information of the data and prior knowledge.

To calculate predictive bounds on simulated streamflow, uncertainties in both model parameters and predictive skill need to be integrated. A time series of predicted percentiles, $\boldsymbol{Q}_\alpha$ Q_α, for the 1-α non-exceedance level can be constructed for the vector of true streamflows $\boldsymbol{Q}$ as follows [Schoups and Vrugt, 2010]:

$$P(\boldsymbol{Q} \le \boldsymbol{Q}_\alpha \mid \boldsymbol{X}) = freq([M(\boldsymbol{\theta}_{M,j}, \boldsymbol{X}) + \boldsymbol{\varepsilon}(\boldsymbol{\theta}_{\varepsilon,j})]_{j=1,\ldots,J} \le \boldsymbol{Q}_\alpha \mid \boldsymbol{X}) = \alpha \qquad (3.3)$$

where, j=1,…, J is the number of parameter sets sampled from the posterior distributions of $\boldsymbol{\theta}_M$ and $\boldsymbol{\theta}_\varepsilon$. That is, J samples of model estimates, $M(\boldsymbol{\theta}_{M,j},\boldsymbol{X})$, and model errors, $\boldsymbol{\varepsilon}(\boldsymbol{\theta}_{\varepsilon,j})$, are generated for each simulated time step to produce a pdf of predicted values from which the predicted percentile can be inferred. The notation *freq()* is used to acknowledge that the probability of the true vector of streamflows $\boldsymbol{Q}$ falling below the vector of percentiles $\boldsymbol{Q}_\alpha$ is approximated using the frequency with which the sum of model predictions and errors fall below those percentiles. A 95% predictive bound around the time series of true streamflows $\boldsymbol{Q}$ can be formed with the bounded region [$\boldsymbol{Q}_{.025}$, $\boldsymbol{Q}_{.975}$]$Q_{0.025}, Q_{0.975}$]. If $\boldsymbol{\varepsilon}(\boldsymbol{\theta}_\varepsilon)$ is set to zero, then model error associated with parameter uncertainty can be isolated.

3.3.2. Integrating uncertainties from the hydrologic model and future climate projections

Uncertainty in future climate must be integrated with errors from the hydrologic model to develop an appropriate range of possible hydrologic alterations. These uncertainties arise

primarily from errors inherent to GCM simulations, which have been shown to exhibit poor skill at predicting even mean climate conditions at sub-continental scales [Wood, 1997; Stainforth et al., 2005]. Additional uncertainty stems from the downscaling technique used to transfer coarse GCM climate fields into meaningful climate changes at the local scale [Fowler et al., 2007]. The climate science literature is ripe with studies exploring different methods to quantify future climate uncertainty. This study does not aim to thoroughly review all of these approaches or examine the merit of each. Rather, a brief overview of common methods is presented and then one method is chosen to demonstrate how future climate uncertainties can be nested in a framework aimed at quantifying the total uncertainty in future hydrologic projections.

The most common approach relies on an ensemble of future climate scenarios to bracket possible climate changes. These scenarios are developed using climate simulations from multiple GCMs that have been forced with several emission scenarios and initiated with different starting conditions, often downscaled with only one technique. Some studies have attempted to address downscaling uncertainty by using multiple downscaling methods [Wilby and Harris, 2006]. Other studies have attempted to assign non-uniform probabilities to different projections, using measures of bias and convergence to inform the choice of probabilities [Tebaldi et al., 2005]. No matter how they are used, however, direct use of downscaled, multi-model GCM output as forcing data can only generate a lower bound on the maximum range of future climate uncertainty [Stainforth et al., 2007a]. Since GCM simulations over the historic record do not fully explore the multiple sources of uncertainty at play, it is difficult, if not impossible, to develop a satisfying

error model and bracket the true uncertainty of future climate projections. The quantification of future climate uncertainty remains largely intractable at present, as expectations for future experiments is that the uncertainty will increase. This study considers the simplest and most common quantification of future climate uncertainty where an ensemble of Z projections of future climate developed from several GCMs and emissions scenarios are downscaled to the region of interest using one downscaling technique. This represents a minimum range of climate uncertainty but allows a comparison of the range of GCM projections to hydrologic modeling uncertainty. The framework presented in Section 3.4 can easily be extended to accommodate more complex quantifications of future climate uncertainty.

3.4. Framework to quantify hydrologic uncertainties under future climate scenarios

The Bayesian hydrologic model described in the previous section can be used to help quantify the uncertainty of important streamflow statistics, **Y**, generated under baseline and future climate conditions. Here, **Y** is a time series of a statistic of interest (e.g. average annual flows, average monthly flows, annual peak flows, etc.) calculated from a simulated time series of streamflow. We present an approach that quantifies uncertainty in inferred quantiles of **Y** stemming from future climate projections, the hydrologic model, and sampling error.

Assume that Z climate change projections are available to provide a model-based range of possible future climate changes. For each climate change projection $z \in Z$, streamflow

simulations are generated from one of two sequences of climate drawn from z: 1) a baseline series, $\boldsymbol{X}_z^b$, which is generated from a downscaled time series of historical (1950-1999) conditions, or 2) a future series, $\boldsymbol{X}_z^f$, generated from a downscaled time series of future conditions (2050-2099). Hereafter, all baseline (b) and future (f) variables will be denoted with superscripts. After the hydrologic model is calibrated in the Bayesian framework to historic observations, Monte Carlo resampling is used to select K parameter sets from the posterior parameter space over which to simulate an ensemble of K streamflow traces for both baseline and future climates. These ensembles capture the parameter and residual uncertainties in the hydrologic model. The simulation procedure can be repeated for each climate sequence $z \in Zz \in Z$, producing a total of K×Z streamflow simulations for both baseline and future climates. Time series of streamflow statistics, $Y_{z,k}^b$ and $Y_{z,k}^f$, can then be developed from these K×Z baseline and K×Z future streamflow projections.

To make an inference on the p^{th} quantile, Y_p, of the statistic Y, sampling error in the estimation of Y_p must also be propagated through the analysis. If the climate projections are of limited length, then sampling error could contribute significantly to uncertainties in quantiles of projected hydrologic statistics and therefore need to be accounted. For each climate projection z and posterior parameter sample k, appropriate pdfs can be fit to the baseline $\boldsymbol{Y}_{z,k}^b$ and future $\boldsymbol{Y}_{z,k}^f$ time series. Since both time series are of limited length, the true parameter values of the fitted pdfs will be unknown, but their uncertainty can be described using their sampling distributions. D samples of the p^{th} quantiles $(Y_{z,k}^b)_p$ and

$(\gamma^{f}_{z,k})_{p_{d=1,\ldots,D}}$ can be estimated using D draws from the sampling distributions of the fitted probability model parameters. Predictive bounds for the quantiles $(\gamma^{b})_{p}$ and $(\gamma^{f})_{p}$ can then be estimated at some confidence level (1-α)100% using the $\frac{\alpha}{2}$ and $(1-\frac{\alpha}{2})$ percentiles of the ensemble of Z×K×D estimates of the p^{th} quantile under both baseline and future climate conditions. We note here that an alternative approach to splitting the climate into pseudo-stationary baseline and future time periods would be to fit a non-stationary probability model [Khaliq et al., 2006] to a transient climate over the entire timeframe (1950-2099).

The methodology proceeds as follows (Figure 3.1):

1. Calibrate the hydrologic model over a set of historic climate and streamflow observations as stated in Section 3.3 to develop posterior distributions of hydrologic and error model parameters. Evaluate the model using a split-sampling testing procedure. If possible, conduct a differential split-sample test to determine the capacity of the model to adequately model changes in climate [Klemes, 1986].

2. Sample K hydrologic and error model parameter sets, $\boldsymbol{\Theta}_M = \{\boldsymbol{\theta}_{M,1}, \boldsymbol{\theta}_{M,2}, \ldots \boldsymbol{\theta}_{M,K}\}$ and $\boldsymbol{\Theta}_\varepsilon = \{\boldsymbol{\theta}_{\varepsilon,1}, \boldsymbol{\theta}_{\varepsilon,2}, \ldots \boldsymbol{\theta}_{\varepsilon,K}\}$ from their posterior distributions developed in step 1.

3. For the i^{th} scenario of future climate, z_i, selected from an ensemble of projections **Z**, develop a baseline climate sequence of length N, $\boldsymbol{X}^{b}_{z_i} = \{\boldsymbol{x}^{b}_{z_i,1}, \boldsymbol{x}^{b}_{z_i,2}, \ldots, \boldsymbol{x}^{b}_{z_i,N}\}$, that represents historic climate. Climate information in $\boldsymbol{X}^{b}_{z_i}$ can be downscaled from a historic

period (i.e. 1950-1999) simulated in z_i, and can include variables such as temperature, precipitation, potential evapotranspiration, etc.

4. Generate a future climate sequence, $\boldsymbol{X}^{f}_{z_i}$, that is downscaled from the same climate projection but is representative of some future time period (i.e. 2050-2099).

5. To develop the k[th] time series of baseline streamflow, $\boldsymbol{Q}^{b}_{z_i,k}$, sample N perturbations $\boldsymbol{\varepsilon}^{b}(\boldsymbol{\theta}_{\varepsilon,k}) = \{\varepsilon^{b}_{k,1}, \varepsilon^{b}_{k,2}, ..., \varepsilon^{b}_{k,N}\}$ from the error model $\square(\boldsymbol{\theta}_{\varepsilon,k})$ using the k[th] error model parameter set $\boldsymbol{\theta}_{\varepsilon,k}$. Then drive the hydrologic model with the baseline climate sequence using the k[th] hydrologic model parameter set $\boldsymbol{\theta}_{M,k}$ and add the output to the error series $\boldsymbol{\varepsilon}^{b}(\boldsymbol{\theta}_{\varepsilon,k})$:

$$\boldsymbol{Q}^{b}_{z_i,k} = M(\boldsymbol{\theta}_{M,k}, \boldsymbol{X}^{b}_{z_i}) + \boldsymbol{\varepsilon}^{b}(\boldsymbol{\theta}_{\varepsilon,k}) \qquad (3.4)$$

The k[th] time series of future streamflow $\boldsymbol{Q}^{f}_{z_i,k}$ can be generated in the same fashion by substituting $\boldsymbol{X}^{b}_{z_i}$ with $\boldsymbol{X}^{f}_{z_i}$ and $\boldsymbol{\varepsilon}^{b}(\boldsymbol{\theta}_{\varepsilon,k})$ with a new sequence of errors $\boldsymbol{\varepsilon}^{f}(\boldsymbol{\theta}_{\varepsilon,k})$.

6. Repeat Step 5 K times to develop K time series of baseline and future streamflow.

7. Calculate the time series of streamflow statistics $Y^{b}_{z,k}$ and $Y^{f}_{z,k}$ for each of the K parameter samples for both baseline and future climate conditions.

8. Fit an appropriate probability model to each time series of streamflow statistics. An estimate of the pth quantile, $(Y^{b}_{z_i,k})_p$ and $(Y^{f}_{z_i,k})_p$ $(Y^{b}_{z,k})_p$, can be inferred from the fitted probability models for both baseline and future statistics. For instance, if the streamflow statistic (or its logarithms) are normally distributed, the pth quantile can be estimated as $Y_p = \mu_y + \xi_p \times \sigma_y$, where μ_y is the mean of the statistic, σ_y is its standard deviation, and ξ_p is the 100p percentile of the standard normal distribution. To account for sampling error, draw D estimates of probability model parameters (i.e. $\mu_{y,d}$ and $\sigma_{y,d}$ with d=1,…,D) from their sampling distributions to produce D estimates of the pth quantile, $(Y^{b}_{z_i,k})_{p\,d=1,\ldots,D}$ and $(Y^{f}_{z_i,k})_{p\,d=1,\ldots,D}$. In this study, sampling distributions were taken as the posterior distributions of probability model parameters developed via a Bayesian fit of the probability model to the streamflow statistics $Y^{b}_{z,k}$ and $Y^{f}_{z,k}$. Vague distributions (e.g. uniform distributions) can be used as priors for probability model parameters in the Bayesian fit. Alternatively, estimates of sampling distributions for different probability models are often available in the literature.

9. Repeat steps 3-8 for each climate projection $z \in Z$. This will produce K×Z×D different estimates of the pth quantile for both baseline and future climate conditions. The expected value of the pth quantile of Y for baseline and future conditions can be calculated by taking the mean across all K×Z×D quantile estimates, $\overline{(Y^b)_p}$ and $\overline{(Y^f)_p}$. Similarly, predictive intervals $[(Y^b)_{p,\frac{\alpha}{2}}, (Y^b)_{p,(1-\frac{\alpha}{2})}]$ and $[(Y^f)_{p,\frac{\alpha}{2}}, (Y^f)_{p,(1-\frac{\alpha}{2})}]$ for the pth quantile can be developed using the $\frac{\alpha}{2}$ and $(1-\frac{\alpha}{2})$ percentiles of the ensemble of Z×K×D quantile estimates. These two intervals quantify the total considered uncertainty in estimates of

the p^{th} quantile of a streamflow statistic Y for both baseline and future conditions. They can be directly compared to provide a reliable depiction of how distinct the future hydrologic alteration for that statistic will be after accounting for all sources of uncertainty considered.

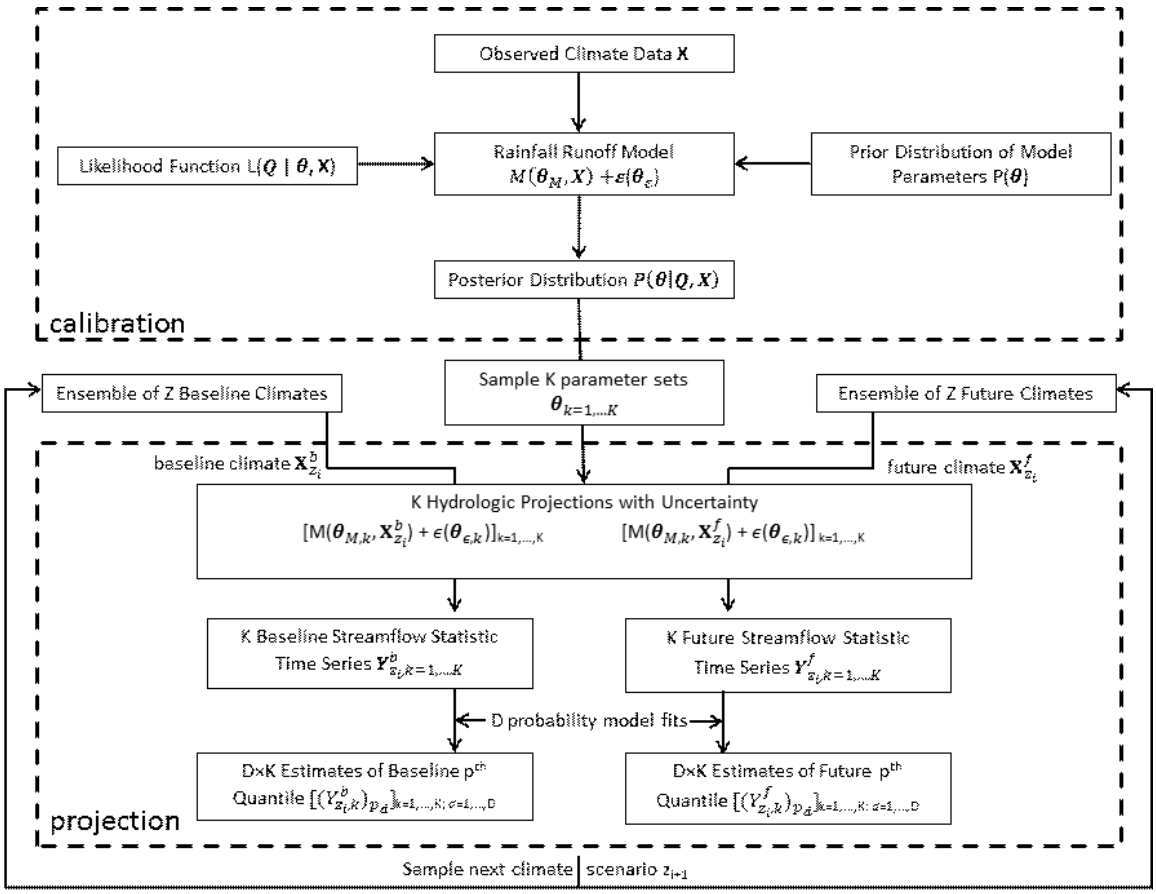

Figure 3.1. Flow chart of the statistical framework for a hydrologic uncertainty analysis under climate change.

3.5. Application of Statistical Framework in a Climate Impacts Assessment

An application of the statistical framework described above is presented for the White River Basin, located in central Vermont. Records of monthly precipitation, temperature, and potential evapotranspiration are used to drive a Bayesian calibration of a conceptual

rainfall-runoff model of the basin. An adaptation of the ABCD conceptual hydrologic model that incorporates a new snow modeling scheme is chosen for this purpose. After calibration, posterior distributions of both hydrologic and error model parameters are examined for convergence, and a probabilistic evaluation of the error model is presented to ensure the distribution of model residuals is well characterized. After the model is evaluated, the framework for climate impact assessments is applied to an ensemble of transient GCM climate scenarios.

3.5.1. White River Basin

The White River is a major tributary of the Connecticut River in New England, draining 1,790 square kilometers in the east-central portion of Vermont (Figure 3.2). Running 97.6 km from the Green Mountains to the Connecticut River Valley below, the White River is the largest gaged basin in the Connecticut River Watershed without significant regulation from upstream reservoirs or land use changes. Precipitation rates are relatively constant throughout the year, averaging approximately 100 mm/month. Regional estimates suggest about 70% of all winter precipitation falls as snow [Huntington et al., 2004]. Seasonal variations in temperature drive snow accumulation and melt processes that dominate hydrologic response throughout the winter and spring months. Streamflow is lowest during the summer and early fall months when evapotranspiration rates reach their peak.

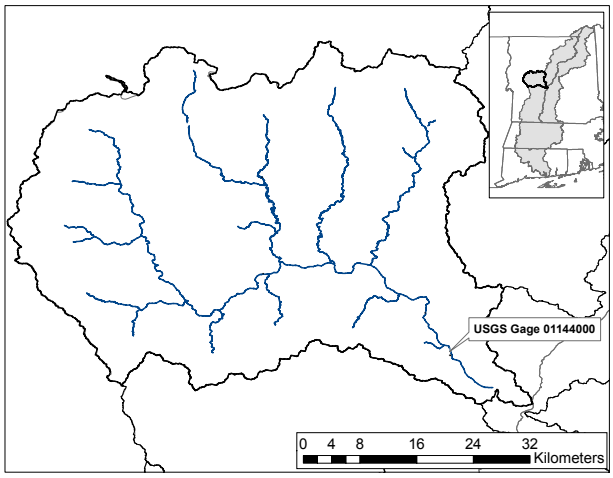

Figure 3.2. Schematic of the White River Basin - Vermont, U.S.

3.5.2. ABCD Hydrologic Model

An altered version of the ABCD hydrologic model is considered to model monthly streamflow in the White River Basin. The original ABCD model is a four parameter (*a,b,c,d*), conceptual rainfall-runoff model designed through a control volume analysis on upper soil moisture zone storage [Thomas, 1981]. The model converts monthly averaged precipitation and potential evapotranspiration into estimates of monthly streamflow by diverting water between two soil storage zones, losses to evapotranspiration, and the stream. The model has been recommended as an effective parsimonious model with physically meaningful parameters capable of efficiently reproducing monthly water balance dynamics in both theory [Vogel and Sankarasubramanian, 2003] and practice [Alley, 1984; Vandewiele et al., 1992]. A detailed review of the original ABCD model formulation can be found in [Fernandez et al., 2000].

A snow component similar to that in *Martinez and Gupta* [2010] was added to the ABCD model to simulate the snow accumulation/melt processes that dominate much of the hydrologic cycle in northern latitude watersheds. A snow storage zone is added that stores all incoming precipitation as snow water equivalent during times of year when the temperature falls below a threshold T_{Snow}. A second threshold, T_{rain}, delimits the temperature above which all precipitation falls as rain. When temperatures rise above T_{rain}, all water held in the snow storage zone melts and is added to incoming precipitation for that month. This threshold melt process is highly representative of springtime hydrology seen in northern New England rivers. When monthly temperatures fall between T_{rain} and T_{snow}, a fraction of the incoming precipitation for that month enters the snow storage component, and the remainder falls as rain. In addition, a fraction of the water held as snow is available for melt and is added to the effective rainfall for that month. The rate of melt is given by the parameter e. The total snow melt in time t is given by

$$Melt_t = \begin{cases} S_{t-1} & T_t \geq T_{rain} \\ (S_{t-1} + frac_t \times P_{tot,t}) \times e \times (1 - frac_t) & T_{snow} < T_t < T_{rain} \\ 0 & T_t \leq T_{snow} \end{cases} \qquad (3.5)$$

where S_{t-1} is the water stored as snow in the previous month, $P_{tot,t}$ is the total precipitation, T_t is the mean monthly temperature, and $frac_t$ is the fraction of precipitation that falls as snow, equal to $\frac{T_{rain} - T_t}{T_{rain} - T_{snow}}$. The water stored as snow in month t is given by

$$S_t = \begin{cases} 0 & T_t \geq T_{rain} \\ (S_{t-1} + frac_t \times P_{tot,t}) - Melt_t & T_{snow} < T_t < T_{rain} \\ P_{tot,t} + S_{t-1} & T_t \leq T_{snow} \end{cases} \quad (3.6)$$

The effective precipitation input to the model (precipitation available for runoff, soil zone storage, ET, etc.) is then given by

$$P_{eff,t} = \begin{cases} P_{tot,t} + Melt_t & T_t \geq T_{rain} \\ (1 - frac_t) \times P_{tot,t} + Melt_t & T_{snow} < T_t < T_{rain} \\ 0 & T_t \leq T_{snow} \end{cases} \quad (3.7)$$

In total, three parameters are used to represent snow accumulation/melt processes, bringing the total number of model parameters to seven (*a, b, c, d, e,* T_{rain}, T_{snow}). During calibration, the parameter T_{snow} is not directly calibrated because its prior distribution would have to be conditioned on the value of T_{rain} to ensure it took a smaller value. To circumvent this issue, a non-negative parameter $dif = T_{rain} - T_{snow}$ is used, from which T_{snow} can be directly computed.

Martinez and Gupta [2010] performed a thorough analysis on the suitability of a similar snow-augmented ABCD model structure for catchments throughout the United States, testing the model using several diagnostic statistics including Nash-Sutcliffe efficiency, bias, and variance error. That study found that the snow-augmented ABCD model structure significantly improves results for snow-dominated watersheds in New England and is a suitable structure for many catchments in the region, supporting its use in this study.

3.5.3. Bayesian Calibration and Evaluation

Historic, monthly averages of precipitation and maximum, minimum, and mean daily temperatures were gathered for the basin over the period of January 1980 to December 2005 from the gridded observed meteorological dataset produced by Maurer et al. [2002]. Average monthly streamflows were collected from the United States Geological Survey (USGS) West Hartford gage (ID #01144000) located at the mouth of the White River. Monthly averages of maximum, minimum, and mean daily temperatures were combined with estimates of monthly extraterrestrial solar radiation to produce a time series of potential evapotranspiration using the Hargreaves method [Hargreaves and Samani, 1982]. Solar radiation was calculated using the method presented in *Allen et al.* [1998].

Based on past hydrologic modeling experience for monthly flows in the New England region, a normal distribution with mean zero and standard deviation σ was initially

chosen to characterize the sampling distribution of the residuals of the natural logarithms of observations and model predictions (hereafter referred to simply as model residuals)

$$f(\varepsilon_{\ln}) = \frac{1}{\sqrt{2 \times \pi \times \sigma^2}} \times \exp(-\frac{\varepsilon_{\ln}^2}{2 \times \sigma^2}) \tag{3.8}$$

where $\varepsilon_{ln} = \ln(Q) - \ln(\hat{Q})$. The likelihood function for the observed streamflow values, **Q**, is then given by

$$L(\boldsymbol{Q} \mid \boldsymbol{\theta}, \boldsymbol{X}) = (2 \times \pi \times \sigma^2)^{\frac{-n}{2}} \prod_{t=1}^{n} \exp(-\frac{(\ln(Q_t) - \ln(\hat{Q}_t))^2}{2 \times \sigma^2}) \tag{3.9}$$

The prior for the unknown parameter σ was set to a gamma distribution with known shape λ=1 and scale ζ=2.5 parameters. The posterior of this parameter characterizes the level of uncertainty in hydrologic model estimates. A verification of the chosen sampling distribution for model residuals is described below.

Past studies were used to inform prior distributions for the hydrologic model parameters *a, b, c,* and *e* [Alley, 1984; Vandewiele et al., 1992; Fernandez et al., 2000; Martinez and Gupta, 2010], and the remaining model parameters (*d,* T_{rain}*, dif*) were given vague priors in the form of uniform distributions or normal distributions with large variances. Initial states were also calibrated in the model to avoid any parameter biases from incorrect initial conditions. The slice sampler was chosen for the MCMC sampling and was implemented in the JAGS programming language [Plummer, 2011]. Three chains were used in the sampling, and the Gelman and Rubin factor was used to test for convergence [Gelman and Rubin, 1992]. Calibration was implemented over the period between

January 1980 and December 1999, leaving six years of data for evaluation. Table 3.1 summarizes the prior and posterior distributions for all parameters inferred in the MCMC sampling, as well as allowable ranges for each parameter. Figure 3.3a shows the history plots of parameter *a* for the three chains, and Figure 3.3b presents histograms of the prior and posterior distributions of parameter *a*. For all model parameters, the Gelman and Rubin convergence factor was within 0.005 of 1, suggesting that convergence was reached for all calibrated parameters.

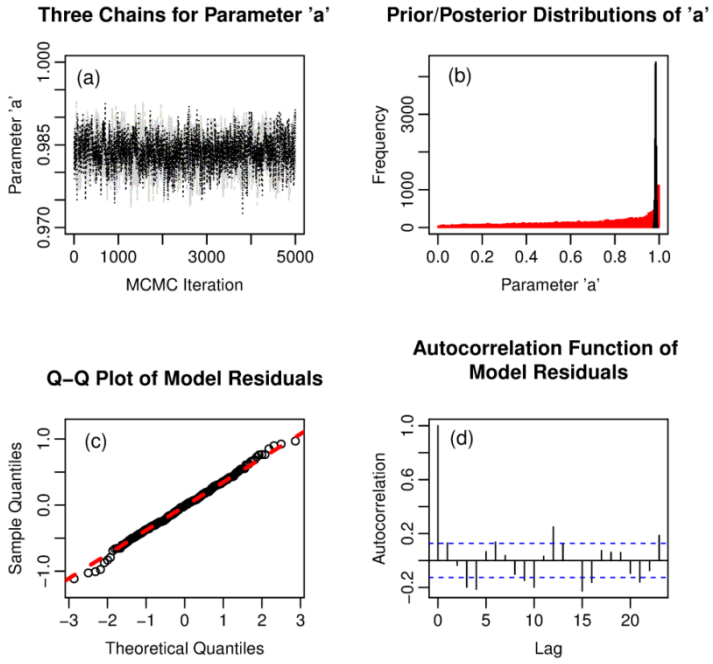

Figure 3.3. MCMC and model error diagnostics, including a) the history plot for parameter a shown for the three MCMC chains, b) a histogram of the prior (red) and posterior (black) distribution for parameter a, c) a Q-Q plot showing sample quantiles of model error εln against theoretical quantiles of a standard normal distribution, and d) the autocorrelation function of model errors ε_{ln}.

Table 3.1. Summary of prior and posterior distributions for all model parameters. Normal priors are given mean (μ) and standard deviation (φ) hyperparameters. Gamma priors have shape (λ) and scale (ζ) hyperparameters.

			Posterior Distribution			
Parameter [unit]	Allowable Range	Prior Distributions	1st Quartile	Median	Mean	3rd Quartile
a [-]	(0,1)	Beta(a=1.2,b=0.6)	0.982	0.984	0.984	0.986
b [mm]	(0,∞)	Normal(μ=300,φ=100)	303	310	310	316
c [-]	(0,1)	Beta(a=0.6, b=1.2)	0.14	0.18	0.18	0.22
d [-]	(0,1)	Uniform(a=0, b=1)	0.45	0.66	0.74	0.90
e [-]	(0,1)	Beta(a=0.8, b=1.8)	0.141	0.205	0.206	0.268
Train [°C]	(-∞,∞)	Normal(μ=0, φ=4)	-1.65	-1.48	-1.47	-1.31
dif [°C]	(0,∞)	Uniform(a=.01, b=20)	12.9	13.9	14.0	15.0
σ [ln(mm)]	(0,∞)	Gamma(λ=1, ζ=2.5)	0.13	0.14	0.14	0.15

Figure 3.3c presents a normal probability plot of the model errors $\boldsymbol{\varepsilon}_{ln}$ generated from the hydrologic simulation under the median posterior parameter set over the evaluation period (January 2000 to December 2005), and Figure 3.3d shows their autocorrelation coefficients. Results from the Q-Q plot suggest that model residuals follow a normal distribution relatively well. Most autocorrelation coefficients in Figure 3.3d are insignificant, including that at lag 1. There are some coefficients that exhibit small but significant values, particular at seasonal lag times. An autocorrelation component could be added to the error model, but this would require additional parameters to be estimated in the calibration, creating a tradeoff between problem dimensionality and error model accuracy. The seasonal autocorrelation seen in Figure 3.3d is rather low and not considered worth the increased dimensionality needed to model its behavior. Therefore, the original choice of a normal error model with no autocorrelation component for $\boldsymbol{\varepsilon}_{ln}$ was considered adequate for this modeling exercise.

Figure 3.4 shows the observed monthly streamflow for the last five years of calibration and the entire evaluation period, as well as model estimates generated by the median values of the posteriors for hydrologic model parameters. The Nash-Sutcliffe efficiency (NSE), mean flow bias, and variance error for simulated streamflow using the median parameter set equals 0.82, -1.4%, and +6.6% for the calibration period and 0.67, -5.2%, and -15.1% for the evaluation period. The bias and variance errors are expressed as a percentage of observed values. These performance statistics are considered either "good" or "acceptable" in other hydrologic modeling studies [Martinez and Gupta, 2010]. Also shown in Figure 3.4 are error bounds consistent with the 2.5^{th} and 97.5^{th} percentiles of streamflow estimates, calculated according to equation 3.3. Observed data from the calibration and evaluation periods fell outside the 95% predictive interval 3.3% and 6.7% of the time, respectively, again suggesting that the error model adopted is appropriate for this application.

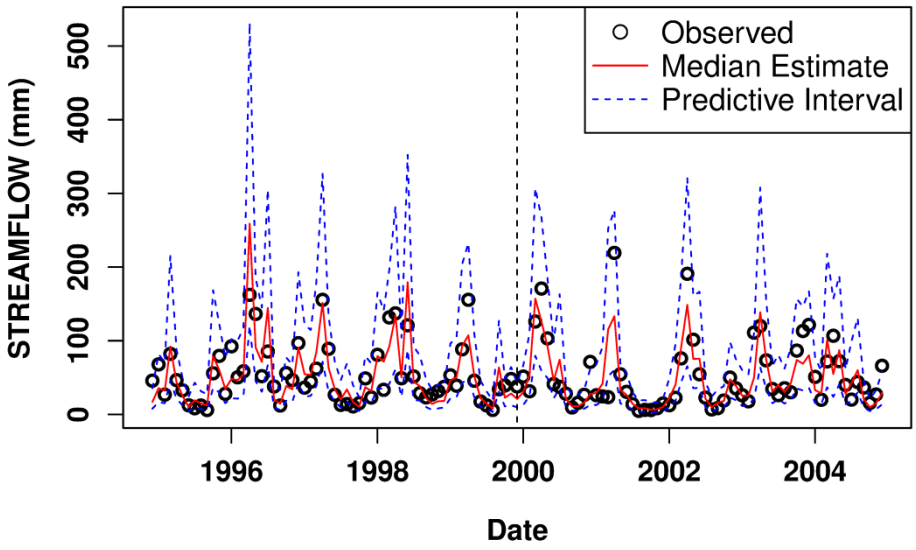

Figure 3.4. Time series of streamflow during calibration (left of vertical dashed line) and evaluation phases (right of vertical dashed line). Only a portion of the calibration time period is shown for clarity.

An additional evaluation procedure was conducted to further evaluate the adequacy of the error model. The details of the procedure can be found in *Laio and Tamea* [2007]. In brief, the procedure tests whether probabilistic predictions for a set of streamflow observations are adequate in a statistical sense. To conduct the test, the cumulative distribution function of predicted streamflow at time t is evaluated with respect to the observation q_t at t via a probability integral transform, $v_t = P_t(q_t)$. If the probabilistic predictions of streamflow are suitable then the v_t values will be mutually independent and distributed uniformly between 0 and 1. To test uniformity, a probability plot can be employed to graphically examine how well the distribution of v_t values matches a U(0,1) distribution. The condition of mutual independence can be tested using the Kendall's tau test of independence.

The probability plot of v_t values versus a theoretical uniform distribution are shown in Figure 3.5, along with Kolmogorov confidence bands at the 95% confidence level. The distribution of v_t values match that of a U(0,1) distribution very well, satisfying the first condition of the test. In addition, the condition of mutual independence was met under the Kendall's tau test of independence (p-value of 0.81), satisfying the second condition of the test. These results provide further support for the error model chosen in this application.

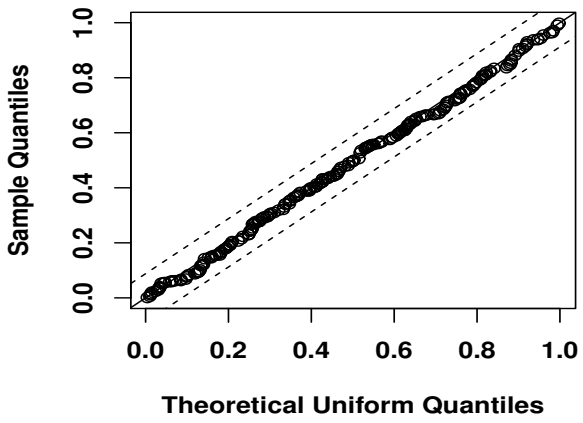

Figure 3.5. Q-Q plot of the sample quantiles of the vt values versus those of a U(0,1) distribution. Kolmogorov confidence bands (dashed) at the 95% confidence level are also shown.

3.5.4. Future Climate Scenarios

Seventy-three transient future climate simulations, running from 1950 to 2100 and sampled across the A1b, A2, and B1 emission scenarios, were gathered from the World Climate Research Programme's (WCRP's) Coupled Model Intercomparison Project Phase 3 (CMIP3) multi-model dataset. GCM simulations were downscaled according to the

bias-correction and statistical downscaling (BCSD) approach described in *Maurer et al.* [2007]. For each GCM simulation, a baseline and future climate scenario (i.e. time series of mean monthly temperatures and total monthly precipitation) was taken from fifty-year windows of downscaled climate data centered about the years 1975 and 2075, respectively. Figure 3.6 shows the absolute and percent difference between mean annual temperatures and mean annual precipitation, respectively, for these two periods across all seventy-three projections. We note here that maximum and minimum monthly temperatures are not provided in the downscaled CMIP3 dataset but are required for calculations of potential evapotranspiration. To generate maximum and minimum monthly temperature fields for baseline and future scenarios, the average differences between maximum and mean monthly temperature and minimum and mean monthly temperature were calculated for each month over the historic record. These average differences were then added to each time series of mean monthly temperature for all projections from the CMIP3 dataset to generate the maximum and minimum monthly temperature fields.

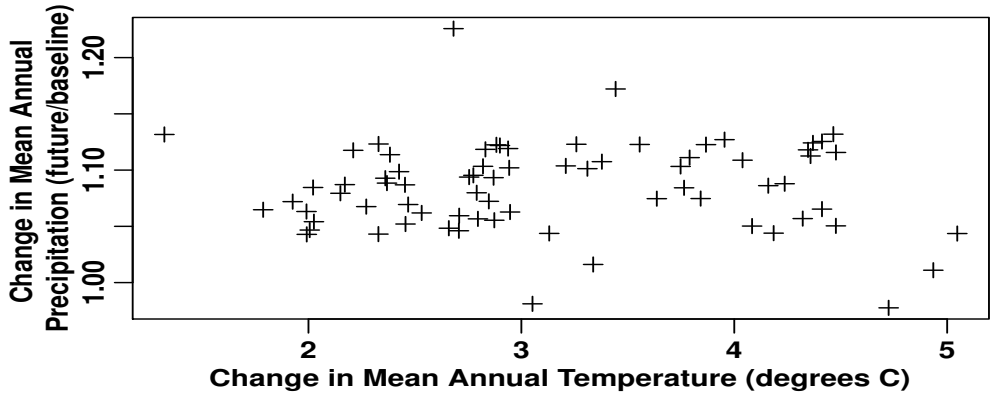

Figure 3.6. The change in mean annual precipitation and mean annual temperature between baseline and future time slices across all seventy-three climate scenarios.

3.5.5. Projections of Hydrologic Response with Uncertainty

The Z=73 baseline (1950-1999) and future (2050-2099) climate scenarios taken from the CMIP3 dataset were each used to drive an ensemble of K=5,000 hydrologic model simulations, each with different parameter sets drawn from the posterior distributions developed in Section 3.5.3. Four different annual streamflow statistics (Y) were considered in the analysis, including average January, March, April, and October streamflows. These monthly statistics were chosen because they exhibit a wide range of changes under future climate and highlight the importance of including hydrologic model error in climate impact assessments. These statistics were assumed to follow a lognormal distribution, similar to the observed historic streamflow data. This assumption was validated for each of these statistics under a large sample of climate scenarios and parameter sets using probability plots. Sampling error in the quantiles of these statistics was estimated using D=1,000 different estimates for the mean and standard deviation of

the fitted lognormal distributions drawn from their posterior distributions. Results are presented as follows. The isolated effects of hydrologic model residual error on the estimation of these statistics are considered first. The integration of uncertainties from the range of climate projections, model residual error, model parameterization, and sampling uncertainty are then addressed. An analysis of alteration in different monthly statistics is then presented in the context of their integrated uncertainty estimates.

Figure 3.7 presents the pdf of a fitted lognormal distribution to January monthly streamflows developed from one GCM scenario over the baseline period forced with one sample of hydrologic and error model parameters. Two pdfs are shown, one developed from the original streamflow trace, and a second developed from the same trace after being perturbed with noise generated from the error model. The variability in both future climate and parameter estimates is omitted by considering only one climate trace and parameter set, therefore isolating the effects of residual error on the distribution of the January flow statistic. As expected, the addition of residual error to the simulated streamflow trace causes the spread in January flows to increase. Addition of residual uncertainty to the model output appropriately adjusts the data so that it better represents the actual precision with which we can estimate characteristics of the streamflow statistic. Since the error model is logarithmic, the spread increases more at higher streamflow values than it does at lower values, suggesting different levels of precision for different magnitudes of flow. Interestingly, this highlights one of the difficulties in the choice of error model. While a transformation might make the data more tractable for a given error

model, the application of that error model may lead to asymmetric uncertainty estimates after the transformation is reversed.

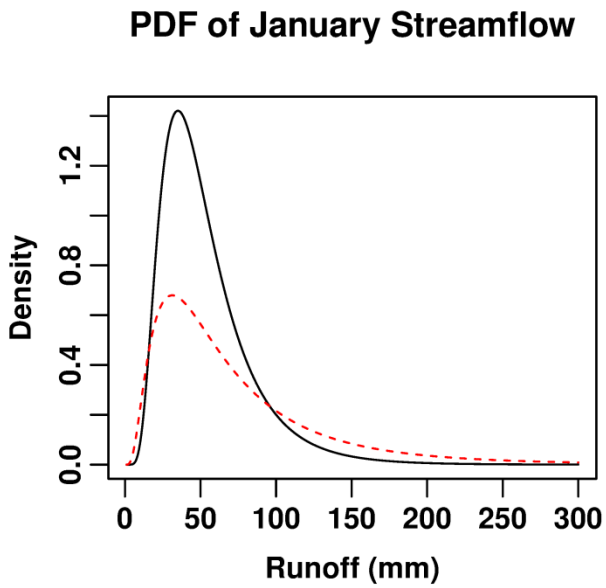

Figure 3.7. Probability density functions of baseline January monthly streamflow with (red dashed) and without (black solid) a perturbation with noise generated from the error model. Only one GCM scenario (z=1) and parameter set (k=1) were used to generate the streamflow trace.

To develop comprehensive uncertainty bounds around future hydrologic statistics, the residual error of the hydrologic model needs to be integrated with uncertainties in model parameterization, future climate projections, and sampling error. Figure 3.8 shows 95% predictive intervals for quantile estimates of baseline-period January streamflow plotted against non-exceedance probabilities for different considerations of uncertainty. Figure 3.8(a-c) shows the isolated contributions of climate uncertainty, hydrologic model parameter and residual error, and sampling error to the uncertainty of quantile estimates, respectively. The range of quantile estimates in Figure 3.8a stems from the ensemble of

baseline climate scenarios run over the median hydrologic model parameter set without the addition of residual noise. The range in Figure 3.8b was developed for only one ensemble member of baseline climate, but both parameter and residual uncertainties from the hydrologic model were considered. The influence of hydrologic model parameter and residual errors are aggregated and presented together in Figure 3.8b in order to represent the total added uncertainty from the hydrologic model. In Figure 3.8c, one baseline climate scenario was used to drive the hydrologic model with the median parameter set and no additional noise, but sampling uncertainty was calculated for each quantile. We note that the ranges of uncertainty in Figure 3.8(a-c) are dependent on the climate ensemble member or parameter set that was held constant during their development and are thus only used to illustrate the range of isolated uncertainty bounds. Figure 3.8(d-f) shows the predictive bounds for quantile estimates when climate, hydrologic model, and sampling uncertainties are integrated together. Figure 3.8d is the same as in Figure 3.8a, but Figure 3.8e shows the uncertainty bounds for quantile estimates when climate uncertainty, parameter uncertainty, and residual uncertainty are considered simultaneously. Figure 3.8f shows the total integrated uncertainty with sampling error considered as well.

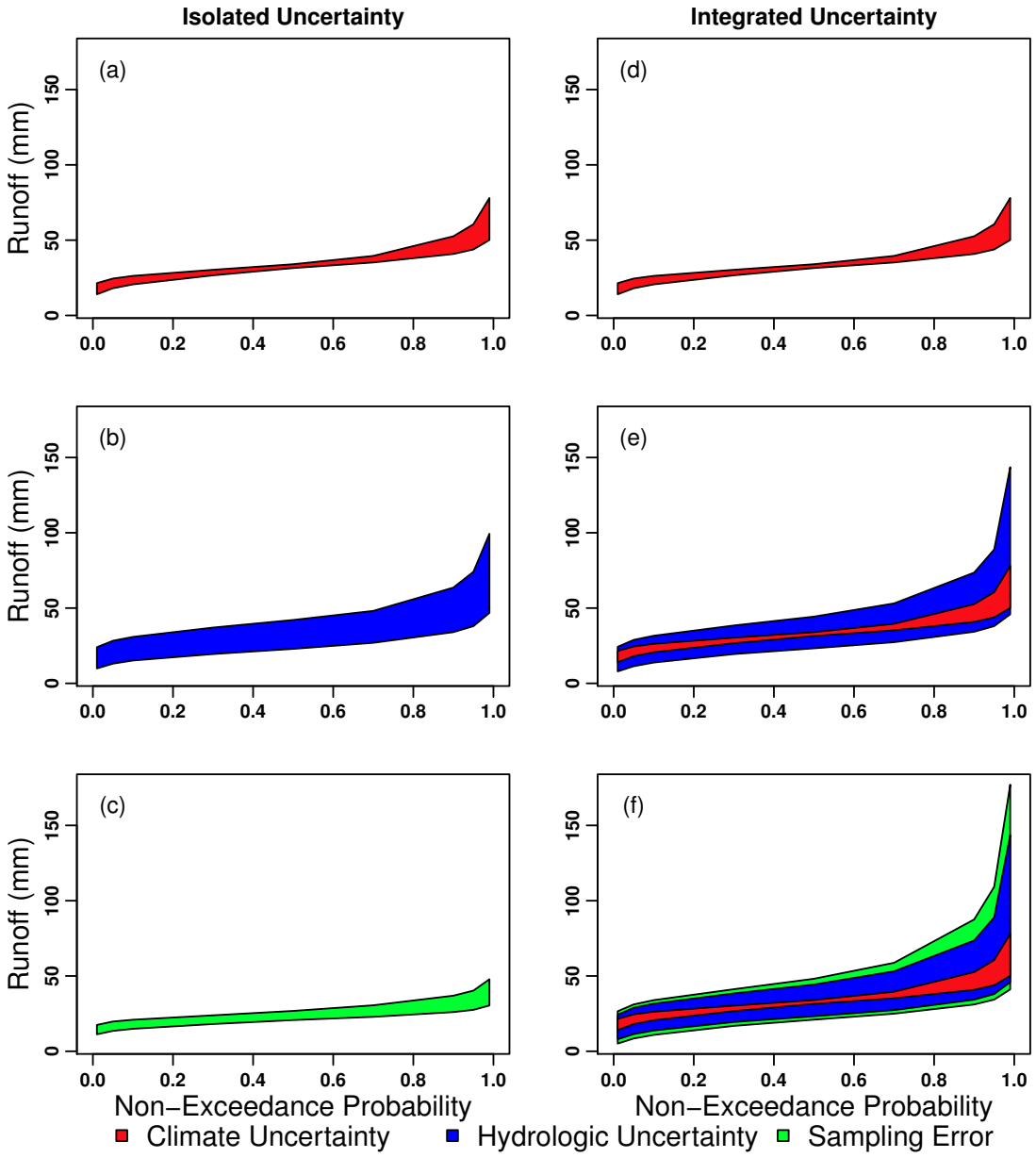

Figure 3.8. Isolated (a-c) and integrated (d-e) 95% predictive intervals for quantiles of January streamflow over the baseline period. Uncertainty originating from a range of climate scenarios, parameter and residual errors in the hydrologic model, and sampling error are shown in isolation in (a), (b), and (c), respectively. Climate uncertainty in (a) is repeated in (d), the integration of climate, parameter, and residual uncertainties is presented in (e), and (f) shows the cumulative uncertainty after sampling error is also considered.

When comparing isolated and integrated uncertainties, it immediately becomes clear that uncertainties from climate projections, hydrologic model parameter and residual error, and sampling error cannot be independently added to generate reliable predictive bounds for estimates of hydrologic statistics and their properties. This is seen in Figure 3.8e and Figure 3.8f, in which the range of uncertainty for many quantiles, particularly the larger ones, is greater than the sum of the uncertainties of their component parts (Figure 3.8(a-c)). This property highlights the dependence of uncertainty bounds on the interactions between the different sources of uncertainty.

This is a particularly important point, so we present a simplified example to emphasize it here. Consider a normalized streamflow quantile, Y_p, with zero mean and a variance conditional on either isolated climate uncertainty ($\sigma^2_{Y_p,c}$) or hydrologic modeling uncertainty ($\sigma^2_{Y_p,h}$). Assuming Y_p is normally distributed, a (1-α) predictive interval under isolated climate uncertainty and isolated hydrologic modeling uncertainty could be respectively written as $[-\sigma_{Y_p,c} \times \xi_{\frac{\alpha}{2}}, \sigma_{Y_p,c} \times \xi_{\frac{\alpha}{2}}]$ and $[-\sigma_{Y_p,h} \times \xi_{\frac{\alpha}{2}}, \sigma_{Y_p,h} \times \xi_{\frac{\alpha}{2}}]$, where $\xi_{\frac{\alpha}{2}}$ is the $(1-\frac{\alpha}{2})$ percentile of the standard normal distribution. Now assume that Y_p can be expressed under the simple additive model $Y_p = \varepsilon_c + \varepsilon_h$, where $\varepsilon_c \sim N(0, \sigma^2_{Y_p,c})$ and $\varepsilon_h \sim N(0, \sigma^2_{Y_p,h})$. Assuming that variations in Y_p stemming from climate and hydrologic modeling uncertainty are independent, we would expect that the total variance of Y_p would equal the sum of the isolated variances, $\sigma^2_{Y_p} = \sigma^2_{Y_p,c} + \sigma^2_{Y_p,h}$. However, the predictive interval for

Y_p under integrated climate and hydrologic modeling uncertainty would be given as [$-\sqrt{\sigma^2_{Y_p,c} + \sigma^2_{Y_p,h}} \times \xi_{\frac{\alpha}{2}}, \sqrt{\sigma^2_{Y_p,c} + \sigma^2_{Y_p,h}} \times \xi_{\frac{\alpha}{2}}$], which does not correspond to the sum of the two isolated intervals above because $\sqrt{\sigma^2_{Y_p,c} + \sigma^2_{Y_p,h}} \neq \sigma_{Y_p,c} + \sigma_{Y_p,h}$. Therefore, even under the simplifying assumption that variations in Y_p can be described by the simple additive model above, we would not expect uncertainty intervals to be additive. Thus, there is no reason to believe that uncertainty intervals would be additive given a more complex situation in which variations in Y_p can be influenced by the interactions of different sources of uncertainty within a hydrologic modeling framework.

The dependence of variations in Y_p on interactions between different sources of uncertainty can be traced to several contributing factors. First, the hydrologic model being considered is nonlinear, so different parameterizations of that model will result in nonlinear responses to a given climate. When those various parameterizations are used to simulate hydrologic response over a range of climates, there is the potential that the combination of an extreme climate ensemble member and parameter set will lead to significantly different streamflow responses than that seen under just climate or parameter uncertainty alone. Another source of dependency arises from the interaction between the error model and the ensemble of climate members. Because the error model used in this application is based on a logarithmic transformation, the uncertainty of large quantile values becomes highly skewed to the right after residual uncertainty is accounted. If an ensemble climate member leads to slightly larger quantile values for the streamflow statistic being considered, the residual error estimated for those larger

quantiles could lead to the significant expansion of their predictive bounds. Finally, there are significant interactions between sampling error estimation and both hydrologic and climate model uncertainties. Sampling error uncertainty bounds will grow with the uncertainty in the parameters of the distribution used to model the streamflow statistic. The sampling distributions of these parameters will likely change when climate and hydrologic model uncertainties are considered, causing the magnitude of sampling error to change with respect to its range when considered in isolation.

After aggregating the uncertainties from climate scenarios, the hydrologic model, and sampling error, it becomes evident that some quantile values for certain streamflow statistics can only be estimated with limited precision. This is shown for the cumulative error under baseline climate conditions in Figure 3.8f. In the case of future climate conditions, the range of climate projections becomes far more significant. Figure 3.9 compares the cumulative uncertainty of January monthly flows evaluated over the historic and future climate conditions. Figure 3.9a is the same as in Figure 3.8f, but Figure 3.9b now shows the uncertainty in future climate projections.

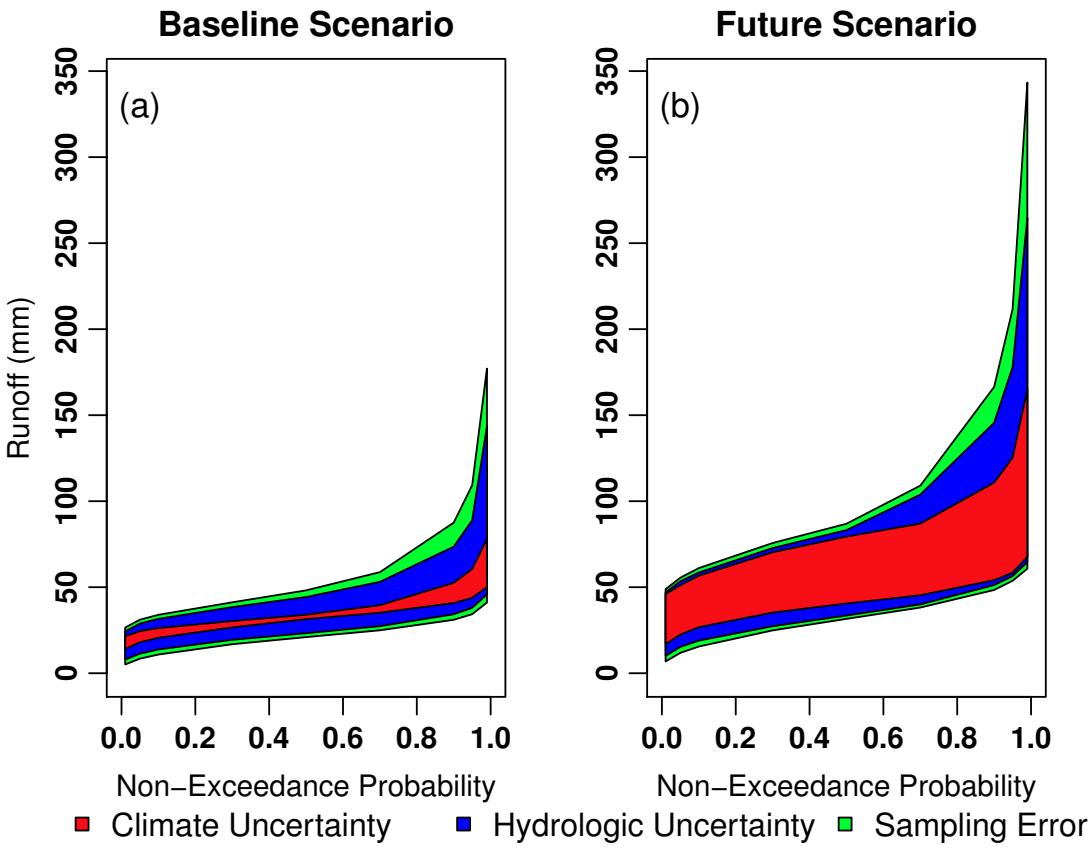

Figure 3.9. Integrated 95% predictive bounds for January flow quantiles under the a) baseline and b) future periods.

Two primary differences arise between the baseline and future cumulative uncertainties for January flow quantiles. First, the underlying climate uncertainty is far greater under the future scenarios than those of the baseline. This is expected because the baseline climate projections are all directly mapped to the historical trace of temperature and precipitation via downscaling. Thus, the range of historical projections does not model climate uncertainty or even climate model uncertainty but rather is an artifact of the bias correction method. Consequently, the range of future projections also does not model the uncertainty of future climate or even the model uncertainty of future climate projections.

Nonetheless, the range of climate projections is commonly used to provide some sense of the uncertainty in the projections that arise due to model error and internal variability and are used for that purpose. That range, albeit a minimum range of climate uncertainty, significantly increases the uncertainty in the quantile estimates relative to hydrologic modeling uncertainty as shown in the comparison between Figure 3.9a and Figure 3.9b. Second, the sampling error for larger quantiles is vastly greater for the future scenarios than for the baseline. This is due to the greater spread of January flows under future conditions and its influence on sampling error estimates. Overall, it is clear that the cumulative uncertainty for quantile estimates of this statistic is much greater for the future than it is under baseline conditions.

Quantile estimates can be directly compared between baseline and future scenarios in the context of their cumulative uncertainties to help determine the level of confidence that can be associated with their possible alteration under climate change. Figure 3.10 presents the cumulative uncertainty of quantile estimates of monthly streamflow statistics in the White River for future and baseline conditions. Here, no distinction is made between the different sources of uncertainty (e.g. climate, hydrologic, or sampling errors). Rather, the cumulative 95% predictive intervals for flow quantiles under baseline and future conditions are overlaid on each other to provide a representation of whether changes in streamflow under climate change exceed the range of uncertainty that arises during the modeling process. Less overlap between predictive intervals of flow quantiles under baseline and future conditions provides greater confidence that the flow quantile will actual differ under future climate conditions. Figure 3.10a shows that there are

significant differences between the distributions of January flows in the baseline and future periods even after accounting for cumulative modeling uncertainties. Results suggest that climate projections of January flows are significantly higher in the future than in the present, likely due to a shift in the snowfall to precipitation ratio driven by increased wintertime temperatures. Over most January quantiles, approximately half of the bounded region for future conditions lies completely outside the range of baseline uncertainty. This suggests that this range of climate changes rises to a level that is well above the baseline uncertainty.

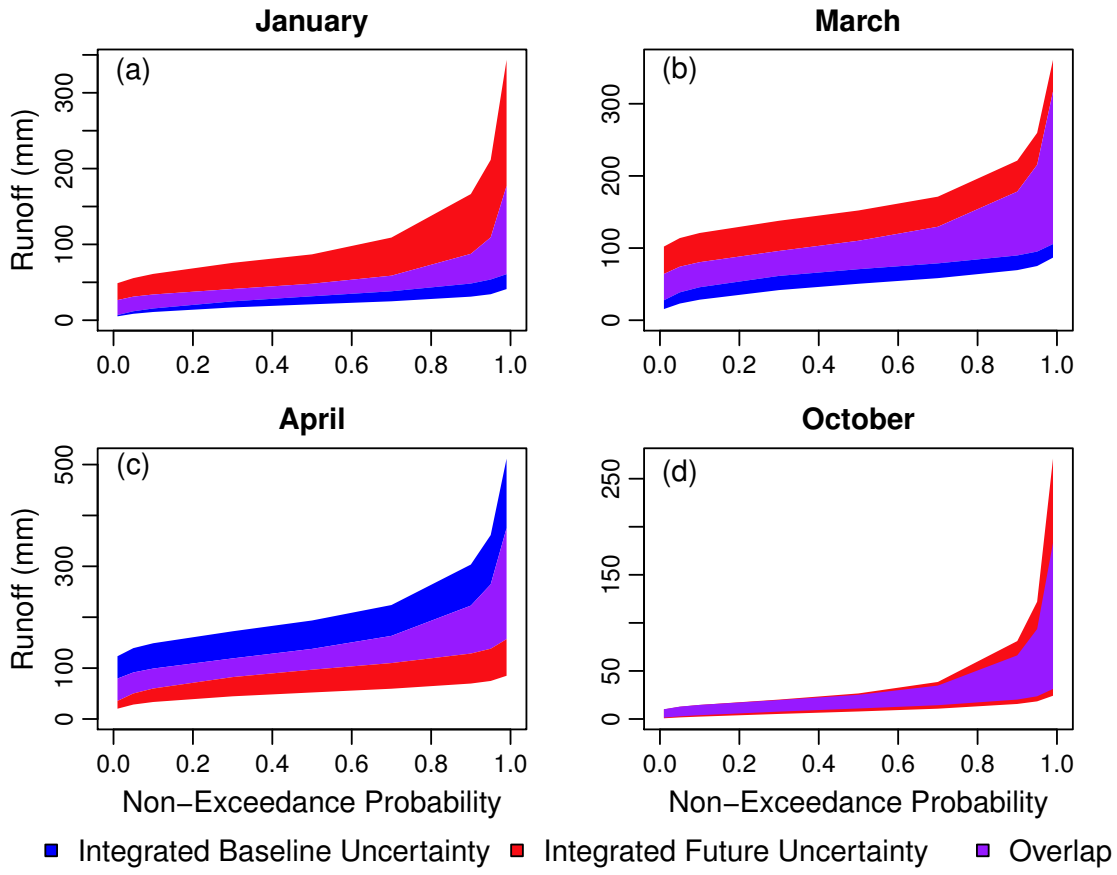

Figure 3.10. Integrated 95% predictive bounds in flow quantiles for baseline and future periods for the months of a) January, b) March, c) April, and d) October.

Figure 3.10b and Figure 3.10c show results for March and April average streamflows, respectively. The range of climate projections show March flows increasing in the future while April flows decrease. These changes are consistent with earlier snowmelt occurrences and decreases in snowpack storage that historically have persisted into the later spring. Interestingly, the highest quantiles of March flows for the future period show minor departures from those of the baseline; differences become more noticeable for flows below the 95th percentile. This is not the case for April flows, which show more significant departures between baseline and future flows at the highest quantiles. This suggests that more confidence can be associated with shifts in the highest flows during April than in March. This is likely because snowpack, a driving factor of the largest spring flows, is consistently reduced in April under all future hydroclimatic projections, but is more variable across the projections in the month of March.

Figure 3.10d shows results for the month of October. The range of climate projections exceeds only minutely the baseline uncertainty bounds for October quantiles. The spread in the future period for most quantiles extends both below and above that of the baseline period, although the changes are extremely small except for the higher quantiles. These results suggest that no real change in most October flow quantiles are projected in this set of CMIP3 climate changes.

3.6. Discussion of Future Research Needs

The framework in this study addresses many types of uncertainty in future hydrologic alterations and integrates them together to form a more comprehensive expression of the

total uncertainty surrounding future hydrologic variables. Nevertheless, several simplifying assumptions were made regarding the quantification of hydrologic model uncertainty in this analysis. Various challenges still hinder a complete quantification of this uncertainty, including source separation of uncertainties, choice of error model, and model structural errors. A discussion of each of these issues follows to highlight further research needed to bolster the framework presented in this study.

3.6.1. Input and response data uncertainties in future hydrologic projections

To simplify the modeling approach this study aggregated all errors associated with input data measurements, response data measurements, and model structure into one error term ε, but the aggregation of different types of error into one term can have significant implications for the quantification of uncertainties in future hydrologic projections [Thyer et al., 2009]. Errors in forcing data sets (e.g. input precipitation data, temperature data, etc.) and observations (e.g. streamflow measurements) are particular to the historic record. Their influence on uncertainty estimates for streamflow predictions should be isolated to the historic period and removed from uncertainty estimates of future streamflow projections. Approaches have been proposed to quantify and separate different sources of uncertainty in hydrologic modeling through Bayesian methods [Kavetski et al., 2006a, Kavetski et al., 2006b, Huard and Mailhot, 2008]. These approaches represent possible contributions of uncertainty from input and output measurement errors using prior distributions chosen by the modeler. Prior distributions for input and output data permit corruptions in those measurements to be filtered out of the calibration process, allowing for more robust and unbiased estimation of hydrologic and error model parameters, along with their associated uncertainties. While not

employed in this study, methodologies for separating input and response data uncertainties from uncertainties in future hydrologic projections are promising tools that should be explored in future applications of the proposed framework.

3.6.2. Error model identification and associated challenges

The choice of error model used to represent the probabilistic structure of model residuals also plays a critical role in accurately assessing uncertainties in future hydrologic projections. In the vast majority of hydrologic applications, model errors violate assumptions of normality, independence, and homoscedasticity [Kuczera, 1983]. If the error model is incapable of capturing these characteristics, parameter estimates can become biased and inferences of parameter and residual uncertainty can degrade [Thyer et al., 2009]. This could significantly impede efforts to accurately propagate hydrologic modeling uncertainty through a climate change impacts analysis.

Previous studies have proposed many alterations to the error model to capture different characteristics of residual error. Several studies have employed autoregressive-moving average (ARMA) models and various transformations to model auto-correlated, non-Gaussian, and heteroskedastic errors [Kuczera, 1983; Bates and Campbell, 2001; Thiemann et al., 2001]. Perhaps the most inclusive error model is proposed in *Schoups and Vrugt* [2010], in which residual errors were modeled using an autoregressive polynomial, a time-variant standard deviation linearly related with mean predicted flow, and a random noise component described by a skew exponential power distribution. The

three components allowed the error model to simultaneously model residuals exhibiting autocorrelation, heteroskedasticity, and non-normality, respectively, without the use of a transformation. A flexible parameterization, inferred through Bayesian techniques, allowed the structure of model errors to be determined during calibration, circumventing the difficulties of specifying error structure *a priori*. Overall, the advances in explicitly representing the stochastic nature of hydrologic model error are promising and suggest that Bayesian methods to quantify predictive uncertainty may be reliable for complex, high temporal resolution (e.g. daily) models often used in climate change impact analyses. Further research is needed to test this hypothesis.

3.6.3. Structural errors in hydrologic modeling

Structural errors in conceptual hydrologic modeling arise because spatially and temporally averaged representations of a catchment are often unable to simulate the true dynamics of a distributed and heterogeneous watershed. Structural errors may present one of the biggest challenges to the use of hydrologic models in predicting catchment response to climate change, especially when those responses fall outside the range of historic variability. Efforts to accurately characterize structural error in hydrologic models have met with only moderate success. Many studies assume input and output data are known and lump structural errors into a residual error term [Bates and Campbell, 2001; Marshall et al., 2004; Stedinger et al., 2008]. This was the approach taken in this study. Other approaches consider fluxes in rainfall-runoff models as stochastic, using state space approaches [Vrugt et al., 2005] and time-varying parameter values [Kuczera et al., 2006; Reichert and Mieleitner, 2009] to compensate for structural deficiencies

stemming from spatial and temporal averaging. The use of several different model structures is also a popular choice [Boorman and Sefton, 1997; Wilby and Harris, 2006; Jiang et al., 2007; Kay et al., 2009; Prudhomme and Davies, 2009a; Prudhomme and Davies, 2009b], and methods like Bayesian model averaging have recently been employed to help generate more reliable predictive intervals from these ensembles [Duan et al., 2007; Marshall et al., 2007]. However, it is often difficult to determine whether enough model structures are considered to develop a complete accounting of structural uncertainties. These different approaches and their underlying assumptions are summarized in more detail in *Renard et al.* [2010]. The formal characterization of structural model uncertainty remains a primary challenge to the hydrologic modeling community, especially as the need for insight about future hydrologic alterations under previously unseen climate forcings increases.

3.7. Conclusions

There is a growing recognition that advancements in climate change alteration studies are required to inform water resource planners and managers of the magnitude and sources of uncertainty in future hydrologic projections. In particular, of interest is whether projected changes in streamflows are important relative to the baseline error of the hydrologic modeling process. A statistical framework for investigating this question was presented here. Our approach was able to propagate uncertainty from a hydrologic model into future streamflow projections and integrate that uncertainty with other sources, producing a more complete uncertainty analysis of future hydrology under climate change.

This study employed a very simple but common approach for quantifying future climate uncertainty based on an ensemble of future climate projections. More comprehensive approaches exist, including those that treat climatological uncertainty with formal probability distributions [Tebaldi et al., 2005]. These approaches present an interesting possibility of recasting the entire cascade of model results in a probabilistic framework. However, GCM simulations are projections, not predictions, and therefore a limit likely exists for how useful direct GCM output will be in developing reliable bounds on future climate. It is difficult to compare the raw projections with observations in meaningful ways to assess skill and error, and current practices that rely on a comparison of the marginal distributions of GCM simulations against those of the observations provide "only a limited kind of confidence" [Stainforth et al., 2007b]. In addition, the downscaling methods are often calibrated over the entire historic record, leaving cross-validation approaches impossible. Nonetheless, the framework presented here allows a comparison of the range of climate projections with hydrologic modeling uncertainty.

The application to the White River Basin demonstrates how a comprehensive treatment of uncertainty can reveal varying levels of precision associated with hydrologic alterations across a spectrum of hydrologic responses. This information could be very valuable in assisting water resource managers with decisions regarding adaptation measures to possible climate changes. Depending on the projected direction and severity of climate change impacts on regional hydrology, water resources investments for adaptation can be quite expensive. The possible regret associated with those investments

increases rapidly with the uncertainty surrounding future hydrologic alterations, particularly key design flow statistics. Since the minimization of regret is often used to govern decisions regarding large capital investments, a reliable quantification of future hydrologic uncertainty is critical for a robust application of decision theory to climate change adaptation investments in the water sector. This study provides a meaningful contribution towards that end. Future work will propagate future hydrologic uncertainties developed in this study through systems and environmental models to understand the impacts of integrated hydrologic and climate uncertainties on decision-making in fields like water resources and ecohydrology.

CHAPTER 4

THE INTEGRATED EFFECTS OF CLIMATE AND HYDROLOGIC UNCERTAINTY ON FUTURE FLOOD RISK ASSESSMENTS

4.1. Abstract

This work examines future flood risk management performance within the context of integrated climate and hydrologic modeling uncertainty. The research questions investigated are 1) whether hydrologic uncertainties are a significant source of uncertainty relative to other sources such as climate variability and change in the final assessment of future flood risk management, and 2) whether a statistical characterization of uncertainty using a lumped, conceptual hydrologic model is sufficient to account for hydrologic uncertainties in the modeling process. To investigate these questions, an ensemble of climate simulations are propagated through hydrologic models and then through a reservoir simulation model to delimit the range of flood protection under a wide array of climate conditions. Mean climate changes and internal climate variability are explored using a stochastic weather generator. Two hydrologic models are considered, a conceptual, lumped parameter model that preserves the water balance and a distributed, physically-based model that preserves both water and energy balances. In the conceptual model, parameter and structural uncertainties are quantified and propagated through the analysis using a Bayesian modeling framework with an innovative error model. The approach is demonstrated in a case study for the Coralville Reservoir on the Iowa River, where intense flooding over the past several decades has raised questions about potential impacts of climate change on flood protection adequacy.

4.2. Introduction

It has long been understood by water resource planners that an analysis of planning uncertainties is critical when appraising the utility of any water resources project. For future flood risk under climate change at the local watershed scale, these uncertainties can be sizeable. Recent studies have stressed the extreme difficulty in projecting changes in local storm intensity and frequency [Liang et al., 2001; Dai, 2006; Knutson et al., 2010], with some arguing that reliable conclusions about these changes are not yet available [Barsugli et al., 2009; Hirsch, 2011]. In addition, modeling hydrologic system response during large floods can be very challenging [Uhlenbrook et al., 1999; Todini, 2004; Brath et al., 2006]. These uncertainties hinder a straightforward appraisal of flood protection infrastructure under climate change. This study presents a framework in which these uncertainties are accounted for and propagated through the analysis, providing information that can inform flood protection planning decisions even when the planning process is marred by deep future uncertainties. The approach relies on a large ensemble of climate projections across a wide range of potential changes to account for the irreducible uncertainty in future climate, while hydrologic uncertainties are accounted for using two model structures and a formal, Bayesian calibration approach that accommodates complex errors in daily hydrologic modeling.

Traditionally, flood risk planning has relied on the underlying assumption that hydroclimatic variables of interest follow a time-invariant probability distribution (i.e. they are stationary). An estimate of this distribution could be used to estimate the expected benefits and costs associated with different flood management projects. Over

the last two decades, however, the validity of the stationary assumption has been strongly challenged [Solomon et al., 2007]. While empirical evidence supporting nonstationarity in flood series has been mixed [Lins and Slack, 1999; Villarini et al., 2009], some argue that enough evidence has been presented to preclude the use of stationarity as a justifiable, default assumption for water resource planning [Milly, 2008].

The water resources planning community has largely accepted that nonstationary climate needs to be considered moving forward. Over the past decade there has been an increasing emphasis in the literature on better accounting of climate change uncertainty in long-term planning efforts, with recent work emphasizing the need for risk-based approaches. Risk-based planning methods attempt to provide probabilistic information about potential impacts using scenario ensembles and relative scenario probabilities [Brekke et al., 2009], allowing decision-makers to choose a level of acceptable risk and discount impacts that do not exceed that threshold. The goal of these planning efforts is often to identify robust decisions - those that provide an adequate level of performance across a range of climate change uncertainty - provided that some of that uncertainty can be characterized with probabilistic information drawn from climate information sources (e.g. global circulation model (GCM) projections).

Most studies adopting risk-based approaches rely on downscaled future climate projections from GCMs to provide an ensemble of climate scenarios and then use the relative frequencies of those downscaled projections to inform the probability analysis of

future change [Dessai and Hulme, 2007; Brekke et al., 2009; Lempert and Groves, 2010]. One potential issue with this approach is that these scenarios may not correctly bound future climate uncertainty [Stainforth et al. 2007a, Stainforth et al. 2007b]. Also, the computational expense of GCMs often hinders a thorough exploration of internal climate variability, despite its importance [Stainforth et al., 2005; Deser et al., 2012]. To circumvent this issue, *Brown et al.* [2012] introduced the methodology of Decision-Scaling, a risk-based planning approach that employs climate scenarios that are independent of and extend beyond the range of GCM projections to identify system vulnerabilities. Future climate projections produced by GCMs and other climate information sources (e.g. historic trends, paleodata, expert opinion) can then be used to provide insight on likely future changes in order to estimate risk. By separating the identification of system vulnerabilities from the assessment of likelihoods of future change, this method is arguably less sensitive to climate model uncertainties because it can identify system vulnerabilities potentially unrealized under downscaled GCM projections. This approach can also utilize more computationally efficient climate generation tools to better explore the effects of internal climate variability.

In addition to the uncertainty surrounding future climate, there is also significant uncertainty surrounding our ability to estimate the hydrologic response, especially the flood response, of a local watershed. There have been significant efforts to explore the uncertainty of future river flows stemming from both climate and hydrologic model uncertainties. Some work has considered these uncertainties in isolation [Arnell 1999; Prudhomme and Davies, 2009a,b; Kay et al. 2009], while others have explored their

integrated effects [Wilby and Harris 2006; Chen et al., 2011], but all of these studies utilize a small sample of conceptual hydrologic model structures (≤ 3) and lumped parameter sets (≤ 10) to quantify hydrologic modeling uncertainty. This literature generally concludes that hydrologic uncertainties are much less significant than those originating from climate change.

The insignificant effects of hydrologic modeling uncertainty can in part be attributed to the limited sampling schemes used. While a small set of hydrologic model structures and parameter sets can provide some insight regarding this source of error, they may be insufficient to quantify the full uncertainty range [Renard et al. 2010]. In fact, the efforts mentioned above have largely ignored recent methodological advancements used to formally quantify hydrologic model uncertainty, particularly Bayesian techniques that account for both predictive and parameter uncertainties [Beven and Freer, 2001; Bates and Campbell, 2001; Marshall et al., 2004; Stedinger et al., 2008; Schoups and Vrugt, 2010], input and response data errors [Kavetski et al., 2006a, 2006b; Thyer et al., 2009; Renard et al., 2010], and structural model uncertainty [Duan et al., 2007; Marshall et al., 2007]. There have been a handful of studies that have utilized some of these advanced methods to more fully explore the influence of hydrologic uncertainty on future streamflow projections [Cameron et al., 2000; Cameron, 2006; Khan and Coulibaly, 2010; Kwon et al., 2011; Steinschneider et al., 2012], and this work generally shows that hydrologic uncertainty has a more substantial influence on the total uncertainty than previously thought. It is possible that an effective statistical characterization of structural and parametric hydrologic uncertainty for a single conceptual, lumped parameter model

may effectively capture the hydrologic uncertainties in future streamflow projections for impact assessments.

To this point the discussion has focused on the progress and gaps of previous work exploring integrated uncertainty assessments in hydrologic impacts studies under climate change. Yet, there is an even more disparate gap in the literature exploring the propagation of these uncertainties through a water resource systems analysis. Most studies that have attempted to account for hydrologic modeling uncertainty in water resources planning mainly focus on short-term (daily-seasonal) decision-making timescales [Georgakakos et al., 1998; Faber and Stedinger, 2001; Yao and Georgakakos, 2001; Alemu et al., 2011] and do not utilize the most recent advances in hydrologic modeling uncertainty methods referenced above. Two notable exceptions include the work presented in *Ajami et al.* [2008] and *Muleta et al.* [2013], which used more recent uncertainty methods to propagate hydrologic error into planning studies for a water supply and urban storm water system, respectively. These studies did not, however, consider planning uncertainties related to climate change. To the authors' knowledge, there have been no attempts to formally quantify hydrologic model uncertainty using recent statistical approaches, couple it with an analysis of climate change uncertainty, and assess their integrated impact on long-term water resource planning decisions.

There are two primary contributions of this work: 1) to present a framework for revealing whether a water resource system is robust under integrated uncertainties from both

climate variability and change and hydrologic modeling capabilities, and 2) to provide a novel exploration of whether it is sufficient to characterize the parametric and structural uncertainty in a simple lumped hydrological model and propagate that uncertainty through a climate risk assessment, or if more complex hydrologic models are required to better explore hydrologic uncertainty under climate change impacts. The proposed framework is considered particularly relevant for flood risk planning studies because hydrologic uncertainties can be substantial and therefore should not be ignored. The approach is demonstrated for the Coralville Reservoir, a flood control facility in Iowa. The remainder of the paper will proceed as follows. Section 4.3 introduces the uncertainty framework, specifying the different strategies used to address climate change and hydrologic model uncertainty. An application of the framework is described in Section 4.4. Results are presented in Section 4.5, and the paper concludes with a discussion of future research needs in Section 4.6.

4.3. Uncertainty Framework

The uncertainty framework proposed in this study is comprised of two primary components. The first utilizes the Decision-Scaling methodology [Brown et al. 2012] to explore the irreducible uncertainties in the climate system and estimate risk based on the most up-to-date climate information available. The second component employs two hydrologic model structures and, for one of those structures, a formal Bayesian calibration framework to characterize the parametric and predictive uncertainty in the hydrologic modeling process. These two components are described further below.

4.3.1 Future Climate Uncertainty and Decision-Scaling

To account for the uncertainty of future climate change in flood risk planning, this study utilizes the Decision-Scaling methodology previously introduced in *Brown et al.* [2012]. At its core, the Decision-Scaling methodology can be characterized by two primary steps: 1) the identification of climate conditions that lead to unacceptable flood control problems (i.e. a vulnerability assessment), and 2) an examination of different sources of climate evidence to determine whether those problematic climate changes are likely to occur. By separating the vulnerability assessment from the analysis of likely climate changes, the approach ensures that the performance of the system is tested over a sufficiently wide range of possible futures to identify important vulnerabilities. When coupled with information regarding the likelihood of different climate changes, the vulnerability analysis provides the decision-maker with an assessment of climate-based risks.

This study utilizes a stochastic weather generator to produce the climate time series over which to conduct the vulnerability analysis. Stochastic weather generators are computer algorithms that produce long series of synthetic weather data. The parameters of the model can be systematically changed to produce new sequences of weather variables that exhibit a wide range characteristics, enabling detailed climate sensitivity analyses [Semenov and Porter 1995, Mearns et al. 1996, Wilks and Wilby 1999, Confalonieri 2012]. The scenarios created by the weather generator do not have to be dependent on any climate projections, allowing for a wide range of possible future climates to be generated. Furthermore, climate scenarios exhibiting the same mean climate changes can

be stochastically generated many times to explore the effects of internal climate variability. These climate scenarios are used to drive hydrologic models to generate time series of streamflow, which in turn are used as input to a reservoir simulation model. The output of the reservoir simulations under each climate time series is used to create a functional link between system flood risk and a set of mean climate conditions.

We note that it is not critical that all climate scenarios generated by the weather generator are plausible when first identifying system hazards, as long as implausible changes are discounted or disregarded later in the analysis when developing estimates of climate risk. The important factor is to determine how far the climate must change before the system no longer functions properly so that the analyst is aware of the potential climate hazards. Initially, the range of climate changes explored using the weather generator should be made wide enough to stress the system to failure. When those failures emerge, judgments can be made regarding the plausibility of the conditions causing them using available climate information (e.g., GCM projections, paleodata records, historic trends); they need not be made earlier.

4.3.2 Hydrologic Modeling Uncertainty

This study considers hydrologic model uncertainty to avoid underestimating flood risk associated with this source of error. There are several sources of uncertainty that obscure hydrologic model predictions, including model structure, parameterization, climate data quality, and streamflow data quality. This study explicitly focuses on the first two of these sources, structural and parameter uncertainties. Model structural uncertainties arise

because spatially and temporally averaged representations of a catchment are often unable to simulate the true dynamics of a distributed and heterogeneous watershed [Renard et al., 2010]. Parameter uncertainty is complicated by the issue of equifinality, characterized by the phenomenon where multiple parameter sets produce relatively indistinguishable streamflow responses under historic climate conditions [Beven and Freer, 2001; Beven, 2006] but can lead to diverging flow projections under alternative climate regimes [Wilby, 2005]. This can make it difficult for an analyst to choose one parameter set in a climate change analysis.

This study accounts for structural hydrologic modeling uncertainties in two ways: 1) the use of two model structures, one lumped and conceptual and the other physical-based and distributed, and 2) using a stochastic representation of model errors for the conceptual model. For the second approach, an innovative error model is chosen to accommodate the non-Gaussian, auto-correlated, and heteroscedastic nature of daily hydrologic model errors [Schoups and Vrugt, 2010]. All input data uncertainties are implicitly lumped together with structural uncertainties in the error model. For the conceptual model, parametric uncertainty is also quantified using Bayesian methods. The structural and parametric uncertainty within the conceptual model is juxtaposed against the output of the physically-based, distributed hydrologic model to determine whether the statistical uncertainty of the simpler model can account for the predictive differences between the two model structures in a climate risk analysis. The components of the conceptual hydrologic uncertainty analysis are described in detail below.

4.3.2.1. An innovative error model to account for structural uncertainty

To model the structural uncertainty of the conceptual model used in the hydrologic analysis, the distribution of model residuals is estimated using an innovative error model. The error model is taken directly from *Schoups and Vrugt* [2010] and is only briefly reproduced here. The reader is directed to the original study for more details on the method.

Assume that daily streamflow observations $\boldsymbol{Q}$ can be modeled as the sum of hydrologic model estimates $\widehat{\boldsymbol{Q}}(\boldsymbol{X}, \boldsymbol{\theta}_M)$ and an error term $\boldsymbol{\varepsilon}$:

$$\boldsymbol{Q} = \widehat{\boldsymbol{Q}}(\boldsymbol{X}, \boldsymbol{\theta}_M) + \boldsymbol{\varepsilon} \tag{4.1}$$

Here, the hydrologic model response is a function of the forcing data $\boldsymbol{X}$ and a set of hydrologic model parameters, $\boldsymbol{\theta}_M$. To account for potential non-normality, auto-correlation, and heteroscedasticity in the residuals, the following model is proposed:

$$\Phi_p(\boldsymbol{B})\varepsilon_t = \Psi_q(\boldsymbol{B})(\sigma_t a_t) \tag{4.2.1}$$

$$\sigma_t = \sigma_0 + \sigma_1 \hat{Q}_t \tag{4.2.2}$$

$$a_t \sim SEP(0,1,\xi,\gamma) \tag{4.2.3}$$

where $\Phi_p(\boldsymbol{B}) = 1 - \sum_{i=1}^{p} \varphi_i B^i$ is an autoregressive modeling structure, $\Psi_q(\boldsymbol{B}) = 1 + \sum_{i=1}^{q} \psi_i B^i$ is a moving average modeling structure, B^i is the backshift operator for the i^{th} term ($B^i \varepsilon_t = \varepsilon_{t-i}$), φ_i is the i^{th} autoregressive coefficient, ψ_i is the i^{th} moving average coefficient, p is the order of the autoregressive model, q is the order of the moving average model, σ_t is a time-varying standard deviation, σ_0 and σ_1 are the intercept and slope of a linear regression of the standard deviation as a function of the predicted flow, and a_t is a normalized error term that is independently and identically distributed according to a skew exponential power (SEP) distribution with mean 0, unit standard deviation, and parameters ξ and γ to account for skew and kurtosis. The parameters of the error model can be lumped into the vector $\boldsymbol{\theta_\varepsilon} = \{\boldsymbol{\varphi}_{1:p}, \boldsymbol{\psi}_{1:q}, \sigma_0, \sigma_1, \xi, \gamma\}$.

There are three major components of the proposed error model. First, the auto-regressive moving average (ARMA) modeling structure of order (p,q) allows for hydrologic model residuals to exhibit persistence. Second, the time-varying standard deviation, σ_t, permits the variance of model errors to rise as the flows being predicted increase. The last component of the error model is the SEP distribution used to describe the frequency of normalized, uncorrelated, and homoscedastic errors a_t. The SEP distribution enables the representation of skewed and fat-tailed residuals, a common situation in daily hydrologic modeling.

This study utilizes a Bayesian calibration scheme to calibrate both the hydrologic and error model parameters $\boldsymbol{\theta} = \{\boldsymbol{\theta_M}, \boldsymbol{\theta_\varepsilon}\}$ and estimate their relative uncertainties. Before

performing the calibration, all previous knowledge about the hydrologic model parameters is summarized in a prior distribution, denoted $P(\boldsymbol{\theta})$, which can be made vague if no prior information is available. The joint posterior distribution of all model parameters can be described using Bayes' Theorem, which states that the joint posterior is proportional to the product of the likelihood function, $L(\boldsymbol{Q}|\boldsymbol{\theta},\boldsymbol{X})$ (see Schoups and Vrugt, 2010), and the prior distribution for each parameter:

$$P(\boldsymbol{\theta}|\boldsymbol{Q},\boldsymbol{X}) = \frac{L(\boldsymbol{Q}|\boldsymbol{\theta},\boldsymbol{X})\times P(\boldsymbol{\theta})}{\int_{\boldsymbol{\theta}} L(\boldsymbol{Q}|\boldsymbol{\theta},\boldsymbol{X})\times P(\boldsymbol{\theta})\times d\boldsymbol{\theta}} \tag{4.3}$$

The integral in the denominator of equation 4.3 is a constant of proportionality required to ensure that $P(\boldsymbol{\theta}|\boldsymbol{Q},\boldsymbol{X})$ is a proper pdf. This integral is often extremely complicated in form and cannot be solved using available analytical methods. However, this challenge has been largely overcome using Markov Chain Monte Carlo (MCMC) techniques that allow for an exhaustive sampling of parameter values that can be used to describe the posterior space.

4.3.2.2. Propagating hydrologic uncertainty into Decision-Scaling

The Decision-Scaling methodology is coupled with the quantification of hydrologic modeling uncertainty to explore how integrated climate and hydrologic uncertainty influences estimates of system adequacy. In the Decision-Scaling approach, stochastic time series of weather representative of different types of climate change are used to

drive hydrologic and reservoir models to estimate system adequacy. Multiple stochastic simulations of the same climate change are used to explore the effects of internal climate variability.

For the physically-based, distributed hydrologic model, only a single time series of streamflow is estimated for each climate sequence. To propagate the conceptual hydrologic model uncertainty through this analysis, an ensemble of hydrologic traces is produced for each climate sequence. For a given climate time series of length *T*, a large number of hydrologic model time series predictions $\widehat{\boldsymbol{Q}}_{1:T}$ can be generated using posterior samples of $\boldsymbol{\theta_M}$. Posterior samples of ξ and γ can then be used generate an ensemble of time series $\boldsymbol{a}_{1:T}$ drawn from the SEP distribution. An ensemble of standard deviations $\boldsymbol{\sigma}_{1:T}$ can be estimated in equation 4.2.2 using the hydrologic model flow estimates $\widehat{\boldsymbol{Q}}_{1:T}$ and posterior samples of σ_0 and σ_1. Using equation 4.2.1, ensembles of $\boldsymbol{a}_{1:T}$ and $\boldsymbol{\sigma}_{1:T}$ can be combined with posterior samples of $\boldsymbol{\varphi}_{1:p}$ and $\boldsymbol{\psi}_{1:q}$ to estimate a large sample of errors $\boldsymbol{\varepsilon}_{1:T}$. The ensemble of residual time series $\boldsymbol{\varepsilon}_{1:T}$ are added to the ensemble of predicted flows $\widehat{\boldsymbol{Q}}_{1:T}$, and this ensemble of streamflow traces is then used to drive the systems model, producing a distribution of system performance measures associated with each climate sequence. The range of performance measures under hydrologic uncertainty for any given climate sequence can then be compared against the range of average performance across different realizations of climate variability and change to determine the relative influential of hydrologic modeling and climate uncertainties on estimated system robustness.

4.4. Application: Coralville Reservoir in Iowa

The framework described in Section 4.3 is applied to a case study of the Coralville Reservoir in Iowa. The Iowa River is located in eastern Iowa, joining the Mississippi River north of Burlington, Iowa and south of Iowa City (Figure 4.1). Coralville Reservoir, completed in 1958 and operated by the United States Army Corps of Engineers (USACE), drains 8,070 km^2 of predominantly agricultural land and is operated for flood control on the Iowa and Mississippi Rivers downstream of the reservoir. In addition to its primary objective of flood risk reduction, Coralville Reservoir is also operated for low flow augmentation, recreation, and fish and wild life management.

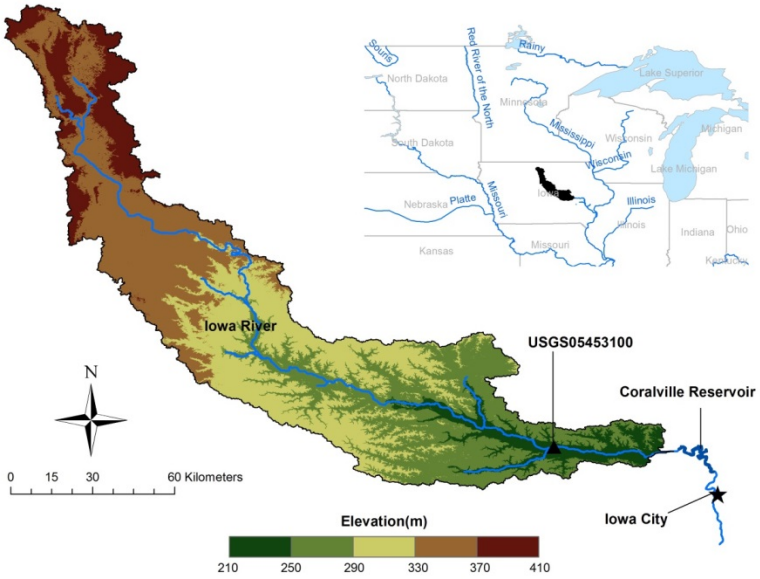

Figure 4.1. Map of the Coralville River basin.

The basin has a humid continental climate with extremes of both heat and cold. Annual precipitation averages 854 mm and snowfall is common. Heavy spring rains coupled with

snowmelt often leads to high flow events in the basin, but the largest floods occur in the summer. The floods of record for the Iowa River occurred in the summers of 1993 and 2008, when peak daily inflows into the Coralville Reservoir reached 1,010 cms and 1,365 cms, respectively (USGS gage ID#: 05453100). Both of these floods resulted from large summer rainfall events that were preceded by extended periods of anomalously wet weather that saturated the basin. In fact, the six months preceding the floods in 1993 and 2008 were the wettest on record. This evidence, along with previous research [Kunkel et al., 1994; Coleman and Budikova, 2010; Nakamura et al., 2013], suggests that the worst flooding in the area is not caused by short duration, high intensity storms in isolation, but rather long periods of wet weather that lead to highly saturated antecedent basin conditions. It is under these conditions that large rainstorms can result in extreme floods that challenge the flood risk reduction capabilities of the dam.

4.4.1. Data

Historic daily climate data, including precipitation, maximum, minimum, and mean temperatures, and wind speeds, were gathered for the Coralville basin area over the period of January 1, 1949 to December 31, 2010 from the 1/8 degree resolution (~140 km^2) gridded observed meteorological dataset produced by *Maurer et al.* [2002]. A total of 70 grid cells covering the Coralville watershed are used as climate input for two hydrologic models: the lumped, conceptual HYMOD model [Boyle et al., 2000; Kollat et al., 2012] and the distributed, physically-based Variable Infiltration Capacity (VIC) model [Liang et al., 1994]. In addition to the climate data, soil texture data for the basin

gathered from *USDA-NRCS* [2000] and vegetation cover data gathered from *Hansen et al.* [1998] are used as VIC forcing inputs.

Daily observed streamflow data were gathered from the United State Geologic Survey (USGS) Iowa River at Marengo gage (ID# 05453100) upstream of the Coralville Reservoir for the period of January 1, 1958 to December 31, 2010. These data were scaled by the drainage area ratio between the gage and the reservoir to produce a time series of observed reservoir inflows.

Time series of observed reservoir releases between October 1, 1992 and September 30, 2010, as well as operating rules for the Coralville Dam, were gathered from our partners in the USACE.

Finally, mean changes in precipitation and temperature were gathered from the World Climate Research Programme's (WCRP's) Coupled Model Intercomparison Project Phase 3 (CMIP3) and Phase 5 (CMIP5) multi-model data sets. GCM simulations were downscaled using the bias correction and statistical disaggregation (BCSD) method [Maurer et al., 2007]. These projected changes in precipitation and temperature were averaged across the entire Coralville watershed and developed using 30-windows between 1970-2000 (baseline) and 2041-2070 (2050 target year).

4.4.2. Stochastic Weather Generator

For this application, a relatively new, data-driven weather generator was chosen to drive the vulnerability assessment [Steinschneider and Brown, 2013]. The weather generator couples a Markov Chain and K-nearest-neighbor (KNN) resampling scheme to generate appropriately correlated multi-site daily weather variables [Apipattanavis et al., 2007] with a Wavelet Autoregressive Modeling (WARM) framework to preserve low-frequency variability at the annual time scale [Kwon et al., 2007]. A quantile mapping technique is used to post-process simulations of precipitation and impose various distributional shifts under possible climate changes; temperature is changed using simple additive factors. More details on the model structure can be found in *Steinschneider and Brown* [2013].

To conduct the climate vulnerability assessment, 62-year, daily simulations of climate are run several times in the weather generator with different climate changes imposed at each simulation (62 years was chosen to match the historic record length). Three types of climate change are examined here, including alterations to the mean of non-zero daily precipitation, its coefficient of variation (CV), and the mean of daily temperatures. Changes to the precipitation mean are ranged from ±30% of historic monthly averages using increments of 10% (7 increments). An alteration to the mean increases all events across the precipitation distribution. The precipitation CV is also changed by month from ±30% of historic monthly values using increments of 15% (5 increments). These changes alter daily precipitation extremes without changing the total amount of water falling each month. Temperature shifts are ranged from 0 to 4 degrees Celsius by 2°C increments (3

increments in total). All possible combinations of these changes are considered, leading to a total of 105 = 7×5×3 different climate change scenarios. These changes were chosen to range beyond the changes suggested by GCM climate projections for this region to ensure the identification of climate changes that cause system failure.

To explore the effects of internal climate variability, several weather generator simulations were considered for each type of climate change. Without parallel computing capabilities, the number of climate simulations for each climate change must be limited to manage the computational burden of the entire vulnerability assessment. However, many weather generator simulations are needed to adequately explore the possible climate fluctuations that may emerge due to chance. To circumvent this issue, we initially generated 500 different 62-year, daily simulations of climate and then chose 10 of those simulations to be used for the vulnerability assessment. Those 10 simulations, or trials, were chosen so that the 100-year basin-averaged annual maximum precipitation event across those trials (estimated using a fitted General Extreme Value (GEV) distribution) varied uniformly above and below the historic estimated 100-year event and spanned the entire range of 500 simulated 100-year events. With 10 trials for each climate change, a total of 1050 = 105×10 different climate runs are considered for the vulnerability assessment.

A brief evaluation is presented here to demonstrate the ability of the model to reproduce historic climate conditions (see *Steinschneider and Brown* [2013] for a more thorough

evaluation). In this evaluation, the 10 weather generator simulations with no climate changes imposed are compared against observed statistics. The mean, standard deviation, and skew of daily precipitation and maximum and minimum temperatures are examined, as well as daily statistics regarding precipitation spells and extreme precipitation events. All of these comparisons are conducted for the basin-averaged time series.

Figure 4.2 shows the mean, standard deviation, and skew of non-zero daily precipitation amounts, daily maximum temperature, and daily minimum temperature for each calendar month. The results suggest good performance for all variables and statistics, with observed values always falling within the range of the simulations. The average lengths of wet and dry spells are also well simulated, which can be a very important statistic with respect to the generation of floods (Figure 4.3). We also note that the cross-correlations of all variables across sites are almost perfectly preserved (not shown), which is expected given the resampling techniques used to generate the daily weather sequences. Finally, a GEV distribution is fit to the basin-averaged annual maximum precipitation time series and different return interval events are examined (Figure 4.4). These extreme event statistics are also well preserved in the model. Overall, the performance of the model for most statistics is either good or adequate, suggesting that the weather generator can produce suitable climate simulations for a flood impacts study in this location.

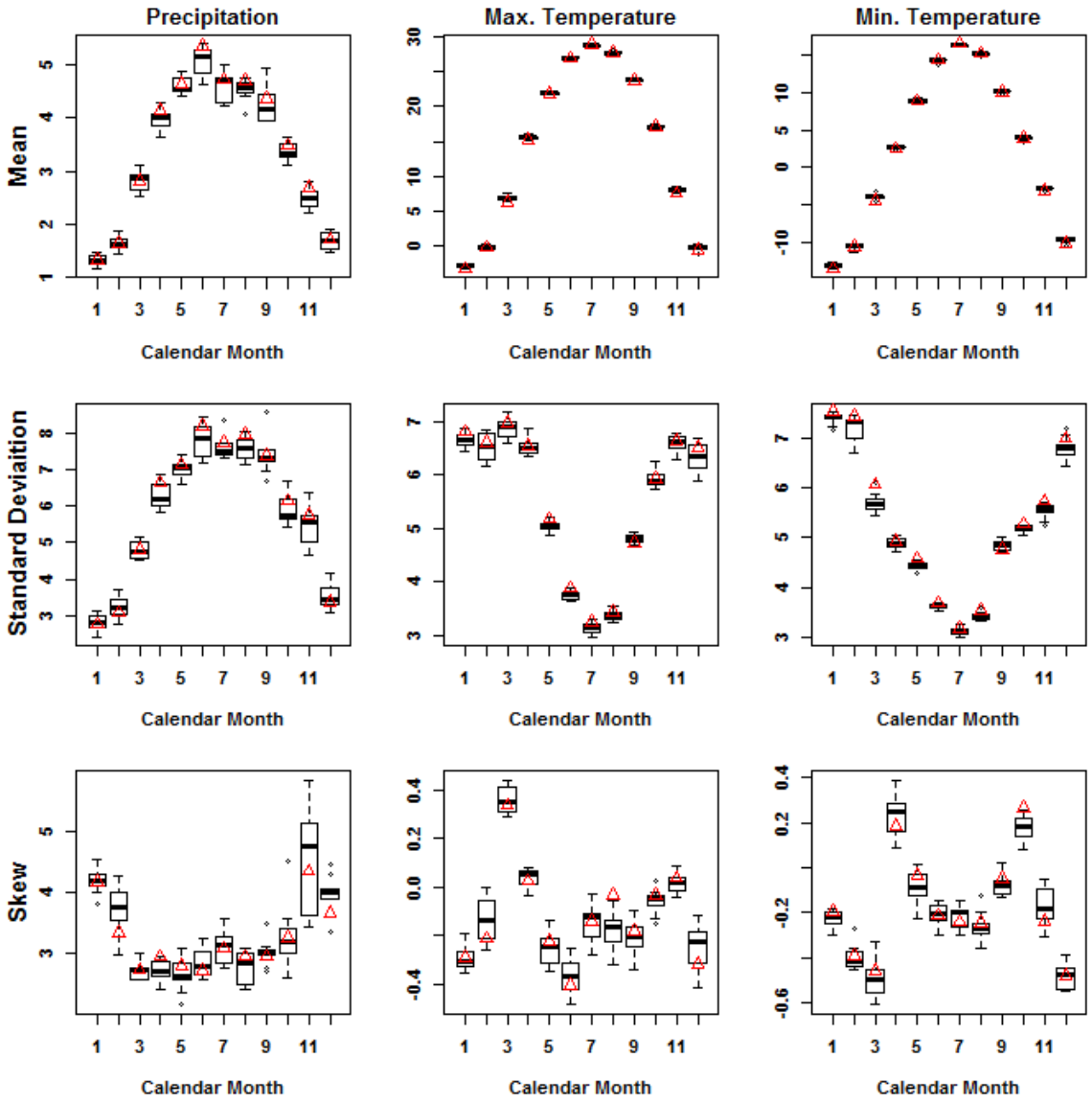

Figure 4.2. Daily performance statistics by month, including the mean, standard deviation, and skew of precipitation, maximum temperature, and minimum temperature. The observed statistics (red triangles) are shown against the distribution of statistics across the 10 trials.

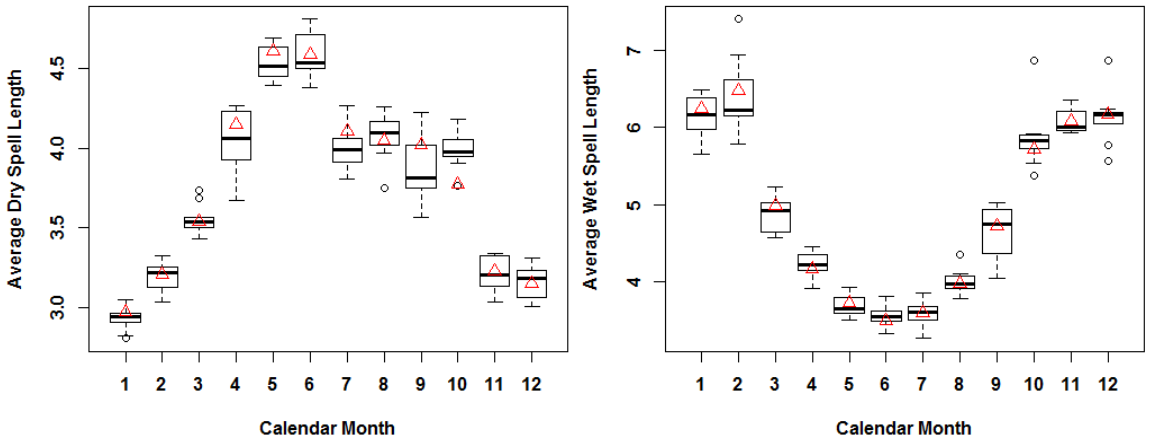

Figure 4.3. The average dry and wet spell length for basin-averaged precipitation by month. The observed statistics (red triangles) are shown against the distribution of statistics across the 10 trials.

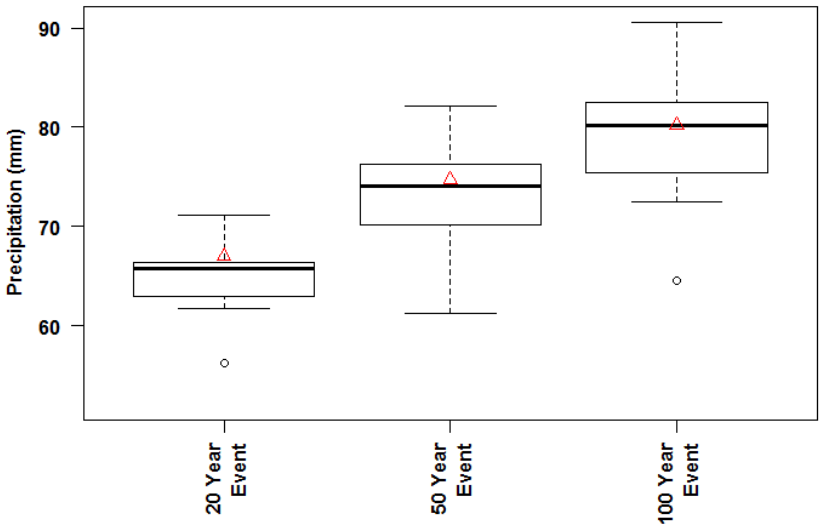

Figure 4.4. The 20-, 50-, and 100-year annual daily maximum precipitation event. The observed statistics (red triangles) are shown against the distribution of statistics across the 10 trials.

4.4.3. Conceptual Hydrologic Model - HYMOD Model

HYMOD is a lumped-parameter model composed of a soil-moisture accounting module, a snow module, and a routing module (see Figure 4.5). The soil-moisture accounting

module utilizes a storage capacity distribution function with parameters C_{max} and β for the storage elements of the catchment [Moore, 1985]. The snow module uses a simple degree-day method for calculating snowmelt [Bergstrom, 1975], with a temperature threshold T_s to determine rain/snow separation, a second threshold T_m to initiate snowmelt, and a melt rate defined by the degree-day factor (DDF). The routing module divides excess water using the split parameter (α) and routes it through parallel conceptual linear reservoirs. The quick (K_q) and slow (K_s) reservoir depletion rates controls the flow from each routing reservoir. Multiple (N_q) quick flow reservoirs can be utilized. The outputs from quick and slow reservoirs are summed to simulate streamflow.

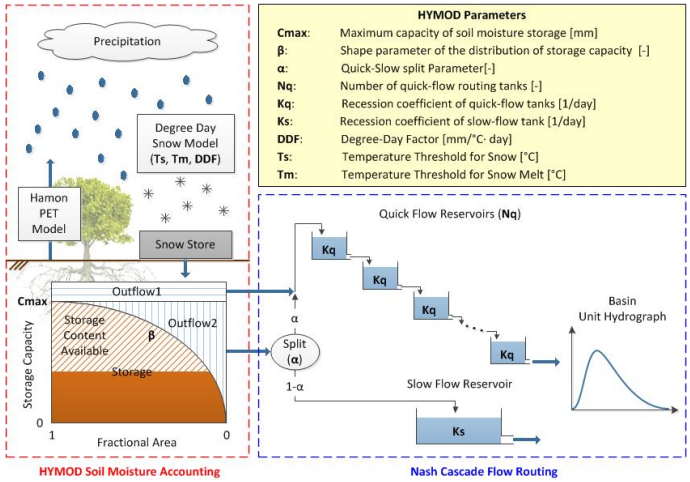

Figure 4.5. Schematic representation of the HYMOD model with a description of its parameters.

During the calibration process, a total of 15 parameters require estimation, including 9 hydrologic model parameters ($\boldsymbol{\theta}_M = \{C_{max}, \beta, \alpha, K_q, K_s, N_q, DDF, T_s, T_m\}$) and 6 parameters for the residual error model ($\boldsymbol{\theta}_\varepsilon = \{\varphi_1, \psi_1, \sigma_0, \sigma_1, \xi, \gamma\}$). One autocorrelation coefficient and one moving average coefficient are included because diagnostic tests (not

shown) suggest autocorrelation in the residuals can be represented using a simple ARMA(1,1) process. The model calibration was conducted over 10 years of historic climate, from October 1, 1989 to September 30, 1999. Daily precipitation and mean temperature were averaged across the catchment to produce a basin-averaged input time series. Daily potential evapotranspiration was calculated using the Hamon method [Hamon, 1963]. Before proceeding with the Bayesian calibration, all prior distributions were set to vague uniform distributions to allow the data to drive model calibration. The DREAM(ZS) MCMC algorithm was chosen to explore the joint posterior distribution of all model parameters [Vrugt et al., 2011]. Three sampling chains were used for each parameter, and convergence was verified using the Gelman and Rubin factor [Gelman and Rubin, 1992]. The feasible range, prior distribution, and posterior distributions for all model parameters are presented in Table 4.1.

Table 4.1. Summary of prior and posterior distributions for all model parameters.

			Posterior Distribution			
Parameter	Feasible Range	Prior Distribution	First Quartile	Median	Mean	Third Quartile
c_{max} (mm)	(0,∞)	Uniform (a=0,b=1000)	171	173	174	175
b	(0,2)	Uniform (a=0,b=2)	0.382	0.390	0.391	0.399
α	(0,1)	Uniform (a=0,b=1)	0.803	0.809	0.809	0.814
n_q	N	Discrete Uniform (a=1,b=7)	3.226	3.480	3.491	3.736
k_q	(0,1)	Uniform (a=0,b=1)	0.237	0.242	0.241	0.245
k_s	(0,1)	Uniform (a=0,b=1)	0.002	0.002	0.002	0.002
DDF	(0,1)	Uniform (a=0,b=1)	2.905	3.216	3.747	4.749
Ts (°C)	(-∞,∞)	Uniform (a=-4,b=4)	-1.667	-1.631	-1.638	-1.611
T_m (°C)	(-∞,∞)	Uniform (a=-10,b=10)	-1.046	-0.976	-0.961	-0.863
σ_0	(0,∞)	Uniform (a=0,b=1)	0.000	0.000	0.000	0.000
σ_1	(0,∞)	Uniform (a=0,b=1)	0.117	0.119	0.119	0.121
ξ	(0,∞)	Uniform (a=0.1,b=10)	1.128	1.140	1.141	1.155
γ	(-1,1)	Uniform (a=-1,b=1)	0.997	0.999	0.998	0.999
φ_1	(-1,1)	Uniform (a=-1,b=0.85)	0.850	0.850	0.850	0.850
ψ_1	(-1,1)	Uniform (a=-1,b=1)	0.455	0.465	0.465	0.476

Diagnostic plots are used to verify the skill of the model and ensure that the fitted error model is consistent with the observed distribution of the residuals. Figure 4.6a shows the fitted SEP density function using the two parameters γ and ξ, as well as the empirical density of the normalized residuals, a_t, of the HYMOD model. Both parameters γ and ξ were estimated very close to unit, suggesting that the error distribution is highly peaked ($\gamma \approx 1$) and symmetric ($\xi \approx 1$). There is a slight bias in the fitted distribution because the errors are not exactly centered about zero, but this bias is small. Figure 4.6b shows the autocorrelation function of the residuals from the model. While the original residuals of the HYMOD model exhibited very significant autocorrelation at several lags, this autocorrelation has largely been removed from the normalized residuals. Finally, Figure 4.6c shows the relationship between predicted streamflow values and the errors associated with those predictions. While the variance of the original residuals varied significantly with predicted flow, the normalized errors maintain a much more constant spread across the range of flow predictions. These three diagnostic plots suggest that the error model used adequately captures the non-Gaussian, auto-correlated, and heteroscedastic nature of the daily residuals.

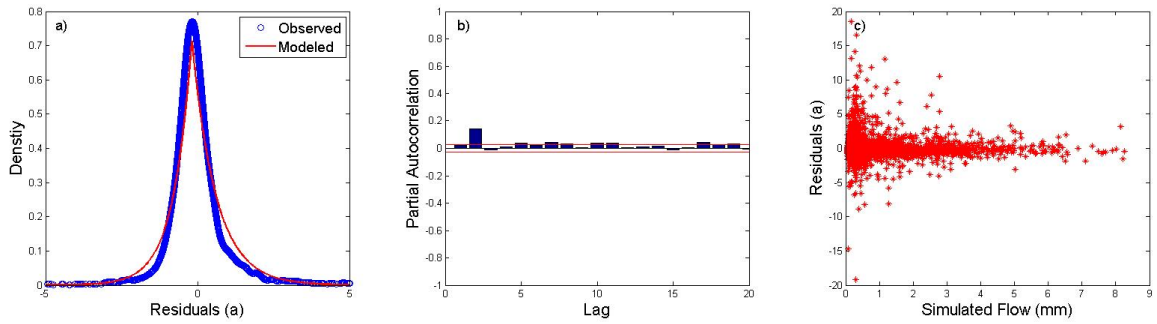

Figure 4.6. a) The fitted SEP density (red line) and empirical density (blue points) of normalized errors. b) The partial autocorrelation function of the normalized errors. c) The residuals plotted against the mean predicted flow from the HYMOD model.

Figure 4.7 shows the observed daily hydrograph for a calibration and validation period, as well as the model estimates generated by the mean of the posteriors for all HYMOD parameters. The Nash-Sutcliffe efficiency (NSE) for the mode prediction over the calibration and validation period is 0.86 and 0.85, respectively. The Kling-Gupta efficiency (KGE) for these two periods is 0.91 and 0.90. This measure is a relatively new objective function that provides an alternative (and potentially improved) balance between mean bias, variability bias, and correlation compared to the NSE [Gupta et al., 2009]. Also shown in Figure 4.7 are 90% predictive bounds for the modeled flows. These bounds are calculated using 500 time series of model predictions and randomly generated residuals. We note that even though the mean model prediction fails to capture some of the highest peak flows in the record, including the 1993 and 2008 floods, the 90% predictive bounds do contain these extreme events.

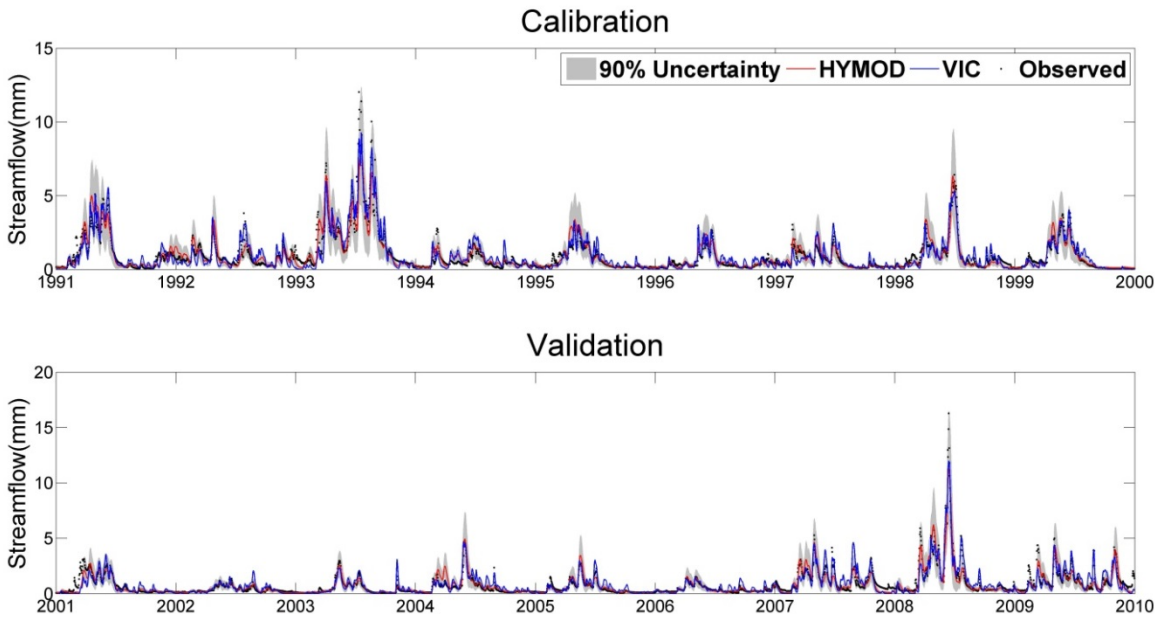

Figure 4.7. Streamflow hydrographs for the calibration and validation period for the HYMOD model. Observations (black dots) are shown against the median HYMOD prediction (red) and the associated 90% confidence bounds (shaded grey). The VIC model predictions are also shown (blue).

4.4.4. Physical, Distributed Hydrologic Model - Variable Infiltration Capacity Model

Unlike the conceptual HYMOD model, that only maintains the water balance, the VIC model accounts for balances in both water and surface energy. We employ a gridded version of VIC that utilizes the same grid as the driving weather input files. A majority of the area within the Coralville watershed is covered by cropland, and two types of land cover (cropland and grassland) account for over 90 percent of the basin area. More details about the most up-to-date VIC modeling processes, including evapotranspiration, soil moisture, and runoff processes, can be found in *Gao et al.* [2010].

The parameters of the VIC model were calibrated by maximizing the KGE over the calibration period. A genetic evolutionary algorithm was utilized to conduct the optimization [Houck et al., 1995]. Six parameters are considered in the calibration: one parameter related to the variable infiltration curve shape, two parameters related to thickness of soil layers, and three parameters related to the baseflow scheme. One parameter set was considered for all grid cells in the model. Therefore, the calibration of the VIC model is parsimonious while still leveraging the benefits of improved routing via the distributed structure. The optimized streamflow simulation is shown in Figure 4.7. The NSE (KGE) for the calibration period is 0.87 (0.93) and for the validation period is 0.82 (0.90). This model skill is comparable to the simulation result from the HYMOD model, suggesting that VIC provides a valuable alternative model to explore future flood response under climate change.

4.4.5. Coralville Reservoir Systems Model

A water resources systems simulation model of the Coralville Reservoir was developed to emulate the flood risk reduction capacity of the system. The systems model was developed to mimic the documented operating policies used by the USACE to manage the reservoir. The model is formulated to switch between different operating rules conditional on the storage in the reservoir. A rectangular weir equation is used to simulate releases made through the spillway.

The primary objective of the Coralville Reservoir is the reduction of flood-related damages downstream. Expected annual flood damage (EAD) is the metric chosen to measure whether or not this objective is being met. This metric is required for risk analysis within the USACE and thus is highly pertinent in the decision-making process [USACE Report ER-1105-2-101, 2006]. Flood damages are calculated from modeled reservoir discharge rates using flow-stage and stage-damage relationship curves previously created for the region (Figure 4.8a).

The systems model was run with historic inflows, and historic releases were used to evaluate the systems model performance (Figure 4.8b). The model was able to capture and closely resemble the reservoir behavior well, especially during peak flow events. Overall, the calibration of the systems model is adequate for the purposes of this study.

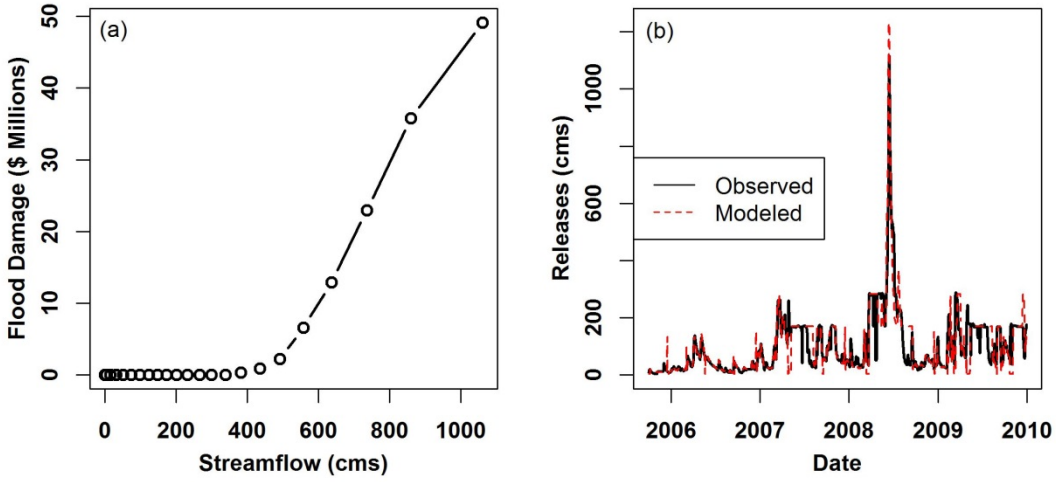

Figure 4.8. Relationship between streamflow downstream of the Coralville Reservoir and resulting flood damage. b) Observed (black) and modeled (red) releases from the Coralville Reservoir simulation model under historic, observed inflows.

4.4.6. Conducting the Climate Vulnerability Assessment

Two modeling experiments are considered in this assessment. First, both the HYMOD and VIC models are run under all 1050 generated climate time series to produce an ensemble of inflow sequences, which are then used to force the Coralville simulation model and estimate an ensemble of expected annual damages. Here, HYMOD is run using the mean posterior parameter set and no residual errors are added to the simulated streamflow time series. These 1050 simulations explore the effects of both internal climate variability and change (10 internal variability trials, 105 changes per trial) on system performance. The results are compared against an ensemble of GCM projections to provide insight on climate risks facing the system.

The use of both VIC and HYMOD above partially explores the effects of structural hydrologic uncertainty. To more fully examine this uncertainty source, as well as parametric uncertainty, we conduct a second experiment in which 500 HYMOD simulations are run using random samples from parameter posterior distributions and are added to randomly generated time series of residuals. This procedure is carried through for 70 climate time series, including all 10 internal variability trials and the 7 changes in mean precipitation (with the precipitation CV and temperature held at baseline levels). The range in flooding upstream of the reservoir and downstream EAD under hydrologic model uncertainty is juxtaposed against that from climate variability and change to explore their relative contributions to total uncertainty of system robustness. Also, the results from HYMOD and VIC are compared to determine whether the statistical

uncertainty in the simpler HYMOD model is sufficient to capture the uncertainty across the HYMOD and VIC structures.

4.5. Results and Discussion: Comparison of System Robustness under Hydrologic and Future Climate Uncertainty

Figure 4.9 shows parallel coordinate representations of the EAD climate response of the system using the mean posterior HYMOD model and VIC. The parallel coordinates enable system response to be visualized across multiple dimensions in climate change space. Each point in 3-dimensional climate change space (dimensions associated with changes in mean precipitation, precipitation CV, and mean temperature) is represented by a polyline with vertices on the parallel axes. Climate change combinations that lead to greater EAD are shown using darker, thicker lines. The average EAD values across the 10 internal variability trials are shown, so only the effects of climate change are being considered here. Finally, the distribution of CMIP3 and CMIP5 climate projections for mean precipitation and temperature centered on 2050 are superimposed on the figure; no attempt was made to quantify the distribution of change to daily precipitation variability because GCMs poorly reproduce this aspect of precipitation.

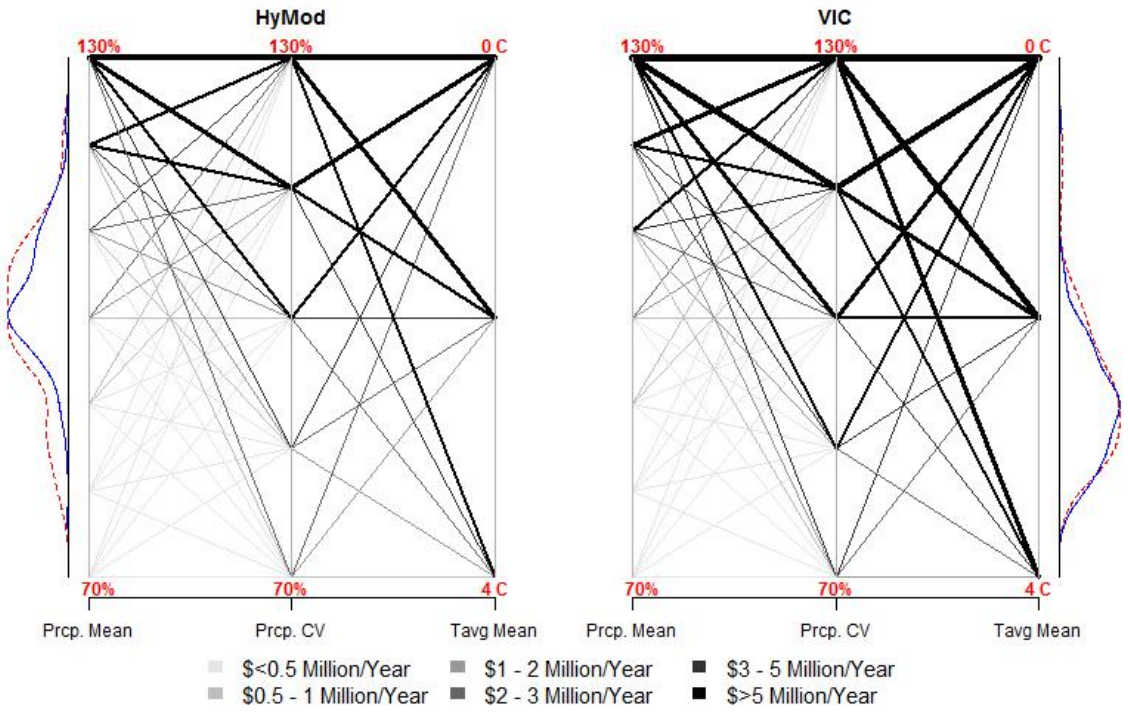

Figure 4.9. Parallel coordinate plots of expected annual damages in a 3-dimensional space of climate changes. Damages are shown for the system forced with HYMOD and VIC streamflows, averaged across all 10 trials. Darker, thicker lines indicate climate change scenarios with greater flood damage. No statistical uncertainty was considered here for the HYMOD model. Empirical density plots of precipitation changes (left) and temperature changes (right) are also shown for CMIP3 (red dashed) and CMIP5 (blue solid) projections.

Figure 4.9 shows the sensitivity of a decision-centric metric to various types of climate change. Several insights emerge from this figure. First, the gradient of EAD values is steepest across changes in mean precipitation, suggesting that this climate change dominates the response of the system. All of the high damage scenarios (>$2 million/year, which is the historic EAD value estimated using observed data) for both hydrologic models require increases in mean precipitation of at least 10%. Interestingly, changes in the CV of precipitation do not dramatically influence EAD values. While greater daily precipitation variability does lead to increased damage, the gradient in EAD

is much less steep across this climate change than for mean precipitation changes. This suggests that the reservoir is able to buffer out the effects of increased daily precipitation variability with storage. The effects of temperature are even less impactful. Even a 4°C temperature increase only marginally reduces damages, likely due to slightly reduced flood risk from shifting snow dynamics and increased evapotranspiration. Finally, CMIP3 and CMIP5 climate projections for mean precipitation suggest a 5% increase on average, although the tails of the projection distribution do include some precipitation increases as high as 22%, which are associated with high EAD values. The very worst EAD scenarios (polylines across the top of the figure) will likely be avoided, however, as indicated by projections suggesting increased temperatures averaging 2.5°C.

The influence of climate change on system performance can also be juxtaposed against that of internal climate variability to determine their relative contribution to total uncertainty. Figure 4.10 shows average EAD responses for both hydrologic models for each type of climate change (with the other two held at baseline levels), as well as the range of responses under internal climate variability (i.e. the 10 trials). Again, no statistical uncertainty in HYMOD is considered in this figure. The results of Figure 4.9 again emerge in Figure 4.10, with changes in the precipitation mean dominating system response. However, what also emerges in Figure 4.10 is the importance of internal climate variability. For all changes and both models, the width of the gray shaded region is of the same order of magnitude as the average change in EAD across adjacent climate changes. For instance, the range of EAD across the 10 trials under the VIC model for a 10% increase in mean precipitation is $2.6 million/year, while the average change in

EAD between 10% and 20% increases in mean precipitation is $2.7 million/year. We also note that the uncertainty in EAD associated with internal climate variability increases as mean climate changes lead to more damage. This suggests that the uncertainty due to internal climate variability cannot be estimated using historic climate conditions, but must be considered in conjunction with mean climate changes.

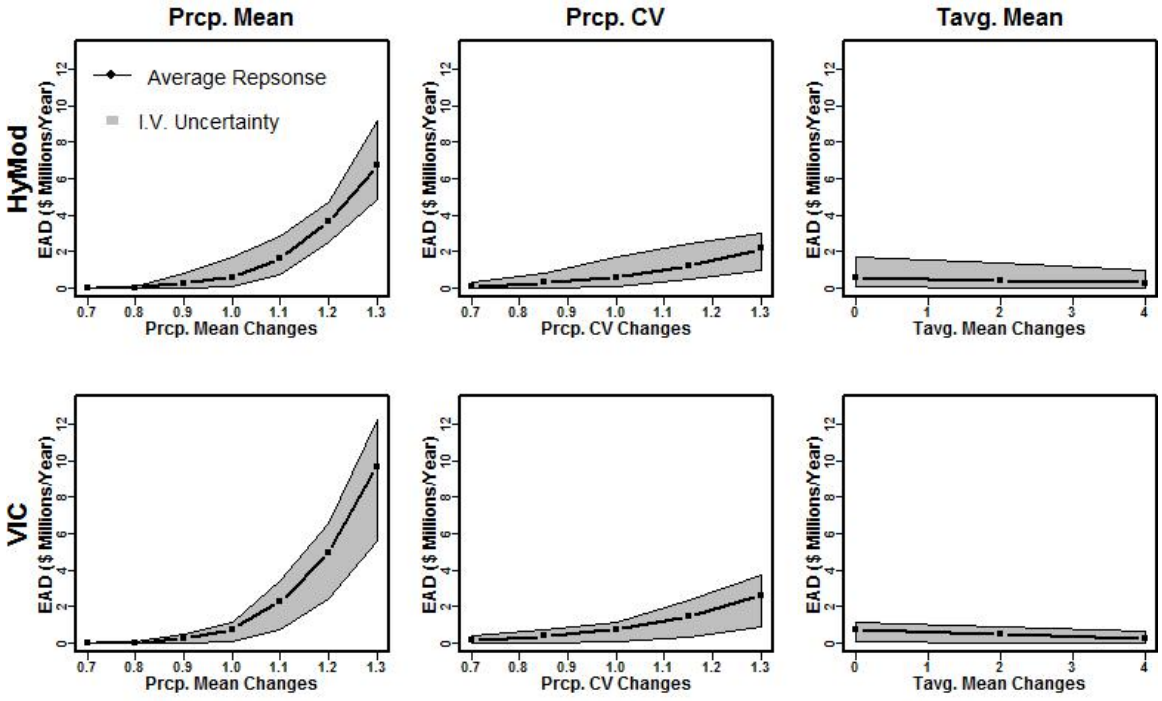

Figure 4.10. Expected annual damage under each type of climate change averaged across all 10 trials, as well as the spread across the range of internal variability (I.V.) uncertainty (grey region). Damages are shown for the system forced with both HYMOD and VIC streamflows. For each climate change, the other two climate factors were held constant at their baseline levels. No statistical uncertainty was considered here for the HYMOD model.

Finally, system vulnerabilities to climate under both hydrologic models can be compared against a backdrop of statistical uncertainty in the conceptual HYMOD hydrologic model. Figure 4.11 shows estimates of the 100-year flood (fitted to annual peaks using a

Log-Pearson Type III distribution) upstream of the Coralville Reservoir (left panel) and EAD estimates downstream of the reservoir (right panel) for all 10 trials and a range of potential changes in mean precipitation. No changes in the CV of daily precipitation or temperature are considered here. For the VIC model, each climate scenario is associated with one sequence of predicted streamflow. For the HYMOD model, these two metrics are shown under climate variability and change uncertainty alone, as well as the additional statistical uncertainty of the conceptual model (i.e. the ensemble of 500 HYMOD traces for each climate sequence).

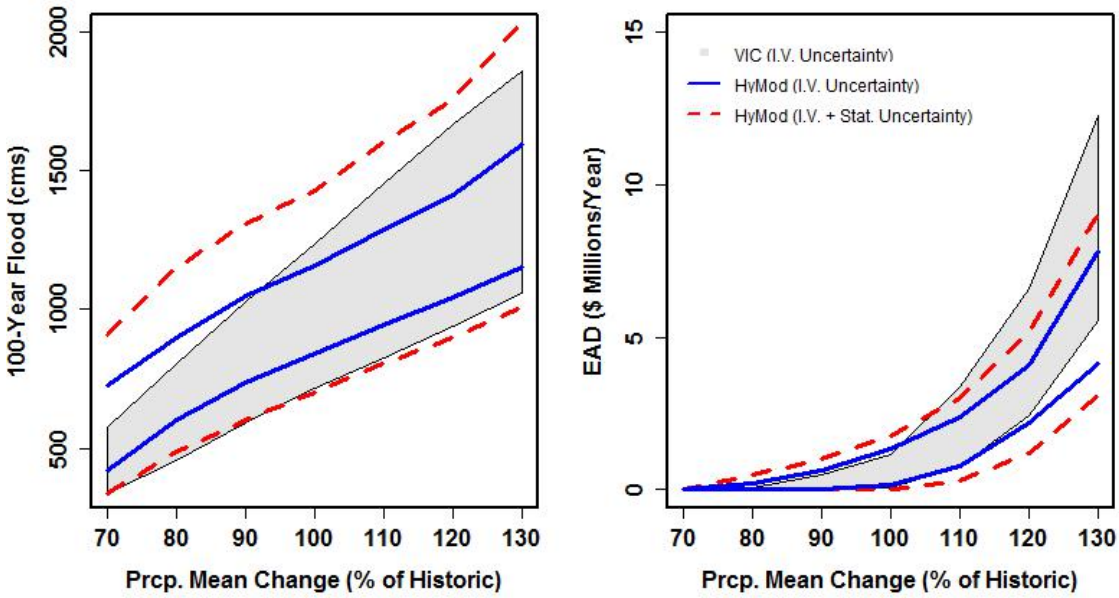

Figure 4.11. The 100-year flood estimate upstream of the Coralville Reservoir (left panel) and expected annual damages downstream of the reservoir (right panel) across a range of uncertainty factors. The range of both metrics under internal climate variability (I.V) uncertainty is shown for VIC (gray shaded) and HYMOD (blue solid). For the HYMOD model, statistical uncertainty associated with parameter calibration and residual error is shown using a 95% predictive interval for the ensemble of Monte Carlo HYMOD runs across all 10 trials (red dashed).

Two main conclusions emerge from these results. First, the statistical uncertainty of the HYMOD model is important for both upstream flooding and downstream damages, as indicated by the width this uncertainty source adds to the range of these two metrics over internal climate variability uncertainty. It is noteworthy, however, that HYMOD uncertainty appears more influential for upstream flooding estimates than for downstream damages. For instance, with a 10% increase in average precipitation, statistical HYMOD uncertainty increases the predictive bounds by 130% over those under internal climate variability alone, whereas the predictive bounds increase by 65% for EAD. This suggests that the effects of statistical hydrologic uncertainty may be dampened if there is storage available to buffer out additional random variability introduced by the hydrologic uncertainty model.

Second, the statistical uncertainty of HYMOD does not account for the structural uncertainties between the two hydrologic models when climate conditions change. For instance, when precipitation increases beyond baseline levels (100% of historic precipitation), downstream damages predicted under the VIC simulations increase above the statistical uncertainty bounds associated with the HYMOD simulations. This occurs even though VIC-simulated damages are contained by the HYMOD statistical uncertainty bounds for baseline climate, and VIC-simulated upstream flooding is contained by HYMOD predictive bounds for almost all climate changes. For the largest precipitation increase (130% of historic precipitation), the range in EAD for the VIC simulations ($6.69 million/year) is 13% larger than the range under HYMOD uncertainty ($5.94 million/year). The distributed structure of VIC can translate alternative, spatially-explicit

climate scenarios produced by the weather generator into new flood generation pathways that lead to a wide array of multi-day flooding events. Even with additional statistical uncertainty embedded in the simulations, the conceptual HYMOD model is more limited in the multi-day floods that are possible under alternative climates because there are fewer ways that water can be stored across the basin and discharged to the reservoir. The results imply that the structural differences between the two models can be more influential with respect to decision-centric metrics than the statistical uncertainty within the conceptual model in a climate risk assessment framework.

4.6. Conclusion

This work presents a framework for assessing the effects of hydrologic modeling uncertainty on the estimation of future flood risk within the context of a changing climate. The influence of hydrologic uncertainty is compared against the effects of other uncertain variables, such as precipitation and temperature. Structural and parameterization uncertainties in the hydrologic modeling process were explored using two hydrologic models and a unique error model couched in a Bayesian calibration approach.

This analysis showed that parametric uncertainty and statistical error in a conceptual hydrologic model can have important, decision-relevant implications when examining water system performance. Furthermore, the structural differences between a simple conceptual hydrologic model and a physically-based, distributed model were very

influential to the climate risk assessment and extended beyond the effects of the statistical uncertainty of the conceptual model alone.

When compared against future climate uncertainty, the hydrologic uncertainty contributed significantly to the imprecision of future flood risk estimates, as did the effects of internal climate variability. Even though uncertainty in future mean precipitation changes emerged as the most influential source of uncertainty on future flood risk, internal climate variability and hydrologic uncertainties led to changes in flood risk of the same order of magnitude, especially for large precipitation changes. This suggests that these additional uncertainties should not be ignored in flood risk studies. In fact, the influence of both internal climate variability and hydrologic modeling uncertainty outweighed that from future temperature changes or daily precipitation variance, indicating that it may be more beneficial to improve our understanding of natural climate fluctuations and hydrologic error characterization than to improve estimates of these types of climate change.

Future work will consider the use of parallel processing to apply the Bayesian uncertainty framework to the more complicated, distributed VIC model for this basin, as this study suggests that a distributed structure appears necessary for an adequate climate risk assessment of flood control infrastructure. The uncertainty framework presented in this report will then be used to explore how integrated climate and hydrologic uncertainties influence investment decisions in climate change adaptation strategies for this system.

Further work will also examine in detail the most recent set of global climate model projections to better understand how information from these projections can be used to better estimate climate-based risks to flood control facilities and inform the decision-making process.

CHAPTER 5

A FRAMEWORK TO IDENTFY ROBUST LONG-TERM WATER SYSTEM PLANS UNDER INTEGRATED UNCERTAINTIES

5.1. Abstract

A framework is presented to identify long-term adaptation plans for a water resources system that are robust to a variety of non-stationary conditions and modeling uncertainties. Several sources of uncertainty were considered, including long-term changes in the underlying distribution of future conditions, sampling variability in realizations of those transient distributions, and uncertainty inherent to transfer functions necessary to convert exogenous conditions into measures of systems performance. The integrated uncertainty analysis is coupled with systems modeling to identify long-term planning alternatives that are robust despite the uncertainties. A new metric is proposed to define robustness in this context. The framework is coupled with a host of long-term projections, including downscaled climate model output and long-term water demand forecasts, in order to understand the likelihood of potential future changes and provide useful guidance for robust planning. The approach is demonstrated in a case study examining dynamic reservoir management as a long-term planning alternative for a dual-purpose surface water reservoir in Texas providing water supply security and flood risk reduction services.

5.2. Introduction

Current long-term planning efforts for complex water systems often depend on a modeling chain that links projections of climate and societal change to models of water supply and demand and finally to a water systems model that can simulate performance under status quo conditions and a variety of planning alternatives. The goal of this exercise is to identify a long-term plan that can ensure adequate system performance under changing future conditions at a reasonable cost. Unfortunately, any insights for decision-making are hindered by a variety of uncertainties that are propagated through each stage of the modeling chain, making it increasingly difficult to select a particular long-term plan without concern that it is vulnerable to uncertainties that were not considered in the modeling process. This study presents a framework to identify effective water system planning alternatives that are robust to nonstationary conditions and a variety of modeling uncertainties. The approach utilizes Monte Carlo methods to explore and propagate the uncertainty at several modeling stages and uses systems modeling to determine which of a variety of planning alternatives provide adequate performance despite the uncertainties.

Nonstationarity in the underlying distribution of future conditions are a primary concern in long-term water systems planning efforts. Global climate change may introduce nonstationarity into local temperature and precipitation through a variety of physical mechanisms, including shifts in large-scale synoptic circulation [Sheridan and Lee, 2010] and thermodynamic boundary conditions, i.e., an increase in the water holding capacity of the atmosphere [Muller et al., 2011; Romps, 2011]. Population growth, economic

development, and shifts in per capita water use will drive nonstationarity in other important variables like municipal and agricultural water demands and land use changes, among others. In both cases, nonstationarity can manifest as shifts in the mean or as changes in variability (i.e. the coefficient of variation, serial correlation, etc.), and in both cases, these shifts are very difficult to predict with confidence.

General circulation models (GCMs) of the Earth's ocean-atmosphere system are the best tools available to understand shifts in the distribution of local climate, but moderate biases in modeled regional circulation under baseline greenhouse gas concentrations makes the interpretation of future projections difficult. For example, if a GCM exhibits erroneous circulation under baseline conditions, how should shifts in those patterns (and resulting shifts in regional precipitation) be perceived? It is unclear whether statistical bias correction used to correct the effects of erroneous baseline circulation can be directly applied to a future climate where that circulation has changed due to global warming; the bias and the future change become very difficult to disentangle. Dynamical downscaling procedures do not provide a satisfying remedy if they are constrained by erroneous large-scale circulation and by design must propagate the error forward [Xu et al., 2005]. In addition to nonstationary climate, projections of future water demands, development, and economic activity linked to the water sector are also ridden with uncertainty. For instance, projections of municipal water demand are notorious for miscalculating long-term trends in water use [Osborn et al., 1986; Fullerton and Molina, 2010]. For all of these reasons, many have claimed that nonstationary variables influencing the water sector exhibit 'Knightian' uncertainty, i.e., we are unable to estimate the probability

distributions necessary to characterize the risk of future events. As such, standard decision-theory methodologies that seek to minimize measures of risk dependent on the characterization of probabilities may be inadequate.

In response to this challenge, alternative robustness-based approaches have been proposed that shift the focus of the analysis to identifying adaptation strategies that provide satisfactory performance over a wide range of plausible future conditions. Prominent methodologies include Robust Decision Making (RDM) [Lempert et al., 2006], Scenario-Neutral Planning [Prudhomme et al., 2010], Info-Gap Analysis [Ben-Haim, 2006], and Decision-Scaling [Brown et al., 2012]. The different methodologies employ a variety of strategies and procedures, but they all recognize that no one future scenario or small subset of scenarios can adequately encapsulate the uncertainty in long-term changes to the underlying distribution of important exogenous variables. Often these approaches attempt to minimize some measure of regret, defined as the difference in terms of expected losses between a given design or plan and the optimal plan for a specific scenario. These methods adapt the planning process to cope with a state of deep uncertainty in future conditions and provide decision-relevant information despite this uncertainty.

While deep uncertainty in nonstationary factors can substantially impact the outcome of long-term water system planning efforts, it is not the only source of uncertainty. Even if the transient distributions of nonstationary variables were known, individual time series

drawn from these distributions are often required in order to estimate the impact on complex water systems, particularly when simulation models are used to understand system response. A limited number of time series realizations can lead to substantial sampling error that can bias our understanding of system response to long-term changes in the underlying distribution of nonstationary variables. This challenge is especially relevant for climate variables like precipitation that follow complex distributions with significant variability, persistence, and fat tails. The climate science community often refers to this form of uncertainty as internal climate variability [Deser et al., 2012]. Any long-term planning process must account for sampling error and ensure that the choice of adaptation plan is not vulnerable to risks that were underexplored with a limited number of scenarios drawn from underlying nonstationary distributions of influential variables.

Finally, future realizations of exogenous conditions exhibiting long-term change generally need to be passed through one or more transfer functions that relate these conditions to measures of systems performance useful for planning purposes. For long-term water systems planning, the primary transfer function is often a hydrologic model that converts future climate into streamflow time series that can be used directly by water systems planning models. While other transfer functions are also relevant, such as the systems model itself, the uncertainty in the hydrologic modeling effort often dominates. To ensure unbiased estimation of performance across different water system plans, relevant transfer function uncertainties need to be propagated through the planning process along with the other uncertainties discussed above.

This study presents a framework to identify effective water system planning alternatives that are robust to uncertain, nonstationary conditions. The Decision-Scaling methodology is utilized to manage the uncertainty in long-term change exhibited by exogenous conditions. This approach has been employed previously in a limited number of studies [Brown et al., 2011; Moody & Brown, 2012; Brown et al., 2012] but has never been embedded in an integrated uncertainty assessment as proposed here. Sampling error in climate time series is addressed using a large ensemble of stochastically generated climate simulations, while transfer function uncertainty associated with the hydrologic modeling process is quantified using a Bayesian approach. All of these uncertainties are integrated with a systems analysis approach to characterize the robustness of different planning alternatives for a dual-purpose surface water reservoir in Texas providing water supply security and flood risk reduction services. The results are coupled with a host of long-term projections, including climate model output downscaled using a variety of methods, in order to understand the likelihood of potential future changes and provide useful guidance for robust planning. The remainder of the paper will proceed as follows. Section 5.3 will introduce the framework used in this study. The case study application and specific models used in the analysis are presented in section 5.4, and results are presented in section 5.5. The paper will conclude with a discussion in section 5.6.

5.3. Methods

The proposed framework embeds systems modeling in an integrated uncertainty analysis that utilizes Monte Carlo methods to explore whether different adaptation strategies

provide adequate performance despite a wide array of uncertainties. These uncertainties include long-term changes in the underlying distribution of future conditions, sampling variability in realizations of those transient distributions, and uncertainty inherent to transfer functions necessary to convert exogenous conditions into measures of systems performance. Different planning alternatives are considered robust to a particular type of long-term change if they provide adequate performance for a certain proportion of Monte Carlo simulations associated with that long-term change. Different water demand projections and climate change projections processed through a variety of downscaling methods are used to determine the likelihood of those long-term changes. An overview of the framework is given in Figure 5.1 and is described in further detail below.

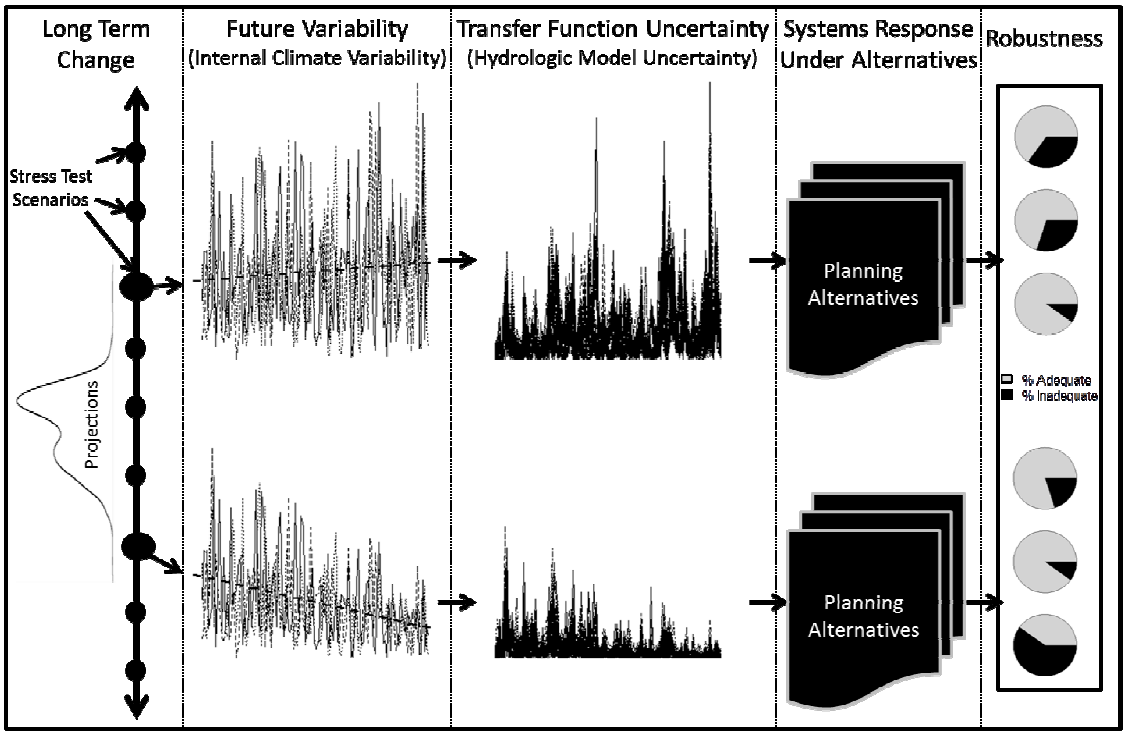

Figure 5.1. Flow chart of robustness assessment for planning alternatives under different sources of uncertainty.

5.3.1. Long-Term Change

To account for Knightian uncertainty in long-term exogenous change, this study utilizes the Decision-Scaling methodology [Brown et al., 2012]. This methodology is explained in detail elsewhere [Brown et al., 2011; Moody & Brown, 2012; Brown et al., 2012], so it is only briefly reviewed here. Decision-Scaling can be characterized by two primary steps: 1) the identification of future conditions that lead to unacceptable systems performance (i.e. a vulnerability assessment), and 2) an examination of different sources of evidence to determine whether those problematic changes are likely to occur. By separating the vulnerability assessment from the analysis of likely change, the approach ensures that the performance of the system is tested over a sufficiently wide range of possible futures to identify important vulnerabilities. When coupled with information regarding the likelihood of different futures, the vulnerability analysis provides the decision-maker with an assessment of risk facing the system.

It is not critical that all scenarios of future long-term change be plausible when first identifying system hazards, as long as implausible changes are discounted when estimating risk. The goal of the vulnerability analysis is to determine how far exogenous factors must change before the system no longer functions properly, so initially the range of future changes should be made wide enough to stress the system to failure. When those failures emerge, judgments can be made regarding the plausibility of the conditions causing them using all available sources of evidence (e.g., downscaled GCM projections, water demand projections, paleo-data records, expert opinion).

The sources of evidence used to infer the likelihood of future long-term change will certainly change as new observations emerge and projections evolve over time. A major benefit of Decision-Scaling is that new evidence regarding the likelihood of future long-term change can easily be superimposed on the vulnerability analysis as it becomes available to quickly determine if that new information suggests a substantial shift in long-term risk. A new ensemble of system model runs is not required. Furthermore, since the vulnerability analysis is decoupled from any projections, there is no need to decide on a particular set of future climate projections prior to the analysis. In fact, by coupling different types of climate evidence (e.g. different climate models or downscaling techniques) with the vulnerability analysis, the approach makes evident whether the decision to select a particular plan is sensitive to the choice of climate evidence, information that can be instructive from a decision-making perspective.

5.3.2. Future Realizations of Variability - Stochastic Climate Generation

While all system stressors are likely susceptible to some sampling error, realizations of future climate time series are particularly sensitive, especially for variables like precipitation [Deser et al., 2013]. A stochastic climate model is used to develop future climate scenarios that represent different realizations of potential futures under long-term change. This model can produce multiple climate time series that exhibit the same mean climate statistics, allowing the analyst to explore the effects of internal climate variability. Details on the specific stochastic climate model used in this study are provided later; the framework is generalizable to any model available in the literature.

In section 5.3.1, many different long-term changes are considered to explore the often-irreducible uncertainty in nonstationary variables. For climate variables, several stochastic generation runs are needed to explore internal climate variability under each long-term change, requiring the generation of hundreds if not thousands of time series simulations. When coupled with transfer function uncertainty methods (described below), the number of ensemble members that need to be run through the systems model can quickly exceed 10^6, pushing the computational limits of standard desktop computers. Multiprocessor and parallel computing offer one solution to manage the computational burden. This study presents an alternative solution appropriate for less intensive computing methods. Ex-post scenario development is utilized to select a handful of realizations of future climate that span the range of the distribution of future climate events. The following procedure is used:

1. Develop a large ensemble (1,000s - 1,000,000s) of climate simulations using the stochastic climate generator.

2. Identify decision-relevant climate events. For instance, if the system of interest is a water supply reservoir with over-year storage sufficient to meet two years of demand, then the minimum 2-year or 3-year moving average of precipitation would be an appropriate, decision-relevant climate event on which to focus. If the system meets multiple objectives, such as water supply and flood risk reduction, then several types of events (e.g. minimum 2-year precipitation total and maximum 3-day precipitation total)

can be selected. In some cases these quantities are known or can be intuited but in other cases preliminary modeling runs may be necessary to identify them.

3. Select a subset (5-10) of climate simulations that span the range of the empirical distribution of climate events. If more than one type of event is selected, then multivariate methods are required. For instance, if X is the minimum 2-year precipitation total and Y is the maximum 3-day precipitation total, then 5 time series can be selected such that $P(X \leq X_0 \cap Y > Y_0) = p$, where $p = \{0.05, 0.25, 0.50, 0.75, 0.95\}$ are probabilities and X_0 and Y_0 are drought and flood thresholds, respectively. These 5 time series represent increasingly difficult realizations of future climate variability for the system to manage (deeper droughts and more intense floods). If X and Y are independent, then by letting $X_0 = X_{\sqrt{p}}$ and $Y_0 = Y_{1-\sqrt{p}}$ be quantiles from the marginal distributions of X and Y associated with nonexceedance probabilities $\sqrt{p}$ and $1 - \sqrt{p}$, we have $P(X \leq X_0 \cap Y > Y_0) = P\left(X \leq X_{\sqrt{p}}\right)\left(1 - P\left(Y \leq Y_{1-\sqrt{p}}\right)\right) = \left(\sqrt{p}\right)\left(1 - \left(1 - \sqrt{p}\right)\right) = p$. The quantiles $X_{\sqrt{p}}$ and $Y_{1-\sqrt{p}}$ can be estimated empirically using probability plotting positions or analytically by fitting distributions to the data. If X and Y are not independent, then appropriate joint distributions, copula functions, or nonparametric methods [Serfling, 2002] can be used for the joint estimation of quantiles for X and Y.

The selected subset of future climate simulations associated with each long-term change will be used to quantify the impact of internal climate variability on systems performance.

5.3.3. Transfer Function Uncertainty – Hydrologic Modeling

We account for the uncertainty in one particular transfer function, the hydrologic model, using a stochastic representation of model parameters and residual errors. A modified version [Evin et al., 2013] of the error model presented in *Schoups and Vrugt* [2010] is chosen to accommodate the non-Gaussian, auto-correlated, and heteroscedastic nature of hydrologic model errors. Input data uncertainties are considered separately using rainfall multipliers. Parametric uncertainty is quantified using Bayesian methods.

Assume that streamflow observations $\boldsymbol{Q}$ can be modeled as the sum of hydrologic model estimates $\widehat{\boldsymbol{Q}}(\boldsymbol{X}, \boldsymbol{\theta}_M)$ and an error term $\boldsymbol{\varepsilon}$:

$$\boldsymbol{Q} = \widehat{\boldsymbol{Q}}(\boldsymbol{X}, \boldsymbol{\theta}_M) + \boldsymbol{\varepsilon} \tag{5.1}$$

Here, the hydrologic model response is a function of the forcing data $\boldsymbol{X}$ and a set of hydrologic model parameters, $\boldsymbol{\theta}_M$. To account for potential non-normality, auto-correlation, and heteroscedasticity in the residuals, the following model is proposed:

$$\varepsilon_t = \sigma_t a_t \tag{5.2.1}$$

$$\sigma_t = \sigma_0 \hat{Q}_t^{\ \sigma_1} \tag{5.2.2}$$

$$\Phi_\rho(\boldsymbol{B}) a_t = \Psi_q(\boldsymbol{B})(a_t) \tag{5.2.3}$$

$$a_t \sim SEP(0,1,\xi,\gamma) \tag{5.2.4}$$

where σ_t is a time-varying standard deviation, σ_0 and σ_1 are the parameters of a power regression of the standard deviation as a function of the predicted flow, and a_t is a normalized error term that is independently and identically distributed according to a skew exponential power (SEP) distribution with mean 0, unit standard deviation, and parameters ξ and γ to account for skew and kurtosis. The novel power regression structure presented here allows for residual variance to grow nonlinearly and can prevent unreasonably wide predictive error bounds for the largest flow values. Following *Evin et al.* [2013], the original residuals ε_t are first normalized and then corrected for autocorrelation, where $\Phi_\rho(\boldsymbol{B}) = 1 - \sum_{i=1}^{\rho} \varphi_i B^i$ is an autoregressive modeling structure of order ρ, $\Psi_q(\boldsymbol{B}) = 1 + \sum_{i=1}^{q} \psi_i B^i$ is a moving average modeling structure of order q, and B^i is the backshift operator for the i^{th} term ($B^i \varepsilon_t = \varepsilon_{t-i}$). The parameters of the error model can be lumped into the vector $\boldsymbol{\theta_\varepsilon} = \{\boldsymbol{\varphi}_{1:p}, \boldsymbol{\psi}_{1:q}, \sigma_0, \sigma_1, \xi, \gamma\}$.

Both the hydrologic and error model parameters $\boldsymbol{\theta} = \{\boldsymbol{\theta_M}, \boldsymbol{\theta_\varepsilon}\}$ are estimated in a Bayesian framework. According to Bayes' Theorem, the joint posterior distribution of all model parameters can be described as:

$$P(\boldsymbol{\theta}|\boldsymbol{Q}, \boldsymbol{X}) \propto L(\boldsymbol{Q}|\boldsymbol{\theta}, \boldsymbol{X}) \times P(\boldsymbol{\theta}) \tag{5.3}$$

where $L(\boldsymbol{Q}|\boldsymbol{\theta}, \boldsymbol{X})$ is the likelihood function and $P(\boldsymbol{\theta})$ is a prior distribution. Markov Chain Monte Carlo (MCMC) techniques allow for an exhaustive sampling of parameter values that describe the posterior space. Hydrologic model uncertainty can then be characterized by simulating a large number of hydrologic model time series predictions $\widehat{\boldsymbol{Q}}$ using posterior samples of $\boldsymbol{\theta_M}$ and adding them to time series samples of $\boldsymbol{\varepsilon}$ using posterior samples of $\boldsymbol{\theta_\varepsilon}$ [see Schoups and Vrugt, 2010].

5.3.4. Robustness of Planning Alternatives

The uncertainty analysis presented above employs Monte Carlo sampling to explore a wide range of long-term distributional changes in nonstationary exogenous variables, sampling error associated with those transient distributions, and transfer function uncertainty inherent to each realized future. Systems analysis is used to identify alternative adaptation plans that are robust despite these uncertainties. A plan or design is considered robust to a set of long-term changes in the distributions of exogenous variables if that plan can provide adequate or satisfying performance across all important system objectives for a certain proportion of Monte Carlo simulations associated with

those long-term changes. The following procedure is used to define the robustness of different planning alternatives:

1. Choose quantitative performance indicators or metrics that accurately characterize all system objectives and set thresholds for all performance metrics in order to indicate unacceptable systems performance.

2. Compile an inventory of z=1,…,Z available options or plans for adapting the system to future threats.

3. For each of i=1,..,N conditions of long-term change, simulate the water system under Monte Carlo ensemble members that explore j=1,…,J future realizations of variability and b=1,…,B realizations of transfer function uncertainty. For example, assume we are interested in exploring system response over a 50-year planning horizon to a set of long-term changes including: a 10% decline in mean precipitation, a 10% increase the coefficient of variation of annual precipitation, a 2°C temperature increase, and a 30% increase in average annual water demands. For that particular set of future conditions, J=5 stochastic climate scenarios can be coupled with B=1,000 hydrologic model simulations to produce 5,000 Monte Carlo ensemble members exhibiting that long-term change. Run these simulations through the systems model under each adaptation plan.

4. For the z^{th} adaptation plan, assign a binary performance score, X_z, to each Monte Carlo simulation. Let X_z=1 if all performance metrics are maintained above their thresholds, and X_z=0 otherwise.

5. An adaptation plan is considered robust to the i^{th} long-term change if all performance metrics are maintained above their thresholds for a certain proportion of Monte Carlo simulations associated with that long-term change. This proportion can be evaluated by averaging binary performance $X_z(i,j,b)$ across all *J* realizations of future variability and *B* realizations of transfer function uncertainty. Because a limited set of realizations of future variability were selected with different relative likelihoods, these realizations can be assigned an additional weight w_j if desired. For instance, letting and Φ be the density and distribution function of a standard normal random variable, respectively, $w_j = \frac{\phi\left(\Phi^{-1}(p_j)\right)}{\sum_{j=1}^{J}\phi\left(\Phi^{-1}(p_j)\right)}$ assigns normalized weights based on the probabilities of occurrence p for each realization of future variability (where $p = \{0.05, 0.25, 0.50, 0.75, 0.95\}$ for example). A robustness score, or *R-Score*, can then be assigned to the z^{th} plan under the i^{th} long-term change as:

$$R\text{-}Score_z^i = \sum_{j=1}^{J} w_j \frac{1}{B}\sum_{b=1}^{B} X_z(i,j,b) \qquad (5.4)$$

A plan might be considered robust to a particular set of long-term changes if the *R-Score* is greater than some threshold, say 0.75. If all realizations of future climate variability were considered equal (i.e. $w_j = {}^{1}/_{J}$), this is equivalent to the situation where 75% of the Monte Carlo runs (3,750 of the 5,000 simulations) have all metrics maintained above their threshold values. This percentage threshold should reflect the risk managers are willing to assume and therefore needs to be selected through a dialogue between analysts and system operators.

5. Steps 1-4 provide a mapping that indicates which plans provide adequate performance under different combinations of long-term change despite a variety of sampling error and transfer function uncertainties. This mapping can be combined with a variety of projection-based data or other evidence sources (e.g. climate or water demand projections, paleo-data records, expert opinion) to indicate which long-term changes are more likely than others to occur. Decision-makers can then review the robustness of different planning alternatives against a backdrop of evidence suggesting the most likely future changes when selecting a long-term plan. This process can be updated quickly as new sources of evidence (e.g. new projections, recent observations) become available.

5.4. Case Study Application

5.4.1. Study Site and Planning Alternatives

This study examines the operation of Belton Lake located in the Brazos G Regional Water Planning Area in Texas, which consists of 37 counties, 30 major reservoirs, and covers over 81,800 square kilometers [BGRWPG, 2010]. Region G is separated into several sub regions and Belton Lake is located in the IH-35 Corridor sub region. The IH-35 Corridor consists of five counties and has been subject to rapid population growth, averaging 3.9 percent annually since 1970 [Jenicek at al., 2011], raising concerns about adequacy of water supply for the region. Belton Lake drains approximately 9,680 square kilometers and is owned by the U.S. Army Corps of Engineers (USACE). The lake serves two primary purposes, downstream flood risk reduction and water supply to a host of communities, including Temple and Kileen, TX. The capacity of the reservoir is approximately 1,357 million cubic meters (MCM), with 537 MCM of water (nearly 40%

of capacity) below the conservation pool allocated for supply (Figure 5.2). The Brazos River Authority (BRA) and the military installation at Fort Hood own the only water rights to Belton Lake, with the BRA owning the majority (123 MCM of the total authorized use per year of 138 MCM). The 821 MCM of empty space above the conservation pool is managed exclusively by the USACE for flood risk reduction. In addition to serving Fort Hood and other surrounding counties, Belton Lake is also operated in conjunction with other reservoirs in Region G to supply water to users far downstream, including the metropolitan area of College Station.

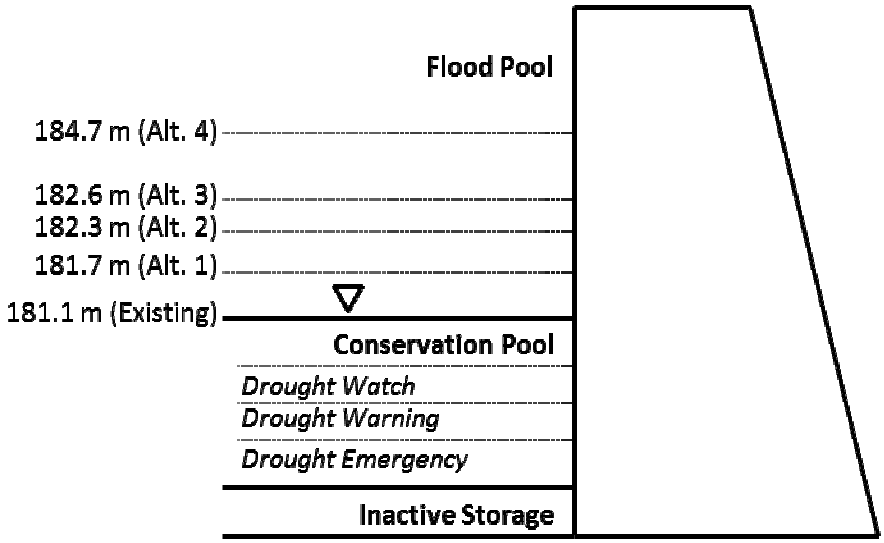

Figure 5.2. Schematic of Belton Lake with the existing and four alternative conservation pool elevations. The figure is not to scale.

The current allocation of storage in Belton Lake for flood risk reduction and water supply objectives is based on a historical analysis of inflows, water demands, and water rights. Under future climate regimes and increasing water demand, this storage allocation may be suboptimal and either or both of the primary objectives of Belton Lake may no longer

be adequately met. A potential low-regret strategy to adapt the system to changing conditions would be to dynamically reallocate reservoir storage across flood control and water supply objectives as water use and climate regimes evolve over time [Wurbs, 1987]. This type of adaptation falls into a class of dynamic management actions recently explored as a potential strategy to mitigate the effects of climate change [Georgakakos et al., 2012; Steinschneider and Brown, 2012]. In response to the regional population increase and potential need for additional water, the USACE, in conjunction with BRA, recently examined the feasibility of storage reallocation across multiple reservoirs in the Brazos River Basin using an integrated water resources model of the entire river basin, with a particular focus on the impact to safe yield under historic inflow conditions [BGRWPG, 2010]. For Belton Lake, four alternative storage allocations were considered (see Table 5.1 and Figure 5.2). This study presents a stylized extension of the USACE feasibility study that only considers one reservoir (Belton Lake) in the Brazos River Basin system but examines alternative operations for nonstationary conditions over a 50-year planning horizon. The framework presented in section 2 is used to examine system robustness under 5 alternative storage allocations (the current allocation and 4 alternatives) across a wide range of possible futures and modeling uncertainties. The analysis is coupled with a set of climate and water demand projections to frame the results in terms of the best available evidence regarding likely future long-term change.

Table 5.1. The alternative management plans for Belton Lake.

Alternative	Conservation Pool Elevation (meters above mean sea level)	Water Supply Storage (MCM)	Percent of Total Capacity (%)
Existing	181.1	537	40
Alternative 1	181.7	568	42
Alternative 2	182.3	598	44
Alternative 3	182.6	615	45
Alternative 4	184.7	739	54

5.4.2. Data

Historic daily climate data, including precipitation and maximum, minimum, and mean temperatures, were gathered for the Belton Lake watershed over the period of October 1, 1949 to September 30, 2010 from the gridded observed meteorological dataset produced by *Maurer et al.* [2002]. These data have a spatial resolution of approximately 144 km^2. All climate fields were spatially averaged across the entire watershed and were aggregated to a monthly time step to produce a single, monthly time series. Monthly potential evapotranspiration was calculated for the basin using the temperature-based Hargreaves method [Hargreaves and Samani, 1982].

Daily observed streamflow data were gathered from two United State Geologic Survey (USGS) gages, the Leon River near Belton gage (ID# 08102500) directly downstream of Belton Lake and the Little River near Little River gage (ID# 08104500) farther downstream. These data were gathered for the period of October 1, 1953 to September 30, 2010. Historical daily inflows, storages, and releases for Belton Lake were gathered online from the USACE Fort Worth District Data Center [http://www.swf-

wc.usace.army.mil/cgi-bin/rcshtml.pl?page=Hydrologic] for the same period. All data were also aggregated to a monthly time step.

Current demand data for all municipal water users served by Belton Lake were taken from county/utility annual demand estimates by BRA [BGRWPG, 2010]. No data was available for the region regarding monthly use, so monthly demand data from a nearby watershed [Griffin and Chang, 1990] was used to generate monthly demand factors and parse annual demands into a monthly time series. Two sets of projections for water demand across all counties served by Belton Lake were gathered from BRA [BGRWPG, 2010] and the Texas Water Development Board (TWDB) [https://www.twdb.texas.gov/waterplanning/data/projections/index.asp]. These two data sources project water demands to increase above current (2010) levels by 48% (BRA) and 62% (TWDB) by the year 2060. We note that while demand is likely related to temperature changes, no such link exists in the water demands used in this study.

Projections of climate change for the Belton Lake watershed were gathered from the North American Regional Climate Change Assessment Program (NARCCAP) [Mearns et al., 2009]. Four AOGCMs (CCSM, CGCM3, GFDL, and HadCM3) were used to drive 6 RCMs (CRCM, ECP2, HRM3, MM5I, RCM3, WRFG) in a variety of combinations (Table 5.2). All simulations were conducted under the A2 SRES emission scenario. Changes in mean annual precipitation and mean annual temperature were calculated for both raw AOGCM output and downscaled RCM simulations. Precipitation and

temperature changes were assessed as a ratio and difference, respectively, between the future (2041-2070) and baseline (1971-2000) periods, averaged across the Belton Lake watershed.

Table 5.2. Climate projections used in this study.

	Driving AOGCM			
RCM	CCSM	CGCM3	GFDL	HadCM3
CRCM	X	X		
ECP2			X	
HRM3			X	X
MM5I	X			X
RCM3		X	X	
WRFG	X	X		

5.4.3. Belton Lake Reservoir Simulation Model

A water resources systems model was developed to simulate the operations of Belton Lake under different management plans. The systems model is a simple monthly mass-balance model that tracks inflows (e.g., river inflows and precipitation directly onto the lake) outflows (e.g., controlled and uncontrolled releases downstream, demand withdrawals directly from the lake, and evaporation directly off of the lake), and storage accumulation. In accordance with the BRA 2012 Drought Management Plan [unpublished report, 2012], water withdrawals and releases for downstream demands are increasingly curtailed as reservoir storage falls below a set of successively lower trigger levels (see Figure 5.2). A simple release rule for flood control is imposed whereby gated downstream releases are used to maintain reservoir levels at the conservation pool but

cannot exceed 220 MCM/month in high inflow months. This threshold is equivalent to an entire month of flow at the bank-full flooding threshold set for the downstream Leon River flood control checkpoint gage (USGS ID# 08102500).

We note that given the flood risk reduction objectives of Belton Lake, a daily model was initially deemed necessary. This requirement would have added significant computational burden to the methods used in this work, requiring the use of parallel computing. However, further investigation suggested that floods can be adequately modeled at a monthly time step, primarily because the capacity of the reservoir is so large that significant reservoir spill events and subsequent downstream damages only occur if inflows are substantially above average for several weeks at a time. Figure 5.3a shows the annual maximum storage levels in Belton Lake between 1980 and 2010 versus cumulative inflows preceding the date of the annual storage maxima. Several accumulation periods are considered, including 1-day, 15-day, and 30-day cumulative inflows. Figure 5.3a shows that maximum annual storage levels are most closely related to 30-day cumulative inflow periods, while there is significant noise in the relationship with the 1-day period. This suggests that a single day of large inflow is generally not sufficient to threaten the flood control objectives of Belton Lake, but rather weeks of high inflows are needed to cause the reservoir to spill. Therefore, a monthly time step model was deemed acceptable for representing the flood control objective of the reservoir. Figure 5.3b shows that the historic simulation of the reservoir under observed inflows compares well against observed monthly storage levels, suggesting that the monthly model adequately captures current operations.

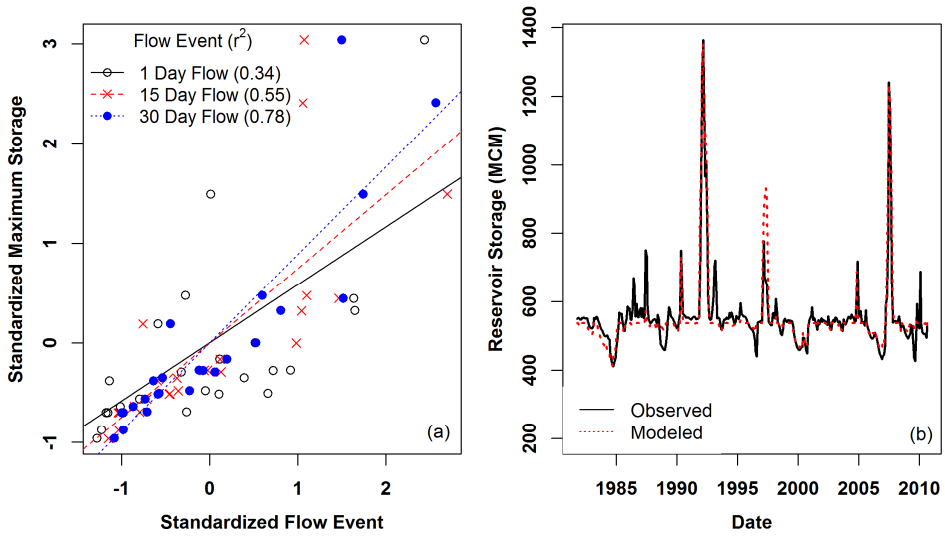

Figure 5.3. a) Annual storage maxima in Belton Lake versus cumulative inflows prior to the annual maxima date of occurrence. Three accumulation periods are shown, including 1-day, 15-day, and 30-day inflows. All storage and flow events are normalized for comparison. b) Observed and simulated monthly average storage in Belton Lake.

5.4.4. Future Climate Scenarios

The stochastic climate model used in this analysis is a monthly version of a previously developed daily model [Steinschneider and Brown, 2013] and therefore is only described briefly here. The proposed model has three primary components, including 1) a wavelet decomposition coupled to an autoregressive model (the wavelet auto-regressive modeling (WARM) approach, see *Kwon et al.* (2007)) to simulate structured, low-frequency oscillations in the aggregate climate (e.g. seasonal precipitation), 2) a temporal disaggregation approach to convert seasonal climate to a monthly time step, and 3) a post-processing adjustment to enforce long-term distributional shifts in climate variables under climate change. These components allow the model to generate time series of climate variables that exhibit realistic characteristics at long-term (inter-annual) and mid-

term (seasonal) timescales, and they also enable the creation of many climate change scenarios over which to stress system performance.

The climate over the Belton Lake watershed exhibits a dual peak in precipitation, with one peak occurring in the spring between January and June (JFMAMJ) and another occurring in the autumn between the months of July and December (JASOND). The stochastic climate generator simulates, disaggregates, and adjusts climate variables from these two seasons separately and then combines them to form a continuous time series of climate data. This is justified since the correlation between seasonal precipitation is very low (Pearson r ~ -0.03). The WARM approach is used to extract low frequency signals in each time series of seasonal precipitation using wavelet decomposition and then stochastically simulates each signal using autoregressive time series models. The power spectrum for spring precipitation exhibits a near-significant 2-4 year signal at the 90% confidence level (as compared to a background white noise spectrum), while autumn precipitation exhibits a dual spectral peak at 1-2 and 5-7 year periods (Figure 5.4). Some of this low-frequency variability is likely related to the well-established influence of El-Nino-Southern Oscillation (ENSO) on the climate of central Texas [Piechota and Dracup, 1996]. The WARM approach is used to simulate the time-series characteristics of these modes of variability as well as residual noise, and in both seasons, the WARM model is capable of reproducing much of the spectral signature of the observed data, particularly for the autumn season. An anomalous spectral peak does emerge in the WARM simulations for spring precipitation near a 7-year period, but this bias is not drastic. Also, the 95% predictive bounds encompass most of the observed power spectrum except for

long periods in the autumn season. Seasonal average, minimum, and maximum temperatures are predicted from simulated seasonal precipitation values using linear regressions of observed temperature variables against observed seasonal precipitation.

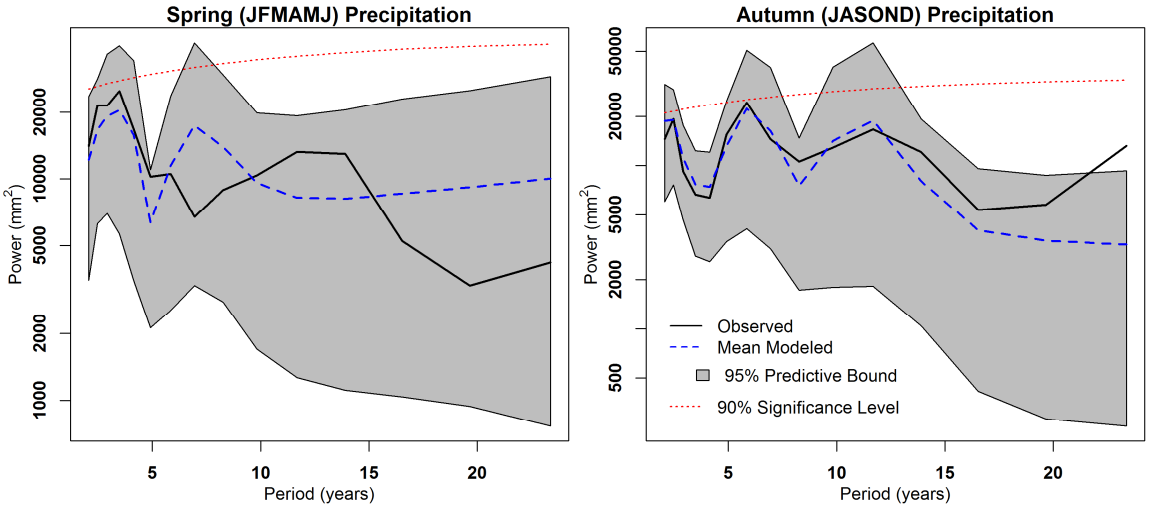

Figure 5.4. Power spectra for seasonal precipitation. The observed spectra (black solid) are compared against the mean power spectra (dashed blue) of 500 simulations, along with the 95% predictive range of the ensemble (grey). Also shown is the 90% significance level (red dotted) developed from a white noise background process.

All seasonal climate variables are then disaggregated to a monthly time step using the method of fragments [Srikanthan and McMahon, 2001]. This is accomplished through a k-nearest-neighbor resampling of monthly disaggregation factors based on the simulated precipitation values. For each simulated value of seasonal precipitation, the nearest k=7 observed values from the observational record are selected. The value k is selected as the square root of the record length, as suggested by *Lall and Sharma* [1996]. Each of the k values is assigned a probability using the discrete kernel function $[d^l] = \frac{1/l}{\sum_{l=1}^{k} 1/l}$, where l indexes the k selected values according to their Euclidean distance d^l from the simulated

seasonal value. One of the k observed values is selected based on the kernel weights, and the monthly precipitation and mean, minimum, and maximum temperature values are gathered for that season. Multiplicative disaggregation factors are derived by dividing the monthly precipitation values by the observed seasonal total. Similarly, additive temperature factors are derived by subtracting the observed seasonal temperature average from the monthly values. The multiplicative and additive factors and then applied to the simulated seasonal precipitation and temperature values to disaggregate them to a monthly time step. This is repeated for every season of simulated climate from the WARM model.

Figure 5.5 compares a range of statistics for monthly precipitation and mean temperature, including the mean, standard deviation, and skew, across a 500-member ensemble of stochastic climate model simulations and the observed values. The ensemble of stochastic simulations reproduces most of the statistics for all months rather well. The standard deviation and skew of precipitation are somewhat underestimated for June, November, and December, while the skew for mean temperature is somewhat overstated in February, October, November, and December. However, the overall performance of the model for both precipitation and temperature variables is satisfactory and considered adequate for the purposes of this study.

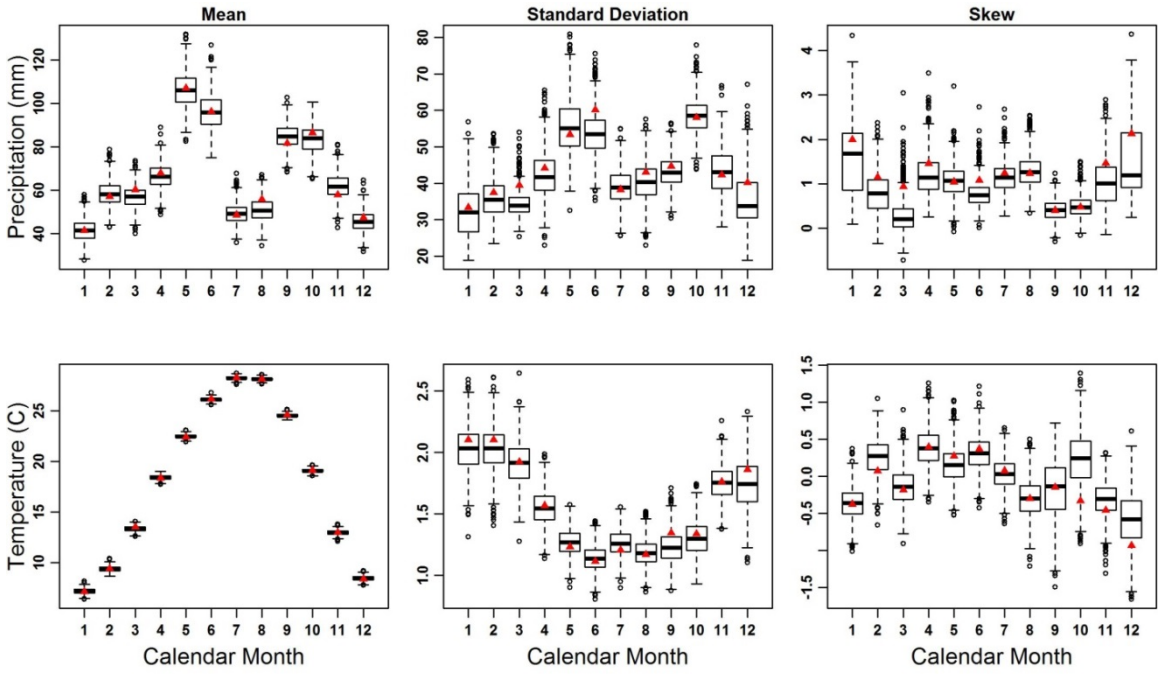

Figure 5.5. Performance statistics by month, including the mean, standard deviation, and skew of precipitation and average temperature. The observed statistics (red triangles) are shown against the distribution of statistics across the 500 stochastic climate simulations.

To impose various climate changes in the simulated time series, transient additive delta factors are used to adjust temperature values over the simulation period. These factors increase linearly by month for the entire simulation period, starting at 0 and ending at the level of specified change (e.g. 2ºC). For precipitation, quantile mapping is utilized. The quantile mapping procedure can be used to adjust the simulated precipitation over time so that it follows a new distribution with transient characteristics (e.g. time-varying mean, coefficient of variation, etc.) (see *Steinschneider and Brown* [2013] for details). In this study, three types of climate change are examined, including alterations to the mean of seasonal precipitation, its coefficient of variation (CV), and the mean of seasonal temperatures. Even though different seasonal changes are possible using the climate

generator, the same change is always applied to both seasons so no shifts in seasonality are explored here. This choice was made because the storage in Lake Belton is large enough to modulate intra-annual variability. Changes to the precipitation mean are ranged from ±30% of historic seasonal averages using increments of 15% (5 increments). The precipitation CV is also changed by season from ±30% of historic monthly values using increments of 30% (3 increments). Temperature shifts are ranged from 0 to 4 degrees Celsius by 2°C increments (3 increments). All possible combinations of these changes are considered, leading to a total of 45 = 5×3×3 different climate change scenarios. These changes were chosen to ensure the identification of climate changes that cause system failure. Each of these 45 combinations of long-term change was applied to 5 different 50-year monthly climate simulations, producing 225 climate simulations altogether. These 5 realizations of internal climate variability were selected amongst 10,000 original runs according to the methods in Section 2.2 using 5 nonexceedance probabilities $p = \{0.05, 0.25, 0.50, 0.75, 0.95\}$ from the joint distribution of the minimum 1-year precipitation total (drought metric) and the maximum 1-month precipitation total (flood metric). The joint distribution was assumed equivalent to the product of the two marginal distributions because these two variables are independent (Pearson $r < 0.001$ across the 10,000 original runs).

5.4.5. Future Water Demand Scenarios

Demands are linearly increased for all municipal users throughout the system over time in order to evaluate the vulnerability of the water supply of Fort Hood to population growth across the region. Demand changes are applied to the local community users

only, i.e. the utilities and nearby cities with access to and authorized use of Belton Lake. Three different demand scenarios are considered, including 100%, 130%, and 160% of current demand levels. The maximum demand scenario was capped at 160% of current demand because this increase is equivalent to the total legal allocation available to the two water-rights holders for Belton Lake, BRA and Fort Hood. All scenarios of future demand are derived by multiplying 2010 demand levels by a time series of transient, multiplicative factors over the 50-year planning horizon for all nearby municipal users. Manufacturing, steam-electric, mining, irrigation, and livestock demand changes are minor in comparison and are not considered in this study. Downstream user demands (e.g. sporadic releases for College Station and other major water rights contract holders) are included in the analysis, but are only applied at historical levels. When combined with the 225 different climate change scenarios, these three demand changes lead to 675 different scenarios of future long-term change and internal climate variability. Each of the 675 scenarios will be coupled with an ensemble of hydrologic model simulations to complete the integrated uncertainty analysis.

5.4.6. Belton Lake Hydrologic Model

The conceptual, lumped-parameter HYMOD model [Kollat et al., 2012] was selected to translate future climate scenarios into inflows for Belton Lake. This model is composed of a soil-moisture accounting module that utilizes a storage capacity distribution function with parameters C_{max} and β and a routing module that divides excess water using a split parameter (α) and routes it through parallel conceptual linear reservoirs. The quick (K_q) and slow (K_s) reservoir depletion rates control the flow from each routing reservoir. The

outputs from quick and slow reservoirs are summed to simulate streamflow. In additional to the hydrologic model parameters, two additional rainfall multipliers M_s and M_a for the spring (January-June) and autumn (July-December) seasons, respectively, were included because there are a limited number of rain gages in this region that were available to develop the gridded data product used in this study.

During the calibration process, a total of 13 parameters require estimation, including 7 hydrologic model parameters ($\boldsymbol{\theta}_M = \{C_{max}, \beta, \alpha, K_q, K_s, M_s, M_a\}$) and 6 parameters for the residual error model ($\boldsymbol{\theta}_\varepsilon = \{\varphi_1, \psi_1, \sigma_0, \sigma_1, \xi, \gamma\}$). One autocorrelation coefficient and one moving average coefficient are included because diagnostic tests (not shown) suggest autocorrelation in the residuals can be represented using a simple ARMA(1,1) process. The model calibration was conducted over 40 years of historic climate, from October 1953 to September 1989, leaving 10 years for validation. All prior distributions were set to vague uniform distributions. In this study, the DREAM$_{(ZS)}$ MCMC algorithm was used to explore the posterior parameter space [Vrugt et al., 2011]. Convergence across three MCMC sampling chains used in the Bayesian calibration was verified using the Gelman and Rubin factor [Gelman and Rubin, 1992]. The feasible range, prior distribution, and posterior distributions for all model parameters are presented in Table 5.3.

Table 5.3. Summary of prior and posterior distributions for all model parameters.

			Posterior Distribution			
Parameter	Feasible Range	Prior Distribution	First Quartile	Median	Mean	Third Quartile
c_{max} (mm)	(0,∞)	Uniform (a=0,b=1000)	291.44	324.51	313.33	340.81
b	(0,2)	Uniform (a=0,b=2)	0.25	0.28	0.28	0.31
α	(0,1)	Uniform (a=0,b=1)	0.47	0.56	0.59	0.73
K_q	(0,1)	Uniform (a=0,b=1)	0.84	0.91	0.89	0.96
K_s	(0,1)	Uniform (a=0,b=1)	0.29	0.39	0.37	0.46
M_s	(0,∞)	Uniform (a=.9,b=1.1)	1.07	1.08	1.08	1.09
M_a	(0,∞)	Uniform (a=.9,b=1.1)	0.93	1.00	0.99	1.06
σ_0	(0,∞)	Uniform (a=0,b=1)	0.58	0.59	0.59	0.60
σ_1	(0,∞)	Uniform (a=0,b=1)	0.99	0.99	0.99	1.00
ξ	(0,∞)	Uniform (a=0.1,b=10)	1.58	1.67	1.68	1.76
γ	(-1,1)	Uniform (a=-1,b=1)	0.89	0.94	0.92	0.97
φ_1	(-1,1)	Uniform (a=-1,b=1)	0.32	0.37	0.37	0.41
ψ_1	(-1,1)	Uniform (a=-1,b=1)	-0.13	-0.07	-0.08	-0.03

Figure 5.6a shows the observed daily hydrograph for part of the calibration period, as well as the model estimates generated by the mean of the posteriors for all HYMOD parameters. The Nash-Sutcliffe efficiency (NSE) for the mean prediction over the calibration and validation period is 0.70 and 0.60, respectively. Also shown in Figure 5.6a are 90% predictive bounds for the modeled flows. These bounds are calculated by sampling 5,000 parameter sets from their posterior distributions and summing 5,000 simulated time series of model predictions and randomly generated residuals. These bounds contain 87% of the observations, suggesting that hydrologic uncertainty is being adequately characterized.

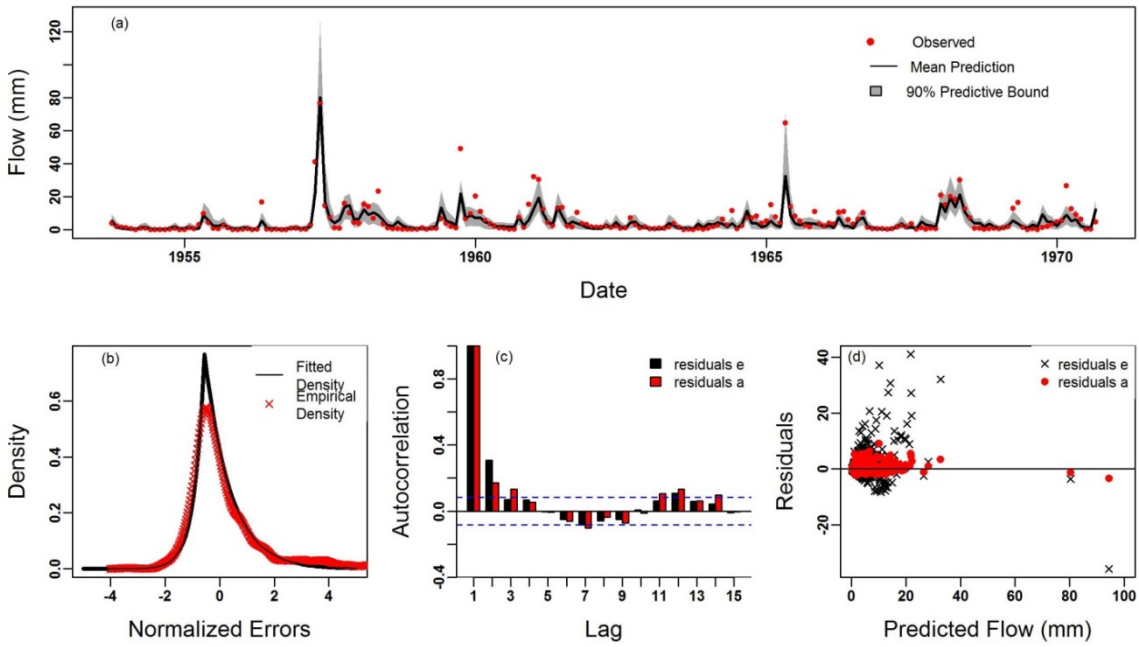

Figure 5.6. a) Streamflow hydrograph for a portion of the calibration period. Observations (red dots) are shown against the mean HYMOD prediction (black line) and the associated 90% confidence bounds (shaded grey). b) The fitted SEP density (black line) and empirical density (red points) of normalized errors. c) The autocorrelation function of the original and normalized errors. d) The residuals plotted against the mean predicted flow.

Diagnostic plots are used to further verify that the fitted error model is consistent with the observed distribution of the residuals. Figure 5.6b shows the fitted SEP density function using the two parameters γ and ξ, as well as the empirical density of the normalized residuals, a_t, of the HYMOD model. The parameter γ is estimated very close to unity, suggesting that the error distribution is highly peaked, and $\xi \approx 1.57$, suggesting positive skew in the distribution. Figure 5.6c shows the autocorrelation function of the residuals from the model. While the original residuals of the HYMOD model exhibited significant autocorrelation at lag 1, this autocorrelation has largely been removed from the

normalized residuals. Finally, Figure 5.6d shows the relationship between predicted streamflow values and the errors associated with those predictions. While the variance of the original residuals varied significantly with predicted flow, the normalized errors maintain a much more constant spread across the range of flow predictions. These three diagnostic plots suggest that the error model used adequately captures the non-Gaussian, auto-correlated, and heteroscedastic nature of the residuals.

In each future climate scenario for Belton Lake, an ensemble of 200 HYMOD simulations are run using samples from the posterior parameter space and additional residuals added to each simulation, as described in section 2.3. When coupled with the 675 scenarios of long-term change and internal climate variability, a total of 135,000=675×200 Monte Carlo ensemble members are generated over which to run the Belton Lake simulation model.

5.4.7. Measures of Performance and Robustness

System performance is tracked using four different metrics and thresholds (Table 5.4). Flood risk reduction is measured using the frequency of reservoir spill events and the maximum flow directly downstream of the reservoir. Any simulation is considered a failure if more than 2 independent spill events occur during the 50 year simulation or if the magnitude of downstream flow rises above 730 MCM (i.e. the flow required to cause flooding for an entire month at the Little River flood control checkpoint gage (USGS ID# 08104500)). Water supply is measured using the frequency of drought watch events and

the minimum storage achieved during the simulation. A simulation is considered unsatisfactory if the frequency of drought watch events is greater than 1 in every 5 years or minimum storage falls below 216 MCM (the designated emergency storage level). A particular management plan z is considered robust to the i^{th} long-term change if $R\text{-}Score_z^i > 0.75$ across the Monte Carlo simulations that represent future variability and hydrologic uncertainty. This threshold, arbitrarily chosen here, should reflect the risk managers are willing to assume when deciding whether a plan is adequate for a particular set of long-term conditions.

Table 5.4. Objectives, metrics, thresholds, and robustness criteria.

Objective	Metric	Threshold of Acceptable Performance	Robustness Threshold
Flood Risk Reduction	Frequency of Reservoir Spills	2 in every 50 years	75%
	Magnitude of Downstream Flows	200 MCM	75%
Water Supply Security	Frequency of Drought	1 in every 5 years	75%
	Minimum Storage	216 MCM	75%

5.4.8. Testing the robustness of dynamic reservoir management

The framework presented in this study is used to demonstrate the robustness of a dynamic management strategy for Belton Lake in which the conservation pool is adjusted through time to manage different types of long-term change and maintain an adequate level of flood control and water supply service. In each 50-year simulation, the conservation pool elevation is kept at its current level for the first 25 years and then is switched to an alternative pool level for the remaining 25 years. Two modeling experiments are conducted to explore the robustness of the adaptive management strategy. In the first experiment, all 5 management plans - maintaining the present conservation pool or

switching to one of the 4 alternative levels at year 25 - are run for each of the 135,000 Monte Carlo ensemble members. For each management plan and long-term change (a specified precipitation mean and CV, temperature mean, and mean water demand), the *R-Score* is calculated and robustness (e.g. *R-Score* > 0.75) mapped across the space of all long-term changes. This will show whether dynamic management of Belton Lake could mitigate the effects of long-term change *if* managers knew which change was going to occur and thus which plan to select.

During the true realization of future climate and water demands, however, operators of Belton Lake will not know which path of long-term climate and water demand change is actually occurring. The second modeling experiment will explore whether the dynamic management strategy is still effective given the lack of perfect foresight. Here, for each of the 135,000 Monte Carlo simulations, a decision rule is implemented to choose whether the lake level should be adjusted during the simulation. For each Monte Carlo ensemble member, the reservoir system model is run under the current conservation pool level for the first 25 years. If the current lake level provides satisfactory performance for that first period, then no change is made for the second 25 years. However, if performance is unsatisfactory, then the lake level is iteratively raised to the next highest elevation and back-tested over the first 25 years until performance is satisfactory for that period. The lake level that first provides adequate back-tested performance is then used for the next 25 years of the simulation. If no lake level provides satisfactory performance for the first 25 years, then the existing plan is maintained. The decision rule used in the second

experiment essentially tests whether dynamic management can "see" past the noise of any particular realization of the future and effectively react to emerging trends.

5.5. Results

Figure 5.7 displays the robustness of two management plans (the status quo and Alternative 4) for changes in municipal water demand and average precipitation, with temperature and the precipitation CV held at baseline levels. The existing conservation pool elevation is capable of providing adequate performance for both flood control and water supply objectives for all demand levels when precipitation averages are at 100% and 85% of historic averages, but when precipitation drops significantly to 70% of historic levels, the existing plan is only viable at current demand levels; increases in demand lead to unacceptable water supply performance. Conversely, Alternative 4 (a very high pool elevation) can robustly manage all futures where precipitation falls below 85% of historic averages, but does not perform well under current precipitation levels because downstream flooding becomes too severe. This demonstrates the tradeoffs across the different system objectives that emerge when the conservation pool elevation is changed.

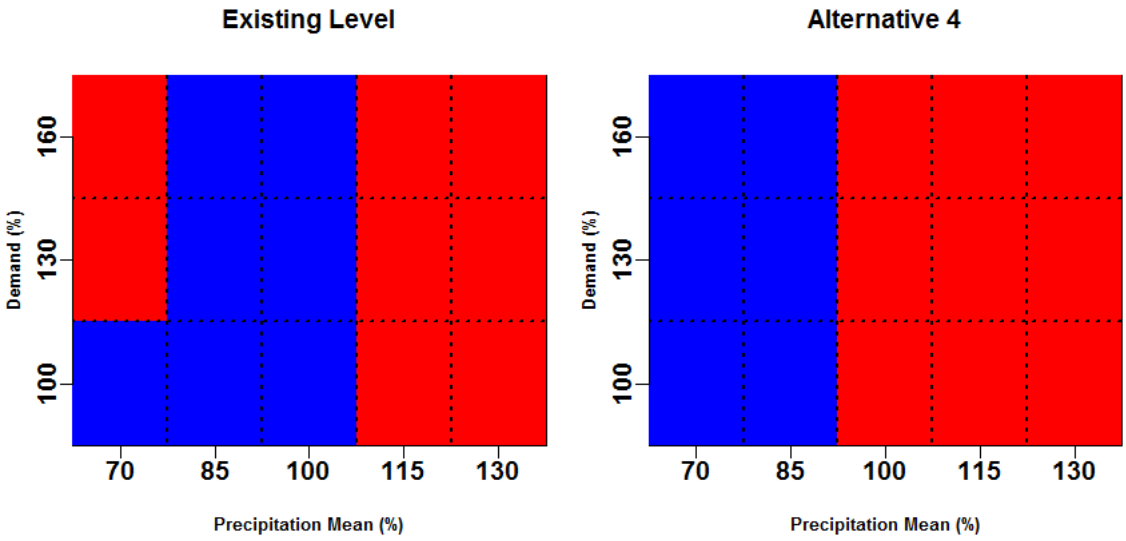

Figure 5.7. Robustness of the existing conservation pool and Alternative 4 under future changes in water demands and mean precipitation, with mean temperature and precipitation CV held at baseline levels. Blue regions indicate futures under which plans are robust (*R-Score > 0.75*).

Figure 5.8 summarizes the results of the first modeling experiment for all of the combinations of long-term change considered. Here, each long-term change is assigned the management plan that provides robust performance. When multiple plans are robust for a particular type of change, the plan closest to the existing conservation pool elevation is chosen because there are costs associated with raising the conservation pool (e.g. flooding out infrastructure and conservation lands surrounding the lake). Also designated are long-term changes for which no plan provides robust performance. Several results emerge from Figure 5.8. First, the existing conservation pool provides adequate performance across a wide swath of future changes in mean precipitation, mean temperature, and water demands as long as the variability (i.e. CV) of precipitation does not rise substantially. This includes all futures with current precipitation averages, as well

as the majority of futures with a 15% decline in precipitation. As temperatures or demands rise, higher lake levels are required to maintain adequate performance. In most instances, the next two lake levels (Alternatives 1 or 2) are sufficient to maintain performance, but when demands rise in conjunction with drier and warmer futures, the highest lake level (Alternative 4) is required. During some the warmest, driest futures with substantial water demand increases, no lake level performs adequately. When the CV of seasonal precipitation increases, there are very few futures that any plan can manage well. This suggests that when precipitation variability rises, droughts and floods both occur more frequently and cannot be managed effectively simply by reallocating storage between flood risk and water supply objectives. In these futures, some additional infrastructure or other investment would be required to manage the increased variability and maintain adequate performance.

Projections of future water demands (BRA and TWDB) and climate (raw AOGCM and RCM-downscaled) for the year 2060 are also superimposed on Figure 5.8. All climate projections exhibited temperature changes close to 2ºC and no substantial change in the CV of seasonal precipitation. Since climate and demand projections were generated independent of one another, all climate projections were coupled with both the BRA and TWDB demand projections. The majority of projections suggest that the status quo plan is robust to future changes, but a handful of projections indicate that Alternative 1 may need to be adopted. A few projections also suggest that in a wetter future, no plan can sufficiently provide robust performance. Importantly, in this case study, the changes suggested by raw AOGCM output do not differ substantially from the downscaled RCM

results, suggesting that the process of downscaling for this application is not critically important for the decision-making process.

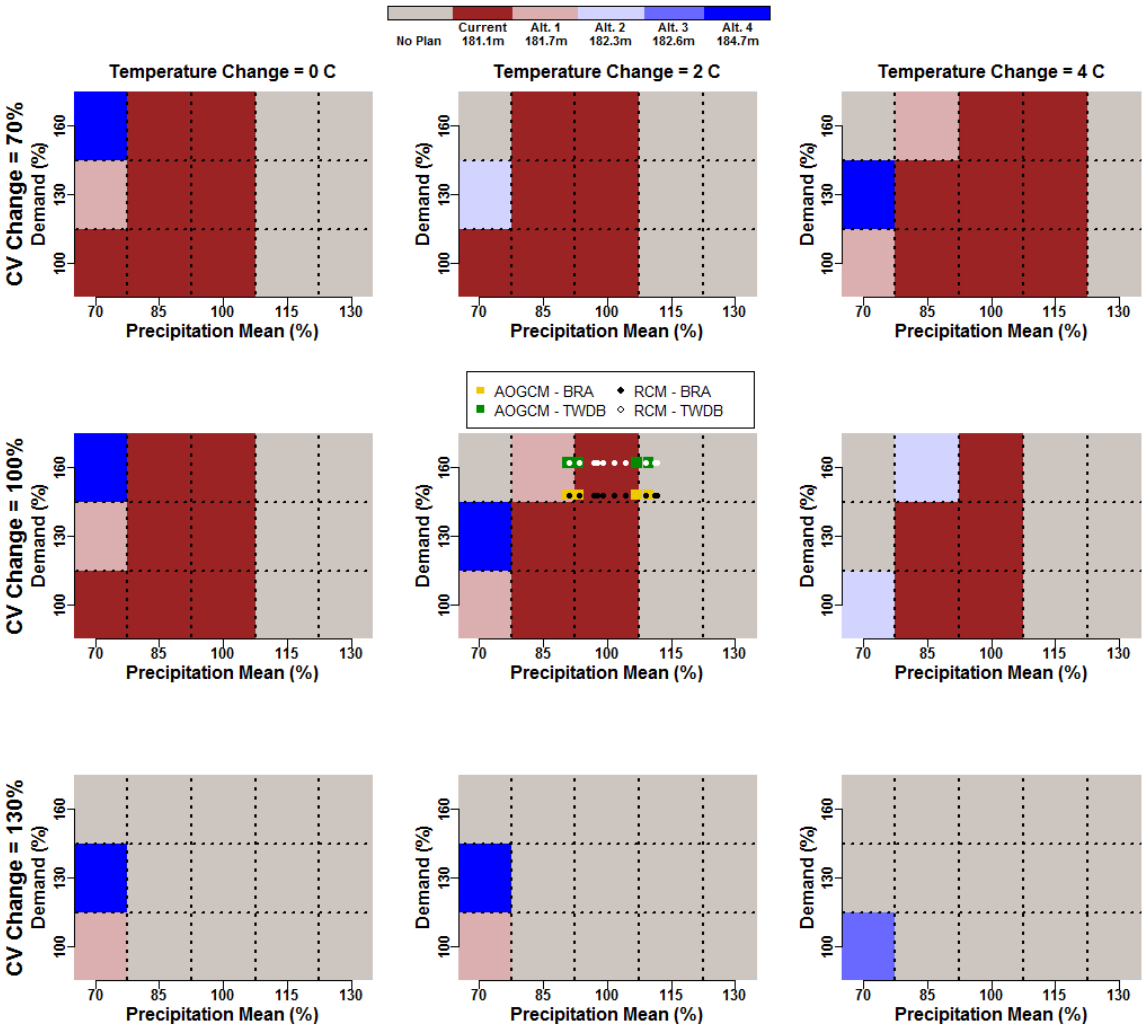

Figure 5.8. The management plan that provides robust performance (*R-Score* > *0.75*) for long-term changes in mean precipitation (x-axis), water demands (y-axis), mean temperature (columns), and precipitation CV (rows). Also shown are 2060 projections of climate from the NARCCAP project (raw AOGCMs and downscaling RCMs) combined with 2060 projections of water demand from BRA and TWDB. All projections exhibited temperature increases near 2ºC and thus were plotted in the 2ºC column.

Figure 5.8 shows which plans would provide robust performance across a range of future changes and modeling uncertainties *if* system operators knew to implement them half way through the planning period. However, in a real-time operational setting, operators may not know which future path is emerging, especially if hydroclimatic noise obscures their ability to detect long-term trends. Figure 5.9 shows the results of the second modeling experiment, in which lake levels are dynamically changed by back-testing their performance over the first 25 years. The results of this experiment highlight the utility of dynamic reservoir management as an adaptation to climate change when based on a naïve forecast, i.e., the past several decades are indicative of the trends to emerge in the next several decades. Here, different plans may be chosen for different Monte Carlo simulations under a particular long-term change, but the general strategy of dynamic management is considered robust to a long-term change if *R-Score* > 0.75. The robustness of the status quo plan is also shown for comparison in Figure 5.9 to determine if dynamic management provides substantial advantages over the current plan. Figure 5.9 shows that without perfect foresight, the dynamic management plan loses much of its utility, and provides almost no additional benefit above just maintaining the current conservation pool level. The dynamic management strategy provides robust management for only one additional long-term change condition over the status quo plan (30% decline in precipitation, 2□C increase in temperature, no change in precipitation CV or demand). In part, this is because the dynamic management strategy chooses to maintain the current conservation pool elevation for over 50% of all Monte Carlo simulations across all long-term change conditions. Even under futures with substantial drying and higher water demands, there is enough hydroclimatic noise in the first 25 years of the simulation to

prevent the performance degradation needed to trigger another management plan. These results suggest that in order for dynamic reservoir management to be a viable adaptation strategy for many long-term changes, some forecast skill beyond a naïve back-testing method may be necessary.

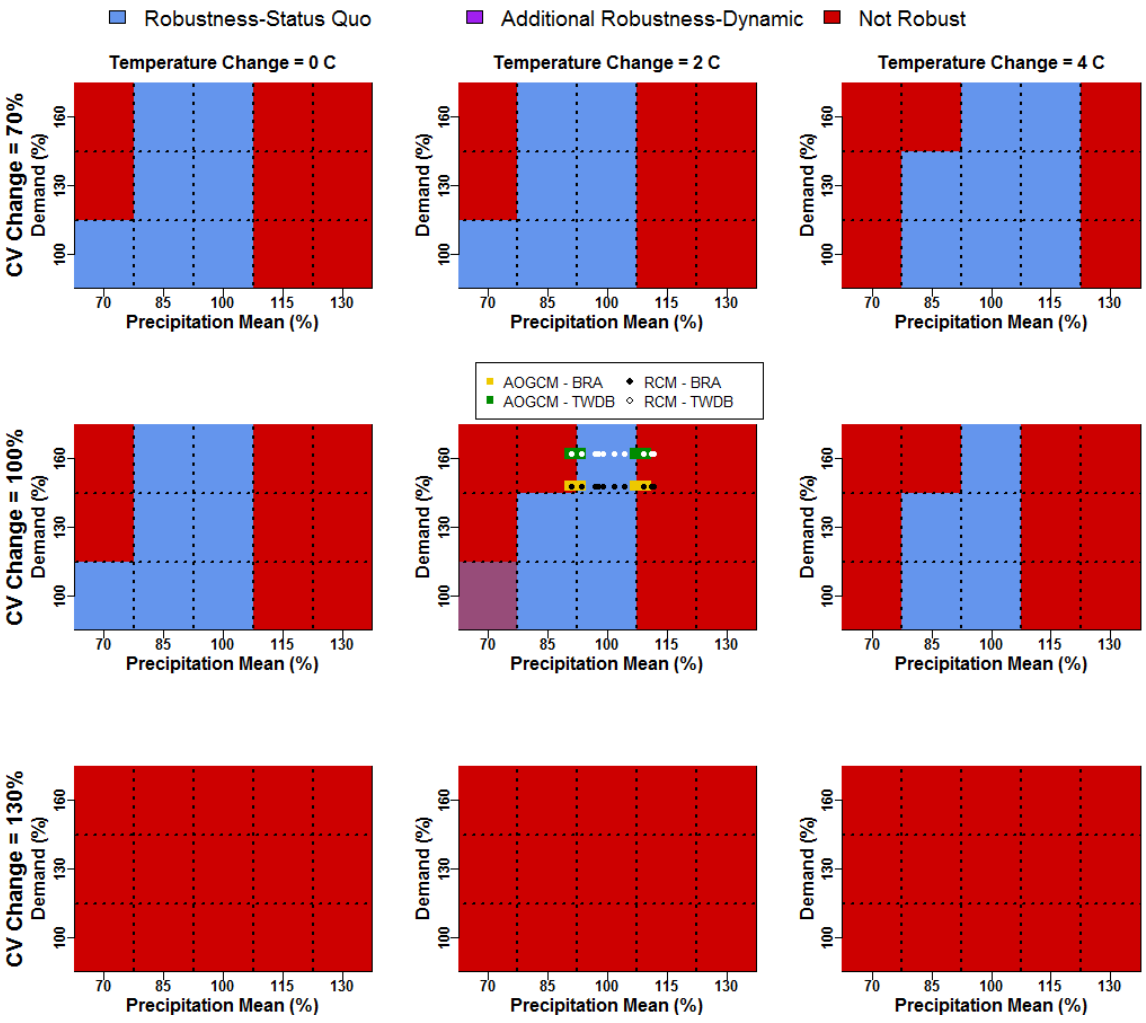

Figure 5.9. The robustness of the status quo plan and the dynamic management strategy for long-term changes in mean precipitation (x-axis), water demands (y-axis), mean temperature (columns), and precipitation CV (rows). Also shown are 2060 projections of climate from the NARCCAP project (raw AOGCMs and downscaling RCMs) combined with 2060 projections of water demand from BRA and TWDB. All projections exhibited temperature increases near 2ºC and thus were plotted in the 2ºC column.

When coupled with the Figure 5.9, a subset of the projections suggests that dynamic management will not provide robust performance. One raw AOGCM (HadCM3) and one downscaled RCM (ECP2/GFDL) both indicate that reduced precipitation will lead to inadequate performance, while another raw AOGCM (CGCM3) and three RCMs (HRM3/HadCM3, MM5I/HadCM3, and WRFG/CGCM3) indicate that wetter conditions will lead to poor flood control performance. Interestingly, while the raw HadCM3 indicates poor performance due to substantial drying, two RCMs forced by this GCM indicate inadequate system performance due to additional flooding. In either case, 5 out of the total 15 climate projections included in this analysis suggest that the dynamic management strategy may be insufficient as a long-term planning strategy that can provide robust performance.

5.6. Conclusion

This study presented a framework to identify long-term adaptation plans for a water resources system that are robust to a variety of non-stationary conditions and modeling uncertainties. Several sources of uncertainty were considered, including long-term changes in the underlying distribution of future conditions, sampling variability in realizations of those transient distributions, and uncertainty inherent to transfer functions. The integrated uncertainty analysis was coupled with systems analysis to identify long-term planning alternatives that are robust despite the uncertainties. A new robustness metric, the *R-Score,* was presented for this purpose.

The case study focused on dynamic reservoir management as an adaptation to long-term climate and societal changes. Results suggested that *if* the path of long-term change were known, an adaptive management strategy could effectively manage many nonstationary futures. However, without foreknowledge about the future, the particular dynamic management strategy employed in this study was unable to see past hydroclimatic noise and effectively adapt to emerging trends. Additional forecast skill may be required to make adaptive management a more viable adaption approach; decadal hydroclimatic forecasts provide a promising path of future research towards this end [Meehl et al., 2013].

The framework presented in this work utilized Monte Carlo methods to explore a variety of uncertainties. The approach is computationally expensive, even after employing methods to reduce the computational burden. While advanced computing and large Monte Carlo ensembles are becoming more popular for integrated risk assessments of water resource systems [Matrosov et al., 2013; Kasprzyk et al., 2013], many local water utilities do not possess the resources or expertise to employ such methods. What may be required are new methods that can effectively approximate and propagate the uncertainty at each stage of the modeling chain with greater efficiency to reduce the computational burden. This is an important avenue of future work to emerge from this research.

CHAPTER 6

CONCLUSION

The primary goal of this dissertation was to present methods that can better estimate the precision associated with future projections of water resource system performance under climate change, and through this provide information that can guide the development of adaptation strategies that are robust to these uncertainties. A series of methods were presented to quantify the interactions, propagation, and relative contributions of different sources of uncertainty in a water resources impacts assessment under climate change. Several sources of uncertainty were considered throughout the work, including uncertainty in long-term climate changes, internal climate variability, and the hydrologic modeling process. A stochastic weather generator was presented to capture uncertainty in long-term climate changes and internal climate variability, while hydrologic modeling uncertainty was addressed using Bayesian methods.

The uncertainties from mean climate changes, internal climate variability, and hydrologic modeling errors were integrated in climate change analyses of a New England river basin, a flood control facility in Iowa, and a multi-purpose surface reservoir system in Texas. The results of these case studies indicated that the precision of impact analyses under climate change significantly declines when multiple sources of uncertainty are accounted for and propagated through the modeling exercise. This was true for hydrologic statistics of interest (Chapter 3), as well as for performance metrics associated with water resource

systems (Chapters 4). However, despite the reduction in precision, it is still possible to provide useful planning guidance with the modeling chains often used for climate change assessments. Chapter 5 presented an example of how such guidance could be developed under an integrated uncertainty assessment.

Several future research needs emerge from the work presented in this dissertation. First, the uncertainty framework adopted for all case studies utilized Monte Carlo methods to explore the uncertainty space. This approach, while effective, is computationally expensive, particularly when using complex hydrologic or water resource system models. If the uncertainty space for any one of the case studies were to be expanded, the computational burden would have required more advanced parallel processing techniques. State-of-the-art computing is indeed a valid path to follow in order to improve the treatment of uncertainty in climate change impact studies. However, utilities, state agencies, and other stakeholder groups often lack access or the expertise necessary to adopt such methods. Therefore, future research needs to explore alternative methods to effectively approximate and propagate the uncertainty at each stage of the modeling chain with greater efficiency to reduce the computational burden.

Secondly, future work is needed to develop an appropriate statistical characterization of hydrologic model uncertainty, as this is not a straightforward process. The results in Chapter 4 highlighted that the structural differences between a simple and complex hydrologic model could extend beyond the predictive bounds associated with the simple

model. One potential approach to resolve this issue is to apply a similar Bayesian uncertainty analysis to more complex, distributed hydrologic structures, but the computational burden would be great. Again, parallel processing methods would be required. Alternatively, the statistical uncertainty analysis applied to simpler hydrologic models may need to be improved and thoroughly tested under different climate regimes to ensure they still function appropriately in a climate change context.

Finally, substantial work is still needed to integrate climate change projections with the uncertainty methods presented in this work. It remains unclear how best to use highly uncertain and difficult-to-verify climate models for long-term water system planning. The work presented in this dissertation proposed methods to account for and propagate future climate change uncertainty into impact analyses. In Chapters 4 and 5, climate projections were superimposed on the uncertainty analyses to put the projections in context. However, the value of this approach for decision-making is still unclear. It remains an open question how projections should be coupled with the methods proposed in this work to best provide decision support for long-term water resources planning. This provides an exciting and important avenue for future research.

APPENDIX

The wavelet transform utilizes a generalized local base function, or wavelet $\Psi(t)$, that can be dilated and translated into a set of elementary functions, $\Psi(\frac{t-b}{a})$, for use in exploring different signals in the original data across various frequencies and time scales. Mathematically, a signal of a particular frequency (or wavelet scale, *a*) is extracted from the data at a localized time (*b*) via the continuous convolution

$$W(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \tilde{x}(t) \times \Psi^*(\frac{t-b}{a}) \times dt \qquad \text{(A1)}$$

where $W(a,b)$ is a wavelet spectrum and Ψ^* indicates the complex conjugate of the wavelet function. In this study, we employ the Morlet wavelet, given by $\Psi(t) = \pi^{-1/4} \times \exp(i \times \omega_0 \times t) \times \exp(-t^2/2)$, where $\omega_0 = 6$ is a nondimensional frequency [Torrence and Campo, 1998] and $i = \sqrt{-1}$. Since climate data are only available at discrete time steps, the continuous convolution in equation A1 is often estimated using a discrete, *N* times convolution, where *N* equals the number of annual observations.

$$W_n(a) = \sum_{n'=0}^{N-1} \tilde{x}_{n'} \times \Psi^*(\frac{(n'-n)\partial t}{a}) \qquad \text{(A2)}$$

Here, ∂t equals the time step. Computationally, it is more convenient to carry out the N times convolution in equation A2 in Fourier space. The Fourier transform of the original annual data is given by

$$\hat{x}_g = \frac{1}{N} \sum_{n=0}^{N-1} \tilde{x}_n \times \exp(\frac{-2 \times \pi \times i \times g \times n}{N}) \qquad \text{(A3)}$$

Here, g=0,1,…,N-1 indexes the frequencies. Given that the Fourier transform of the wavelet function $\Psi(\frac{t}{a})$ is given by $\hat{\Psi}(a\omega)$ in its continuous limit, the convolution theorem can be employed to calculate the wavelet transform of the data as the inverse Fourier transform of the product:

$$W_n(a) = \sum_{g=0}^{N-1} \hat{x}_g \times \hat{\Psi}^*(a\omega_g) \times \exp(i \times \omega_g \times n \times \partial t) \qquad \text{(A4)}$$

Once the wavelet transform has been calculated over a set of scales for each time step, low-frequency component series, $\mathbf{z}_h$, can be reconstructed by summing the real part of the wavelet transform over a subset of scales associated with that component:

$$z_{h,n} = \frac{\partial_o \partial t^{1/2}}{C_\partial \Psi_0(0)} \times \sum_{o \in O_h} \frac{\Re\{W_n(\mathsf{a}_o)\}}{\mathsf{a}_o^{1/2}} \tag{A5}$$

where ∂_o is the scale increment, C_∂ is a reconstruction factor, and $\Psi_0(0)$ is a factor that removes the energy scaling. The term O_h indexes all of the scales associated with the h^{th} low-frequency component. The statistical significance of wavelet power spectra can be tested at all of the scale values by comparing the power spectra against that of a white-noise or red-noise process. Scales at which the power spectra appear significant can be bundled together into groups, O_h. The modeler should use judgment in the development of these groups, as they should reflect known structure in the climate system for that location (i.e. significant wavelet scales associated with periods between 3-7 years could be grouped and associated with ENSO, while significant wavelet scales associated with periods on the order of 10-20 years could be grouped and associated with the PDO).

BIBLIOGRAPHY

Ajami, N. K., G. M. Hornberger, and D. L. Sunding (2008), Sustainable water resource management under hydrological uncertainty, Water Resour. Res., 44, W11406, doi:10.1029/2007WR006736.

Ajami, N. K., Q. Duan, and S. Sorooshian (2007), An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction, *Water Resour. Res., 43, W01403, doi:10.1029/2005WR004745.*

Akaike, H. (1974), A new look at the statistical model identification, *IEEE Transactions on Automatic Control, 19,* 716-723.

Alemu, E.T., R.N. Palmer, A. Polebitski, and B. Meaker (2011), Decision support system for optimizing reservoir operations using ensemble streamflow predictions, *Journal of Water Resources Planning and Management, 137,* 72-82.

Allen, R., L. Pereira, D. Raes, and M. Smith (1998), Crop evapotranspiration guidelines for computing crop water requirements, *FAO Irrig. Drain. Pap. 56,* Food and Agric. Organ. Of the U.N., Rome.

Alley, W.M. (1984), On the treatment of evapotranspiration, soil moisture accounting, and aquifer recharge in monthly water balance models, *Water Resources Research,20,* 1137-1149.

Apipattanavis, S., G. Podesta, B. Rajagopalan, and R. W. Katz (2007), A semiparametric multivariate and multisite weather generator, *Water Resour. Res., 43*, W11401, doi:10.1029/2006WR005714.

Arnell, N.W. (1999), The effect of climate change on hydrologic regimes in Europe: a continental perspective, *Global Environmental Change, 9,* 5-23.

Barsugli, J., C. Anderson, J.B. Smith, and J.M. Vogel, 2009. Options for Improving Climate Modeling to Assist Water Utility Planning for Climate Change. Published by the Water Utility Climate Alliance, 146 pp. http://www.wucaonline.org/assets/pdf/ actions_whitepaper_120909.pdf, accessed November 2013.

Bates, B.C., and E.P. Campbell (2001), A Markov chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall-runoff modeling, *Water Resources Research*, *37 (4)*, 937-947.

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

Ben-Haim, Y. (2006). Info-gap decision theory: Decisions under severe uncertainty (2nd ed.). Oxford: Academic Press.

Bergström, S. (1975), The development of a snow routine for the HBV-2 model, *Nord. Hydrol., 6(2)*, 73–92.

Beven, K. J., and J. Freer (2001), Equifinality, data assimilation, and uncertainty estimation in mechanistic modeling of complex environmental systems, *J. Hydrol., 249*, 11–29.

Beven, K.J., 2006. A manifesto for the equifinality thesis. *Journal of Hydrology 320(1)*, 18-36.

Beven, K.J., and A.M. Binley (1992), The future of distributed hydrologic models: Model calibration and uncertainty prediction, *Hydrol. Process., 6,* 279-298.

Boorman, D.B., and C.E.M. Sefton (1997), Recognizing the uncertainty in the quantification of the effects of climate change on hydrologic response, *Climatic Change, 35,* 415-434.

Box, G.E.P., and G. Jenkins (1970), *Time Series Analysis, Forecasting, and Control*, Holden-Day, Boca Raton, Fla.

Boyle, D., H. Gupta, and S. Sorroshian (2000), Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods, *Water Resour. Res., 36*, 3663-3674.

Brath, A., A. Montanari, and G. Moretti (2006), Assessing the effect on flood frequency of land use change via hydrological simulation (with uncertainty), *Journal of Hydrology, 324*(1), 141-153.

Brazos G Regional Water Planning Group (BGRWPG). The 2011 Brazos G Regional Water Plan, September 2010.

Brekke, L. D., E. P. Maurer, J. D. Anderson, M. D. Dettinger, E. S. Townsley, A. Harrison, and T. Pruitt (2009), Assessing reservoir operations risk under climate change, *Water Resour. Res., 45(4)*, W04411, doi:10.1029/2008WR006941.

Brown, C., and R.L. Wilby (2012), An alternate approach to assessing climate risks, *Eos Trans. AGU, 93(41),* 401-402, doi: 10.1029/2012EO410001.

Brown, C., W. Werick, W. Leger, and D. Fay (2011), A decision-analytic approach to managing climate risks: Application to the Upper Great Lakes, *Journal of the American Water Resources Association, 47 (3),* 524-534, doi: 10.1111/j.1752-1688.2011.00552.x.

Brown, C., Y. Ghile, M. Laverty, and K. Li (2012), Decision scaling: Linking bottom-up vulnerability analysis with climate projections in the water sector, *Water Resour.Res.,48*, W09537, doi:10.1029/2011WR011212.

Buishand, T.A., and T. Brandsma (2001), Multisite simulation of daily precipitation and temperature in the Rhine basin by nearest-neighbor resampling, *Water Resources Research, 37 (11)*, 2761-2776.

Cameron, D. (2006), An application of the UKCIP02 climate change scenarios to flood estimation by continuous simulation for a gauged catchment in the northeast of Scotland, OK (with uncertainty), *Journal of Hydrology, 328,* 212-226.

Cameron, D., K. Beven, and P. Naden (2000), Flood frequency estimation by continuous simulation under climate change (with uncertainty), *Hydrology and Earth System Sciences, 4(3),* 393-405.

Cameron, D., K. Beven, and P. Naden (2001), Flood frequency estimation under climate change (with uncertainty), *Hydrol. Earth Syst. Sci., 4(3)*, 393 – 405.

Chao, P. (1999), Great Lakes water resources: climate change impact analysis with transient GCM scenarios, *Journal of the American Water Resources Association, 35 (6)*, 1499-1507.

Chen, J., F. P. Brissette, A. Poulin, and R. Leconte (2011), Overall uncertainty study of the hydrological impacts of climate change for a Canadian watershed, *Water Resour. Res., 47,* W12509, doi:10.1029/2011WR010602.

Chen, J., F.P. Brissette, and R. Leconte (2010), A daily stochastic weather generator for preserving low-frequency of climate variability, *Journal of Hydrology, 388,* 480-490.

Christensen, N.S., A.W. Wood, N. Voisin, D.P. Lettenmaier, and R.N. Palmer (2004), The effects of climate change on the hydrology and water resources of the Colorado River Basin, *Climate Change, 62,* 337-363.

Clark, M.P., and J.A. Vrugt (2006), Unraveling uncertainties in hydrologic model calibration: Addressing the problem of compensatory parameters, *Geophysical Research Letters*, *33*, L06406, doi: 10.1029/2005GL025604.

Coleman, J.S.M., and D. Budikova (2010), Atmospheric aspects of the 2008 Midwest floods: a repeat of 1993?, *International Journal of Climatology, 30 (11),* 1645-1667.

Collins, M. (2000), Understanding uncertainties in the response of ENSO to greenhouse warming, Geophys. Res. Lett., 27(21), 3509–3512, doi:10.1029/2000GL011747.

Confalonieri, R. (2012), Combining a weather generator and a standard sensitivity analysis method to quantify the relevance of weather variables on agrometeorological model outputs, *Theo. Appl. Climatol., 108,*19-30, doi: 10.1007/s00704-011-0510-0.

Dai, A. (2006), Precipitation characteristics in eighteen coupled climate models, *J. Clim., 19*(18), 4605-4630.

Deser, C., A. Phillips, M.A. Alexander, and B.V. Smoliak (2013), Projecting North American climate over the next 50 years: uncertainty due to internal variability, *Journal of Climate,* doi: http://dx.doi.org/10.1175/JCLI-D-13-00451.1.

Deser, C., A. Phillips, V. Bourdette, and H. Teng (2012), Uncertainty in climate change projections: the role of internal variability, *Clim. Dyn., 38*(3-4), 527-546.

Dessai, S., and M. Hulme (2007), Assessing the robustness of adaptation decisions to climate change uncertainties: A case study on water resources management in the East of England, *Global Environmental Change, 17 (1),* 59-72.

Duan, Q, N.K. Ajami, X. Gao, and S. Sorooshian (2007), Multi-model ensemble hydrologic prediction using Bayesian model averaging, *Advances in Water Resources, 30 (5),* 1371-1386.

Duan, Q., S. Sorooshian, and V. Gupta (1992), Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resources Research 28(4)*, 1015-1031.

Dubrovsky, M., J. Buchtele, and Z. Zalud (2004), High-frequency and low-frequency variability in stochastic daily weather generator and its effect on agricultural and hydrologic modeling, *Climatic Change, 63,* 145-179.

Dubrovsky, M., Z. Zalud, and M. Stastna (2000), Sensitivity of ceres-maize yields to statistical structure of daily weather variables, *Climatic Change, 46,* 447-472.

Evin, G., D. Kavetski, M. Thyer, and G. Kuczera (2013), Pitfalls and improvements in the joint inference of heteroscedasticity and autocorrelation in hydrological model calibration, *Water Resour. Res., 49,* 4518–4524, doi:10.1002/wrcr.20284.

Faber, B.A. and J.R. Stedinger (2001), Reservoir optimization using sampling SDP with ensemble streamflow prediction (EPS) forecasts, *Journal of Hydrology, 249,* 113-133.

Fatichi, S., V.Y. Ivanov, and E. Caporali (2011), Simulation of future climate scenarios with a weather generator, *Advances in Water Resources, 34 (4)*, 448-467.

Fatichi, S., V.Y. Ivanov, and E. Caporali (2013), Assessment of a stochastic downscaling methodology in generating an ensemble of hourly climate time series, *Climate Dynamics, 40 (7-8)*, 1841-1861.

Fernandez, W., R. M. Vogel, and A. Sankarasubramanian (2000), Regional calibration of a watershed model, *Hydrol. Sci. J., 45(5),* 689–707.

Fowler, H., S. Blenkinsop, and C. Tebaldi (2007), Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modeling, *Int. J. Climatol., 27,* 1547–1578.

Fullerton, T. M., Jr., and A. L. Molina Jr. (2010), Municipal water consumption forecast accuracy, *Water Resour.Res., 46*, W06515, doi:10.1029/2009WR008450.

Gao, H., Q. Tang, X. Shi, C. Zhu, T. J. Bohn, F. Su, J. Sheffield, M. Pan, D. P. Lettenmaier, and E. F. Wood, 2010: Water Budget Record from Variable Infiltration Capacity (VIC) Model. In *Algorithm Theoretical Basis Document for Terrestrial Water Cycle Data Records* (in review).

Gelman, A., and D. B. Rubin (1992), Inference from iterative simulation using multiple sequences, *Stat. Sci., 7(4)*, 457-511.

Georgakakos, A. P., H. Yao, M. G. Mullusky, and K. P. Georgakakos (1998), Impacts of climate variability on the operational forecast and management of the Upper Des Moines River Basin, *Water Resour. Res., 34(4),* 799–821, doi:10.1029/97WR03135.

Georgakakos, A.P., H. Yao, M. Kistenmacher, K.P. Georgakakos, N.E. Graham, F.Y. Chen, C. Spencer, and E. Shamir (2012), Value of adaptive water resources management in Northern California under climatic variability and change: Reservoir management, *J. Hydrol, 412–413*, 34–46.

Gleick, P. (1986), Methods for evaluating the regional hydrologic impacts of global climatic changes, *J. Hydrol., 88*, 97–116.

Gober, P., C.W. Kirkwood, R.C. Balling Jr., A.W. Ellis, and S. Deitrick (2010), Water planning under climatic uncertainty in Phoenix: Why we need a new paradigm, *Annals of the Association of American Geographers, 100 (2),* 356-372.

Griffin, R.C. and Chang, C. (1990), Pretest analysis of water demand in thirty communities, *Water Resources Research, 26(10)*, 2251-2255.

Groves, D. G., D. Yates, and C. Tebaldi (2008), Developing and applying uncertain global climate change projections for regional water management planning, Water Resour. Res., 44, W12413, doi:10.1029/2008WR006964

Gupta, H.V., H. Kling, K.K. Yilmaz, and G.F. Martinez (2009), Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrologic modelling, *Journal of Hydrology, 377 (1-2)*, 80-91.

Hamlet, A. F., and D. P. Lettenmaier (1999), Effects of climate change on hydrology and water resources in the Columbia river basin, *Journal of the American Water Resources Association, 35 (6)*, 1597-1623.

Hamon, W.R. (1963), Computation of direct runoff amounts from storm rainfall, *Int. Assoc. Sci. Hydrol. Publ., 63,*52-62.

Hansen, J.W., and T. Mavromatis (2001), Correcting low-frequency variability bias in stochastic weather generators, *Agricultural and Forest Meteorology, 109,* 297-310.

Hansen, M., R. DeFries, J.R.G. Townshend, and R. Sohlberg (1998), UMD Global Land Cover Classification, 1 Kilometer, 1.0, Department of Geography, University of Maryland, College Park, Maryland, 1981-1994.

Hargreaves, G.H., and Z.A. Samani (1982), Estimating potential evapotranspiration, *J. Irrig. Drain. Engng, 108 (3)*, 225-230.

Hirsch, R. M. (2011), A Perspective on Nonstationarity and Water Management1, *JAWRA Journal of the American Water Resources Association, 47*(3), 436-446.

Houck, C. R., Joines, J. A., & Kay, M. G. (1995), A genetic algorithm for function optimization: a Matlab implementation, NCSU-IE TR, 95 (09).

Huard, D., and A. Mailhot (2008), Calibration of hydrologic model GR2M using Bayesian uncertainty analysis, *Water Resour. Res., 44,* W02424, doi:10.1029/2007WR005949.

Huntington, T.G., G.A. Hodgkins, B.D. Keim, and R.W. Dudley (2004), Changes in the proportion of precipitation occurring as snow in New England (1949-2000), *J. Climate, 17,* 2626-2636.

Hurst, H.E. (1951), Long term storage capacities of reservoirs, *Trans. Am. Soc. Civil Engrs, 116,* 776-808.

Hutchinson, M.F. (1995), Stochastic space-time models from ground-based data, *Agricultural and Forest Meteorology, 73,* 237-264.

Intergovernmental Panel on Climate Change (IPCC) (2007), *Climate Change 2007: The Physical Basis, Contributions of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change,* edited by S. Solomon et al., Cambridge Univ. Press, Cambirdge, U.K.

Jenicek, E.M., Carroll, R.A., Curvey, L.E., Hessel, M.S., Holmes, R.M., and Pearson, E., 2011. Water sustainability assessment for ten Army installations, Report TR-11-5, Engineering Research and Development Center (ERDC) / Construction Engineering Research Laboratory (CERL).

Jiang, T., Y.D. Chen, C. Xu, X. Chen, X. Chen, and V.P. Singh (2007), Comparison of hydrologic impacts on climate change simulated by six hydrologic models in the Dongjiang Basin, South China, *Journal of Hydrology*, *336*, 316-333.

Johnson, T.E., and C.P. Weaver (2009), A framework for assessing climate change impacts on water and watershed systems, *Environmental Management, 43,* 118-134.

Jones, R.N. (2000), Analyzing the risk of climate change using an irrigation demand model, *Climate Research, 14,* 89-100.

Jones, R.N. (2001), An environmental risk assessment/management framework for climate change impact assessments, *Natural Hazards, 23,* 197-230.

Karlsson, M., and S. Yakowitz (1987), Nearest-neighbor methods for nonparametric rainfall-runoff forecasting, *Water Resour. Res., 23(7),* 1300–1308, doi:10.1029/WR023i007p01300.

Kasprzyk, J.R., S. Nataraj, P.M. Reed, and R.J. Lempert (2013), Many objective robust decision making for complex environmental systems undergoing change, *Environmental Modelling and Software, 42,* 55-71.

Kavetski, D., G. Kuczera, and S. W. Franks (2006a), Bayesian analysis of input uncertainty in hydrologic modeling: 2. Application, *Water Resour. Res., 42*, W03408, doi:10.1029/2005WR004376.

Kavetski, D., G. Kuczera, and S. W. Franks (2006b), Bayesian analysis of input uncertainty in hydrologic modeling: 1. Theory, *Water Resour. Res., 42*, W03407, doi:10.1029/2005WR004368.

Kay, A. L., H. N. Davies, V. A. Bell, and R. G. Jones (2009), Comparison of uncertainty sources for climate change impacts: Flood frequency in England, *Clim. Change, 92*, 41–63.

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

Khan, M.S., and P. Coulibaly (2010), Assessing hydrologic impact of climate change with uncertainty estimates: Bayesian neural network approach, *J. Hydrometeorol., 11,* 482-495.

Kilsby, C.G., P.D. Jones, A. Burton, A.C. Ford, H.J. Fowler, C. Harpham, P. James, A. Smith, and R.L. Wilby (2007), A daily weather generator for use in climate change studies, *Environmental Modelling and Software, 22,* 1705-1719.

Kim, Y., R.W. Katz, B. Rajagopalan, G.P. Podesta, and E.M. Furrer (2011), Reducing overdispersion in stochastic weather generators using a generalized linear modeling approach, *Climate Research, 53*, 13-24.

Klemes, V. (1986), Operational testing of hydrological simulation models, *Hydrol. Sci. J., 31*, 13– 24.

Knutson, T. R., J. L. McBride, J. Chan, K. Emanuel, G. Holland, C. Landsea, I. Held, J. P. Kossin, A. Srivastava, and M. Sugi (2010), Tropical cyclones and climate change, *Nature Geoscience, 3*(3), 157-163.

Kollat, J. B., P. M. Reed, and T. Wagener (2012), When are multiobjective calibration trade-offs in hydrologic models meaningful?, *Water Resour. Res., 48*, W03520, doi:10.1029/2011WR011534.

Koutsoyiannis, D. (2003), Climate change, the Hurst phenomenon, and hydrologic statistics, *Hydrologic Sciences Journal, 48 (1)*, 3-24.

Kuczera, G. (1983), Improved parameter inference in catchment models, 1. Evaluating parameter uncertainty, *Water Resources Research, 19(5),* 1151-1162, doi:10.1029/WR019i005p01151.

Kuczera, G., D. Kavetski, S. Franks, and M. Thyer (2006), Towards a Bayesian total error analysis of conceptual rainfall-runoff models: Characterizing model error using storm-dependent parameters, *Journal of Hydrology, 331(1-2),* 161-177, doi: 10.1016/j.hydrol.2006.05.010.

Kunkel, K. E., S.A. Changnon, and J.R. Angel, (1994), Climatic Aspects of the 1993 Upper Mississippi River Basin Flood. *Bull. Amer. Meteor. Soc.*, *75*, 811–822.

Kwon, H.H., B. Sivakumar, Y.I. Moon, and B.S. Kim (2011), Assessment of change in design flood frequency under climate change using a multivariate downscaling model and a precipitation-runoff model, *Stoch. Environ. Res. Risk Assess., 25,* 567-581.

Kwon, H.-H., U. Lall, and A. F. Khalil (2007), Stochastic simulation model for nonstationary time series using an autoregressive wavelet decomposition: Applications to rainfall and temperature, *Water Resour. Res., 43*, W05407, doi:10.1029/2006WR005258.

Kwon, H.-H., U. Lall, and J. Obeysekera (2009), Simulation of daily rainfall scenarios with interannual and multidecadal climate cycles for South Florida, *Stochastic Environmental Research and Risk Assessment, 23,* 879-896.

Laio, F., and S. Tamea (2007), Verification tools for probabilistic forecasts of continuous hydrologic variables, *Hydrology and Earth System Sciences, 11,* 1267-1277.

Lall, U., and A. Sharma (1996), A nearest neighbor bootstrap for time series resampling, *Water Resources Research, 32 (3),* 679-693.

Lee, T., T.B.M.J. Ouarda, and C. Jeong (2012), Nonparametric multivariate weather generator and an extreme value theory for bandwidth selection, *Journal of Hydrology, 452-453,* 161-171.

Lempert, R. J., Groves, D. G., Popper, S. W., & Bankes, S. C. (2006). A general, analytic method for generating robust strategies and narrative scenarios. Management Science, 52(4), 514-528.

Lempert, R. J., M. E. Schlesinger, and S. C. Bankes (1996), When we Don't Know the Costs Or the Benefits: Adaptive Strategies for Abating Climate Change, *Climatic Change, 33 (2)*, 235-274.

Lempert, R.J. and D.G. Groves (2010), Identifying and Evaluating Robust Adaptive Policy Responses to Climate Change for Water Management Agencies in the American West, *Technological Forecasting and Social Change*, *77*, 960-974.

Lettenmaier, D. P., A.W. Wood, R.N. Palmer, E.F. Wood, and E.Z. Stakhiv (1999), Water Resources Implications of Global Warming: A U.S. Regional Perspective, *Climatic Change, 43 (3)*, 537-579.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges (1994), A Simple hydrologically Based Model of Land Surface Water and Energy Fluxes for GSMs, *J. Geophys. Res.*, *99(D7)*, 14415-14428.

Liang, X., K. E. Kunkel, and A. N. Samel (2001), Development of a regional climate model for US Midwest applications. Part I: Sensitivity to buffer zone treatment, *J. Clim., 14*(23), 4363-4378.

Lins, H. F. and J. R. Slack (1999), Streamflow trends in the United States, *Geophys. Res. Lett., 26*(2), 227-230.

Lopez, A., F. Fung, M. New, G. Watts, A. Weston, and R. L. Wilby (2009), From climate model ensembles to climate change impacts and adaptation: A case study of water resource management in the southwest of England, Water Resour. Res., 45, W08419, doi:10.1029/2008WR007499.

Marshall, L., D. Nott, and A. Sharma (2004), A comparative study of Markov chain Monte Carlo method for conceptual rainfall-runoff modeling, *Water Resources Research*, *40*, W02501, doi: 10.1029/2003WR002378.

Marshall, L., D. Nott, and A. Sharma (2007), Towards dynamic catchment modeling: a Bayesian hierarchical mixture of experts framework, *Hydrologic Processes*,*21*, 847-861.

Martinez, G. F., and H. V. Gupta (2010), Toward improved identification of hydrological models: A diagnostic evaluation of the "abcd" monthly water balance model for the conterminous United States, Water Resour. Res., 46, W08507, doi:10.1029/2009WR008294.

Matrosov, E.S., A.M. Woods, and J.J. Harou (2013), Robust decision making and info-gap decision theory for water resource system planning, *Journal of Hydrology, 494,* 43-58.

Maurer, E. P., L. Brekke, T. Pruitt, and P. B. Duffy (2007), Fine-resolution climate projections enhance regional climate change impact studies, *Eos Trans. AGU, 88* (47), 504.

Maurer, E.P., A.W. Wood, J.C. Adam, D.P. Lettenmaier, and B. Nijssen (2002), A Long-Term Hydrologically-Based Data Set of Land Surface Fluxes and States for the Conterminous United States, *J. Climate 15*, 3237-3251.

Mearns, L. O., W. J. Gutowski, R. Jones, L.-Y. Leung, S. McGinnis, A. M. B. Nunes, and Y. Qian (2009), A regional climate change assessment program for North America, *EOS, 90 (36)*, 311-312.

Mearns, L.O., C. Rosenzweig, and R. Goldberg (1996), The effect of changes in daily and interannual climatic variability on ceres-wheat: a sensitivity study, *Climatic Change, 32,* 257-292.

Meehl et al. (2013), Decadal climate prediction: An update from the trenches, *Bulletin of the American Meteorological Society,* doi: 10.1175/BAMS-D-12-00241.1.

Milly, C., et al. (2008), Stationary is dead: Whither water management, *Science, 318*, 573–574.

Moody, P., and C. Brown (2013), Robustness indicators for evaluation under climate change: Application to the upper Great Lakes, *Water Resour. Res., 49,* 3576–3588, doi:10.1002/wrcr.20228.

Muleta, M.K., J. McMillan, G.G. Amenu, and S.J. Burian (2013), Bayesian approach for uncertainty analysis of an urban storm water model and its application to a heavily urbanized watershed, *Journal of Hydrologic Engineering, 18,* 1360-1371.

Muller, C.J., P.A. O'Gorman, and L.E. Back (2011), Intensification of precipitation extremes with warming in a cloud-resolving model, J. Climate, 24, 2784-2800.

Nakamura, J., U. Lall, Y. Kushnir, A. W. Robertson, and R. Seager (2013), Dynamical structure of extreme floods in the U.S. Midwest and the UK. *J. Hydrometeor.,* DOI: 10.1175/JHM-D-12-059.1.

Nijssen, B., G.M. O'Donnell, A.F. Hamlet, and D.P Lettenmaier (2001), Hydrologic Sensitivity of Global Rivers to Climate Change, *Climatic Change, 50 (1)*, 143-175.

Osborn, C. T., Schefter, J. E. and Shabman, L. (1986), The accuracy of water use forecasts: Evaluation and implications. *Journal of the American Water Resources Association, 22*, 101–109. doi: 10.1111/j.1752-1688.1986.tb01865.x

Palmer, T.N., and J. Raisanen (2002), Quantifying the risk of extreme seasonal precipitation events in a changing climate. *Nature, 415*, 512-517.

Piani, C., D.J. Frame, D.A. Stainforth, and M.R. Allen (2005), Constraints on climate change from a multi-thousand member ensemble of simulations. *Geophys. Res. Lett. 32*, L23825.

Piechota, T.C., and J.A. Dracup (1996), Drought and regional hydrologic variation in the United States: Associations with the El Niño-Southern Oscillation, *Water Resources Research, 32(5),* 1359-1373.

Plummer, M. (2011), rjags: Bayesian graphical models using MCMC. R

Prudhomme, C., and H. Davies (2009a), Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK. Part 1: baseline climate, *Climatic Change, 93,* 177-195.

Prudhomme, C., and H. Davies (2009b), Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK. Part 2: future climate, *Climatic Change, 93,* 197-222.

Prudhomme, C., R. L.Wilby, S. Crooks, A. L. Kay, andN. S. Reynard (2010), Scenario-neutral approach to climate change impact studies: Application to flood risk, *J. Hydrol., 390*,198–209, doi:10.1016/j.jhydrol.2010.06.043.

Räisänen, J., and T. N. Palmer (2001), A Probability and Decision-Model Analysis of a Multimodel Ensemble of Climate Change Simulations. *J. Climate*, *14*, 3212–3226.

Rajagopalan, B., and U. Lall (1999), A k-nearest-neighbor simulator for daily precipitation and other weather variables, *Water Resources Research, 35 (10),* 3089-3101.

Reclamation, (2013), Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections: Release of Downscaled CMIP5 Climate Projections, Comparison with preceding Information, and Summary of User Needs, prepared by the U.S. Department of the Interior, Bureau of Reclamation, Technical Services Center, Denver, Colorado. 47pp.

Reichert, P., and J. Mieleitner (2009), Analyzing input and structural uncertainty of nonlinear dynamic models with stochastic, time-dependent parameters, *Water Resources Research, 45,* W10402, doi: 10.1029/2009WR007814.

Renard, B., D. Kavetski, G. Kuczera, M. Thyer, and S. W. Franks (2010), Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors, *Water Resour. Res., 46*, W05521, doi:10.1029/2009WR008328.

Richardson, C.W. (1985), Weather simulation for crop management models, *Transactions of the American Society of Agricultural Engineers, 28,* 1602-1606.

Riha, S.J., D.S. Wilks, and P. Simoens (1996), Impact of temperature and precipitation variability on crop model predictions, *Climatic Change, 32,* 293-311.

Rinderknecht, S.L., M.E. Borsuk, and P. Reichert (2012), Bridging uncertain and ambiguous knowledge with imprecise probabilities. Environmental Modelling and Software 36, 122-130.

Romps, D.M. (2011), Response of tropical precipitation to global warming, J. Atmos. Sci., 68, 123-138.

Rougier, J., D. Sexton, J.M. Murphy, and D. Stainforth (2009), Analyzing the Climate Sensitivity of the HadSM3 Climate Model Using Ensembles from Different but Related Experiments. *Journal of climate*, *22(13)*, 3540–3557.

Schoups, G., and J. A. Vrugt (2010), A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors, *Water Resour. Res., 46*, W10531, doi:10.1029/2009WR008933.

Semenov, M.A., and J.R. Porter (1995), Climatic variability and the modeling of crop yields, *Agricultural and Forest Meteorology, 73 (3-4)*, 265-283.

Serfling, R. (2002), Quantile functions for multivariate analysis: approaches and applications. *Statistica Neerlandica, 56*, 214–232. doi: 10.1111/1467-9574.00195.

Sexton, D., J.M. Murphy, M. Collins, and M.J. Webb (2012), Multivariate probabilistic projections using imperfect climate models part I: outline of methodology, *Climate Dynamics*, *38 (11-12)*, 2513-2542.

Sheridan, S.C., and C.C. Lee (2010), Synoptic climatology and the general circulation model, *Progress in Physical Geography, 34 (1),* 101-109.

Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. M. B. Tignor, and H. L. Miller, Eds., (2007), Climate Change 2007: The Physical Science Basis. Cambridge University Press, Cambridge, UK.

Srikanthan, R., and G.G.S. Pegram (2009), A nested multisite daily rainfall stochastic generation model, *Journal of Hydrology, 371,* 142-153.

Srikanthan, R. and McMahon, T. A., (2001), Stochastic generation of annual, monthly and daily climate data: A review. Hydrol. Earth Syst. Sci., 5, 653-670, doi:10.5194/hess-5-653-2001.

Stainforth, D. A., M. R. Allen, E. Tredger, and L. A. Smith (2007a), Confidence, uncertainty and decision-support relevance in climate predictions, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 365*(1857), 2145-2161.

Stainforth, D. A., T. Aina, C. Christensen, M. Collins, N. Faull, D. Frame, J. Kettleborough, S. Knight, A. Martin, and J. Murphy (2005), Uncertainty in predictions of the climate response to rising levels of greenhouse gases, *Nature, 433*(7024), 403-406.

Stainforth, D. A., T. E. Downing, R. Washington, A. Lopez, and M. New (2007b), Issues in the interpretation of climate model ensembles to inform decisions, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 365*(1857), 2163-2177.

Stedinger, J. R., R. M. Vogel, S. U. Lee, and R. Batchelder (2008), Appraisal of the generalized likelihood uncertainty estimation (GLUE) method, *Water Resour. Res., 44*, W00B06, doi:10.1029/2008WR006822.

Stedinger. J.R. (1983), Confidence intervals for design events, *Journal of Hydraulic Engineering, 109 (1),* 13-27.

Steinschneider, S., A. Polebitski, C. Brown, and B. H.Letcher (2012), Toward a statistical framework to quantify the uncertainties of hydrologic response under climate change, *Water Resour. Res., 48,* W11525, doi:10.1029/2011WR011318.

Steinschneider, S., and C. Brown (2012), Dynamic reservoir management with real-option risk hedging as a robust adaptation to nonstationary climate, *Water Resour. Res., 48*, W05524, doi:10.1029/2011WR011540.

Steinschneider, S., and C. Brown (2013), A semiparametric multivariate, multisite weather generator with low-frequency variability for use in climate risk assessments, *Water Resour. Res., 49,* 7205–7220, doi:10.1002/wrcr.20528.

Tebaldi, C., R. L. Smith, D. Nychka, and L. O. Mearns (2005), Quantifying Uncertainty in Projections of Regional Climate Change: A Bayesian Approach to the Analysis of Multimodel Ensembles, *J. Climate*, *18*, 1524–1540.

Thiemann, M., M. Trosset, H. Gupta, and S. Sorooshian (2001), Bayesian recursive parameter estimation for hydrologic models, *Water Resources Research*, *37 (10)*, 2521-2535.

Thomas, H.A. (1981), Improved methods for national water assessment. Report, Contract WR 15249270, US Water Resources Council, Washington, DC, USA.

Thyer, M., B. Renard, D. Kavetski, G. Kuczera, S.W. Franks, and S. Srikanthan (2009), Critical evaluation of parameter consistency and predictive uncertainty in hydrologic modeling: A case study using Bayesian total error analysis, *Water Resources Research*, *45*, W00B14, doi: 10.1029/2008WR006825

Timmermann, A., J. Oberhuber, A. Bacher, M. Esch, M. Latif, and E. Roeckner (1999), Increased El Nino frequency in a climate model forced by future greenhouse warming, *Nature, 398,* 694-697.

Todini, E. (2004), Role and treatment of uncertainty in real-time flood forecasting, *Hydrol. Process., 18*(14), 2743-2746.

Torrence, C., and G.P. Compo (1998), A practical guide to wavelet analysis, *Bull. Amer. Meteor. Soc, 79,* 61-78.

U.S. Army Corps of Engineers (2006), Report: Risk Analysis for Flood Damage Reduction Studies - ER 1105-2-101, Department of the Army.

Uhlenbrook, S., J. Seibert, C. Leibundgut, and A. Rodhe (1999), Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure, *Hydrological Sciences Journal, 44*(5), 779-797.

USDA-NRCS, 2000. Soil Survey Geographic (SSURGO) database for Johnson County, Iowa, http://www.nrcs.usda.gov/products/datasets/sssurgo/.

Vandewiele, G. L., Xu, C.- Y. & Ni-Lar-Win (1992) Methodology and comparative study of monthly water balance models in Belgium, China and Bunna. J. Hydrol. 134, 315-347.

Villarini, G., F. Serinaldi, J. A. Smith, and W. F. Krajewski (2009), On the stationarity of annual flood peaks in the continental United States during the 20th century, *Water Resour. Res., 45*(8), W08417.

Vogel, R.M., and A. Sankarasubramanian (2003), Validation of a watershed model without calibration, *Water Resources Research, 39 (10)*, 1292, doi:10.1029/2002WR001940.

Vrugt, J. A., E. Laloy, and C.J.F. ter Braak (2011), Differential evolution adaptive Metropolis with sampling from past states, *SIAM J. Optimiz.*, in review.

Vrugt, J.A., C.G.H. Diks, H.V. Gupta, W. Bouten, and J.M. Verstraten (2005) Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation, *41*, W01017, doi: 10.1029/2004WR003059.

Wang, Q.J., and R.J. Nathan (2007), A method for coupling daily and monthly time scales in stochastic generation of rainfall series, *Journal of Hydrology, 346,* 122-130.

Wilby, R. L. (2005), Uncertainty in water resource model parameters used for climate change impact assessment, *Hydrol. Processes, 19*, 3201–3219.

Wilby, R. L., and I. Harris (2006), A framework for assessing uncertainties in climate change impacts: Low-flow scenarios for the River Thames, UK, *Water Resour. Res., 42, W02419, doi:10.1029/2005WR004065.*

Wilby, R.L., and S. Dessai (2010), Robust adaptation to climate change, *Weather, 65 (7)*, 180-185.

Wilby, R.L., O.J. Tomlinson, and C.W. Dawson (2003), Multi-site simulation of precipitation by conditional resampling, *Climate Research, 23,* 183-194.

Wiley, M.W., and R.N. Palmer (2008), Estimating the impacts and uncertainty of climate change on a municipal water supply system, *ASCE Journal of Water Resources Planning and Management, 134 (3),* 239-246.

Wilks, D.S. (1992), Adapting stochastic weather generation algorithms for climate change studies, *Climatic Change, 22,* 67-84.

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

Wilks, D.S. (1999), Simultaneous stochastic simulation of daily precipitation, temperature, and solar radiation at multiple sites in complex terrain, *Agricultural and Forest Meteorology, 96,* 85-101.

Wilks, D.S. (2002), Realizations of daily weather in forecast seasonal climate, *Journal of Hydrometeorology, 3,* 195-207.

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

Wood, A.W., D.P. Lettenmaier, and R.N. Palmer (1997), Assessing Climate Change Implications for Water Resources Planning, *Climatic Change*, *37*, 203-228.

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

Wurbs, R. A. (1987), Analysis of reservoir storage reallocations, *Journal of Hydrology, 92(1-2),* 77-95.

Xu, C., E. Widen, and S. Halldin (2005), Modelling hydrological consequences of climate change – progress and challenges, *Advances in Atmospheric Sciences, 22(6),* 789-797.

Yao, H., and A. Georgakakos (2001), Assessment of Folsom Lake response to historical and potential future climate scenarios 2. Reservoir Management, *Journal of Hydrology, 249,* 176-196, doi: 10.1016/S0022-1694(01)00418-8.