

UC Irvine

UC Irvine Electronic Theses and Dissertations

Title

Frameworks for Assessing Climate Change Impacts on Water Resources: From Uncertainty in Model Simulations to Accounting for Local Resilience

Permalink

<https://escholarship.org/uc/item/0z76c4cc>

Author

Mehran, Ali

Publication Date

2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Frameworks for Assessing Climate Change Impacts on Water Resources: From Uncertainty in
Model Simulations to Accounting for Local Resilience

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Civil Engineering

by

Ali Mehran

Dissertation Committee:
Assistant Professor Amir AghaKouchak, Chair
Professor Soroosh Sorooshian
Assistant Professor Jasper Vrugt
Professor Brett Sanders

2015

DEDICATION

To

My parents

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

TABLE OF CONTENTS

	Page
LIST OF FIGURES.....	v
LIST OF TABLES.....	viii
ACKNOWLEDGEMENTS.....	ix
CURRICULUM VITAE.....	x
ABSTRACT OF THE DISSERTATION.....	xii
CHAPTER 1: Introduction	1
1.1. Problem Statement.....	1
1.2. Objectives	4
1.3. Organization of the Dissertation	8
CHAPTER 2: Performance Metrics for Evaluation of Climate Model Simulations.....	9
2.1 Introduction.....	9
2.2 Methodology and Results	10
2.3 Conclusions.....	17
CHAPTER 3: Evaluation of CMIP5 Precipitation Simulations	20
3.1 Introduction.....	20
3.2 Datasets.....	23
3.3 Methodology and Results	24
3.4 Conclusions.....	37
CHAPTER 4: Improving Global Hydrological Modeling Using a Multi-Objective Calibration Framework 42	
4.1 Introduction.....	42
4.2 Data Resources.....	44
4.3 Methodology.....	45
4.4 Results.....	48
4.5 Summary and Conclusions	56
CHAPTER 5: Climate Change Impacts on Water Resources Accounting for Local Resilience 58	
5.1 Introduction.....	58
5.2 Data.....	64
5.3 Methods.....	66
5.4 Model Calibration.....	67
5.5 Conclusions.....	69
CHAPTER 6: A Hybrid Framework for Water Stress Assessment: Linking Climate Variability and Local Resilience.....	73

6.1 Introduction.....	73
6.2 Classical methods for performance assessment.....	77
6.3 Methodology.....	78
6.4 Results and discussion.....	83
6.5 Conclusions.....	90
CHAPTER 7: Summary and Conclusions.....	93
REFERENCES.....	99

LIST OF FIGURES

	Page
Figure 1.1 Climate model simulations are used to force global or regional hydrologic models to understand the effect of climatic change and variability on hydrologic response.....	2
Figure 2.1 The Contingency Table	12
Figure 2.2 Two synthetic precipitation simulations (SIM_1 and SIM_2) and reference observation (OBS).....	15
Figure 2.3 POD , VHI , FAR , $VFAR$, $MISS$, VMI , CSI , and $VCSI$ values for two daily precipitation data sets (OBS : Stage IV radar-based gauge adjusted data; SIM : PERSIANN [Sorooshian et al., 2000] satellite data; spatial resolution 0.25°).....	16
Figure 3.1 Bias (with optimal value = 1) of selected CMIP5 precipitation simulations, and of their ensemble mean and median, all with respect to GPCP observations. White-colored areas indicate “NaN”—undefined or no values.	26
Figure 3.2 Quantile Bias ($t = 75$ percentile, QB_{75} with optimal value = 1) maps for selected CMIP5 precipitation simulations, and for their ensemble mean and median, all with respect to GPCP observations.	27
Figure 3.3 Volumetric Hit Index (VHI) for $t = 0$ (with optimal value = 1) maps for selected CMIP5 precipitation simulations, and for their ensemble mean and median, all with respect to GPCP observations.	29
Figure 3.4 Volumetric Hit Index for $t = 75$ percentile VHI_{75} (with optimal value = 1) for selected CMIP5 precipitation simulations, and for their ensemble mean and median, with respect to GPCP observations.	32
Figure 3.5 Averaged Bias (B) and 75^{th} and 90^{th} Quantile Bias values (QB_{75} and QB_{90} , all with optimal values = 1) for selected CMIP5 precipitation simulations, and for their ensemble mean and median, both before bias adjustment (solid lines) and after (dashed lines), all with respect to GPCP reference data.	34
Figure 3.6 Quantile Bias ($t=75^{\text{th}}$ percentile, QB_{75} with optimal value =1) maps after bias-adjustment of CMIP5 precipitation simulations, with respect to GPCP observations.	35
Figure 3.7 Averaged Volumetric Hit Index (VHI) and 75^{th} and 90^{th} percent quantile values VHI_{75} and VHI_{90} (with optimal values = 1) for selected CMIP5 precipitation simulations, and for their ensemble mean and median, both before bias adjustment (solid lines) and after (dashed lines), all with respect to GPCP reference data.	36
Figure 3.8 Volumetric Hit Index for $t = 75^{\text{th}}$ percentile (VHI_{75} , with optimal value = 1) after bias-adjustment of selected CMIP5 precipitation simulations, all with respect to GPCP observations.	37
Figure 4.1 Schematic view of a cell in the model forced with precipitation (P) and temperature (T) as inputs. The top soil layer is calibrated with the satellite observations. Stored water in each layer can be exchanged with other layers (red vectors) vertically or evaporate from the soil layers.	46
Figure 4.2 The locations of the 21 major basins from around the globe used for calibration and validation.....	48
Figure 4.3 Monthly runoff [m/month] simulations before and after calibration against composite gridded runoff observations (1987 - 1995).	50

Figure 4.4 Monthly soil moisture [m^3/m^3] before and after calibration against satellite-based soil moisture observations (1987 - 1995).	51
Figure 4.5 Relative errors of soil moisture simulations versus the satellite-based observations before calibration and after calibration (results shown in logarithmic scale for better visualization).	52
Figure 4.6 Relative errors of the simulated runoff against the global gridded runoff observations before and after calibration (results shown in logarithmic scale for better visualization).	53
Figure 4.7 Long-term soil moisture climatology (1986-1995) based on satellite observations [m^3/m^3] (left), and model simulations [m^3/m^3] (right).	55
Figure 4.8 Long-term runoff climatology (1986-1995) based on the composite gridded runoff observations [m/month] (left), and model simulations [m/month] (right).	56
Figure 5.1 (a) Mean monthly inflow to and outflow from Melbourne major reservoirs (A- Thomson, B-Upper Yarra, C- O'Shannassy, and D- Maroondah) shown in panel (b).	61
Figure 5.2 Reservoir water storage anomalies for the projection period (2020-2035) relative to the baseline (1995-2010) for different water demand scenarios. A positive value (blue) indicates that on average the future storage will be more than the baseline, whereas a negative value (red) indicates that the system will expect more water stress relative to the baseline.	63
Figure 5.3 (a) Melbourne water demand anomalies in different scenarios (see Table S1), for the projection period (2020-2035) relative to the baseline (1995-2010). The blue boxplot indicates that on average the future water demand will be less than the baseline, whereas a negative value indicates that future water demand will be more relative to the baseline. (b) Reservoir water storage anomalies for the projection period relative to the baseline for different water demand scenarios. The blue boxplot indicates that on average the future storage will be more than the baseline, whereas a negative value indicates that the system will expect more water stress relative to the baseline. The boxplots show the median (midlines in boxes), 25 th and 75 th percentiles, and whiskers of the summations relative to the baseline.	65
Figure 5.4 Water storage observation compared with reservoir model simulation during baseline period (1995-2010).	71
Figure 6.1 Mean annual inflow to and outflow from Shasta Lake in California.	74
Figure 6.3 Melbourne's major reservoirs; (top panel): Standardized Water Storage Resilience (WSR) Index, Inflow-Demand Reliability (IRD) Index, and Multivariate Standardized Resilience and Reliability Index (MSRRI); (bottom panel) annual Water Storage Resilience Indicator vs. annual Inflow-Demand Reliability Indicator.	83
Figure 6.4 Shasta Lake(top panel): Standardized Water Storage Resilience (WSR) Index, Inflow-Demand Reliability (IRD) Index, and Multivariate Standardized Resilience and Reliability Index (MSRRI); (bottom panel) annual Water Storage Resilience Indicator vs. annual Inflow-Demand Reliability Indicator.	87
Figure 6.5 Lake Oroville; (top panel): Standardized Water Storage Resilience (WSR) Index, Inflow-Demand Reliability (IRD) Index, and Multivariate Standardized Resilience and Reliability Index (MSRRI); (bottom panel) annual Water Storage Resilience Indicator vs. annual Inflow-Demand Reliability Indicator.	88
Figure 6.6 Trinity Lake; (top panel): Standardized Water Storage Resilience (WSR) Index, Inflow-Demand Reliability (IRD) Index, and Multivariate Standardized Resilience and Reliability Index (MSRRI); (bottom panel) annual Water Storage Resilience Indicator vs. annual Inflow-Demand Reliability Indicator.	89

Figure 6.7 Dry and wet conditions in March 2006, 2011, 2012, 2013 and 2014 (based on 6-month standardized precipitation index). D0: Abnormally Dry; D1: Moderate Drought; D2: Severe Drought; D3: Extreme Drought; D4: Exceptional Drought; W0: Abnormally Wet; W1: Moderate Wetness; W2: Severe Wetness; W3: Extreme Wetness; W4: Exceptional Wetness. .. 90

LIST OF TABLES

	Page
Table 3.1 CMIP5 climate models and summary statistics of global Bias (B) and Volumetric Hit Index (VHI) before and after mean-field bias adjustment (after bias adjustment $B=1$ for all models). Optimal values of these metrics are all equal to 1.	30
Table 4.1 Correlation coefficients, relative errors, and Nash-Sutcliff coefficients for runoff and soil moisture simulations relative to the observations in the validation period (Jul. 1986 to Dec. 1995).	54
Table 5.1 Description of water demand scenarios	71
Table 5.2 List of the CMIP5 climate model simulations used for simulating future conditions..	72
Table 5.3 Model efficiency coefficients for the calibration and validation periods.	72
Table 6.1 Four categories of reservoir conditions based on inflow-demand reliability and reliability indicators	85

ACKNOWLEDGEMENTS

I would like to start by expressing my sincerest gratitude to several people for their support and advice during my Ph.D. study. First, I would like to express sincerest gratitude to my Ph.D. adviser, Professor Amir AghaKouchak. I am extremely honored and fortunate to have been given the opportunity to learn from him as his student. During the course of my Ph.D., he helped me discover a new definition of hydrology and shared his experience and invaluable advice on my research topics, papers, and my career. What I learned from him is not only scientific knowledge, also life lessons, such as being patient, respectful, responsible, and creative. Through working with him, I realized the true meaning of friendship in work environment. He is a great mentor, advocating collaboration and friendship to the point that all lab members are not only efficiently working together, but are close friends outside lab. I want to thank him for his wide aspect for future and teaching me to network, collaborate and enjoy my life. I also wanted to thank all my colleagues in my research group who never hesitated helping me. Furthermore, I would like to express my gratitude to Professor Sorooshian, Professor Vrugt, and Professor Sanders. Professor Sorooshian taught me the basics of hydrologic systems through which I could look further into my Ph.D. research. His suggestions are like those a father would give his son, respectful, clear, and kind. Also, during past year I had the chance to work closely with Professor Vrugt. I am grateful for his smart ideas and impressed by the way he investigates complicated problems. Without his help I could not be able to finish my work and my calibration study. I would like to thank Professor Sanders for helping guide me in my research and for taking the time to be one of my committee members. I thank Diane Hohnbaum, who helped with the administrative works. Most of all, I would like to thank my parents who have always believed in me and supported me in all my goals in life. I wish they could both be here to celebrate this day with me. Also, I wish to thank all the past and present students at Hydroclimate Research Group and students in Center for Hydrometeorology and Remote Sensing. Their support and friendship are indeed my lifetime treasure.

Curriculum Vitae

Ali Mehran

Education

PhD, University of California, Irvine, Civil and Environmental Engineering (Hydrology and Water Resources), 2015.

MS, North Carolina State University, Raleigh, NC, Civil Engineering (Systems and Computer aided Engineering), 2010.

MS, Iran University of Science and Technology, Tehran, Iran, Civil Engineering (Water Resources and Hydrodynamics), 2007.

BS, KNT University of Technology, Tehran, Iran, Civil-Environmental Engineering, 2004.

Peer-Reviewed Journal Publications

1. **Mehran A.**, AghaKouchak A., Phillips T.J., 2014, Evaluation of CMIP5 Continental Precipitation Simulations Relative to Satellite-Based Gauge-Adjusted Observations, *Journal of Geophysical Research*, 119, 1695-1707, doi: 10.1002/2013JD021152.
2. **Mehran A.**, AghaKouchak A., 2014, Capabilities of Satellite Precipitation datasets to Estimate Heavy Precipitation Rates at Different Temporal Accumulations, *Hydrological Processes*, 28, 2262-2270, doi: 10.1002/hyp.9779.
3. Liu Z., **Mehran A.**, Phillips T.J., AghaKouchak A., 2014, Seasonal and Regional Biases in CMIP5 Precipitation Simulations, *Climate Research*, 60, 35-50, doi: 10.3354/cr01221.
4. AghaKouchak A., **Mehran A.**, 2013, Extended Contingency Table: Performance Metrics for Satellite Observations and Model Simulations, *Water Resources Research*, 49, 7144-7149, doi:10.1002/wrcr.20498.
5. AghaKouchak A., **Mehran A.**, Norouzi H., Behrangi A., 2012, Systematic and Random Error Components in Satellite Precipitation Data Sets, *Geophysical Research Letters*, L09406, 39, doi: 10.1029/2012GL051592
6. AghaKouchak A., Norouzi H., Madani K., Mirchi A., Azarderakhsh M., Nazemi N., Nasrollahi N., **Mehran A.**, Farahmand A., Hasanzadeh E., 2014, Aral Sea Syndrome Desiccates Lake Urmia: Call for Action, *Journal of Great Lakes Research*, doi: 10.1016/j.jglr.2014.12.007
7. **Mehran A.**, Mazdiyasn O., AghaKouchak A., 2015, A Hybrid Framework for Water Stress Assessment Accounting for Water Resources Resilience, *Journal of Geophysical Research*, in Review.
8. **Mehran A.**, AghaKouchak A., Assessing Melbourne Water Availability in a Changing Climate, in Preparation.
9. **Mehran A.**, AghaKouchak A., A Nested Modeling Framework for Evaluating Climate Change Impacts on Water Resources, in Preparation.

Conference Papers and Presentations

1. **Mehran A.**, AghaKouchak A., 2014, Water Resources Vulnerability Assessment Accounting for Human Influence, Oral Presentation at AGU 2014 Fall Meeting, Sustainable Water Quantity and Quality in the built Environment.

2. **Mehran A.**, AghaKouchak A., Phillips T.J., 2013, Performance Metrics for Climate Model Evaluation: Application to CMIP5 Precipitation Simulations, Poster at AGU 2013 Fall Meeting.
3. AghaKouchak A., **Mehran A.**, 2013, Global Terrestrial Hydrologic Modeling: Roadblocks, Challenges and Opportunities, Invited Talk, AGU 2013 Fall Meeting.
4. Liu Z., **Mehran A.**, Phillipps T. J., AghaKouchak A., 2013, Seasonal and Regional Biases in CMIP5 Precipitation Simulations, Poster at AGU 2013 Fall Meeting.
5. **Mehran A.**, Nakhjiri N., AghaKouchak A., 2012, Integrating Satellite Data into a Nested Global-Local Hydrological Model for Large Scale Flood Forecasting, Poster at AGU 2012 Fall Meeting.
6. Mohammadi M., **Mehran A.** A New Modified Boundary Element Method (MBEM) for Boundary Domain Integral Method. ICHE 2008, Nagoya, Japan.
7. **Mehran A.**, Baktiary A., Hajivali F., 2007, 3-D Numerical Simulation of a Topographic Gradient on Coastal Currents. VF 2007, Fukuoka, Japan.
8. **Mehran A.**, Baktiary A., 2007, 3-D Numerical Simulation of Cubic Cavity Flow with Corner Vortices, VF 2007, Fukuoka, Japan.

Teaching Experience

University of California, Irvine, CA

Engineering Practicum II (CEE 81B) Lab Instructor (Winter 2012 and Winter 2013);

Responsible for teaching lab sessions of Engineering Practicum (undergraduate) course. The course included hands-on Geographic Information System (GIS) instruction to 40 students (4 sessions per week). Responsible for preparing course materials, assignments, exams, and the final project. Used a range of interactive teaching aides and instructions (as template in AutoCAD and ArcGIS) to enhance learning.

Grants and Proposals

- Wrote a successful proposal for *UCI Water-PIRE* project 2013; This National Science Foundation funded Partnerships for International Research and Education (NSF-PIRE) project aims to catalyze, through research and education, the development and deployment of low-energy options (LEOs) for improving water productivity while protecting human and ecosystem health. Received funding for two years of graduate research.
- Wrote a successful proposal to the *Sustainability Science Team 2012* (SST-2012) at University of California, Irvine; an interdisciplinary collaboration of the Environment Institute, the Center for Environmental Biology, and the Urban Water Research Center to study restoration of the Salton Sea.

Honors and Awards

- NCSU ‘*DHS Homeland Security HS-STEM* Career Development Grants.
- Japan’s governmental scholarship (*Monbu-kagaku-shō Shōgaku-kin*) to conduct research at Tokyo Institute of Technology, 2008.

Professional Associations / Services

Reviewer/referee for *Climate Dynamics*

Member of American Society of Civil Engineers

Attended and completed the **Copula for Hydrology and Climate Applications of International Commission of Statistical Hydrology and International Association of Hydrological Sciences (ICSH-IAHS)**.

ABSTRACT OF THE DISSERTATION

Frameworks for Assessing Climate Change Impacts on Water Resources: From Uncertainty in Model Simulations to Accounting for Local Resilience

by

Ali Mehran

Doctor of Philosophy in Civil Engineering

University of California, Irvine, 2015

Professor Amir AghaKouchak, Chair

Numerous studies have highlighted that water resources and hydrologic extremes are sensitive to climate change and variability. An interesting research question is how the hydrologic cycle and water availability respond to the climatic change and variability. In the past decades, numerous methods and models have been developed for assessing climate change impacts on water resources. However, there are still major research gaps from uncertainties in climate model simulations to limitations in the current global hydrologic models. Some of the current research gaps include: (I) high uncertainty of climate model simulations; (II) limitations and high uncertainties of the global hydrologic model simulations because of calibration challenges at the global scale; and (III) lack of frameworks for accounting for the local resilience and man-made infrastructure in climate impact assessment studies. The overarching goal of this dissertation is to address the above mentioned research gaps. First, several novel evaluation metrics are introduced that can be used for evaluation of errors and biases in input data which is a key factor in the overall

uncertainty of climate change studies (Chapters 2 and 3). Furthermore, a modeling framework is introduced for improving global hydrologic models through a comprehensive multi-objective calibration approach (Chapter 4). Then, an additional layer for including man-made reservoirs is integrated in the above mentioned calibrated global hydrologic model. This additional layer is designed to account for local reservoirs that provide resilience against climatic change and variability (Chapter 5). Finally, having global hydrologic simulations and the local response including the effects of reservoirs, a multivariate framework is introduced for water stress assessment using both climatic information and local reservoir conditions (Chapter 6).

CHAPTER 1: Introduction

This chapter corresponds [AghaKouchak and Mehran, 2013]. The presentation has been modified to meet formatting requirements.

1.1. Problem Statement

Numerous studies have highlighted that water resources and hydrologic extremes (e.g., floods and droughts) are sensitive to climate change and variability [Wood *et al.*, 1997b; Merritt *et al.*, 2006; IPCC, 2007; Sivakumar, 2011; Stoll *et al.*, 2011]. In 2010 alone, one of the deadliest flood events over Pakistan, killed over 2000 people, and displaced 2 million people, resulting in USD 9.7 Billion in economic damage. Further, the 2012 drought resulted in more than USD 12 billion in economic loss in the United States that had dramatic impacts over food and commodity prices [Crutchfield, 2012]. The 2011-2014 California drought is another example of a major event affecting not only water resources but also the ecosystem and the agriculture industry [Aghakouchak *et al.*, 2014]. An interesting research question is what the role of climate change is in occurrence of these events. More importantly, how the hydrologic cycle and water availability respond to the climatic change and variability. Therefore, there exists a strong need to study water resources and hydrologic cycle under different climate change scenarios at the global scale [IPCC, 2007; IPCC, 2013].

In the past decades, numerous methods and models have been developed for assessing climate change impacts on water resources [Reed *et al.*, 2004]. The process typically involves forcing hydrologic models with climate model simulations to evaluate the hydrologic response to climatic change and variability (see Figure 1). The Intergovernmental Panel on Climate Change (IPCC), in coordination with numerous international organizations and research institutes, provide model simulations of the historical and future climate [Taylor *et al.*, 2012]. However, climate model simulations are subject to high uncertainties arising from uncertainties in model

physics, initial conditions, boundary conditions, and model parameterization [Feddema et al., 2005; John and Soden, 2007; Reichler and Kim, 2008; Liepert and Previdi, 2012; Brekke, Levi D, Barsugli, 2013; Wehner, 2013; Mehran et al., 2014]. Understanding climate model simulations uncertainty is fundamental to climate impact assessment studies [Brekke, Levi D, Barsugli, 2013].

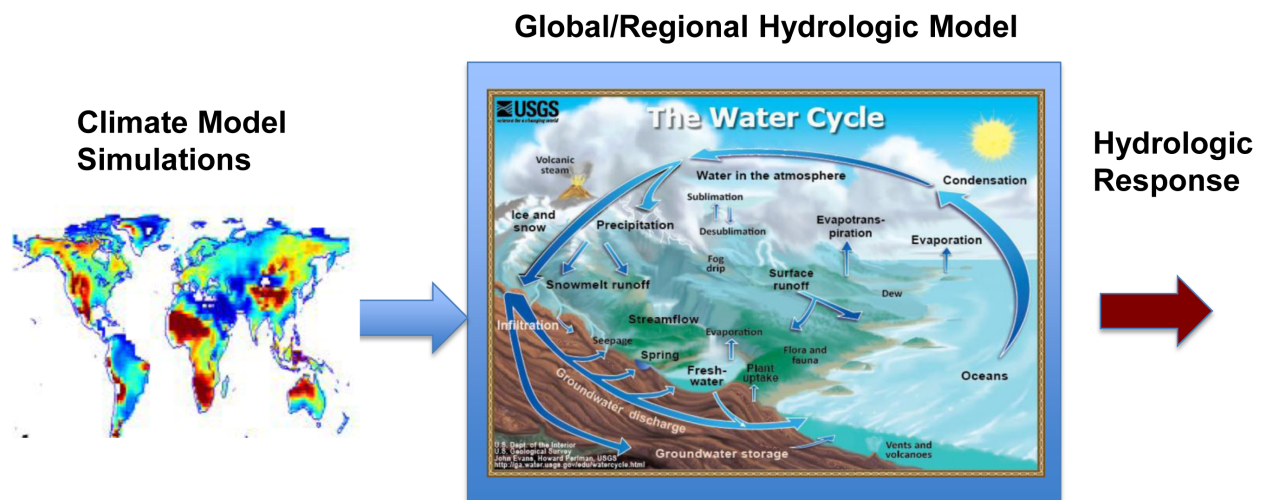


Figure 1.1 Climate model simulations are used to force global or regional hydrologic models to understand the effect of climatic change and variability on hydrologic response.

In recent years, numerous studies have focused on evaluation and quantification of uncertainty in climate model simulations [Feddema et al., 2005; Phillips and Gleckler, 2006; Jiang et al., 2012]. Several efforts have been devoted to development of metrics and tools for validation and uncertainty analysis of and climate model simulations [Gleckler et al., 2008; Mehran et al., 2014]. Some of the common metrics include the Taylor diagram, bias, bias score, dependence metrics such as correlation coefficient, and contingency table metrics [Gleckler et al., 2008]. The latter is used to analyze or validate the relationship between two categorical variables and is the categorical equivalent of the scatterplot (e.g., percent of agreement in rain/no rain detection

between model simulations and observations). In other words, the contingency table metrics describe whether simulations hit or miss the reference observations and/or lead to false estimates. While the contingency table metrics offer invaluable information, they do not provide any insight into biases and errors in the magnitude of the hit, miss and false simulations relative to the observations. To address this gap, the commonly used contingency table has been extended to a set of volumetric indicators that can provide insights into model simulations deficiencies and uncertainties (see Objective I below).

An important element of an impact assessment study is the hydrologic modeling component (Figure 1.1(middle)). Global Hydrological Models (GHM) have been widely used in water availability studies by forcing them with climate model simulations of the past and future. Hydrologic models provide estimates of the past or future runoff, soil moisture and other hydrologic variables. These simulations are also subject to uncertainties due to input data, model structure and parameterization [*van Beek and Bierkens, 2008; Müller Schmied et al., 2014*]. GHMs have thousands of model parameters and because of complexity of global-scale parameter estimation and calibration, in most climate impact studies uncalibrated GHMs have been used [*van Beek and Bierkens, 2008*]. Few studies have attempted to adjust GHM parameters based on observed river discharge and reduce the associated uncertainty [*Liang et al., 1994; Hanasaki et al., 2007*]. In this approach, observed river discharge time series are used to adjust model parameters so that a satisfactory fit is obtained. While this method may result in reasonable runoff estimation, it does not guarantee reliable soil moisture or other hydrologic variables [*Wanders et al., 2014*]. In fact, results presented in this dissertation show that even after calibration based on runoff, model simulations of soil moisture may be significantly biased leading to unrealistic representation of the hydrologic cycle. Substantial biases in simulated

hydrologic variables can lead to misleading information on climate change impacts on water availability. To address this research gap and to narrow the uncertainty in climate impact studies, a modeling framework is introduced that allows multi-objective calibration using observed runoff and top-layer soil moisture information (see Objective II below).

A recent study calls for hydrological models that in addition to surface runoff represent the water balance processes of reservoirs and large water bodies, especially for climate change impact assessment studies [*Reclamation*, 2011]. Reservoirs provide resilience against climatic change and variability, and they affect the water distribution. A major research gap is lack of representation of reservoirs in models used for climate change impact studies [*Reclamation*, 2011]. In this dissertation, an additional layer including man-made reservoirs is integrated in the above mentioned calibrated global hydrologic model. This additional layer is designed to account for local reservoirs that provide resilience against extremes (see Objective III below). Finally, having global hydrologic simulations and the local response including the effects of reservoirs, a framework is introduced for water stress assessment using both climatic information and local reservoir conditions (see Objective IV below).

1.2. Objectives

The overarching goal of this study is to address the above mentioned research gaps. First, several novel evaluation metrics are introduced that can be used for evaluation of errors and biases in input data which is a key factor in the overall uncertainty of climate change studies (Chapters 2 and 3). Furthermore, a modeling framework is introduced for improving global hydrologic models through a comprehensive multi-objective calibration approach (Chapter 4). Then, an additional layer for including man-made reservoirs is integrated in the above mentioned calibrated global

hydrologic model. This additional layer is designed to account for local reservoirs that provide resilience against climatic change and variability (Chapter 5). Finally, having global hydrologic simulations and the local response including the effects of reservoirs, a multivariate framework is introduced for water stress assessment using both climatic information and local reservoir conditions (Chapter 6). These components of the study (described in Chapters 2 - 6) are summarized below:

Objective I (Chapters 2 and 3) - Combining categorical and continuous evaluation metrics: toward a volumetric contingency table

In this chapter, the commonly used categorical metrics (e.g., *hit*, *miss*, *false*) are extended to volumetric measures such that one can investigate both the categorical and their corresponding volumetric errors. The main purpose of the suggested volumetric performance metrics is to decompose the total bias into multiple terms for model diagnosis. These indicators can be used for evaluation of climate model simulations that are often used to investigate changes in the hydrologic cycle. This component of the study focuses on input data uncertainty which is a major component of the total uncertainty. A Validation Toolbox is developed based on the methods presented in this chapter. To demonstrate the performance of the suggested indicators 34 CMIP5¹ historical simulations of monthly mean precipitation are validated against the GPCP² observations to quantify model pattern discrepancies and biases for both the entire distributions and their upper-tails (i.e., extremes) – see Chapter 3. Results showed that most of the CMIP5 simulations are in fairly good agreement with GPCP observations in many areas, but model replication of observed precipitation patterns over desert and certain regions is problematic.

¹ Coupled Model Intercomparison Project Phase 5

² Global Precipitation Climatology Project

Furthermore, significant inter-model variability was observed from model to model which should be considered when evaluating climate model simulations of the future.

Objective II (Chapter 4) Improving global hydrologic modeling using a multi-objective calibration framework

One of the limitations of most, if not all, current global hydrological models is that they are either used uncalibrated or only calibrated based on runoff. Using long-term historical satellite-based observations, this study proposes a framework for multi-objective calibration and parameter estimation. The structure of an existing model has been modified so that it can digest satellite soil moisture observations for calibration. In this framework, in addition to reproducing observed runoff, the top layer of the system is calibrated to reproduce satellite observations of the top soil moisture layer. This leads to model simulations calibrated for two key hydrologic variables. The top soil moisture layer is an important element that affects infiltration and evaporation. Having this secondary constraint based on soil moisture results in improved representation of the hydrologic cycle.

Objective III (Chapter 5) Accounting for local resilience in climate change impact assessments

In this chapter, the global modeling framework, discussed above, is applied for climate change impact assessment. An additional layer for integrating man-made reservoirs is included in the model to account for local reservoirs that provide resilience against extremes. The proposed global-local framework satisfies the recommendation of Reclamation [Reclamation, 2011] that calls for hydrological models that in addition to surface runoff represent the water balance processes of reservoirs and large water bodies (i.e., a global model for large scale and assessment and a local model for modeling the capacity of the system to cope with climatic variability). The

modeling framework is applied over the Melbourne major reservoirs using CMIP5 climate simulations. We consider explicitly the human influence on the water cycle by integrating man-made reservoirs. Furthermore, the study uses 17 different operational water demand scenarios ranging from very optimistic to very unfavorable. These scenarios consider population and industrial growth and involve different consumption behaviors. The Results suggest that for a thorough analysis of climate change impact on water resources, including the effect of local resilience (i.e., reservoirs) and the expected demand in the future are fundamental.

Objective IV (Chapter 6) A hybrid multivariate framework for describing water stress relative to the local capacity to cope with extremes

The global-local modeling framework discussed above provides information on climatic change and variability impacts on not only water availability, but also the local response including the effect of man-made reservoirs. In this dissertation, a hybrid framework, termed Multivariate Standardized Reliability and Resilience Index (hereafter, MSRRI), is introduced that combines global and local scale information for water stress assessment (i.e., combines inflow to the system and water storage relative to the demand for an overall assessment of water stress). This hybrid framework combines: (I) a “top-down” approach that focuses on climate variability and change that cannot be simply controlled or altered by decision makers, and (II) a “bottom-up” methodology that represents the local resilience and societal capacity to respond or adapt to climate extremes and water stress.

The results show that MSRRI is superior to univariate indices because it captures both early onset and persistence of water stress over time. MSRRI provides information on not only inflow deficit, but also how long it takes to recover from an extreme dry condition based on reservoir levels. Example applications of MSRRI for several extreme wet and dry conditions are provided

including the Australian Millennium Drought (1998-2010) and the 2014 California Drought.

1.3. Organization of the Dissertation

This dissertation consists of five journal articles organized as follows: After the introduction (Chapter 1), a series of volumetric indicators are introduced that extend the commonly used contingency table (Chapter 2) and used for evaluating of 34 CMIP5 climate model simulations of precipitation (Chapter 3). Chapter 4 describes a global hydrological modeling framework that includes a multi-objective calibration approach. Chapter 5 introduces a global-local modeling framework for climate change impacts assessment and accounting for local resilience. Chapter 6 outlines a hybrid multivariate approach to investigate water stress relative to the local capacity to cope with extremes. The last chapter provides a summary of the findings and conclusions (Chapter 7).

CHAPTER 2: Performance Metrics for Evaluation of Climate Model Simulations

This chapter corresponds [Aghakouchak and Mehran, 2013]. The presentation has been modified to meet formatting requirements.

2.1 Introduction

Climate model simulations are subject to uncertainties and biases arising from physical and algorithmic aspects. The science of evaluation of model simulations, primarily developed by the remote sensing community and future continued in the climate community has attracted a great deal of attention. Evaluation and uncertainty quantification of remotely sensed data and climate model simulations are fundamental to scientific advancements, algorithm/model developments, and integration of data into applications. For this reason, numerous studies are devoted to evaluation of gridded datasets including remote sensing data (e.g., [Anagnostou *et al.*, 1998; Jackson, T.J., Entekhabi, D., Njoku, 2005; Turk *et al.*, 2008; Pinker *et al.*, 2009; AghaKouchak *et al.*, 2010, 2011; Dorigo *et al.*, 2010; Norouzi *et al.*, 2011; Aghakouchak and Mehran, 2013]), and climate model simulations (e.g., [Feddema *et al.*, 2005; Phillips and Gleckler, 2006; Jiang *et al.*, 2012; Liepert and Previdi, 2012]) versus ground-based observations. [Gleckler *et al.*, 2008] suggested several performance metrics for validation of historical climate model simulations. [Gilleland, 2013] proposed the spatial prediction comparison test for evaluation of precipitation forecasts. [Mehran and AghaKouchak, 2014] developed several indices for evaluation of high quantiles of satellite precipitation observations. [Entekhabi *et al.*, 2010a] introduced a number of metrics for evaluation of remotely sensed soil moisture observations. [Hossain and Huffman, 2008] recommended a set of spatial, retrieval, and temporal error metrics for satellite data sets that can advance hydrologic applications. [Gebremichael and Krajewski, 2007] outlined a framework for validating satellite data sets using ground-based observations. A number of

geometrical and object-oriented metrics are also proposed for spatial validation and verification (e.g., [Brown et al., 2004; Davis et al., 2009; AghaKouchak et al., 2011]). Among the metrics, the contingency table [Wilks, 2006] which includes a number of categorical indices is extensively used in evaluation studies (e.g., [Anagnostou et al., 2010; Hirpa et al., 2010; Behrangi et al., 2011; Gourley et al., 2012; Haile et al., 2013; Hao et al., 2013]). The contingency table is used to analyze or validate the relationship between two categorical variables and is the categorical equivalent of the scatterplot. The contingency table metrics describe whether climate model simulations or even remote sensing observations (hereafter, *SIM*) hit or miss the reference observations (hereafter, *OBS*) and/or lead to false estimates relative to *OBS*. While the contingency table metrics offer invaluable information, they do not provide any insight into biases and errors in the magnitude of *SIM* relative to *OBS*. Hence, errors and biases should be evaluated using additional metrics such as the unbiased root mean square error [Entekhabi et al., 2010a], quantile bias [AghaKouchak et al., 2011], and relative error [Gleckler et al., 2008]. In this chapter, the commonly used categorical metrics are extended to volumetric measure such that one can investigate both the categorical hit, miss, false, and their corresponding volumetric errors. The main purpose of the suggested indices is to decompose the total bias into volumetric errors terms associated with hit, miss, and false components. This chapter is organized into three sections. After the introduction, the commonly used categorical indices are reviewed briefly. The proposed volumetric indices are then described followed by an example application.

2.2 Methodology and Results

The most common form of the contingency table is 2×2 , which is used to evaluate dichotomous variables (see Figure 2.1). In this table, hit (*H*) indicates that both reference observation and simulation detect the event, whereas miss (*M*) refers to events identified by reference observation

but missed by the simulation. False (F), also known as false alarm, represents events identified by the simulation but not confirmed by observations. Based on the contingency table, several metrics are defined as follows [Wilks, 2006]:

1. The Probability of Detection (POD) describes the fraction of the reference observations detected correctly by the simulation: $POD = H / (H + M)$. The POD ranges from 0 to 1; 0 indicates no skill and 1 indicates perfect score.
2. The False Alarm Ratio (FAR) corresponds to the fraction of events identified by simulation but not confirmed by reference observations: $FAR = F / (H + F)$. The FAR ranges from 0 to 1; 0 indicates perfect score.
3. The Critical Success Index (CSI), also known as the Threat Score, combines different aspects of the POD and FAR , describing the overall skill of the simulation relative to reference observation: $CSI = H / (H + M + F)$. The CSI ranges from 0 to 1; 0 indicates no skill and 1 indicates perfect skill.

		<i>Event Occurred?</i>	
		Reference Observations	
		Yes	No
Event Occurred?	Simulations	Yes	Hit (<i>H</i>)
		No	False (<i>F</i>)
		Miss (<i>M</i>)	True Null Event (<i>Q</i>)

Figure 2.1 The Contingency Table

The original contingency table metrics provide categorical measures of performance. For example, a *POD* of 0.8 indicates that the simulation detects 80% of events (e.g., precipitation events). However, it does not provide any information as to what fraction of the volume of precipitation is detected. For most climate variables one may need to go beyond the *POD* and estimate the volume of the variable of interest detected correctly. For this reason, the Volumetric Hit Index (*VHI*) can be defined as follows:

$$VHI = \frac{\sum_{i=1}^n (SIM_i | (SIM_i > t \& OBS_i > t))}{\sum_{i=1}^n (SIM_i | (SIM_i > t \& OBS_i > t)) + \sum_{i=1}^n (OBS_i | (SIM_i \leq t \& OBS_i > t))} \quad (2.1)$$

SIM refers to satellite observations or climate model simulations being evaluated, whereas *OBS* represents reference observations. In equation (2.1), *n* is the sample size and *t* is the threshold above which the *VHI* is computed. A *t* = 0, indicates evaluation of the entire distribution of simulated versus observed variables. A higher threshold can be used to evaluate solely the higher quantiles of simulations relative to observations (e.g., *VHI* of values above 50th percentile of observations). By computing the *VHI* above different thresholds, one can plot the performance of

SIM relative to the magnitude of *OBS*. The *VHI* ranges from 0 to 1 with 1 being the perfect score. A similar threshold concept can be used to derive the Quantile Probability of Detection (*QPOD*) [AghaKouchak *et al.*, 2011] to describe correct detection and identification above a certain threshold (see equation (A1) in Appendix A).

It should be noted that *VHI* is an extension of *POD*; however, it is defined slightly differently. In *POD*, the number of $(SIM_i | (SIM_i > t \& OBS_i > t))$ is the same as the number of $(OBS_i | (SIM_i > t \& OBS_i > t))$. However, the $\sum_{i=1}^n (SIM_i | (SIM_i > t \& OBS_i > t))$ is not identical to $\sum_{i=1}^n (OBS_i | (SIM_i > t \& OBS_i > t))$ as simulated, and observed data sets are often biased against each other. For this reason, the *VHI* is defined as the volume of correctly detected simulations relative to the volume of the correctly detected simulations and missed observations. The *VHI* should be complemented by information on bias, hit bias (defined in [Tian *et al.*, 2009]), and mean quantile bias for a better understanding of the performance of simulations against observations. Figure 2.1 illustrates the differences between *VHI* and *POD* using synthetic precipitation data. In this example, the solid blue line shows the reference observation (*OBS*), whereas the dashed red (*SIM₁*) and green (*SIM₂*) lines represent two sets of model simulations (or satellite observations). The bias values, defined as *SIM/OBS* show that *SIM₁* underestimates by over 50%, while *SIM₂* underestimates by 11% (Table 2.1). Also, a visual comparison indicates that *SIM₂* is in better agreement with *OBS* relative to *SIM₁*. However, both *SIM₁* and *SIM₂* lead to the same *POD* of 0.83 as the number of categorical matches between the two data sets and the observations are the same. The *VHI*, on the other hand, shows 0.74 and 0.84 for *SIM₁* and *SIM₂*, respectively, indicating the *SIM₂* is in better agreement with *OBS* compared to *SIM₁*. Similarly, the Volumetric False Alarm Ratio (*VFAR*) can be expressed as the volume of false *SIM* above the threshold *t* relative to the sum of simulations:

$$VFAR = \frac{\sum_{i=1}^n (SIM_i | (SIM_i > t \& OBS_i \leq t))}{\sum_{i=1}^n (SIM_i | (SIM_i > t \& OBS_i > t)) + \sum_{i=1}^n (OBS_i | (SIM_i > t \& OBS_i \leq t))} \quad (2.2)$$

The denominator of equation (2.2) can be summarized as total volume of simulations ($\sum_{i=1}^n (SIM_i | (SIM_i > t))$). The *VFAR* ranges from 0 to 1, with 0 being the perfect score. It should be noted that similar to *QPOD*, one can define the Quantile False Alarm Ratio (*QFAR*) [AghaKouchak et al., 2011], which describes the categorical ratio of the number of false identifications of SIM relative to the number of exceedance above a certain threshold (e.g., 90% and 95% quantiles), see equation (A2) in Appendix A. In the earlier example, the *FAR* values of both *SIM₁* and *SIM₂* are 0.19, while a visual comparison shows that *SIM₂* exhibits more false precipitation (see Figure 2.2). As shown in Table 2.1, the *VFAR* value of *SIM₂* is higher than *SIM₁* confirming the visual comparison. The fraction of the volume of missed *SIM* relative to *OBS* can be expressed using the Volumetric Miss Index (*VMI*):

$$VMI = \frac{\sum_{i=1}^n (OBS_i | (SIM_i \leq t \& OBS_i > t))}{\sum_{i=1}^n (SIM_i | (SIM_i > t \& OBS_i > t)) + \sum_{i=1}^n (OBS_i | (SIM_i \leq t \& OBS_i > t))} \quad (2.3)$$

The *VMI* ranges from 0 to 1, with 0 being the perfect score. Based on the definition of *MISS* (1-*POD*), the categorical Quantile Miss Index (*QMISS*) can be expressed as 1 - *QPOD*. In the provided example, the *MISS* index for both *SIM₁* and *SIM₂* are 0.17 (indicating 17% of categorical miss). However, Figure 2.2 clearly shows that *SIM₂* is in better agreement with *OBS*. The *VMI* values confirm that *SIM₂* (0.16) is superior to *SIM₁* (0.26) with respect to the volume of missed precipitation (see Table 2.1). Finally, following the original *CSI* concept, the Volumetric Critical Success Index (*VCSI*) is defined as an overall measure of volumetric performance:

$$VCSI = \frac{\sum_{i=1}^n (SIM_i | (SIM_i > t \& OBS_i > t))}{\sum_{i=1}^n ((SIM_i | (SIM_i > t \& OBS_i > t)) + (OBS_i | (SIM_i \leq t \& OBS_i > t)) + (SIM_i | (SIM_i > t \& OBS_i \leq t)))} \quad (2.4)$$

The *VCSI* ranges from 0 to 1, with 1 being the perfect score. While the *CSI* values of *SIM₁* and

SIM_2 are the same (0.70), the $VCSI$ values indicate that the SIM_2 is in better agreement with SIM_1 (Table 2.1). The $QCSI$ (equation (A3) in Appendix A) can be used as the categorical equivalent of $VCSI$.

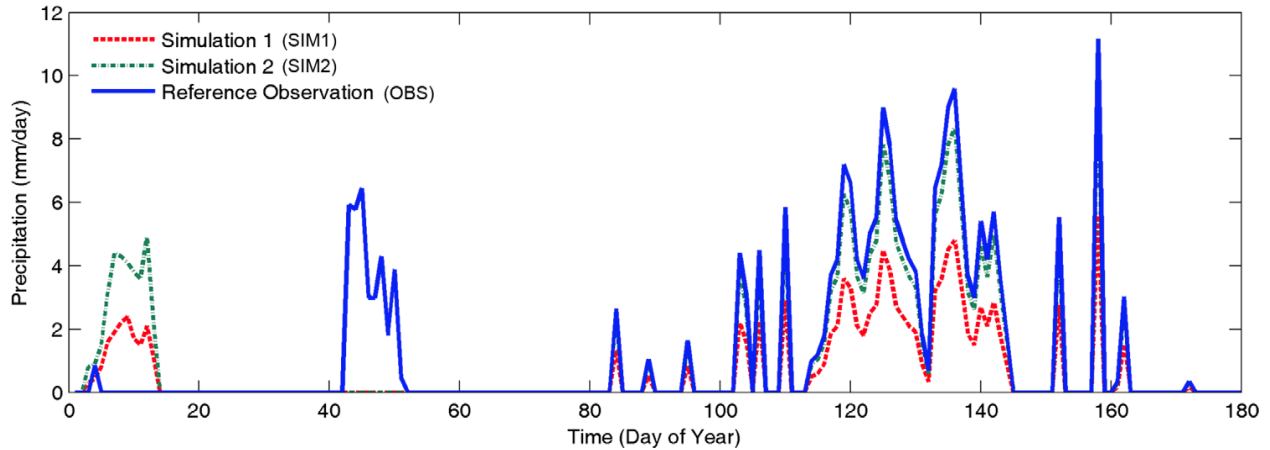


Figure 2.2 Two synthetic precipitation simulations (SIM_1 and SIM_2) and reference observation (OBS).

For two daily precipitation data sets (OBS : Stage IV radar-based gauge adjusted data; SIM : PERSIANN [Hsu et al., 1997; Sorooshian et al., 2000] satellite data; spatial resolution 0.25°), Figure 2.3 displays sample POD , VHI , FAR , $VFAR$, $MISS$, VMI , CSI , and $VCSI$ values. One can see that the volumetric indices provide additional information beyond the contingency table categorical metrics. For example, the POD values range primarily between 0.4 and 0.6, while VHI values indicate that SIM detects more than 80% of the volume of observed precipitation. While the FAR values are relatively high (0.5 in the eastern United States indicating around 50% false precipitation), the $VFAR$ values over the eastern United States show that the false precipitation with respect to volume of precipitation is mainly below 10%. Similarly, the $MISS$ index shows that SIM does not detect a large fraction of precipitation. Based on VMI , however, the fraction of precipitation SIM does not detect is relatively small (compare $MISS$ and VMI in Figure 2.3). This implies that most of the missed events in SIM are light rainfall events. The CSI values show that for example, in the eastern United States the overall performance score of SIM

is between 0.3 and 0.5, whereas the *VCSI* indicates a higher performance score (between 0.5 and 0.8) with respect to the volume of precipitation. This example shows that the volumetric measures provide additional information that cannot be achieved from the original categorical metrics. These indices can also be used to decompose biases at high quantiles of a data set by computing them for different thresholds (e.g., $t = 25^{\text{th}}$, 50^{th} , 75^{th} percentiles). It is worth pointing out that the volumetric indices should be computed along with the categorical metrics for a comprehensive assessment of simulations against observations.

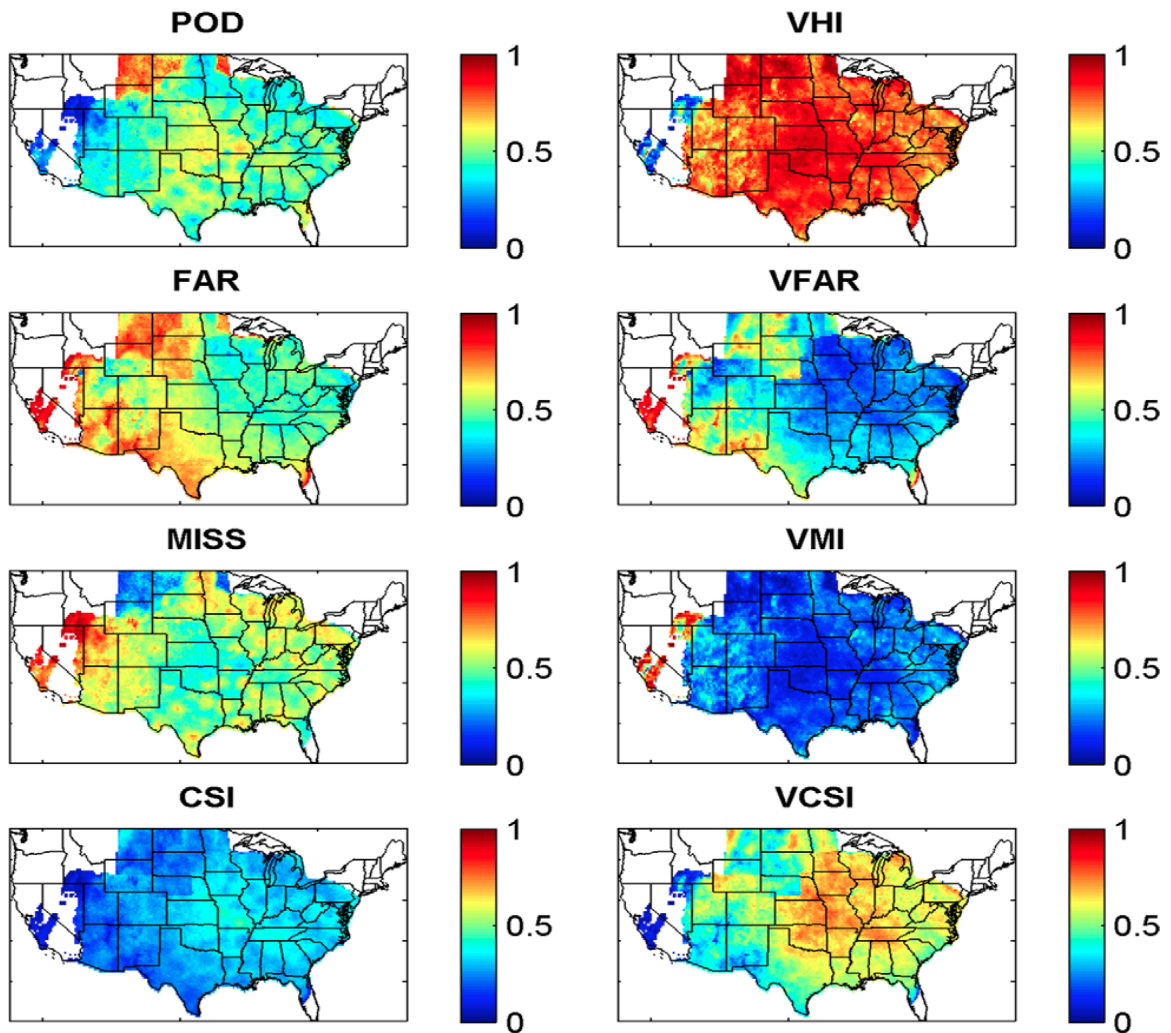


Figure 2.3 *POD*, *VHI*, *FAR*, *VFAR*, *MISS*, *VMI*, *CSI*, and *VCSI* values for two daily precipitation data sets (OBS: Stage IV radar-based gauge adjusted data; SIM: PERSIANN [Sorooshian et al., 2000] satellite data; spatial resolution 0.25°).

2.3 Conclusions

Current weather and climate models have been widely used to simulate historical and future climate over various spatial and temporal scales, validation and uncertainty quantification of gridded climate model simulations are fundamental to future improvements in model developments and climate impact assessments. In this part of the study, the contingency table categorical metrics are extended to volumetric indices for evaluation of gridded data relative to a reference data set. Several indices are introduced including (a) the Volumetric Hit Index (*VHI*) which describes the volume of correctly detected simulations relative to the volume of the correctly detected simulations and missed observations; (b) the Volumetric False Alarm Ratio (*VFAR*) which identifies the volume of false simulations relative to the volume of simulations; (c) Volumetric Miss Index (*VMI*) which expresses the fraction of the volume of missed observations relative to the volume of the correctly detected simulations and missed observations; and (d) the Volumetric Critical Success Index (*VCSI*), defined as an overall measure of volumetric performance including the volumetric hits, false alarms, and misses. The suggested indices decompose the total volumetric error (bias) into volumetric errors terms associated with hit, false, and miss components in simulations. Using two synthetic time series of simulated precipitation, the volumetric indices are evaluated against the contingency table categorical indices. The synthetic example highlights the difference between the commonly used categorical and the volumetric metrics. The volumetric indices are then applied for validation of a gridded satellite data set relative to reference observations. The results show that the volumetric indices provide additional information beyond the commonly used categorical metrics that can be useful in evaluating gridded data sets. Chapter 2 contributes to ongoing metrics development efforts for validation and verification of gridded data sets. It is noted that the introduced volumetric indices

are not meant to replace the commonly used categorical metrics. Rather, they should be viewed as metrics that can provide additional information and complement the contingency table categorical metrics. Furthermore, we do not claim that these indices are sufficient for a thorough evaluation of gridded data sets. Additional metrics such as quantile bias, hit bias, relative error, and unbiased root mean square error should also be used for validation and verification studies.

Appendix A

Quantile Probability of Detection (*QPOD*) [AghaKouchak *et al.*, 2011] is defined as the POD above the threshold t :

$$QPOD = \frac{\sum_{i=1}^n I(SIM_i | (SIM_i > t \& OBS_i > t))}{\sum_{i=1}^n I(SIM_i | (SIM_i > t \& OBS_i > t)) + \sum_{i=1}^n I(OBS_i | (SIM_i \leq t \& OBS_i > t))} \quad (A1)$$

where t is the threshold (e.g., 90% and 95% quantiles); I is the indicator function; and n is the number of exceedances. *QPOD* represents the ratio of the number of correct identifications above a certain threshold (t) relative to the total number of exceedances (n). The *QPOD* ranges from 0 (no detection skill) to 1 (perfect detection). Similarly, the Quantile False Alarm Ratio (*QFAR*) can be expressed as

$$QFAR = \frac{\sum_{i=1}^n I(SIM_i | (SIM_i > t \& OBS_i \leq t))}{\sum_{i=1}^n I(SIM_i | (SIM_i > t \& OBS_i > t)) + \sum_{i=1}^n I(SIM_i | (SIM_i > t \& OBS_i \leq t))} \quad (A2)$$

The *QFAR* ranges from 0 (perfect score) to 1. Quantile Critical Success Index (*QCSI*) is defined as the *CSI* above the threshold t :

$$QCSI = \frac{\sum_{i=1}^n I(SIM_i | (SIM_i > t \& OBS_i > t))}{\sum_{i=1}^n I((SIM_i | (SIM_i > t \& OBS_i > t)) + (OBS_i | (SIM_i \leq t \& OBS_i > t)) + (SIM_i | (SIM_i > t \& OBS_i \leq t)))} \quad (A3)$$

CHAPTER 3: Evaluation of CMIP5 Precipitation Simulations

This chapter corresponds [Mehran *et al.*, 2014]. The presentation has been modified to meet formatting requirements.

3.1 Introduction

Numerous studies have emphasized that water resources are sensitive to climate change, and thus water resources management and planning strategies should be adjusted accordingly (e.g., [Wood *et al.*, 1997a; Seager *et al.*, 2007a; Barnett *et al.*, 2008; Sivakumar, 2011; Stoll *et al.*, 2011]). One of the key climate variables is precipitation, which plays a dominant role in the hydrologic cycle. Developing future water resources management and planning strategies thus requires estimation of current and future precipitation magnitude and variability.

In the past several decades, global climate models have been used to estimate future projections of precipitation [IPCC, 2007]. However, these projections are inherently uncertain and often are difficult for decision makers to interpret (e.g., [Feddema *et al.*, 2005; Min *et al.*, 2007; Reichler and Kim, 2008; Liepert and Previdi, 2012]). Quantification of biases and uncertainties in climate simulations of precipitation thus are fundamental to understanding the reliability of climate simulations for future water resources management. [Gleckler *et al.*, 2008] introduced several metrics for performance analysis of climate models and emphasized the need to go beyond the mean statistics for comprehensive analysis of climate model performance. [Moise and Delage, 2011] and [Schaller *et al.*, 2011] presented alternative approaches and metrics for evaluating seasonal precipitation simulations. [Aghakouchak and Mehran, 2013] introduced a number of volumetric indicators for validation and verification of climate model simulations. Since the inception of the of the Coupled Model Intercomparison Project Phase 3 (CMIP3) by the World Climate Research Programme (WCRP) Working Group on Coupled Modelling (WGCM),

evaluation of coupled ocean-atmosphere simulations of historical climate relative to available observational data has become an especially strong scientific focus [Bony *et al.*, 2006]. Indeed, future developments and improvements in global climate models (GCMs) rely heavily on their rigorous and informative validation.

Generally, climate model simulations are known for poor representation of frontal, convective and meso-scale processes [Weverberg *et al.*, 2013]. Numerous studies have evaluated various aspects of precipitation in the CMIP3 model simulations. [Phillips and Gleckler, 2006], for example, evaluated CMIP3 simulations of seasonal-mean continental precipitation amounts, and concluded that many of these differed markedly from several observational estimates. They also noted that the ensemble-mean model precipitation was generally closer to the observations than that of any individual CMIP3 model. Several studies focused on the common errors and/or frequency and intensity of CMIP3 daily precipitation simulations (e.g. [Dai, 2006; Sun *et al.*, 2007; Brown *et al.*, 2010; Stephens *et al.*, 2010]). [Dai, 2006] evaluated the mean spatial patterns, precipitation intensity, frequency, and diurnal cycle. The results showed that many climate models simulated unrealistic double Intertropical Convergence Zone precipitation patterns, though most models captured the overall precipitation pattern. Furthermore, [Dai, 2006] showed that the CMIP3 simulations produce light rain (1 to 10 mm) too frequently (see [Sun *et al.*, 2007; Wilcox and Donner, 2007; Brown *et al.*, 2010; Stephens *et al.*, 2010]).

[Brown *et al.*, 2010] demonstrated that models capture the synoptic regimes well, and concluded that uncertainties in precipitation simulations are due to problems in simulating the characteristics of precipitation within different synoptic regimes. In a recent study, [Catto *et al.*, 2013], argued that climate models often underestimate frontal precipitation estimates. Biases have been reported in model precipitation simulations from warm clouds, associated with

unrepresentative microphysical parameterizations [Lebsock *et al.*, 2013]. [Wehner *et al.*, 2009] showed that many climate models underestimated 20-year return values of precipitation, and suggested increasing the horizontal resolution could improve estimation of extremes. [Ghan *et al.*, 2002] showed that improving representation of subgrid variability and surface topography has a significant positive impact on model precipitation simulations (see also [Qian *et al.*, 2009]).

Following upon CMIP3, the current Coupled Model Intercomparison Project Phase 5 (CMIP5) includes an unprecedented suite of coordinated simulations of historical and future-climate scenarios [Taylor *et al.*, 2012] that are designed to facilitate consideration of the wide range of scientific issues to be addressed in the forthcoming Intergovernmental Panel on Climate Change (IPCC) 5th Assessment Report. The CMIP5 climate simulations are archived by institutional participants in the Global Organization for Earth System Science Portals that is coordinated by the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison (PCMDI).

A comprehensive description of the detailed numerical, dynamical, and physical properties of the CMIP5 models is now in progress [Guilyardi *et al.*, 2013]. In a recent study, [Liu *et al.*, 2012] evaluated the variability of CMIP5 precipitation simulations and their response to temperature using satellite data and showed that there is generally good agreement (correlation) between model simulations and satellite-inferred observed precipitation anomalies over land, both in the Tropics and globally. In addition, [Sillmann *et al.*, 2013] evaluated models' performance in simulating precipitation extremes at 1-5 day time scales. [Hirota and Takayabu, 2013] investigated reproducibility of observed precipitation distribution in CMIP5 relative to CMIP3 simulations over the tropical oceans. They showed ensemble mean of CMIP5 simulations

exhibited slightly higher skill score compared to CMIP3 ensemble mean. [Gaetani and Mohino, 2013] studied the decadal predictability of CMIP5 simulations of the Sahel precipitation, and concluded that predictive skills of CMIP5 precipitation simulations varies significantly from model-to-model. [Kumar *et al.*, 2013] showed that the CMIP5 ensemble mean precipitation matched very well with that of ground-based observations, while there were substantial biases in the simulation of regional precipitation trends. A number of other studies assessed future changes in precipitation based on CMIP5 simulations at regional or global scales (e.g., [Chadwick *et al.*, 2013; Joetzer *et al.*, 2013]).

The present study evaluates CMIP5 historical simulations of continental precipitation against the Global Precipitation Climatology Project (GPCP) monthly mean observational estimates [Adler *et al.*, 2003] using several quantitative statistical measures. This model evaluation focuses on the years 1979-2005, a period for which long-term and gauge-adjusted satellite observations are available. The remainder of this chapter is organized into three sections. The observational datasets are briefly introduced in the next section, while after that is devoted to methodology and results. Summary remarks and conclusions are included in last section.

3.2 Datasets

The GPCP reference data set [Adler *et al.*, 2003] is derived from merged satellite-precipitation data that are bias-corrected using thousands of continental rain-gauge observations. The data set is available as a monthly time series from 1979 onward. GPCP data sets have been validated and widely used in numerous studies (e.g., [Bolvin *et al.*, 2009; Huffman *et al.*, 2009]). The ground based data are assembled by the Global Precipitation Climatology Centre (GPCC) of the Deutscher Wetterdienst and by the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (CPC).

In this study, 34 CMIP5 historical simulations of monthly mean precipitation for the period 1979-2005, as well as their ensemble mean and median, are validated against the GPCP observations. Because of lack of reference gauge data across oceans, GPCP data over oceans are not bias-adjusted. One cannot evaluate biases in CMIP5 simulations with a reference data set having unknown bias. For this reason, this study is limited to evaluation of CMIP5 simulations overland where GPCP data are bias-adjusted using thousands of ground-based gauge data [Adler *et al.*, 2003].

All CMIP5 precipitation simulations and GPCP data are re-gridded onto a common 2×2 degree grid. Table 3.1 summarizes the CMIP5 models considered in this study. In addition, the results for both the multi-model ensemble mean and median are provided, since the latter is less sensitive to statistical outliers than the former.

It should be noted that the model simulations that are designated, as “_esm” are historical simulations of climate with atmospheric CO_2 *emissions* specified in coupled earth systems models (ESMs) that include a prognostic carbon cycle, but with model-specific dynamical vegetation schemes “turned off”. All the other climate simulations are performed with coupled ocean-atmosphere models in which the historical time series of global atmospheric CO_2 *concentrations* are prescribed. Although the CO_2 concentration time series in the “_esm” runs are not identical to these prescribed values (owing to model-specific differences in converting CO_2 emissions to concentrations), they are effectively constrained to be very similar.

3.3 Methodology and Results

Several statistical measures are employed to assess climate model-based historical precipitation simulations. Figure 3.1 displays the overall biases of climate model simulations relative to observations. The bias B is defined as the sum of monthly precipitation amount P for each

CMIP5 model divided by the sum of the corresponding GPCP observations in each 2×2 degree grid box:

$$B = \frac{\sum_{i=1}^n P_{CMIP5}}{\sum_{i=1}^n P_{GPCP}} \quad (3.1)$$

Here n is the number of exceedances of a specified local monthly (daily) precipitation threshold t (which in this case is set equal to 1 mm/day), while, for simplification, the index i signifying each month in the 1979-2005 time series is not included (i.e., $P_{CMIP5} = P_{CMIP5_i}$ and $P_{GPCP} = P_{GPCP_i}$). A bias value above (below) 1 thus indicates an aggregate model overestimation (underestimation) of the monthly GPCP precipitation amounts for a particular grid box.

In the Figure 3.1 mappings of B , the color green indicates where there is little simulation bias, while red (blue) indicates large positive (negative) bias relative to GPCP data (white areas indicate no data in either observations or model simulations). This pertains to most model simulations of precipitation over the eastern United States and northern Europe and Asia; but many show a large positive bias (as high as ~ 2) in regions of complex topography such as western North and South America, and southern Africa and Asia, as was also noted for CMIP3 models by [Phillips and Gleckler, 2006]. In contrast, most models underestimate precipitation over the Saharan and central Asian deserts. Australia and Amazonia are other locations where there are substantial variations in both the sign and magnitude of the biases across individual simulations. On the other hand, the GPCP data may be subject to biases due to the limited availability of ground-based gauge data for bias-correction, as well as the limitations of satellite data in estimating orographic precipitation [Sorooshian et al., 2011; Mehran and Aghakouchak, 2014].

In Figure 3.1, it is also noteworthy that the spatial patterns of the biases of the ensemble mean and median simulations are similar (bottom right panels), and over several regions (e.g.,

Australia and the Americas) are of lower magnitude than the biases of most of the individual model simulations. From comparing global-average biases B summarized in Table 3.1, however, the overall global bias of a number of the CMIP5 simulations is seen to be closer to the optimum value of 1 than is that of the ensemble mean ($B = 0.89$) and median ($B=0.85$).

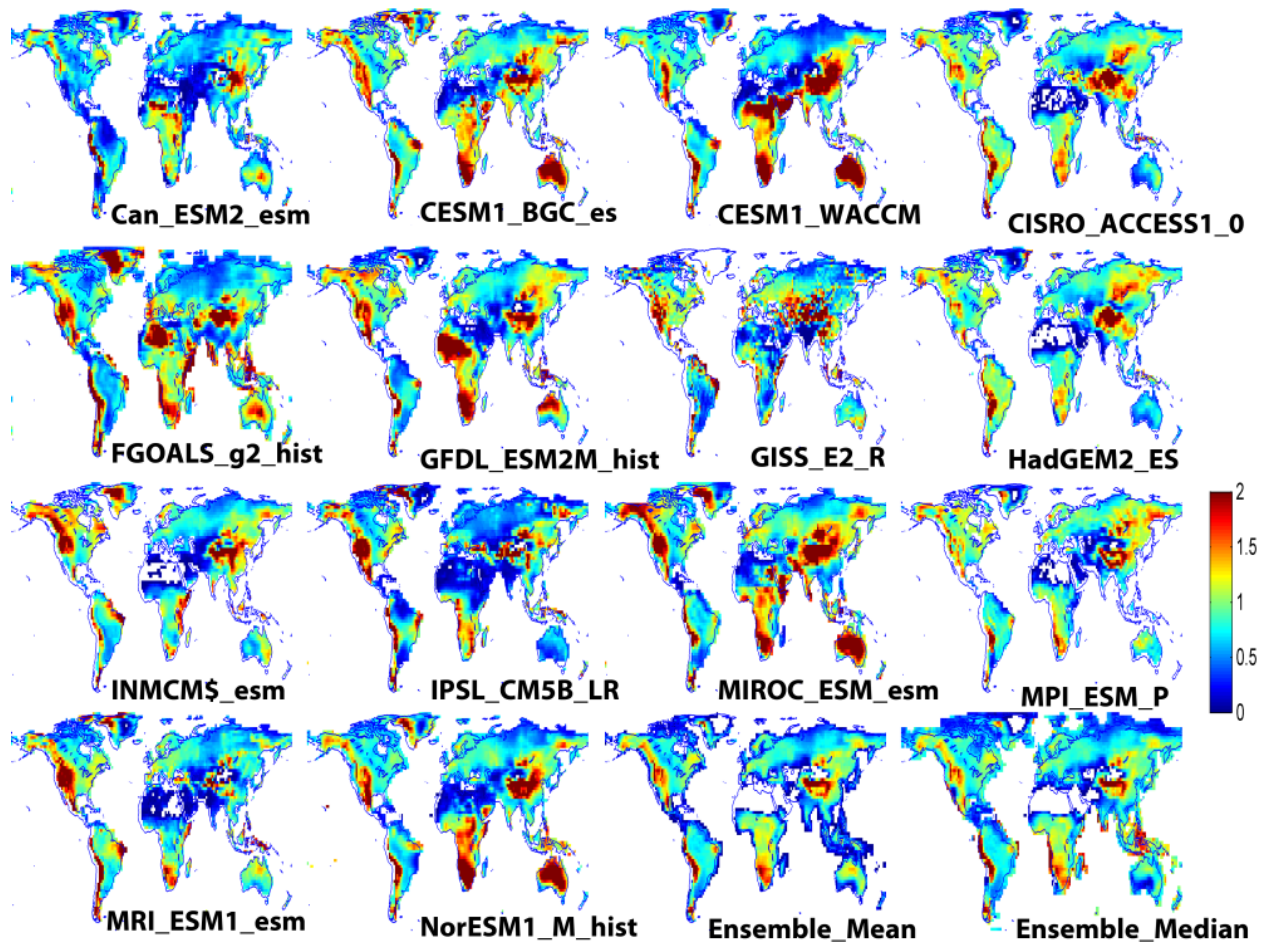


Figure 3.1 Bias (with optimal value = 1) of selected CMIP5 precipitation simulations, and of their ensemble mean and median, all with respect to GPCP observations. White-colored areas indicate “NaN”—undefined or no values.

It is well known that the bias is not necessarily constant throughout a distribution function, but may change at different quantile levels. To further investigate precipitation biases in the CMIP5 simulations, the quantile bias (QB_t), defined as the ratio of monthly precipitation amounts in each

simulation to that of the GPCP observations above a specified threshold t (e.g., the 75th percentile of all the local monthly values), can be calculated in each 2×2 degree grid box:

$$QB_t = \frac{\sum_{i=1}^n (P_{CMIP5} | P_{CMIP5} \geq t)}{\sum_{i=1}^n (P_{GPCP} | P_{GPCP} \geq t)} \quad (3.2)$$

Here, $QB_t = 1$ indicates no bias in the simulations, whereas a value above (below) 1 corresponds to a climate model's overestimation (underestimation) of precipitation amount above the specified threshold t , with respect to that of the GPCP observations.

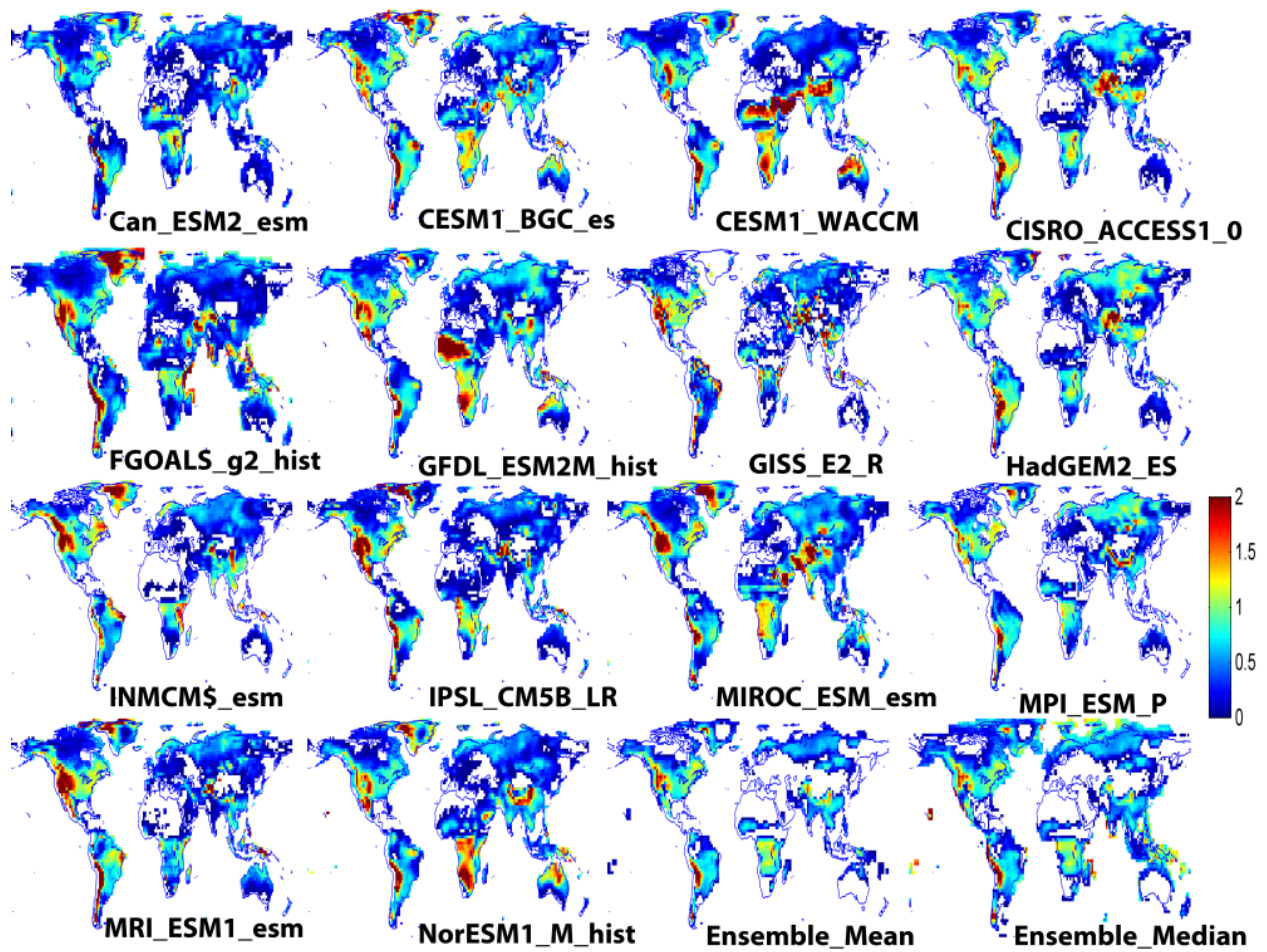


Figure 3.2 Quantile Bias ($t = 75$ percentile, QB_{75} with optimal value = 1) maps for selected CMIP5 precipitation simulations, and for their ensemble mean and median, all with respect to GPCP observations.

Figure 3.2 displays QB_t values, computed for the 75th percentile (precipitation values above $t =$

75% of the reference data), in the CMIP5 simulations. This figure indicates that the climate model biases apparent in Figure 3.1 are generally accentuated in the upper tail (i.e., > 75% quantile) of the GPCP precipitation distribution. While individual CMIP5 models behave somewhat differently from one another, most of their simulations underestimate heavier precipitation amounts over large areas (e.g. Eurasia, Middle East, northern China), while overestimating them only in certain limited regions (e.g., Amazonia, central Africa, United States). Given this general behavior, it is not surprising that the multi-model ensemble mean and median also show large negative biases for observed precipitation amounts > 75% quantile of the distribution. Such negative biases are even more pronounced for P amounts > 90% of the distribution (figure not shown for brevity).

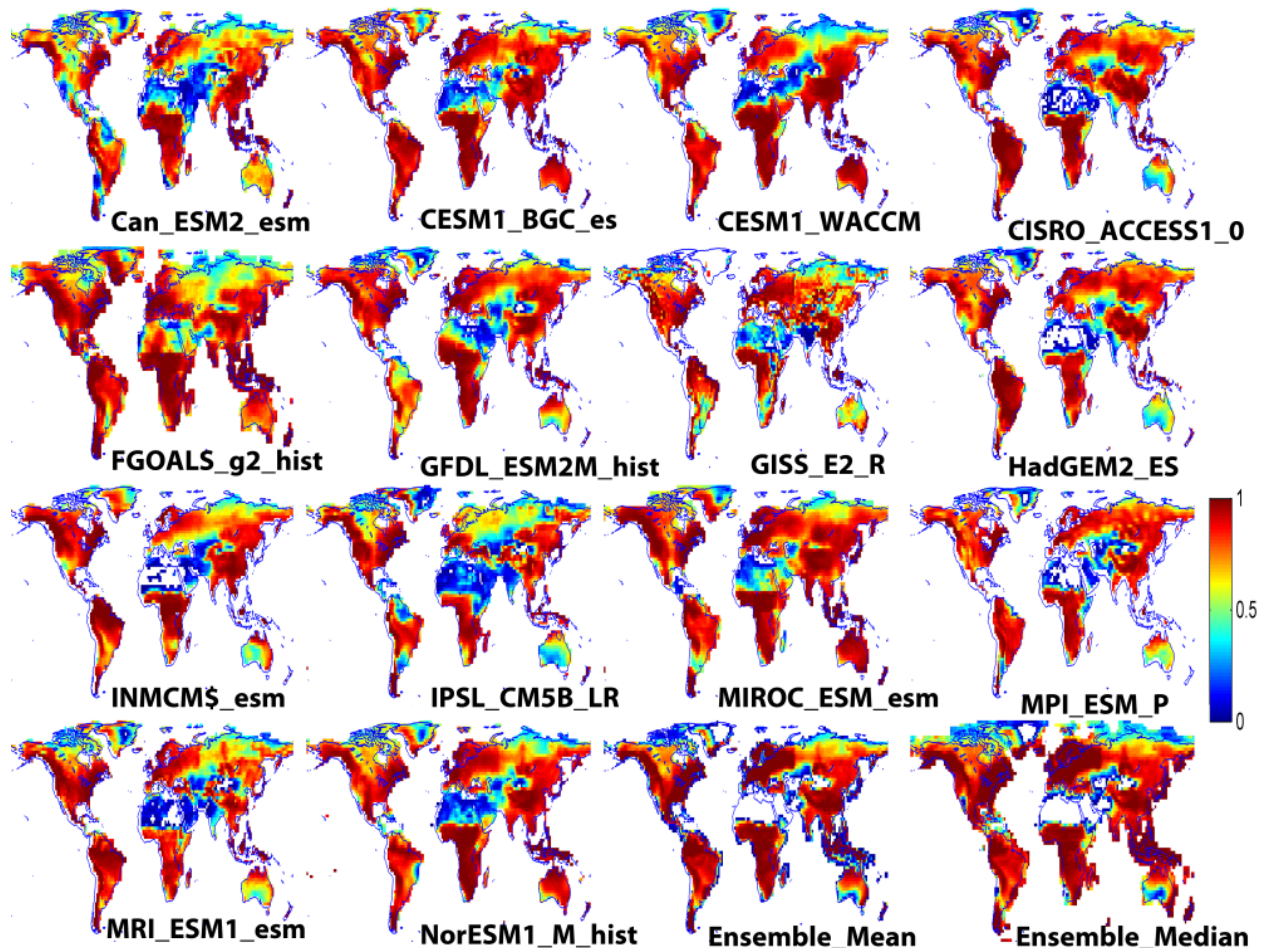


Figure 3.3 Volumetric Hit Index (VHI) for $t = 0$ (with optimal value = 1) maps for selected CMIP5 precipitation simulations, and for their ensemble mean and median, all with respect to GPCP observations.

Bias and Quantile Bias describe the overall ratio of simulations over observations, and do not provide information on the grid-scale matching of simulated precipitation relative to missed precipitation based on reference observations. The Volumetric Hit Index (VHI ; [AghaKouchak *et al.*, 2011]), which measures the volume of precipitation above the threshold (t) detected correctly by climate models with respect to the total simulated and missed precipitation (based on GPCP), can provide such a measure of model performance [Mehran and Aghakouchak, 2014]. For $t = 0$ and $t > 0$, the VHI in each 2×2 degree grid box is defined as [AghaKouchak *et al.*, 2011]:

$$\text{for } t=0, \quad VHI = \frac{\sum_{i=1}^n (P_{CMIP5} | P_{CMIP5} > 0 \ \& \ P_{GPCP} > 0)}{\sum_{i=1}^n (P_{CMIP5} | P_{CMIP5} > 0 \ \& \ P_{GPCP} > 0) + \sum_{i=1}^n (P_{GPCP} | P_{GPCP} > 0 \ \& \ P_{CMIP5} = 0)} \quad (3.3)$$

$$\text{for } t > 0, \quad VHI = \frac{\sum_{i=1}^n (P_{CMIP5} | P_{CMIP5} \geq t \ \& \ P_{GPCP} \geq t)}{\sum_{i=1}^n (P_{CMIP5} | P_{CMIP5} \geq t \ \& \ P_{GPCP} \geq t) + \sum_{i=1}^n (P_{GPCP} | P_{GPCP} \geq t \ \& \ P_{CMIP5} < t)}$$

where P_{CMIP5} = CMIP5 simulations, P_{GPCP} = GPCP observations, n = number of exceedances above threshold t . The ideal VHI score is 1, indicating perfect simulation skill, while 0 corresponds to no skill. In this study, VHI is computed for the entire distribution of precipitation ($t = 0$) and for values above the 75th percentile of the observations ($t \geq 75$ percentile of GPCP). Figure 3.3 presents the VHI when all precipitation data are included in the analysis ($t = 0$) – white areas indicate no data below the choice of threshold t in either climate model simulations or observations. There is generally good agreement between model simulations and GPCP observations over many areas, especially in moist tropical regions such as Amazonia and southern Africa, and in temperate latitudes of Eurasia and North American, consistent with the findings of [Liu *et al.*, 2012]. However, there are obvious discrepancies over arid regions, especially northern Africa and the Middle East, but also the southwestern U.S. and Australia. From Figure 3.3, the ensemble mean and median appear to be superior to the individual climate models in reproducing the main GPCP precipitation patterns (see also Table 3.1 for global-average VHI values of the CMIP5 models and their ensemble mean and median).

Table 3.1 CMIP5 climate models and summary statistics of global Bias (B) and Volumetric Hit Index (VHI) before and after mean-field bias adjustment (after bias adjustment $B=1$ for all models). Optimal values of these metrics are all equal to 1.

Climate Models	Original Data		After Bias Adjustment
	B	VHI	VHI
BCC-CSM1-1	0.97	0.71	0.79
CanESM2_esm	0.76	0.66	0.76

CanESM2	0.77	0.66	0.76
CCSM4	1.08	0.78	0.81
CESM1-BGC_esm	1.06	0.78	0.81
CESM1-CAM5	0.94	0.72	0.79
CESM1-WACCM	1.04	0.73	0.79
CNRM-CM5	0.9	0.71	0.79
CSIRO-ACCESS1-0	0.91	0.73	0.79
CSIRO-Mk3-6-0	0.93	0.69	0.77
FGOALS-g2	1.1	0.78	0.81
FGOALS-s2	0.83	0.67	0.76
GFDL-CM3	0.92	0.72	0.79
GFDL-ESM2G_esm	1.04	0.73	0.78
GFDL-ESM2M_esm	1.05	0.73	0.78
GFDL-ESM2M	1.06	0.74	0.78
GISS-E2-H	0.96	0.69	0.78
GISS-E2-R	0.96	0.71	0.79
HadGEM2-CC	0.88	0.71	0.78
HadGEM2-ES_esm	0.91	0.73	0.79
HadGEM2-ES	0.9	0.72	0.78
INMCM4_esm	0.93	0.74	0.8
IPSL-CM5A-LR_esm	0.77	0.67	0.77
IPSL-CM5A-LR	0.77	0.67	0.77
IPSL-CM5A-MR	0.79	0.68	0.77
IPSL-CM5B-LR	0.83	0.62	0.74
MIROC5	0.99	0.73	0.79
MIROC-ESM_esm	1.14	0.78	0.81
MIROC-ESM	1.17	0.78	0.82
MPI-ESM-LR_esm	1.15	0.78	0.82
MPI-ESM-P	0.88	0.72	0.79
MRI-CGCM3	0.88	0.73	0.79
MRI-ESM1_esm	0.84	0.68	0.77
NorESM1-M	0.85	0.69	0.77
Ensemble Mean	0.89	0.69	0.87
Ensemble Median	0.85	0.76	0.84

It is acknowledged that the *VHI* is a skill score that assumes that a “perfect” model will be able to recreate the observations at each grid. However, the CMIP5 historical runs have large internal variability in precipitation, and one cannot expect the models to precisely reproduce precipitation observations, since they are not forced with prescribed, historical sea surface temperatures. Nevertheless, the *VHI* still provides valuable information as to what extent model simulations can capture historical satellite-based gauge-adjusted observations. Here, the main purpose of using *VHI* is to show whether climate models, relative to each other, are consistent with

observations. As shown, over many regions, many climate model simulations exhibit high VHI scores indicating reasonable consistency with observations.

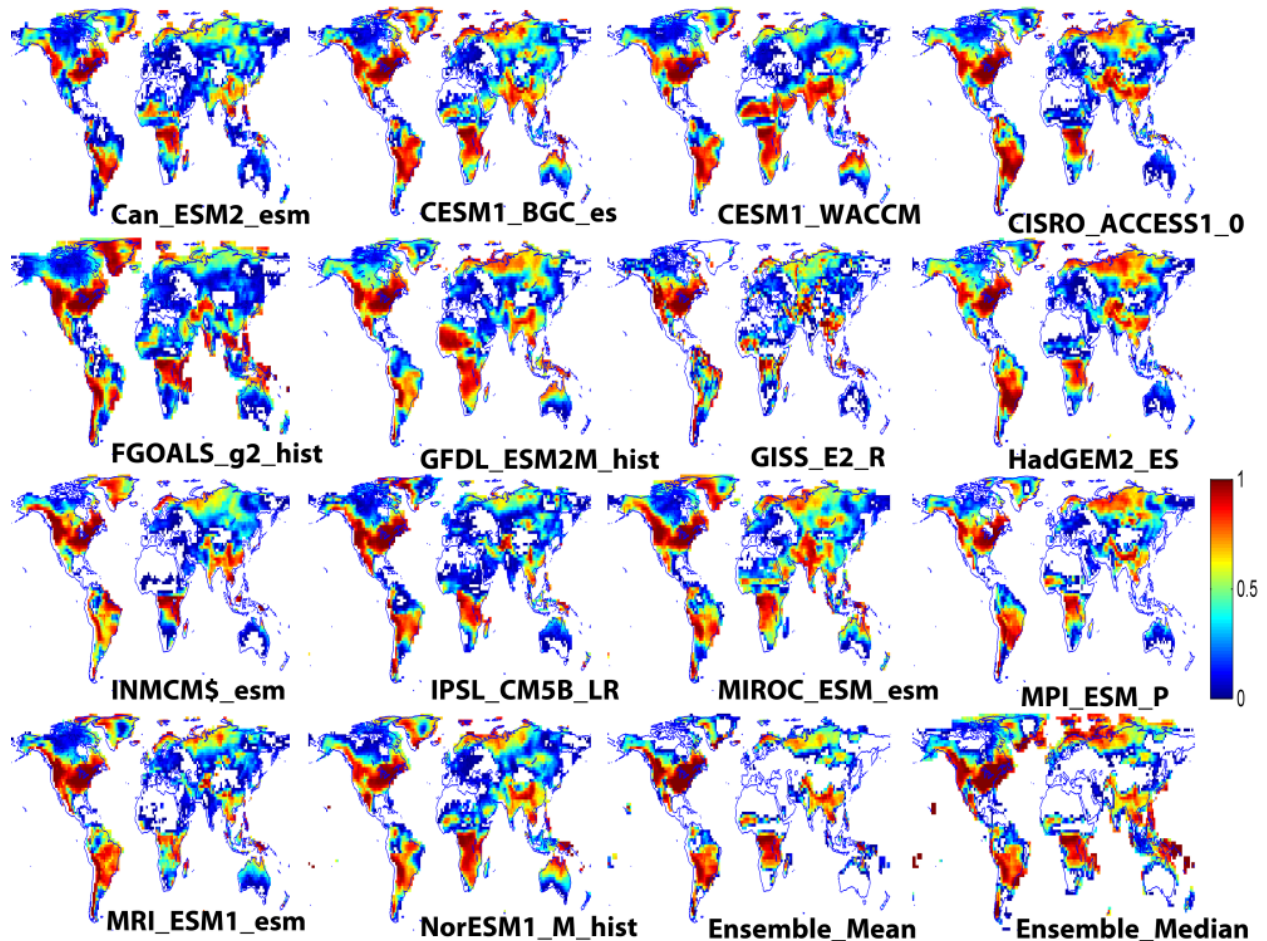


Figure 3.4 Volumetric Hit Index for $t = 75$ percentile VHI_{75} (with optimal value = 1) for selected CMIP5 precipitation simulations, and for their ensemble mean and median, with respect to GPCP observations. Figure 3.4 displays VHI for the 75th percentile threshold of the observations (hereafter, VHI_{75}).

When considering only data above this threshold, the performance of most models is seen to decrease substantially, indicating the presence of systematic biases in the CMIP5 simulations at higher quantiles. The VHI_{75} maps show that except over parts of high-latitude Eurasia, temperate North America, the lower Amazon, southeast Asia and central Africa, the model simulations lack skill above the 75% quantile. This is even more so for VHI at the 90th percentile threshold (figure not shown). It should be emphasized, however, that a low value of VHI (or VHI_{75}) does not

necessarily imply the absence of simulated precipitation, but only that the models are simulating amounts below the local threshold (here, the 75th percentile) of the GPCP reference observations. The results of both the VHI and QB_t analyses confirm that there are biases in climate simulations of precipitation at higher quantiles, implying that more effort should be focused on improving precipitation physics in climate models, so as to more realistically simulate extreme values.

As an alternative to model physics improvements, bias adjustment algorithms have been developed to bring climate simulations closer to observational reference data (e.g., [Christensen et al., 2008; Li et al., 2010; Dosio and Paruolo, 2011; Haerter et al., 2011; Xu and Yang, 2012]). Bias-adjustment of global model simulations is necessary in order to supply more realistic estimates of precipitation (or other climate variables) to regional-scale models that can assess the impacts of climate on hydrology or agriculture, for example. The most common approach involves the removal of the mean-field bias of climate simulations relative to a given observed data set. Figure 3.5 plots global average global bias values of all 35 evaluated CMIP5 models for the thresholds of $t=0$ (all data), $t=75\%$ (QB_{75}) and $t=90\%$ (QB_{90}) percentiles of observations, both before and after removal of the mean-field bias. The models are sorted based on their overall bias values for better visualization. Here the mean-field bias is removed by multiplying the inverse of Equation (3.1) by the original CMIP5 simulations. Considering all the data, the models exhibit global-average biases B between 0.75 to 1.25 (solid black line in Figure 3.5), and after mean-field adjustment, this overall bias can be removed (dashed black line in Figure 3.5). However, while the mean-field bias adjustment eliminates the overall bias, the figure indicates that this adjustment does not necessarily reduce the bias associated with a particular high or low quantile. Figure 3.5 instead confirms that a simple mean-field bias adjustment only marginally reduces such quantile biases.

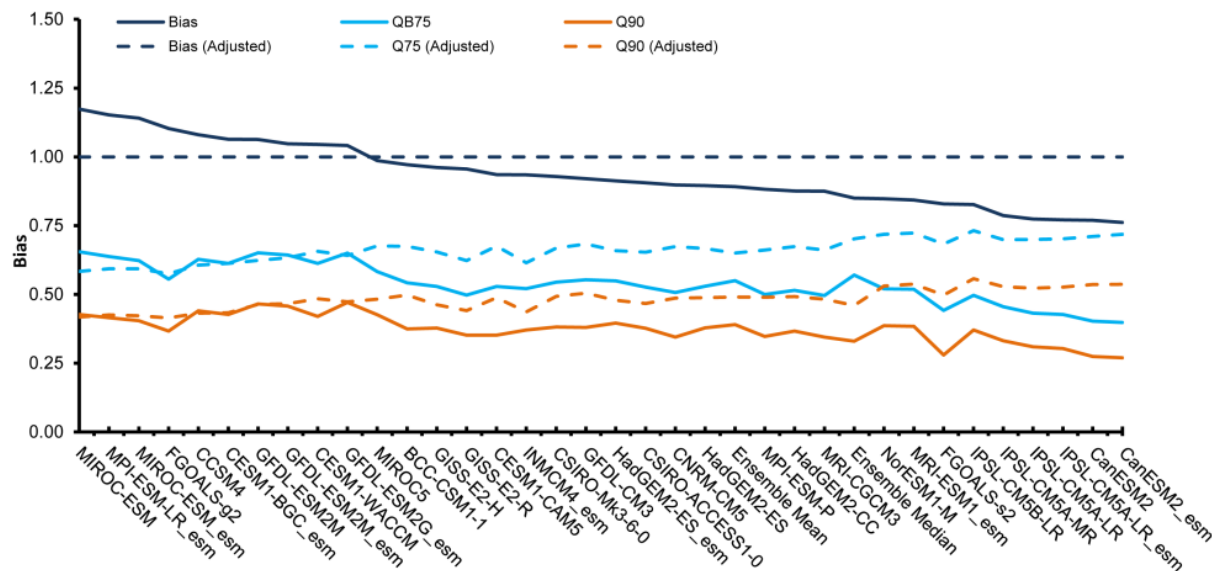


Figure 3.5 Averaged Bias (B) and 75th and 90th Quantile Bias values (QB_{75} and QB_{90} , all with optimal values = 1) for selected CMIP5 precipitation simulations, and for their ensemble mean and median, both before bias adjustment (solid lines) and after (dashed lines), all with respect to GPCP reference data.

In order to compare the spatial patterns of bias after mean-field bias adjustment, Figure 3.6 displays QB_{75} after removing the mean-field bias of model simulations with respect to GPCP data (the bias-adjusted version of Figure 3.2). One can see that, while some improvements in QB can be achieved through such adjustment, this is not the case over several areas such as portions of Australia, Africa, Eurasia, and North America. Thus, on average, CMIP5 models underestimate high quantiles of precipitation even after mean-field bias adjustment. This result underscores the importance of developing more sophisticated precipitation bias-adjustment techniques, such as [Watanabe *et al.*, 2012] and [Mehrotra *et al.*, 2006], that go beyond consideration of only the mean statistics.

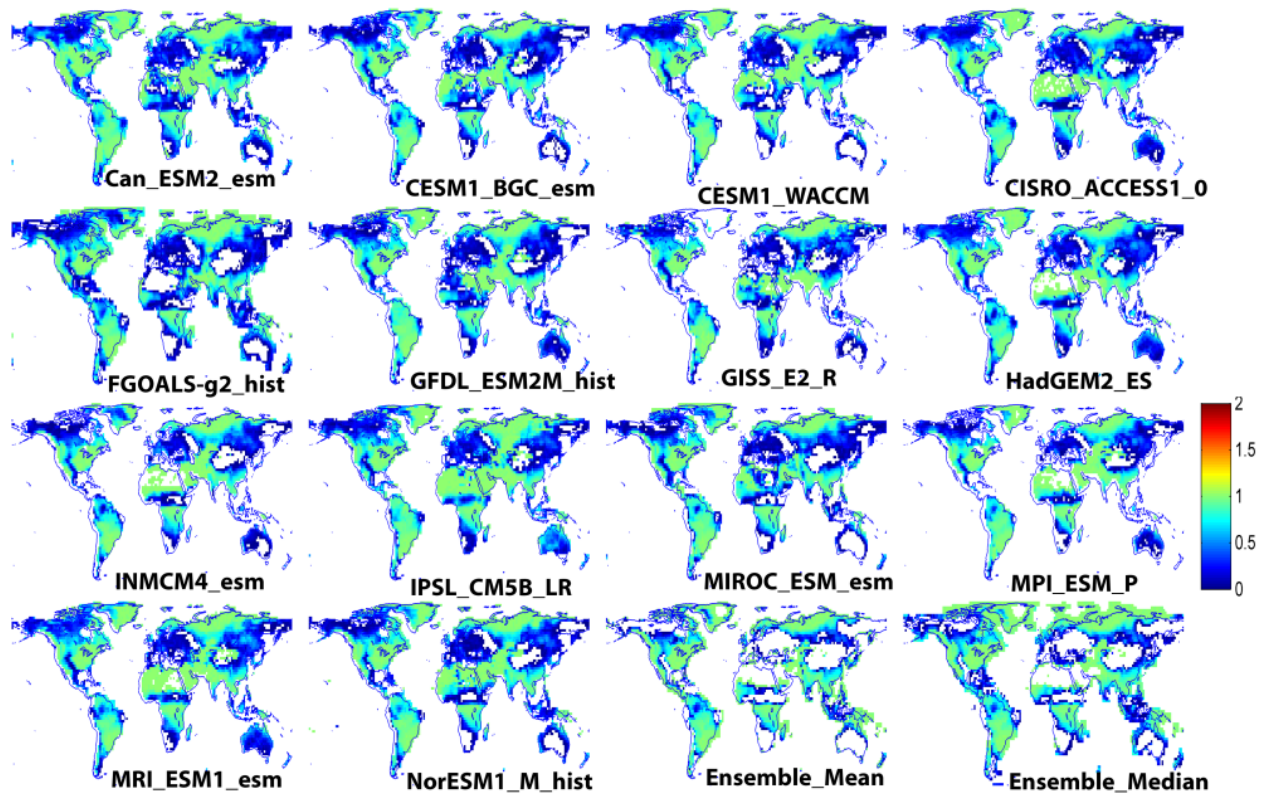


Figure 3.6 Quantile Bias ($t=75$ th percentile, QB_{75} with optimal value =1) maps after bias-adjustment of CMIP5 precipitation simulations, with respect to GPCP observations.

Figure 3.7 displays the globally averaged VHI , VHI_{75} and VHI_{90} of climate model simulations against GPCP data before and after removal of the mean-field bias. The models are ranked on the x-axis based on their overall VHI for better visualization. Similar to the case of QB , it is seen that VHI values drop substantially as the threshold increases, but that the mean-field bias adjustment significantly improves the VHI of the ensemble median. Unlike the bias metrics of Figure 3.5, the VHI values of the ensemble median in Figure 3.7 are consistently higher than those of the ensemble mean.

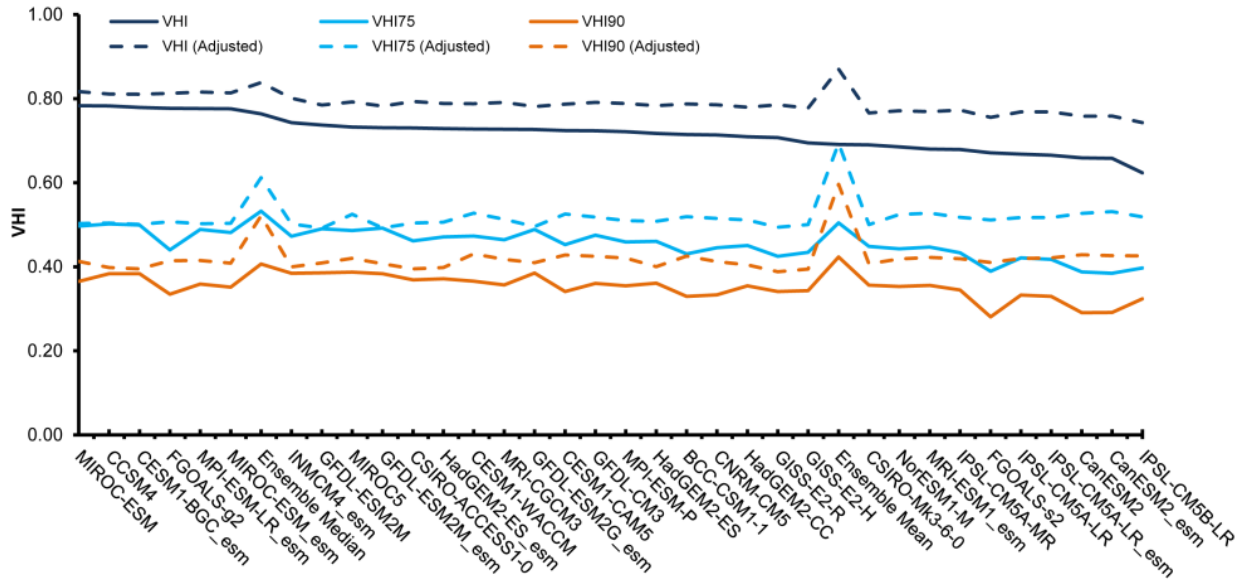


Figure 3.7 Averaged Volumetric Hit Index (VHI) and 75th and 90th percent quantile values VHI_{75} and VHI_{90} (with optimal values = 1) for selected CMIP5 precipitation simulations, and for their ensemble mean and median, both before bias adjustment (solid lines) and after (dashed lines), all with respect to GPCP reference data.

Figure 3.8 displays the spatial patterns of VHI_{75} of climate model simulations against GPCP data after mean-field bias removal (i.e. Figure 3.8 depicts results similar to those shown in Figure 3.4, but after bias adjustment). Compared with Figure 3.4, the VHI values in Figure 3.8 improve over certain areas, such as the United States, Amazonia, and Southeast Asia. In these regions, models such as FGOALS_g2_hist, GFDL_ESM2M_hist, CESM1-BGC_esm and MIROC show high (~ 1) values of VHI_{75} . Overall, the VHI_{75} values of the ensemble mean and median are superior to those of individual models.

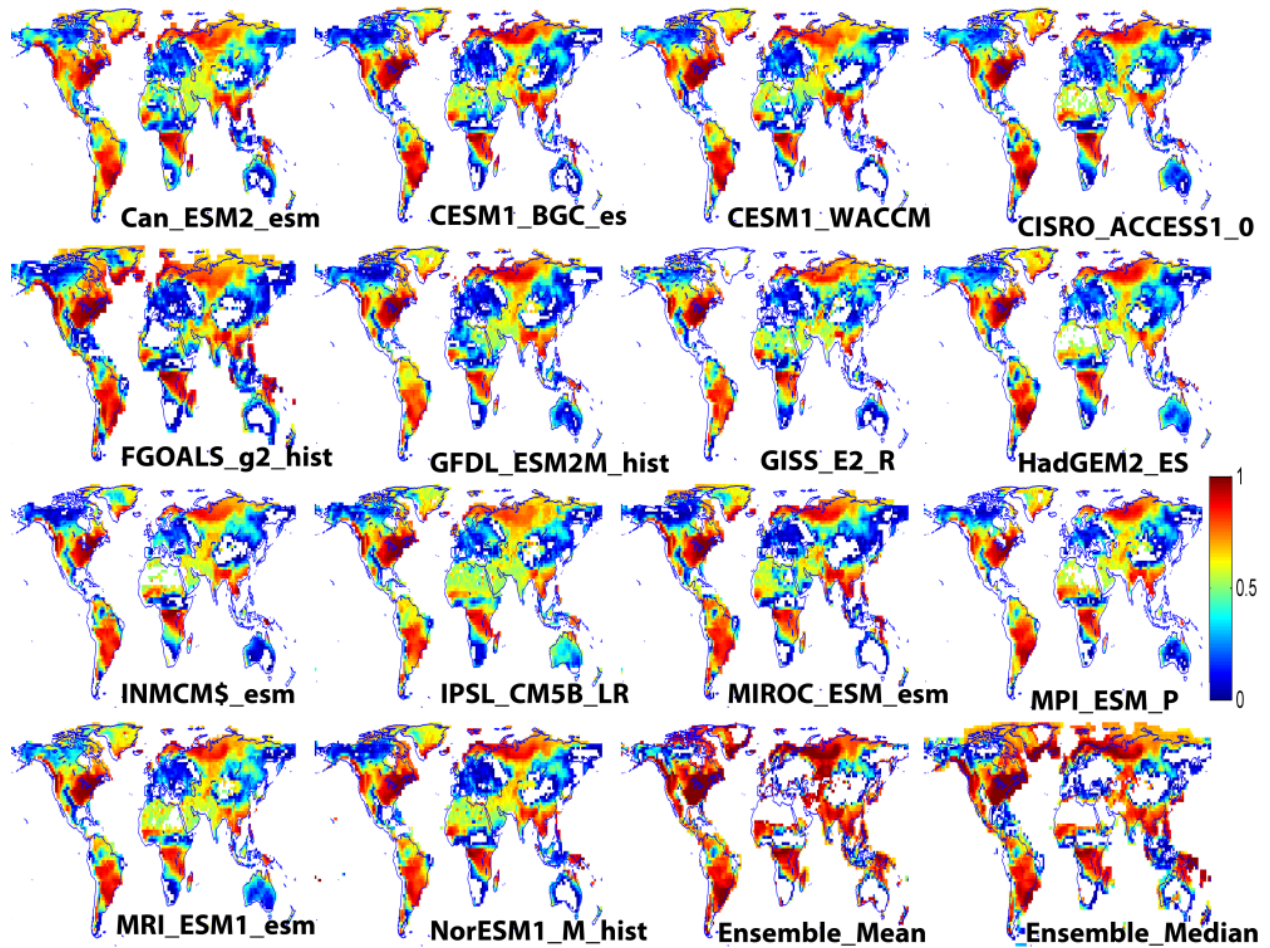


Figure 3.8 Volumetric Hit Index for $t = 75$ th percentile (VHI_{75} , with optimal value = 1) after bias-adjustment of selected CMIP5 precipitation simulations, all with respect to GPCP observations.

3.4 Conclusions

Reliable estimates of precipitation are essential for both research and practical applications. CMIP5 climate simulations provide both historical simulations and future projections of climate variables. Numerous studies have highlighted that climate simulations are subject to various biases and uncertainties (e.g., [Maurer *et al.*, 2007]). The objective of this study is to cross-validate CMIP5 historical simulations of precipitation relative to GPCP reference data, quantifying model pattern discrepancies (VHI metric) and biases (B and QB metrics) for both

entire data distributions and their upper tails. It is acknowledged that observational data sets, like model simulations, are also subject to uncertainties, including systematic and random sampling errors [Aghakouchak *et al.*, 2012]; over land, however, the GPCP data set is bias-adjusted using thousands of rain gauges [Adler *et al.*, 2003], and hence it should serve as an a suitable reference for evaluation of continental precipitation in climate models.

From the results of the Volumetric Hit Index (*VHI*) pattern analysis of the total monthly precipitation amounts, it is found that most CMIP5 simulations are in fairly good agreement with GPCP observations in many areas, but model replication of observed precipitation patterns over deserts and certain sub-continental regions (e.g. northern Eurasia, central Australia) is problematical. The *VHI* of the multi-model ensemble mean and median also are found to be superior to most CMIP5 model simulations overall.

Analyses of total biases (*B*) in CMIP5 simulations reveal that most models overestimate precipitation in regions of steep topography, while underestimating it leeward of the mountains, as well as over many other arid regions. Moreover, while most climate model simulations show low *B* values over Europe, there are considerable inter-model variations in bias over Australia and Amazonia.

At high quantiles ($> 75\%$ and $> 90\%$) of the distribution of monthly precipitation, the Quantile Bias (*QB*) analyses indicate that CMIP5 simulations show more glaring discrepancies in precipitation amounts with respect to the GPCP satellite observations. While continuing to overestimate precipitation in regions of steep topography, the models generally underestimate it in tropical locations such as Amazonia, central Africa, and southern Asia, as well as in broad swaths of the extra-tropics such as Australia, the arid regions of northern Africa and central Asia, and northern China, Russia, and Canada. At high precipitation quantiles also, the CMIP5 models

show substantially reduced agreement with the patterns of the GPCP reference data (e.g., the VHI_{75} metric for precipitation above the 75 % quantile $\ll VHI$ over the entire distribution). Except over North America, Amazonia, and central Africa, most CMIP5 simulations are lacking in predictive skill ($VHI_{75} \sim 0$) for the higher tail of the precipitation distribution. Note, however, that a low VHI_{75} does not necessarily imply the absence of locally simulated precipitation, only that its amount falls below the given reference data's threshold value. In addition, the ensemble-mean and median precipitations at the higher quantiles are found to be superior to the individual climate model simulations when evaluated by the VHI_{75} , and VHI_{90} pattern measures, but not by the QB_{75} and QB_{90} bias metrics.

These results thus suggest that, while today's climate model simulations are generally in agreement with satellite-based gauge-adjusted estimates of total monthly precipitation in many areas, they presently are not well suited for simulating upper quantiles of the precipitation distribution. Such distribution errors, which have persisted across CMIP phases [*Dai*, 2006; *Sun et al.*, 2007; *Sillmann et al.*, 2013], are often characterized by a general tendency for the models to precipitate too frequently in light amounts, but too rarely in the intense downbursts that are occasionally observed. In focusing on the upper tails of the precipitation distributions, the present study reveals such model intensity errors in a particularly stark way (e.g. in Figures 3.2 and 3.4).

The persistence of these upper-tail errors in all the evaluated CMIP5 simulations is indicative of the presence of general deficiencies in the models. For instance, these systematic precipitation distribution errors do not seem to be very sensitive to inter-model differences in horizontal resolution (e.g. the MRI-ESM1_esm model, with a 160×320 grid, does not clearly outperform other coarser-resolution models in Table 3.1). Substantive differences in error structure also are

not apparent between the “_esm” historical simulations with prescribed CO₂ *emissions* and those with prescribed CO₂ *concentrations*. Thus, it is likely that these systematic precipitation errors are due more to general model shortcomings in representing the dynamics or physics of climatic phenomena than to inter-model differences in greenhouse forcings or horizontal resolution.

For example, an ongoing preoccupation of model developers is to improve sub-grid scale parameterizations of convection, since precipitation errors tend to be especially large in the tropics (e.g. [IPCC, 2007]). It is perhaps less widely appreciated that intense precipitation also often originates in frontal systems [Catto *et al.*, 2012; Pfahl and Wernli, 2012] and that representative climate models tend to underestimate these extra-tropical precipitation events in spite of current abilities to adequately simulate the interaction of dynamics and moisture at model grid scale [Catto *et al.*, 2010, 2013].

An underestimation or incorrect placement of intense tropical and extra-tropical precipitation is also clearly displayed by the CMIP5 simulations analyzed in the present study (e.g. in Figures 3.2 and 3.4). Such ubiquitous precipitation errors suggest that improvements, not only in model convective parameterizations, but also in their representation of sub-grid scale cloud microphysical processes that regulate droplet auto-conversion, accretion, and through-fall [Lebsock *et al.*, 2013; Weverberg *et al.*, 2013], may be essential for better simulation of the observed global precipitation distributions.

Once such enhancements of model physics are in place, increases in model resolution are also likely to contribute to more realistic simulation of precipitation [Wehner *et al.*, 2009; CHAMPION *et al.*, 2011]. This may be especially true in mountainous regions, where an accurate representation of the interaction of complex dynamics and steep moisture gradients is difficult to achieve solely through parameterization (e.g. [Ghan *et al.*, 2002; Qian *et al.*, 2009]).

Finally, this study demonstrates that, while a simple mean-field bias removal enhances the overall B and VHI values, it does not yield much improvement at high quantiles (i.e., QB_{75} , QB_{90} , VHI_{75} , and VHI_{90}). Thus, for purposes of climatic impacts studies, it is important to develop more sophisticated techniques for adjusting the upper-tail biases of global precipitation simulations in order to better replicate observed extreme values.

CHAPTER 4: Improving Global Hydrological Modeling Using a Multi-Objective Calibration Framework

4.1 Introduction

Global fresh water resources are under stress because of population growth, increased demand [Vorosmarty, C. J. Green, P. Salisbury, J. Lammers, 2000], and the potential impacts of the climate change on intensification of the hydrological cycle [Huntington, 2006; Stott *et al.*, 2010]. This has led to a growing interest in quantifying freshwater availability under different climate change and emission scenarios not only at the basin scale, but also at the global scale [Döll *et al.*, 2003]. Global hydrological models (GHM), forced with climate model simulations have been widely used for understanding changes in the hydrologic cycle or assessing freshwater availability [Liang *et al.*, 1994; Vorosmarty, C. J. Green, P. Salisbury, J. Lammers, 2000; Nijssen *et al.*, 2001; Oki *et al.*, 2001; Fekete, 2002; Milly and Shmakin, 2002; Sheffield and Wood, 2007; Palmer *et al.*, 2008; Gong *et al.*, 2009; Werth and Güntner, 2009; Haddeland *et al.*, 2011]. However, GHM simulations are subject to uncertainties arising from input data error, model structure uncertainty and model parameterization [van Beek and Bierkens, 2008; Müller Schmied *et al.*, 2014]. Numerous model development, intercomparison and assessment studies have focused on improving models, sources of uncertainty as well as strength and limitations of the GHMs [Henderson-Sellers *et al.*, 1995; Polcher *et al.*, 1996; Arnell, 1999a; Dirmeyer *et al.*, 1999; Alcamo *et al.*, 2003; Hanasaki *et al.*, 2008; Rost *et al.*, 2008; Haddeland *et al.*, 2011; Hagemann *et al.*, 2013].

In most global hydrologic modeling experiments and climate impact assessment studies, uncalibrated GHMs are used primarily because of complexities in model calibration at the global scale [Meigh and Tate, 2002; van Beek and Bierkens, 2008; Stacke and Hagemann, 2012]. Global gridded models have hundreds to thousands of models that makes calibration and validation extremely challenging. Even in calibrated studies, the focus has primarily been on calibration of GHM parameters based on gridded runoff or observed streamflow records [Döll et al., 2003; Hanasaki et al., 2007; Gosling and Arnell, 2011]. The objective of this type of calibration is to adjust all or some of the model parameters to reproduce acceptable runoff. However, this does not guarantee reliable estimates of the other components of the model (e.g., soil moisture, evapotranspiration, recharge). For a study that focuses on flood forecasting or only streamflow conditions, such single objective calibration may be sufficient. However, for analyzing the global water cycle, other components of the model should be reasonable as well.

Specifically, top soil moisture is a variable that controls infiltration, and also affects evapotranspiration. However, limited soil moisture data is a major roadblock in calibration of model-based soil moisture simulations. In recent years, satellite-based soil moisture products have offered a unique avenue for top layer soil moisture research [Entekhabi et al., 2008; Liu et al., 2011; Kerr et al., 2012]. The Soil Moisture and Ocean Salinity (SMOS; [Kerr et al., 2001]) and the recently launched Soil Moisture Active Passive (SMAP; [Entekhabi et al., 2010b]) provide valuable soil moisture information that can potentially be integrated into model. In this paper, a global hydrologic model is modified so that its top soil moisture later can be adjusted with satellite-based observations. Then, a multi-objective calibration approach [Yapo et al., 1998; Boyle et al., 2000; Vrugt and Robinson, 2007; Vrugt et al., 2009b] is employed for simultaneous calibration of the model based on gridded runoff and top soil moisture information. This

modeling framework leads to major improvements in modeling the components of the water cycle.

4.2 Data Resources

The following data sets are used in this study:

(a) ERA-Interim reanalysis [*Dee et al.*, 2011] daily precipitation and temperature data sets are used as inputs for global hydrologic modeling. The ERA-Interim reanalysis uses a data assimilation approach that includes observations with prior information from a dynamical forecast model to simulate the evolving state.

(b) Essential Climate Variable (ECV) multi-sensor satellite-based soil moisture data [*Wagner et al.*, 2012], available from European Space Agency, has been used as the reference data for calibration of the top soil moisture information.

(c) Spatially distributed observed daily runoff fields (available from the Global Runoff Data Center, GRDC) is used as the reference data set for calibration of gridded runoff. GRDC's Composite Runoff Fields V_{1.0} [*Fekete*, 2002] is calculated based on a large number of selected gauging stations from around the world [*Vorosmarty, C. J. Green, P. Salisbury, J. Lammers*, 2000]. This reference data set is only available for 1986 to 1995 and at a 0.5° spatial resolution.

Given that the reference gridded runoff is only available for 1986 to 1995, the model has been calibrated and validated for this period. For consistency, all data sets are regridded onto a 0.5° grid.

4.3 Methodology

In this study, the PCR-GLOBWB [*van Beek et al.*, 2011; *Wada et al.*, 2011a, 2012] hydrologic model is used for the multi-objective calibration experiment. This PCRaster [*Karssenberget al.*, 2001] model is a grid-based global hydrological model based on the leaky bucket concept [*Bergström and Singh*, 1995]. In this study, the top layer is modified as a layer consistent with the microwave-based satellite soil moisture information [*Liu et al.*, 2011; *Wagner et al.*, 2012]. In the version used in this study, the model includes two soil layers on top of the groundwater layer (see Figure 4.1). The stored water in each layer can be exchanged between the layers and the atmosphere. Based on an improved ARNO method [*Hagemann*, 2002], the fraction of saturated soil is determined that is then used for estimation of direct runoff. The actual evaporation is calculated based on potential evaporation and the status of soil moisture [*Allen et al.*, 1998]. For more information about the model, the interested reader is referred to [*van Beek*, 2008; *van Beek and Bierkens*, 2008; *Van Beek et al.*, 2011; *Wada et al.*, 2011b, 2012; *Taylor et al.*, 2013].

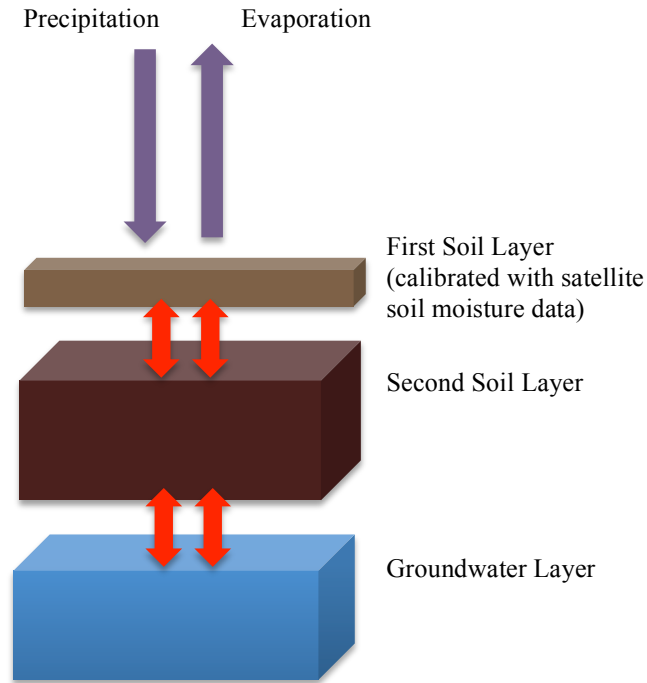


Figure 4.1 Schematic view of a cell in the model forced with precipitation (P) and temperature (T) as inputs. The top soil layer is calibrated with the satellite observations. Stored water in each layer can be exchanged with other layers (red vectors) vertically or evaporate from the soil layers.

In multi-objective calibration, parameter estimation is based on evaluating model simulation results against more than one objective (here, reproducing the observed gridded runoff and satellite-based soil moisture information). In this study, the Multialgorithm Genetically Adaptive Method (AMALGAM; [Vrugt and Robinson, 2007]) multi-objective optimization technique is used for calibrating the global hydrologic model with respect to the satellite soil moisture and gridded runoff information. Evolutionary algorithms such as AMALGAM are able to maintain a diverse population of solutions and exploit any kind of similarities by recombining the populations [Vrugt and Robinson, 2007]. AMALGAM runs multiple evolutionary optimization algorithms at the same time with adaptive offspring creation that helps solving complex problems.

The algorithm starts with a random population (P_0) of size N , which is generated by the Latin hypercube algorithm [Tang, 1993]. Each of the parents in the initial random populations are ranked by using the non-dominated sorting algorithm (FNS) [Deb et al., 2002]. Then, offspring population Q_0 of size N is calculated by implementing AMALGAM method. In AMALGAM, several k individual algorithms (using different adaptive procedures to create offspring from P_0) are used for reproduction of the offspring generation, $Q_0 = \{Q_0^1, Q_0^2, \dots, Q_0^k\}$. A combined ranked population of R_0 of size $2N$ is formed by each of the offspring ($Q_0^1, Q_0^2, \dots, Q_0^k$) and the parents using FNS.

$$R_0 = P_0 \cup Q_0 , \quad (4.1)$$

In the final step, members of the next population P_1 are selected out of the non-dominated fronts of R_0 considering their rank and crowding distance [Deb et al., 2002]. The same procedure will be applied over the new population P_1 to create offspring until convergence is achieved. The method used in adaptive offspring generation, favors the individual algorithms that have the most contribution in offspring reproduction. For this study, four adaptive offspring approaches (NSGA-II [Deb et al., 2002], particle swarm optimization (PSO) [Kennedy et al., 2001], adaptive metropolis search (AMS) [Haario et al., 2001], and differential evolution (DE) algorithms [Storn and Price, 1997]) have been selected based on numerical experiments [Vrugt and Robinson, 2007]. To establish the preference for each of the adaptive procedures, they are weighted (N_t^i) based on the number of offspring members an algorithm contributes to the new population:

$$N_t^i = N \cdot (P_t^i / N_{t-1}^i) / \sum_{i=1}^k (P_t^i / N_{t-1}^i) \quad (4.2)$$

where i is the adaptive procedure rank, t is generation index, P_t^i / N_{t-1}^i is the ratio of the offspring members contributing to the new population and P_t^i is the corresponding number for the previous generation.

The model is calibrated for 21 major basins from across the globe (Figure 4.2). The period of 1986 to 1987 is used for calibration and the rest of the record (1987-1995) for validation. Two objective functions are defined for soil moisture and runoff: (a) the root mean square of monthly soil moisture residuals over each basin; and (b) the sum of squared residuals for annual runoff over each sub-basin. Initially, 40 parameters from each pixel were considered for calibrating the model. Using a sensitivity analysis, the parameters were reduced to 12 parameters to which the model showed more sensitivity including saturated and unsaturated soil parameters. The top soil layer in the model is forced to have a thickness similar to the penetration depth of the microwave in the soil which is typically less than 0.2_m [Loew *et al.*, 2013]. With this modification, the top soil moisture estimates will be consistent with the satellite observations.

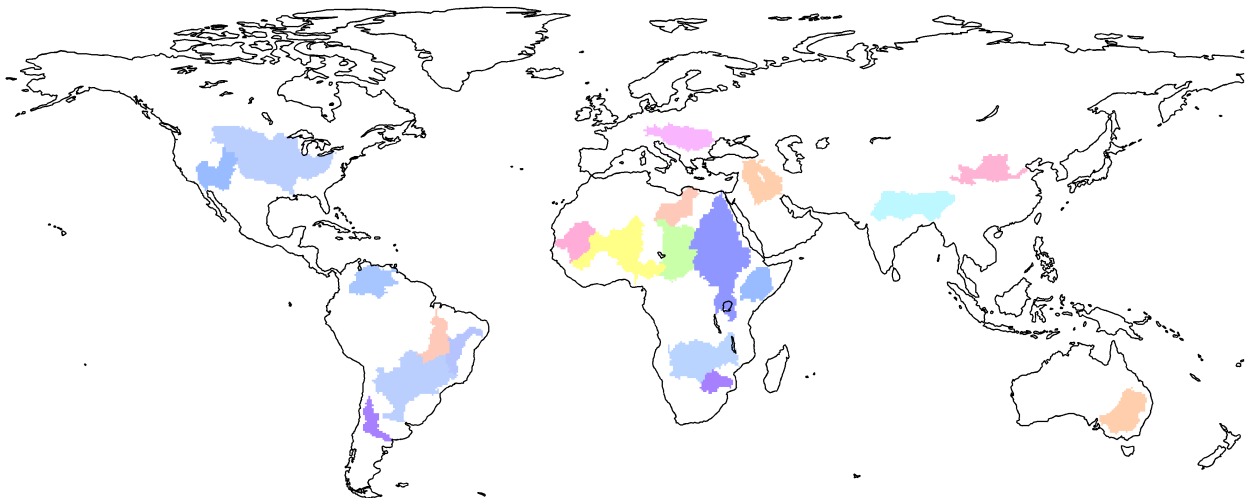


Figure 4.2 The locations of the 21 major basins from around the globe used for calibration and validation.

4.4 Results

For the selected basins shown in Figure 4.2, the global hydrologic model is used to simulate gridded runoff and soil moisture after the calibration procedure. The results including monthly

runoff and soil moisture simulations are provided in Figures 4.3 and 4.4, respectively. One can see that the model simulations overestimate the runoff in many of the basins including in Mississippi, Nile, Colombia and Murray-Darling basins. After model calibration, one can see that the simulated runoff improve substantially (Figure 4.3).

It is interesting to note that before calibration, the model overestimates the soil moisture in many of the basins (Figure 4.4). Calibration of soil moisture results in model simulations consistent with the satellite observations. We acknowledge the uncertainties associated with satellite-based soil moisture estimates. However, as shown in Figure 4.4, having a calibration framework that allows integration of soil moisture in model calibration will significantly improve terrestrial water budget modeling.

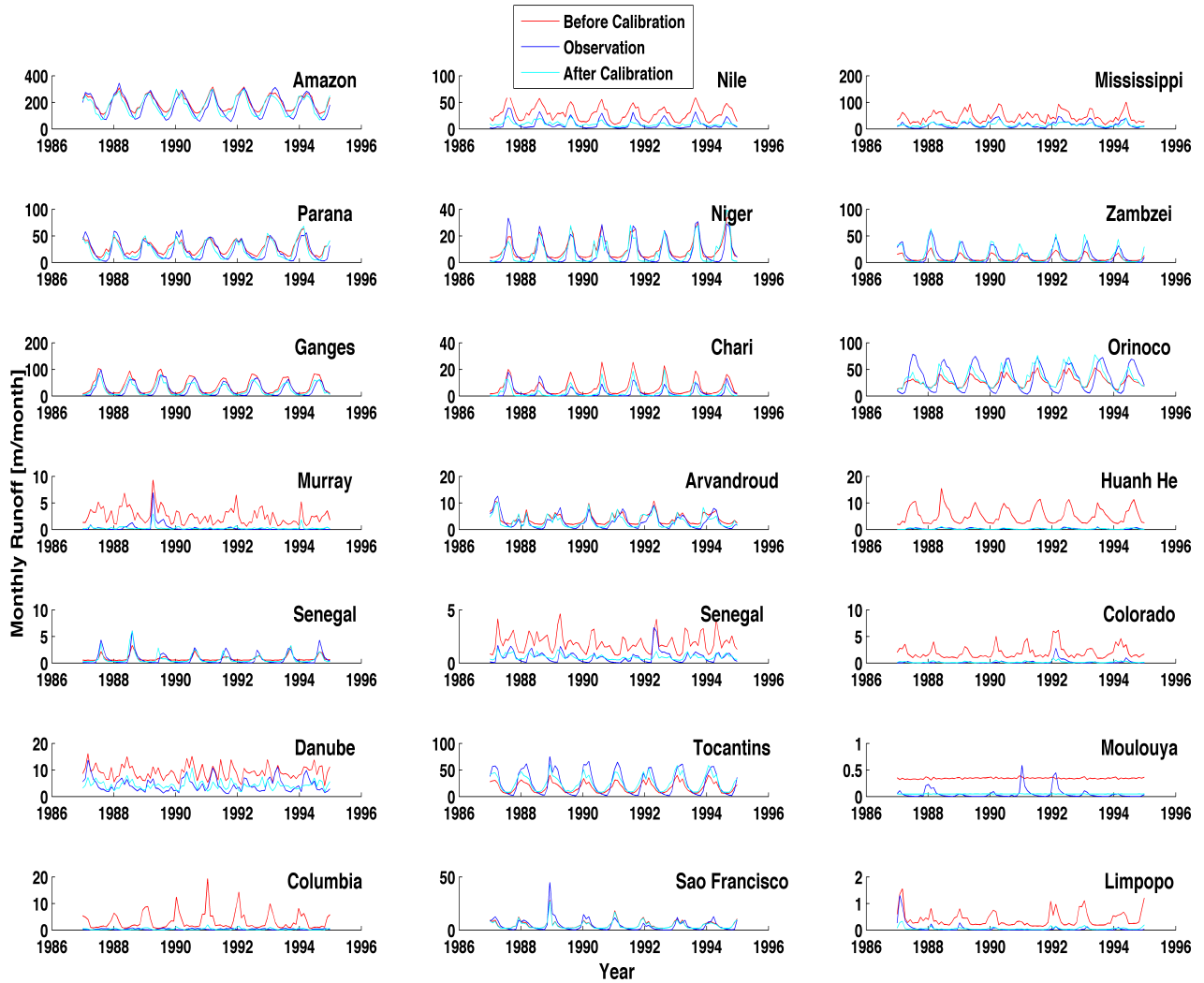


Figure 4.3 Monthly runoff [m/month] simulations before and after calibration against composite gridded runoff observations (1987 - 1995).

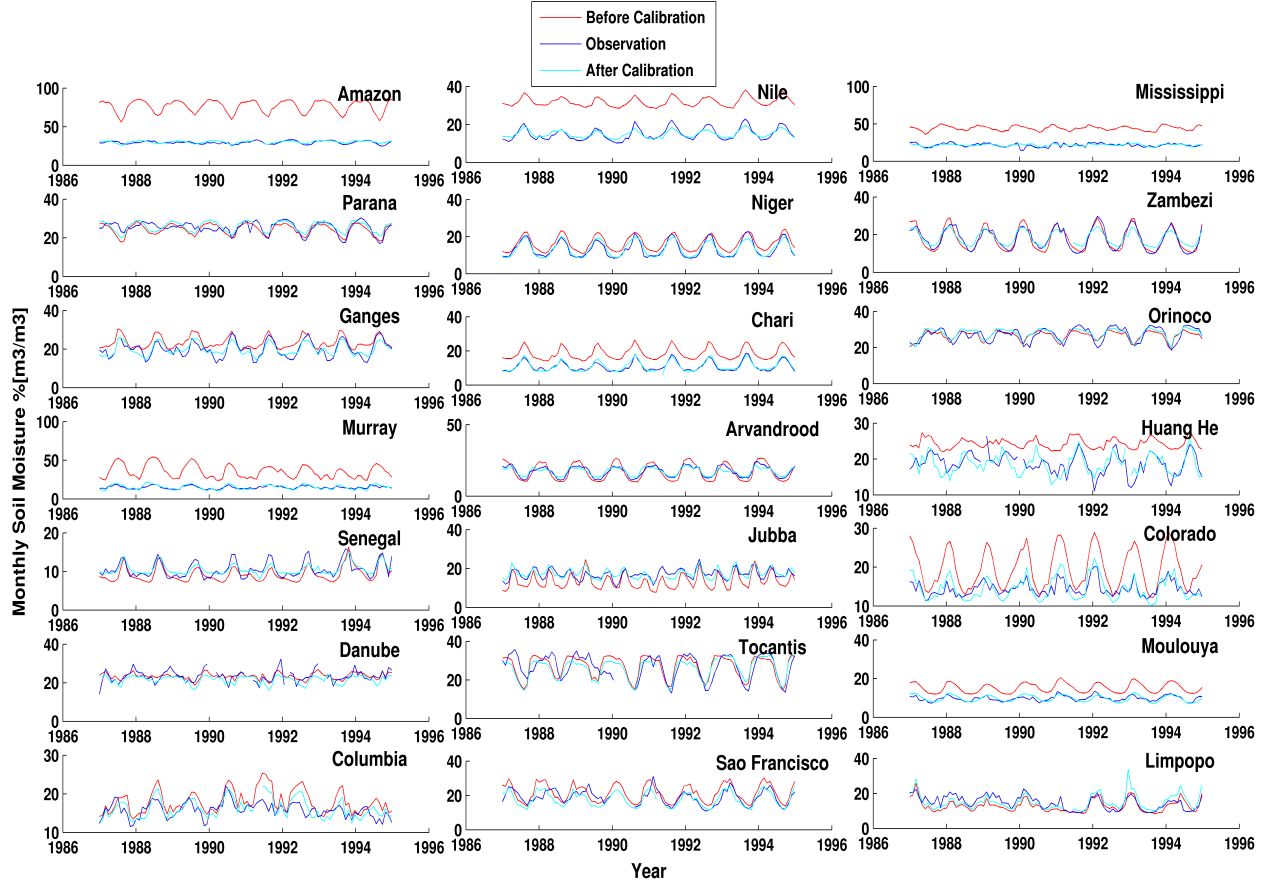


Figure 4.4 Monthly soil moisture [m^3/m^3] before and after calibration against satellite-based soil moisture observations (1987 - 1995).

Figures 4.5, 4.6 and Table 4.1 summarize the statistical performance metrics for before calibration, after calibration, and the validation period, respectively. Three metrics of Correlation Coefficient, Relative Error, and Nash-Sutcliffe are used for model evaluation:

$$\text{Correlation Coeff} (sim, obs) = \frac{cov(sim, obs)}{\sigma_{sim}\sigma_{obs}} \quad (4.3)$$

$$\text{Relative Error}(sim, obs) = 100 \times \frac{1}{n} \times \sqrt{\frac{\sum_{t=1}^n (sim_t - obs_t)^2}{\sum_{t=1}^n (obs_t)^2}} \quad (4.4)$$

$$\text{Nash - Sutcliffe}(sim, obs) = 1 - \frac{\sum_{t=1}^n (sim_t - obs_t)^2}{\sum_{t=1}^n (obs_t - \bar{obs})^2} \quad (4.5)$$

where, *sim* represents model simulations (i.e., runoff, soil moisture), and *obs* corresponds to the observation datasets (i.e., composite runoff observation, satellite soil moisture data), *n* is the length of dataset in month, and *t* is the time-step.

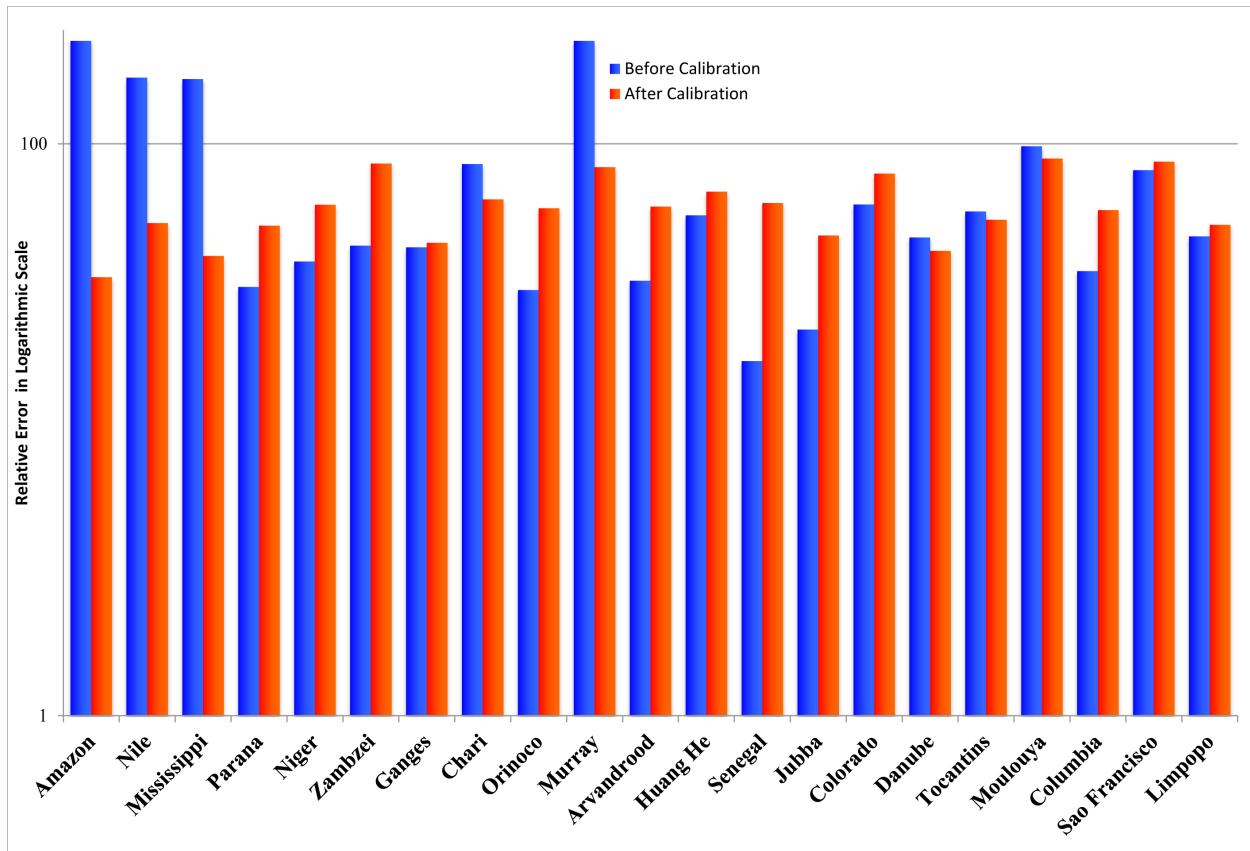


Figure 4.5 Relative errors of soil moisture simulations versus the satellite-based observations before calibration and after calibration (results shown in logarithmic scale for better visualization).

A comparison of the performance metrics presented in Figures 4.5 and 4.6 clearly shows improvements gained by calibrating the model for both runoff and soil moisture. For better visualization, the results presented in Figures 4.5 and 4.6 are provided in logarithmic scale, since the relative error of uncalibrated runoff for several basins including Murray-Darling, Arvandrood and Huang He were very large

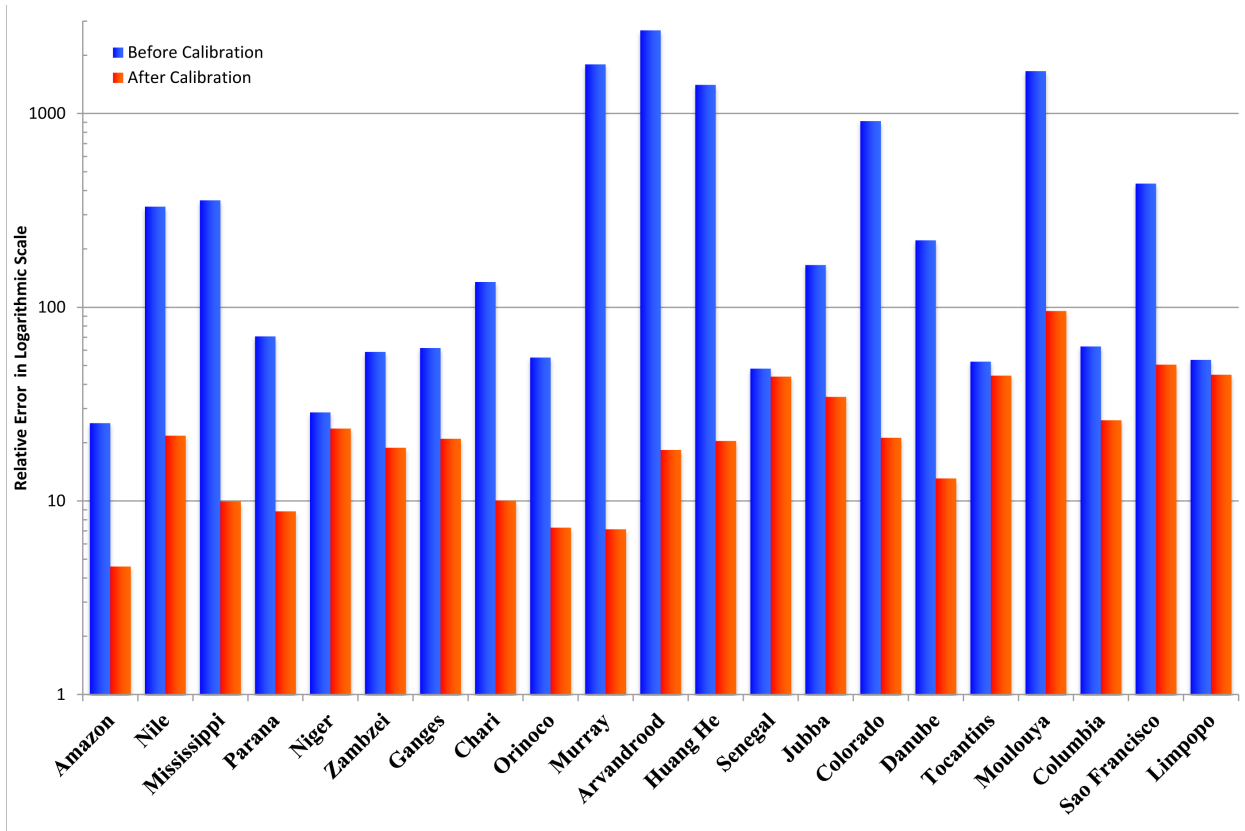


Figure 4.6 Relative errors of the simulated runoff against the global gridded runoff observations before and after calibration (results shown in logarithmic scale for better visualization).

Table 4.1 Correlation coefficients, relative errors, and Nash-Sutcliff coefficients for runoff and soil moisture simulations relative to the observations in the validation period (Jul. 1986 to Dec. 1995).

	After Validation					
	Runoff			Soil moisture		
	Corr.	R.E.	N.S.	Corr.	R.E.	N.S.
Amazon	0.824	2.503	0.679	0.831	0.418	0.528
Nile	0.921	4.634	0.614	0.990	1.706	0.819
Mississippi	0.863	3.447	0.595	0.922	0.422	0.818
Parana	0.818	3.812	0.663	0.837	0.653	0.444
Niger	0.879	4.049	0.772	0.988	1.787	0.971
Zambzei	0.861	4.417	0.715	0.989	1.071	0.900
Ganges	0.855	3.969	0.730	0.963	1.576	0.899
Chari	0.834	5.123	0.671	0.994	3.012	0.981
Orinoco	0.695	4.502	0.480	0.905	0.561	0.787
Murray	0.516	6.763	0.250	0.988	0.624	0.656
Arvandrood	0.774	5.380	0.576	0.315	4.240	0.314
Huang He	0.935	2.596	0.847	0.799	1.914	0.598
Senegal	0.919	4.084	0.786	0.865	6.465	0.697
Jubba	0.834	4.670	0.524	0.832	5.317	0.655
Colorado	0.605	5.656	0.350	0.943	3.932	0.220
Danube	0.290	4.478	0.069	0.712	1.737	0.122
Tocantins	0.902	3.537	0.752	0.796	8.043	0.567
Moulouya	0.279	8.382	0.015	0.821	31.386	0.411
Columbia	0.475	5.226	0.225	0.940	7.090	0.883
Sao Francisco	0.944	3.175	0.884	0.801	30.938	0.595
Limpopo	0.891	3.709	0.759	0.860	14.069	0.445

It should be noted that in multi-objective calibration, there are often tradeoff between the best objective function for one variable and the best for all considered variables (here, runoff and soil moisture). The relative errors reported in Figures 4.5, and 4.6 and Table 4.1, however, shows that model simulations in all basins improve relative to both runoff and soil moisture.

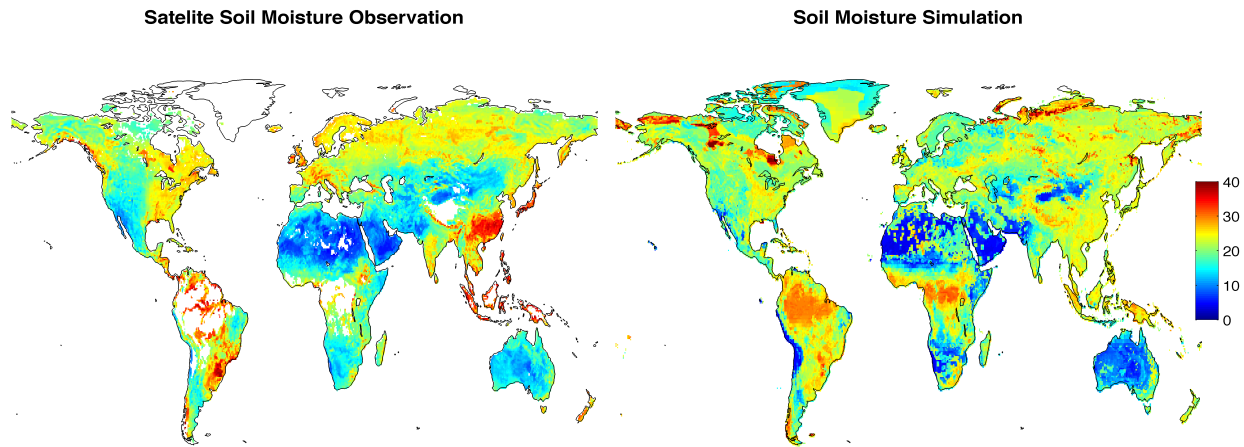


Figure 4.7 Long-term soil moisture climatology (1986-1995) based on satellite observations [m^3/m^3] (left), and model simulations [m^3/m^3] (right).

Figure 4.5 compares the long-term soil moisture climatology based on satellite observations (left) and the calibrated model simulations (right). Figure 4.5 shows very good agreement between model simulations and satellite observations. Note that satellite soil moisture data do not provide soil moisture information in densely vegetated regions including parts of the Amazon basin. The model simulations for the areas of the Amazon for which observations are available show high consistency with the observations. This indicates that the model can be used to derive estimates for the locations where satellite data do not provide observations. Further, comparison of long-term averaged runoff observations against model simulations are displayed in Figure 4.6. Similar to soil moisture estimates, the gridded runoff simulations are consistent with the observations across most of the globe.

In summary, this model reproduces not only observed gridded runoff from ground-based measurements, but also generates soil moisture fields consistent with the satellite observations. Soil moisture affects infiltration, recharge and evapotranspiration. Therefore, reproducing observed soil moisture enhances estimates of the other component of the hydrologic cycle. This

is a unique model that can potentially be used for assessing changes in the water cycle (see Chapter 5).

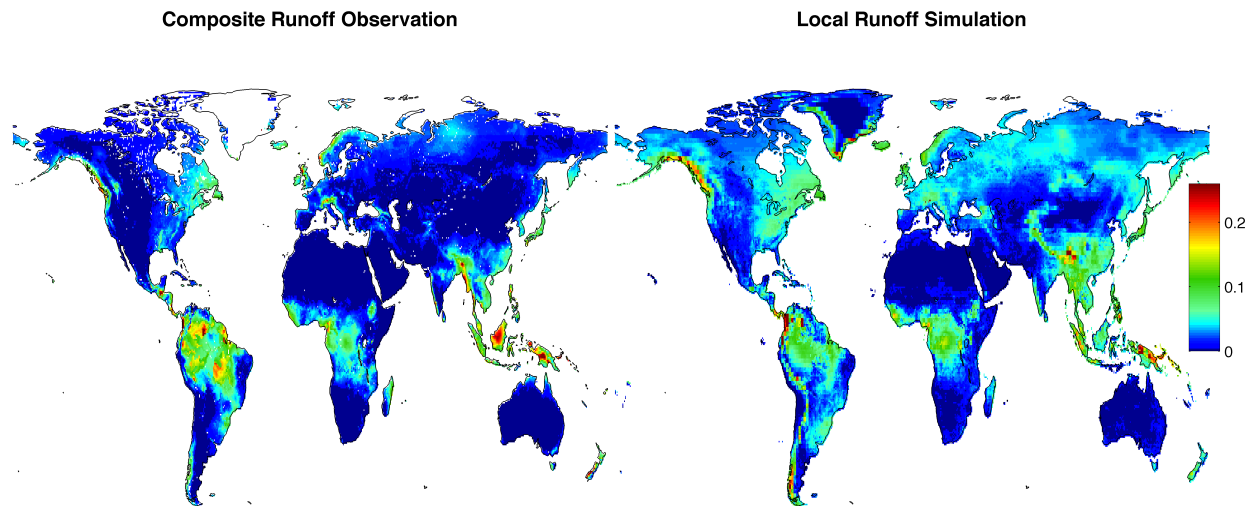


Figure 4.8 Long-term runoff climatology (1986-1995) based on the composite gridded runoff observations [m/month] (left), and model simulations [m/month] (right).

4.5 Summary and Conclusions

One of the limitations of most, if not all, current global hydrological models is that they are either used uncalibrated or only calibrated based on runoff. Using long-term historical satellite-based observations, this study proposes a framework for multi-objective calibration and parameter estimation. In this framework, in addition to reproducing observed runoff, the top layer of the system is calibrated to reproduce satellite observations of the top soil moisture layer. This leads to model simulations calibrated for two key hydrologic variables. The top soil moisture layer is an important element that affects infiltration and evaporation. Having this

secondary constraint based on soil moisture results in improved representation of the hydrologic cycle.

A globally improved consistent simulation of hydrological model was achieved through integrating several key points in a multi-objective calibration approach: 1) considering soil moisture and surface runoff as key factors in the non-saturated and saturated soil layer. 2) Basin-based calibration of the globe by considering 12 most sensitive model parameters. 3) Multi objective calibration approach based on annual surface runoff and monthly soil moisture over each sub-basin. 4) Application of a new multi-objective calibration approach (AMALGAM) to decrease calibration cost (time).

The results show that before calibration, the original model overestimated soil moisture substantially. This issue has significant impacts on assessing climatic change and variability impacts on the water cycle and may lead to misleading results. Applying the multi-objective calibration approach to the global hydrological model resulted in improved simulation results (i.e., significant reduction in the relative errors of both runoff and soil moisture simulations against the observations).

CHAPTER 5: Climate Change Impacts on Water Resources Accounting for Local Resilience

5.1 Introduction

Water resources are sensitive to climate change and variability [Wood *et al.*, 1997b; Trenberth, 2001; Merritt *et al.*, 2006; Sivakumar, 2011; Stoll *et al.*, 2011; Aghakouchak and Mehran, 2013], especially in arid and semi-arid regions [Schlenker *et al.*, 2007; Seager *et al.*, 2007b; Cayan *et al.*, 2010]. Global climate model simulations coupled with regional/global hydrologic models have been widely used to assess changes in water resources [McDonald *et al.*, 2011; Schewe *et al.*, 2014]. However, current modeling frameworks whose output are to issue climate/hydrology projections simulate primarily the natural hydrologic cycle, without consideration man-made infrastructures such as dams and reservoirs that are known to alter significantly the flow and distribution of water [Christensen *et al.*, 2007; Karl *et al.*, 2009]. For instance, about 20% of the global annual river discharge is controlled by man-made surface reservoirs ($\sim 8000 \text{ km}^3$ out of 40000 km^3 ; [Shiklomanov *et al.*, 2000]). Surface water reservoirs not only provide resilience against droughts, but also play a key role in water resource management and energy production [Palmer *et al.*, 2008; Hallegatte, 2009; Madani, 2010]. We argue that changes in future water availability should be presented relative to the local resilience (i.e., primarily surface water reservoirs). The reason is that a system with substantial water storage (which builds extra local resilience) will be less vulnerable to climatic change and variability compared to a system with limited local resilience. This means that in different regions a certain change in water availability will have different impacts depending on the region's local resilience to cope with variability or to adapt to change.

The omission of reservoirs in large scale water cycle models also hinders evaluating the impacts of climate change and variability on hydropower energy production [Tarroja *et al.*, 2014]. More generally, the absence of surface water reservoirs introduces a large source of uncertainty in current assessments of the global water cycle. Indeed, continental scale closure errors of the water budget range between 13% (Europe) to around 21% (Australia) [Trenberth *et al.*, 2007], but can in part be attributed to input data uncertainty, and anthropogenic influences on the water distribution. Reservoirs fundamentally change water availability [Grant *et al.*, 2012] and distribution based on local and regional water demand. In Melbourne, Australia, for example, most of the natural inflow occurs in the period of July-October when the demand of water is relatively low. The water stored in the reservoir during this wet season is subsequently released during summer time when demand significantly exceeds inflow rates. (Compare the natural flow with the outflow of man-made reservoirs in Figure 5.1a). Furthermore, over-year reservoirs [McMahon *et al.*, 2007], designed to satisfy water demands during extended periods of time (e.g. several years) without noticeable precipitation, redistribute water over multiple years.

Thus, an accurate assessment of impact of climate change on water resources availability and distribution necessitates explicit consideration of human influences such as dams and reservoirs (amongst others). Thus, global scale hydrologic models should not be limited to processes such as surface runoff and river discharge, but also include an explicit numerical description of the water balance dynamics of reservoirs (surface and subsurface) and other large water bodies. This necessitates consideration of the movement of water between groundwater and the unsaturated zone, and their interaction with surface water [Reclamation, 2011]. This paper introduces a nested modeling framework that accounts explicitly for (man-made) reservoirs (local resilience) when evaluating the impact of climate change on water resources (see Methods Section). We

focus on the Melbourne metropolitan area in the southeast of Australia where most of the water for consumptive or industrial use comes from large reservoirs in protected areas (Figure 5.1b). In the past century, Melbourne has suffered several major water crisis and severe droughts, the most extreme being the well-known Millennium Drought (1997-2009), which drained the reservoirs and caused major wildfires with significant economic and human losses [*Barker et al.*, 2011; *AghaKouchak and Nakhjiri*, 2012; *Grant et al.*, 2013; *Van Dijk et al.*, 2013]. This Metropolitan area of Melbourne is a quintessential example of a highly regulated environment through a number of reservoirs that enhanced local resilience (though still vulnerable to climate change and variability) but altered the water distribution [*Grant et al.*, 2013].

Our modeling framework is illustrated for the Melbourne Metropolitan area using climate simulations from the fifth phase of the Coupled Model Intercomparison Project (CMIP5). We consider explicitly the human influence on the water cycle by modeling each of the reservoirs using 17 different operational water demand scenarios ranging from very optimistic to very unfavorable (see Data Section). These scenarios consider population and industrial growth and involve different consumption behaviors. Most climate change impact studies focus on changes in future water availability relative to a baseline period in the past, ignoring the effect of potential changes in demand [*Arnell*, 1999b; *Vorosmarty, C. J. Green, P. Salisbury, J. Lammers*, 2000; *Moss et al.*, 2010]. Explicit consideration of water demand scenarios is a major step forward from a more large-scale and common top-down [*Mastrandrea et al.*, 2010] climate change impact assessment approach. Our framework focuses on individual cities and includes not only system resilience but also water consumption behaviors. Our analysis proceeds as follows.

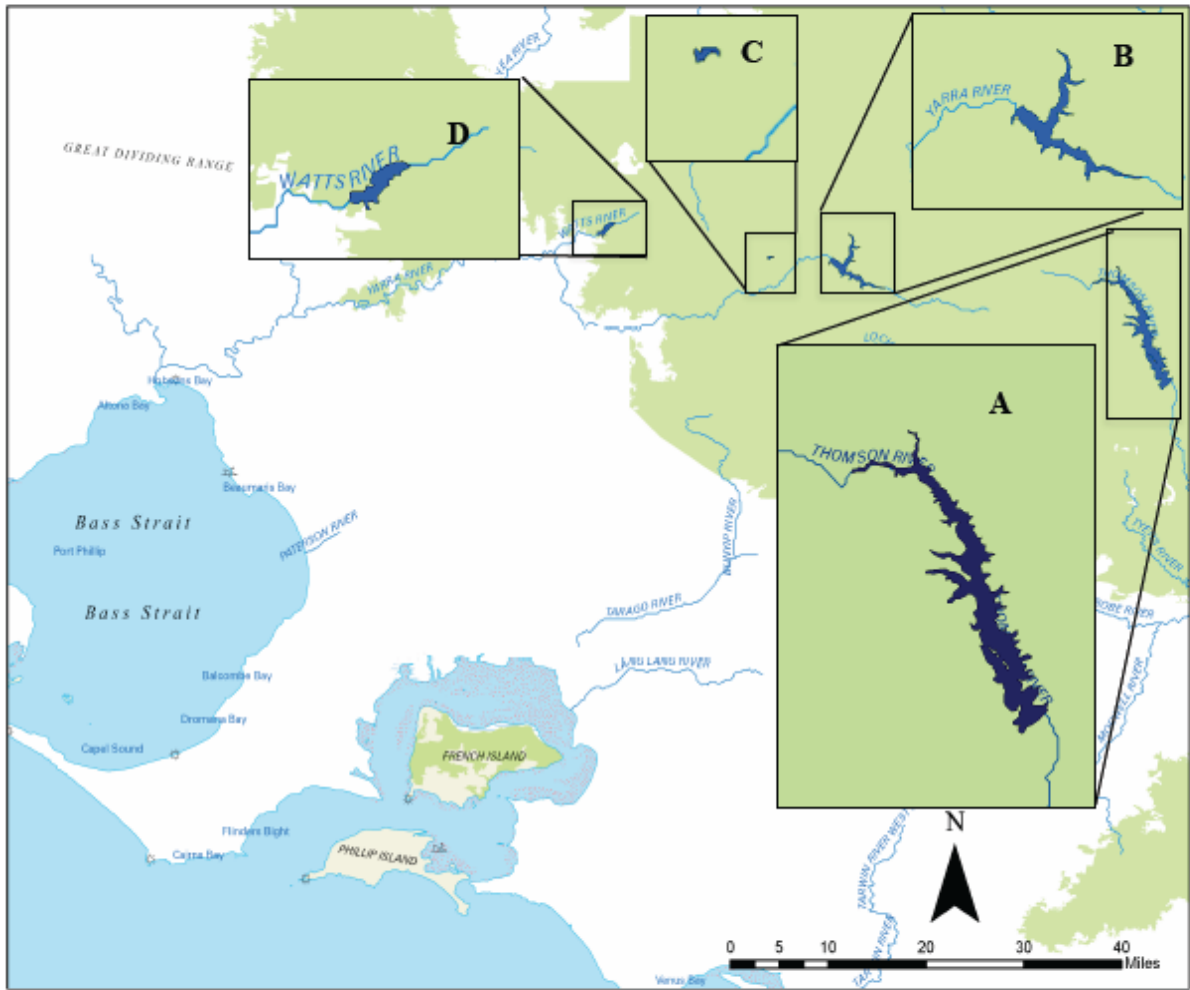
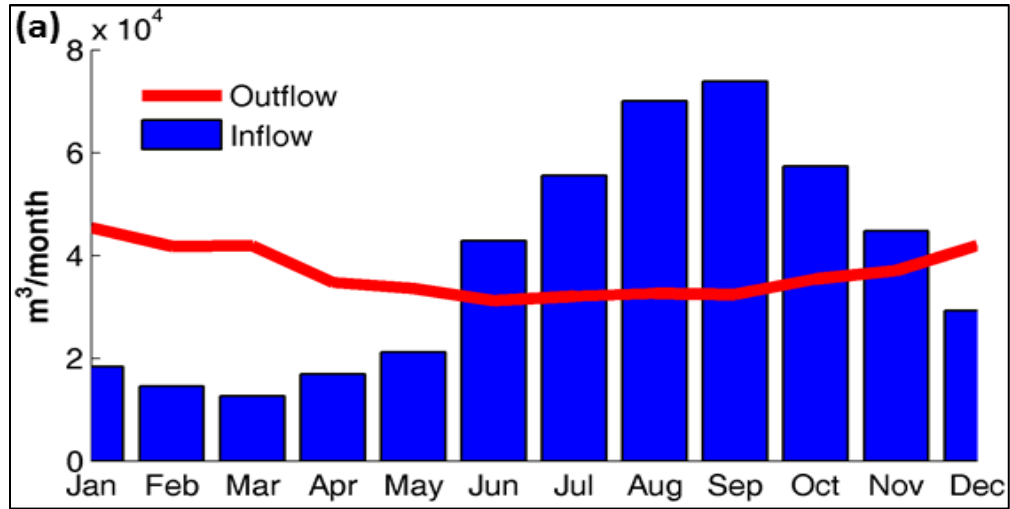


Figure 5.1 (a) Mean monthly inflow to and outflow from Melbourne major reservoirs (A-Thomson, B-Upper Yarra, C- O'Shannassy, and D- Maroondah) shown in panel (b).

We first define and setup a simple water balance model of the major reservoirs in the Melbourne area (Maroondah, O'Shannassy, Upper Yarra, and Thomson), and calibrate this reservoir model using a historical record of inflow, outflow and demand data (Methods Section). Then, a distributed hydrologic model is used to obtain inflow to the major reservoirs using daily projections from the CMIP5 precipitation and temperature simulations. Throughout the study, the results are presented for the projection period (2020-2035) relative to the baseline period (1995-2010).

Figure 5.2 displays reservoir storage anomalies (%) for selected water demand scenarios discussed in Table 5.1, The gray lines represent simulations of our hydrologic model derived from the predictions of different climate models, whereas the red and blue lines denote the ensemble mean. The first two scenarios, referred to as optimistic, assume that future water demands would be less than that of the baseline period. The ensemble mean remains mainly positive indicating more storage in the reservoir during the projection period relative to the baseline period. Yet, for most other demand scenarios (DS, 11-17) the ensemble means are negative (red lines) and the storage of the reservoirs is decreasing.

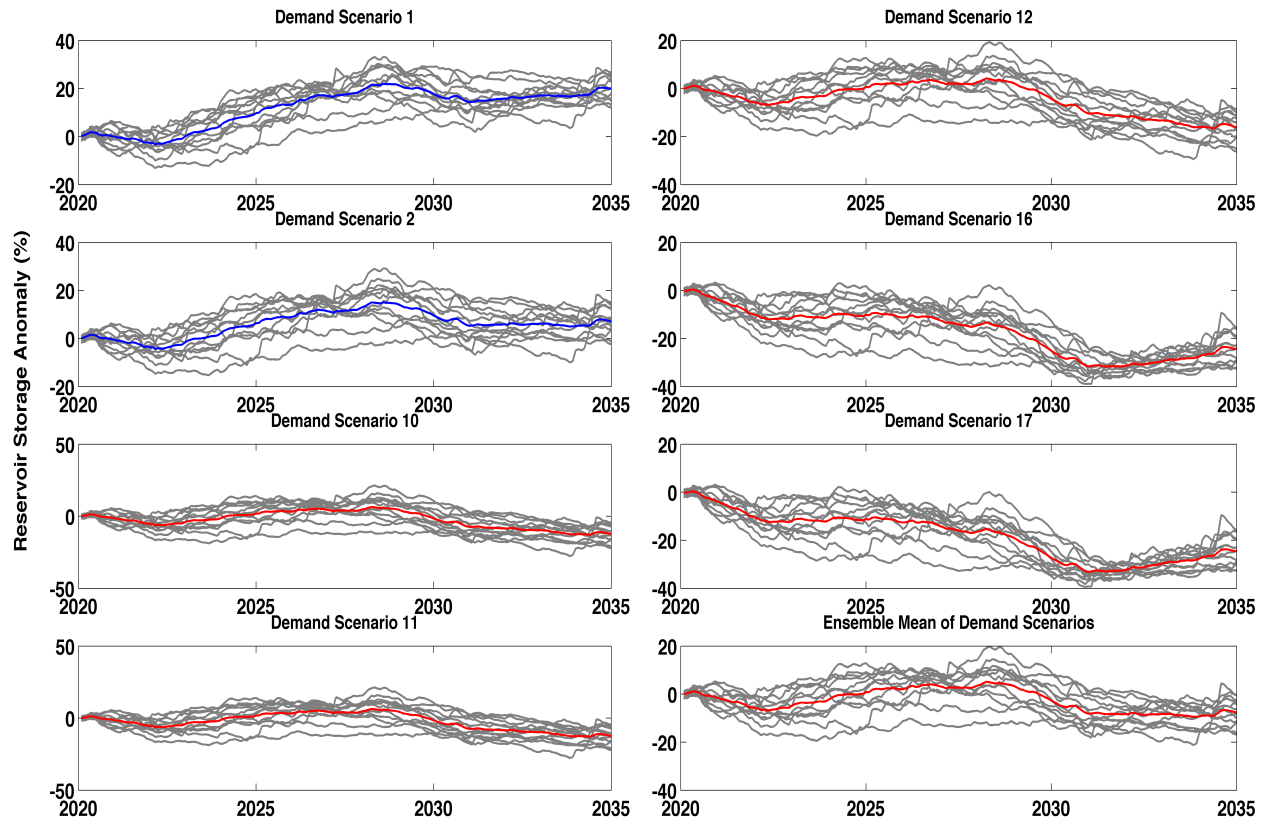


Figure 5.2 Reservoir water storage anomalies for the projection period (2020-2035) relative to the baseline (1995-2010) for different water demand scenarios. A positive value (blue) indicates that on average the future storage will be more than the baseline, whereas a negative value (red) indicates that the system will expect more water stress relative to the baseline.

To provide further insights into the results, please consider Figure 5.3 that presents water demand anomalies for all the scenarios (a) as well as their corresponding reservoir storage anomalies (b). If we ignore the reservoirs in our simulations, and just investigate the effect of future demands relative to the historical record, the ensemble mean of all but three scenarios projects higher demands relative to the historical record (see red boxplots in Figure 5.3a). Their key question is whether the current reservoirs provide sufficient resilience for potential future increases in water demand. This issue cannot be addressed without estimates of future storage conditions. Using the model presented in the Methods Section, total water storage anomalies are

obtained for each of the demand scenarios and the ensemble mean (Figure 5.3b). While 14 scenarios project higher demands relative to that of the baseline period, only 9 of these lead to a storage deficit of the reservoirs (see the red boxplots in Figures 5.3a and 5.3b). In other words, for 9 of the demand scenarios the Melbourne water resources expects more stress than the baseline period. Note that the baseline period includes the Millennium drought that posed significant stress on the system. These results suggest that the water stress for the city of Melbourne and surrounding areas is likely to increase compared to that during the Millennium drought. On the other hand, results suggest that if the demand of Melbourne does not exceed the first 8 scenarios (Very Optimistic to Neutral+), then even with the expected climate change water stress would be less than the baseline period (i.e., net average storage remains above the baseline period).

5.2 Data

A wide range of water demand scenarios ranging from very optimistic to very unfavorable are obtained from Melbourne Water (Table 5.1).

Furthermore, observed inflow and outflow to Melbourne major reservoirs (Maroondah, O'Shannassy, Upper Yarra, and Thomson) are obtained from Melbourne Water, and used for model calibration (see Methods Section). Ensemble simulations of daily precipitation and temperature from the Coupled Model Intercomparison Project Phase 5 (CMIP5; [Taylor *et al.*, 2012]) are used to estimate future water availability. These model simulations are summarized in Table 5.2. CMIP5 includes a suite of historical and future climate scenarios, designed to address a wide range of climate science issues in the fifth assessment report of the intergovernmental panel on climate change [IPCC 2013]. We have used the 8.5 representative concentration pathways.

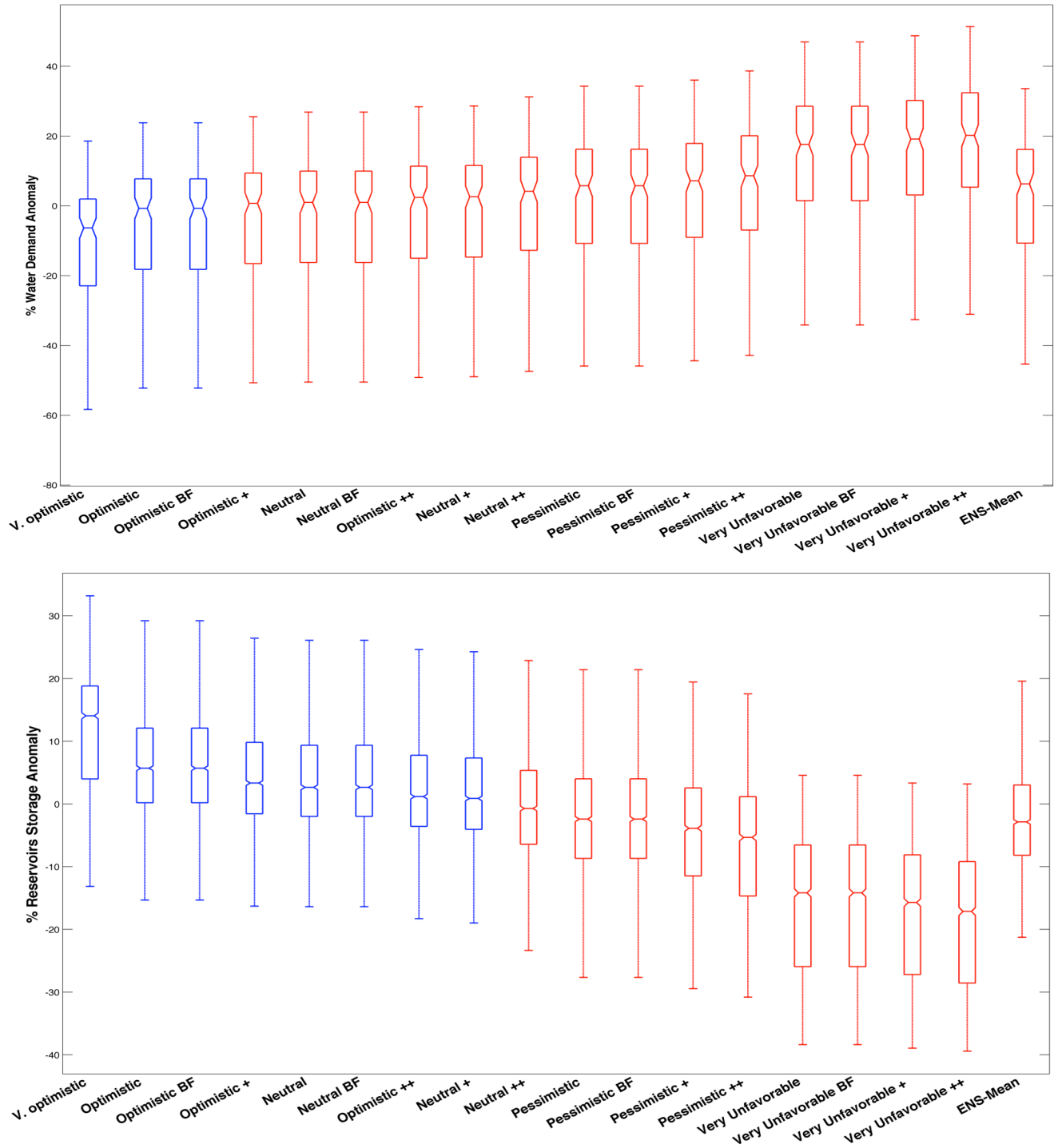


Figure 5.3 (a) Melbourne water demand anomalies in different scenarios (see Table S1), for the projection period (2020-2035) relative to the baseline (1995-2010). The blue boxplot indicates that on average the future water demand will be less than the baseline, whereas a negative value indicates that future water demand will be more relative to the baseline. (b) Reservoir water storage anomalies for the projection period relative to the baseline for different water demand scenarios. The blue boxplot indicates that on average the future storage will be more than the baseline, whereas a negative value indicates that the system will expect more water stress relative to the baseline. The boxplots show the median (midlines in boxes), 25th and 75th percentiles, and whiskers of the summations relative to the baseline.

5.3 Methods

Estimates of local surface runoff are derived from the spatially distributed PCR-GLOBWB [Van Beek *et al.*, 2011], a process-based conceptual hydrologic model that includes a surface water and groundwater component. This model has been validated extensively in other studies using satellite and streamflow observations [Van Beek *et al.*, 2011; Wada *et al.*, 2011b, 2012]. The PCR-GLOBWB model is forced with daily precipitation and temperature estimates (bias corrected) of CMPI5 [Liu *et al.*, 2014] and uses a nested reservoir model to account for water use and redistribution throughout the area.. This nested model simulates the storage, S (L) of the Melbourne reservoirs [Haddeland *et al.*, 2006; Hanasaki *et al.*, 2006; Van Beek *et al.*, 2011] using the following water balance equation:

$$\frac{\partial S}{\partial t} = Q_{in} - Q_{out} - Q_{add} \quad (5.1)$$

where t (T) denotes time, Q_{in} (LT^{-1}) signifies the inflow rate, Q_{out} (LT^{-1}) is the reservoir outflow rate, and Q_{add} (LT^{-1}) represents the additional release from the reservoir. Equation (5.1) is solved numerically using a monthly integration time step. Monthly inflow rates are derived from a historical archive of data of Melbourne Water, and reservoir outflow rates are calculated from the following min-max equation.

$$Q_{out} = \max[\min[Q_d, Q_{lim}], Q_{avg} \cdot g(S_t)] \quad (5.2)$$

where Q_d (LT^{-1}) is the outflow rate required to satisfy only the monthly demand, Q_{lim} (LT^{-1}) denotes a default limit of the discharge (defined operationally), Q_{avg} (LT^{-1}) signifies the long-term average reservoir inflow and $g(S_t)$ is the so-called potential release ratio. The variable Q_d is calculated from:

$$Q_d = \begin{cases} \left(\frac{S(t)}{S_{\min}}\right) D(t) & \text{if } S(t) \leq S_{\min} \\ D(t) & \text{if } S(t) > S_{\min} \end{cases}, \quad (5.3)$$

where S_t (L) denotes the storage at time t (in months), S_{\min} (L) signifies the minimum storage level below which the reservoir fails to operate at full capacity, and D_t (LT^{-1}) is the actual water demand at time t . The potential release ratio, $g(S) \in [0,1]$ is dimensionless factor is a function of the storage at the present time and long-term average inflow rate, S_{\min} and the maximum storage capacity of the reservoir, S_{\max} (L):

$$g(S) = \begin{cases} 0 & \text{if } S \leq S_{\min} \\ \frac{S - S_{\min}}{S_{\max} - S_{\min}} & \text{if } S_{\min} < S \leq S_{\max} \\ 1 & \text{if } S > S_{\max} \end{cases} \quad (5.4)$$

Finally, the additional release is based on flood control and reservoir management and calculated as follows:

$$Q_{\text{add}} = \max\left[0, \left(\frac{S(t) - S_{\max}}{S_{\max} - S_{\max}}\right) (Q_b - Q_{\text{out}})\right] + \max[0, S(t) - S_{\max}], \quad (5.5)$$

where Q_b (LT^{-1}) is the river bank-full discharge.

Basically, Q_{out_t} is estimated based on the buffer between $S_{\max,o}$ and the absolute capacity of the reservoir (S_{\max}).

5.4 Model Calibration

Given that the demand is typically satisfied by water from a combination of reservoirs, the inflow and outflow to the major reservoirs (Maroondah, O'Shannassy, Upper Yarra, and Thomson) are used for calibration of the model. We have available monthly records of the inflow, outflow and demand, and the minimum and maximum operating fill levels S_{\min} and S_{\max} are selected as the calibration parameters. The parameters are optimized using the Differential Evolution (DE) global optimization algorithm.

This method is an efficient global optimizer in continuous search domains [Ter Braak and Vrugt, 2008; Vrugt et al., 2009a]. Sixteen (1995-2010) years of monthly data were available from which 6 years (1995 - 2000) are used for calibration and the remaining 9 years (2001 - 2010) for evaluation. The following objective function is used to compare the model simulations with the data:

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (S_t - O_t)^2}, \quad (5.6)$$

where, S (L) is a N -vector with simulated storages, and O (L) is a N -vector with observed storages.

The DE algorithm starts with evolving a population of initial parameter sets (initial vectors). This initial population (e.g., 20 samples, known as the total number of population (NP)) describes the parameter space uniformly between the parameter criteria (i.e., parameters' lower bound and upper bound) [Storn and Price, 1997]. For instance, the initial vector for Minimum Monthly Operating Fill Level ($S_{min,o}$) at the first generation is estimated by simulating 12 random parameters (for each month) within its criteria ($0.05 < S_{min,o} < 0.25$ – see [Oliveira and Loucks, 1997; Chang et al., 2005]):

$$x_{j_{initial}} = Min_{S_{min,o}} + \varepsilon \cdot (Max_{S_{min,o}} - Min_{S_{min,o}}) \quad (5.7)$$

where, $x_{j_{initial}}$ is the initial vector for the Minimum Monthly Operating Fill Level parameter (j ranges from 1 to NP), $Min_{S_{min,o}}$ and $Max_{S_{min,o}}$ are the lower and upper bounds for this parameter, and ε is the uniform random number generator. After setting up the initial generation, DE creates new parameter vectors by adding the weighted difference between two randomly generated population vectors to a third population vector (mutation).

$$v_{j_{G+1}} = x_{r_{1G}} + F \cdot (x_{r_{2G}} - x_{r_{3G}}) \quad (5.8)$$

where, G is the number of generation, r_1 , r_2 , and r_3 correspond to three vectors of simulations, and $F > 0$. The mutated population vectors are then mixed with another population vector to increase the diversity of the perturbed parameter vectors (cross-over):

$$u_{j_{G+1}} = (v_{1_{G+1}}, v_{2_{G+1}}, v_{3_{G+1}}, \dots, v_{NP-1_{G+1}}, v_{NP_{G+1}}) \quad (5.9)$$

The last operation (selection) is to choose the target vector from $x_{j_{initial}}$ and $u_{j_{G+1}}$ since each population vector should serve only once in each generation. To specify each member of the G+1 generation, both objective functions of parameter vectors (trial vector $u_{j_{G+1}}$ and target vector u_{j_G}) are compared and the one that yields to smaller objective function is set into next generation of the parameter vector. Evaluating new generation of parameter vectors continues to the point that the best parameter vector is set. After calibration, the model is tested for the validation period (Figure 5.1). Table 5.3 (Supplementary Materials) presents the summary statistics for both calibration and validation periods. The table shows that the model reasonably simulates water storage with the relative mean error (i.e., storage error relative to total storage) being 3% and 5%, in the calibration and validation periods, respectively.

5.5 Conclusions

In this chapter, the global modeling framework, discussed in Chapter 4, is applied for climate change impact assessment. An additional layer for integrating man-made reservoirs is included in the model to account for local reservoirs that provide resilience against extremes. The novelty of this global-local modeling framework is to take advantage of models in different scales (global hydrology and local reservoirs). The modeling framework is applied over the Melbourne

major reservoirs using CMIP5 climate simulations as input. We consider explicitly the human influence on the water cycle by integrating man-made reservoirs. Furthermore, the study uses 17 different operational water demand scenarios ranging from very optimistic to very unfavorable. These scenarios consider population and industrial growth and involve different consumption behaviors. The Results suggest that for a thorough analysis of climate change impact on water resources, including the effect of local resilience (i.e., reservoirs) and the expected demand in the future are fundamental. Our results showed explicit consideration of water demand scenarios is a major step forward from a more large-scale and common top-down climate change impact assessment approach toward a bottom-up local assessment. Our framework allows investigating regions, considering the local capacity to cope with extremes and water consumption behaviors which are ignored in large-scale and top-down studies.

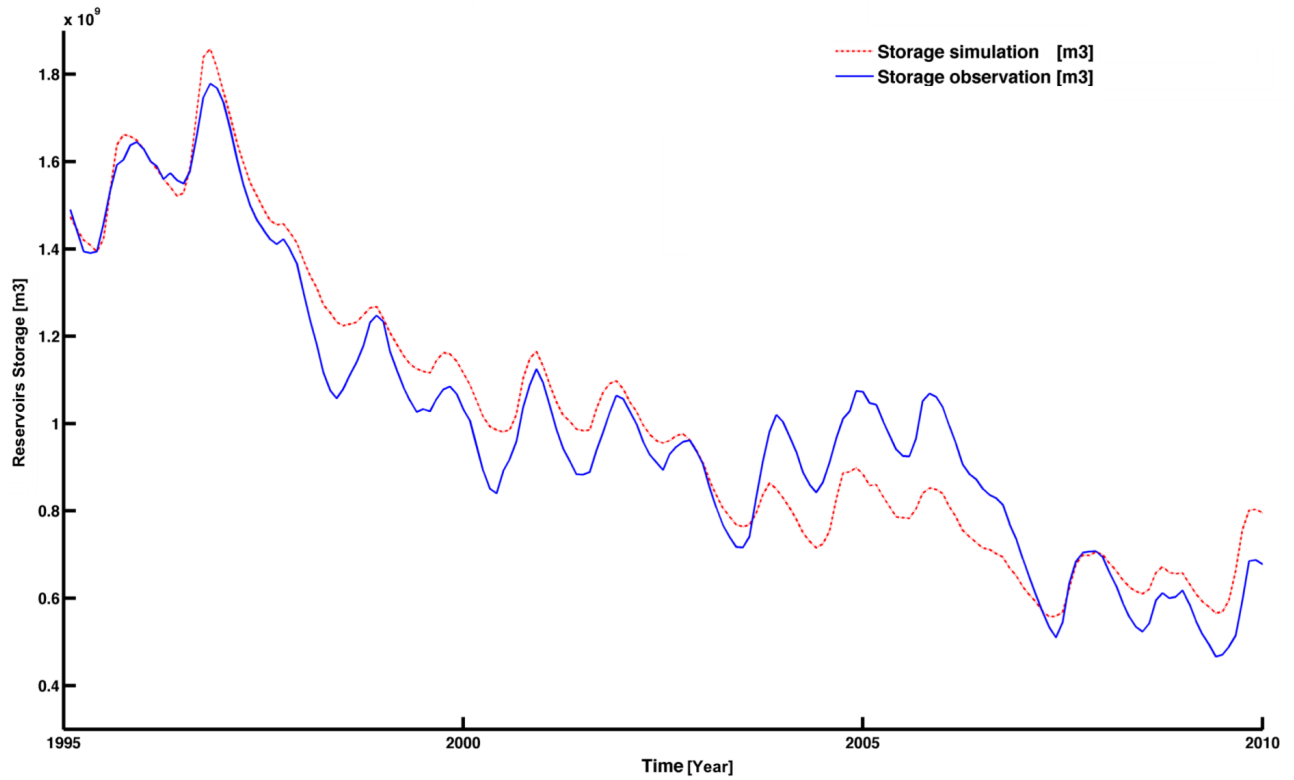


Figure 5.4 Water storage observation compared with reservoir model simulation during baseline period (1995-2010).

Table 5.1 Description of water demand scenarios

	Demand Scenario	Description
1	Very Optimistic	
2	Optimistic	
3	Optimistic BF	
4	Optimistic +	
5	Neutral	
6	Neutral BF	
7	Optimistic ++	
8	Neutral +	
9	Neutral ++	
10	Pessimistic	
11	Pessimistic BF	
12	Pessimistic +	
13	Pessimistic ++	
14	Very Unfavorable	
15	Very Unfavorable BF	
16	Very Unfavorable +	
17	Very Unfavorable ++	

Table 5.2 List of the CMIP5 climate model simulations used for simulating future conditions.

CMIP5 Climate Models		Institution
1	BNU_ESM	College of Global Change and Earth System Science, Beijing Normal University
2	CCSM4	NCAR, USA
3	CESM1_BGC	NSF, DOE, and NCAR, USA
4	FIO_ESM	The First Institute of Oceanography, China
5	MIROC_ESM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies, Japan
6	MPI_ESM_LR	Max Planck Institute for Meteorology (MPI-M), Germany
7	MRI_CGCM3	Meteorological Research Institute, Japan
8	NorESM1_M	Norwegian Climate Centre, Norway
9	IPSL_CM5a_LR	Institut Pierre-Simon Laplace, France
10	GFDL_ESM2M	NOAA Geophysical Fluid Dynamics Laboratory, USA
11	BCC_CSM1	Beijing Climate Center, China
12	INMCM4_ESM	Institute for Numerical Mathematics, Russia

Table 5.3 Model efficiency coefficients for the calibration and validation periods.

Summary Statistics	Calibration 1995-2000	Validation 2001-2010
Spearman Correlation	0.98	0.83
Correlation Coefficient	0.98	0.82
Nash Sutcliffe Coefficient	0.88	0.57
Relative Mean Error	3.35%	5.03%

CHAPTER 6: A Hybrid Framework for Water Stress Assessment: Linking Climate Variability and Local Resilience

6.1 Introduction

Many areas of the world face water scarcity and water availability challenges as a result of deteriorating water quality, increasing water demand, and climatic variability and change [Wood *et al.*, 1997b; Trenberth, 2001; Sivakumar, 2011; Stoll *et al.*, 2011; Schewe *et al.*, 2014]. Semi-arid and arid regions are particularly vulnerable to climatic variability and change impacts on water availability and distribution [Seager *et al.*, 2007a; Cayan *et al.*, 2008]. Reservoirs are one of the main man-made infrastructures providing resilience against extremes (e.g., floods and droughts) and they play a key role in water resources management. Currently, man-made reservoirs [Vorosmarty, C. J. Green, P. Salisbury, J. Lammers, 2000] control approximately 20% of the total global annual river discharge [Shiklomanov *et al.*, 2000; Fekete *et al.*, 2002]. Approximately 70% of global fresh water withdrawal comes from reservoirs [Shiklomanov *et al.*, 2000], which gives some indication of reservoirs' importance in providing resilience for human water use globally [Sillmann *et al.*, 2013]. Human influence on the hydrological cycle and redistribution of water is already apparent, as can be seen in Figure 6.1, where comparison of the natural inflow to Shasta Lake in California and its controlled outflow shows a shift in the maximum mean annual runoff from March to July. The figure highlights the anthropogenic influence on the water supply leading to a substantial change in the distribution of the water relative to the natural conditions. The primary objective of this redistribution is to store water for the time it is needed most, and to build resilience in the system to secure water availability throughout the year. Studying water stress (water supply vulnerability) is likely to grow more

important as the climate changes and the population grows. Changes in seasonality of precipitation or snowmelt combined with population, agricultural and industrial growths can lead to more stress on water security.

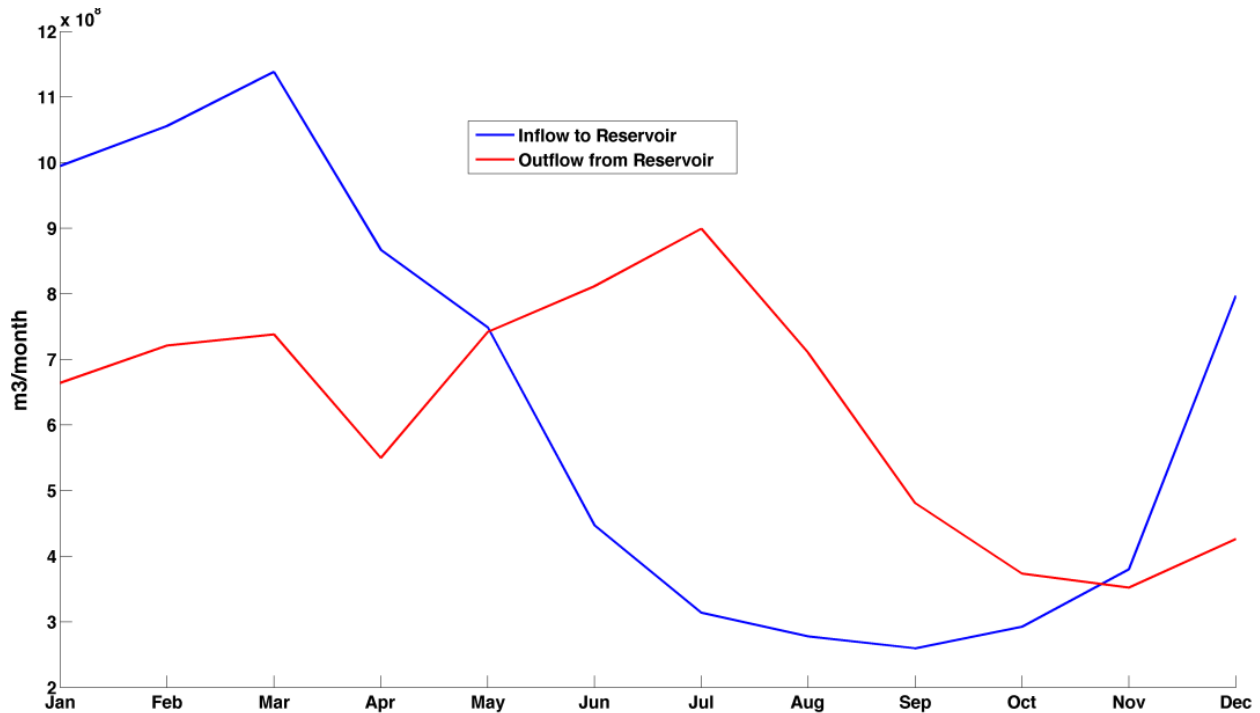


Figure 6.1 Mean annual inflow to and outflow from Shasta Lake in California.

Many statistical performance indices have been introduced to measure or describe a system under stress. One of the first set of indicators designed for this purpose was by [Hashimoto *et al.*, 1982] who elaborated the use of three basic statistical indices: reliability, resilience, and vulnerability. Together, these measures were used to explain the performance of a system (e.g. a water storage reservoir). Reliability was defined as the probability that no failure occurs within a fixed period of time (e.g., planning/operation period). Resiliency was described as how quickly a system is likely to recover or bounce back from failure once failure has occurred. Vulnerability was referred to as the likely magnitude of a failure, if one occurs. These indicators have been

widely used to assess the effect of climate variability on water resources [Kundzewicz *et al.*, 2008]. Other definitions of reliability, resilience and vulnerability as well as different performance metrics have been developed and discussed in the literature (e.g., [Moy *et al.*, 1986; Simonović, 1995; Vogel and Bolognese, 1995; Srinivasan *et al.*, 1999; Merabtene *et al.*, 2002; Steinschneider and Brown, 2012; Moody and Brown, 2013; Ward *et al.*, 2013]). In recent studies, more parametric rules are developed for multi-reservoir systems [Nalbantis and Koutsoyiannis, 1997]. [Miller *et al.*, 2010] identified a major difficulty in integrating resilience and vulnerability performance indices — namely that differences existed between the use of these terms in concept and theory, in methodologies to assess them, and in the real-world practice of addressing climate changes [JAIN and BHUNYA, 2008]. The definition of vulnerability has evolved over time. In the Fourth Assessment Report of Intergovernmental Panel on Climate Change (IPCC), vulnerability was defined as the degree to which a system is susceptible and unable to cope with climate variability and extreme weather events, as well as sensitivity and the adaptive capacity of the system to this variability [IPCC, 2007]. In the more recent IPCC report [Rogelj *et al.*, 2012], social implication of vulnerability is emphasized and is generally referred to as the tendency or susceptibility of a system to be adversely affected, while such susceptibility establishes an internal characteristic of the affected element. Recent definitions of vulnerability indicate that it is a relative concept, and possible statements about vulnerability must clearly specify the entity that is vulnerable, the stimulus to which it is vulnerable, and the preference criteria to calculate the result of the interaction between them [Ionescu *et al.*, 2008]. Although there are many similarities between the performance indices' approaches (e.g., vulnerability, reliability), they have been kept separate from each other and studied in parallel tracks [Preston *et al.*, 2011]. Integrating different aspects of the same crisis — while acknowledging the value of multiple

perspectives — would benefit the different communities involved [Miller *et al.*, 2010]. The integrated nature of water resources problems and hydrological systems has convinced researchers in the fields of resilience and vulnerability to choose hybrid approaches. For example, [Ziervogel *et al.*, 2006] has explored local adaptations due to climate variations in different social and environmental stresses and proved that using a combination of approaches is more successful than single method approaches [Miller *et al.*, 2010]. Another frequent distinction in vulnerability and water stress studies arises from the “top-down” and “bottom-up” methodologies [Dessai and Hulme, 2004]. The top-down methodology comes from climate variability or climate change impacts assessments and focuses on biophysical vulnerability while considering different meteorological or climate conditions and impact models. In other words, “top-down” focuses on climatic and meteorological conditions that cannot be simply controlled or altered by decision makers. A “bottom-up” methodology focuses on the capacity of people, societies and governments to respond or adapt to climate extremes or water availability challenges placed on them. In fact, a bottom-up approach is a local issue addressed by people/societies, and focuses on social conditions, perception of available water vulnerability (here, water stress) and a human dimension (e.g., water reuse, conservation) [Mastrandrea *et al.*, 2010]. In this chapter, the vulnerable entity is defined as the total water storage in reservoirs, and water stress refers to lack of inflow and/or storage to satisfy the demand. The stimulus to which the water storage in reservoirs is vulnerable is the seasonal change in surface runoff. The preference criterion used to assess the result of the interaction between the entity and the stimulus is to satisfy water demand during a certain timeframe by using both the current storage and inflow to the system. Furthermore, the term resilience, in this paper, corresponds to the storage available to cope with climatic variability and water availability challenges. Basically,

building resilience into the system is considered a bottom-up approach to address climatic variability. In this chapter, a methodology is proposed that integrates both top-down and bottom-up approaches for assessing vulnerability to water availability (hereafter, water stress). The approach includes an inflow versus water demand reliability indicator that is dominated by climatic and meteorological conditions (top-down), and a water storage resilience index that considers the man-made infrastructure to cope with climate variability (bottom-up). The two indicators are combined using a multivariate statistical framework. The model offers a unique approach for estimating the overall water stress through water supply vulnerability including the effect of system resilience (here, reservoirs). The final outcome is a hybrid indicator that combines the above-mentioned top-down and bottom-up approaches for water resources vulnerability assessment.

6.2 Classical methods for performance assessment

Performance measures of reliability, resilience and vulnerability have been defined and widely used in the past decades [Harberg, 1997]. There are two main definitions for reliability. In the first approach reliability is defined as the probability of no-failure over the operating period. The other approach considers reliability as probability of system failure. A failure is defined as the inability of reservoir system to deliver the desired water demand [Vogel and Bolognese, 1995]. [Hashimoto *et al.*, 1982] introduced reliability as:

$$\alpha = pr\{X_t \in S_a\} \quad (6.1)$$

where, X is the system's output state/status at time t and S_a is all satisfactory conditions/outputs. Reliability is interpreted as probability that no failure occurs. The concept of resilience in reservoir systems was first introduced by [Hazen, 1914] and later redefined by [Sudler, 1927], [Hurst, 1951], [Matalas C. Nicholas, 1977], [Fiering, 1982], and [Moy *et al.*, 1986]. Basically,

resilience can be expressed as [Hashimoto *et al.*, 1982]:

$$\gamma = \frac{pr\{X_t \in F \ \& \ X_{t+1} \in S_a\}}{pr\{X_t \in F\}} \quad (6.2)$$

where, F represents the failure of the system, X is the system's output/status at times t and $t + 1$ and, S_a is all satisfactory conditions/outputs. Resilience in classical methods was interpreted as the failure occurrence of the system that is followed by satisfactory condition/output. Further, system vulnerability was defined as the likely magnitude of a failure, if one occurs. It can also be defined as the expected maximum severity of a failure state into the set of unsatisfactory states [Hashimoto *et al.*, 1982], shown:

$$\vartheta = \sum_{t \in F} e_t S e_t \text{ and } e_t = pr\{X_t \in F\} \quad (6.3)$$

where, e_t is the probability of a temporal failure state and $S e_t$ is the indicator of its severity at time t . Based on this definition of vulnerability, severity of a failure state is more important than its duration. There are alternative definitions for vulnerability developed for different applications (e.g., [Kundzewicz and Kindler, 1995] and [KJELDSEN and ROSBJERG, 2004]), such as the ratio of maximum deficit to the target water demand [McMahon *et al.*, 2006]. This chapter considers water stress to be the situation when water resources are vulnerable. This hybrid indicator proposed in this study combines the top-down and bottom-up approaches to assess water stress.

6.3 Methodology

In general, there are two major classes of reservoir systems [Vogel and Bolognese, 1995]: over-year and within-year systems. The within-year systems are typically refilled by the end of each year and as a result, this system is very sensitive to seasonality and other temporal variations in

the system (e.g. monthly inflow, monthly water demand, evaporation). In contrast, over-year systems do not refill by the end of each year, and are sensitive to long-term water supply deficit (drought).

The latter is more common in areas where long-term dry conditions are expected frequently. This classification signifies the importance of variations, especially the time period that is affecting the reservoir system. Here, there is a timeframe defined for each reservoir system that is either 6 month (for within-year systems) or 12 months (for over-year systems) depending on the category of the reservoir system. It should be noted that the presented model is general and can be used for assessment over different timeframes.

In this study, the proposed method to investigate water stress is a multivariate approach that relies on two individual (univariate) indicators. At first, a timeframe is defined based on the type of reservoir system (within-year or over-year). After defining the timeframe, two new indicators are defined as: water storage resilience indicator and inflow-demand reliability indicator. Inflow-

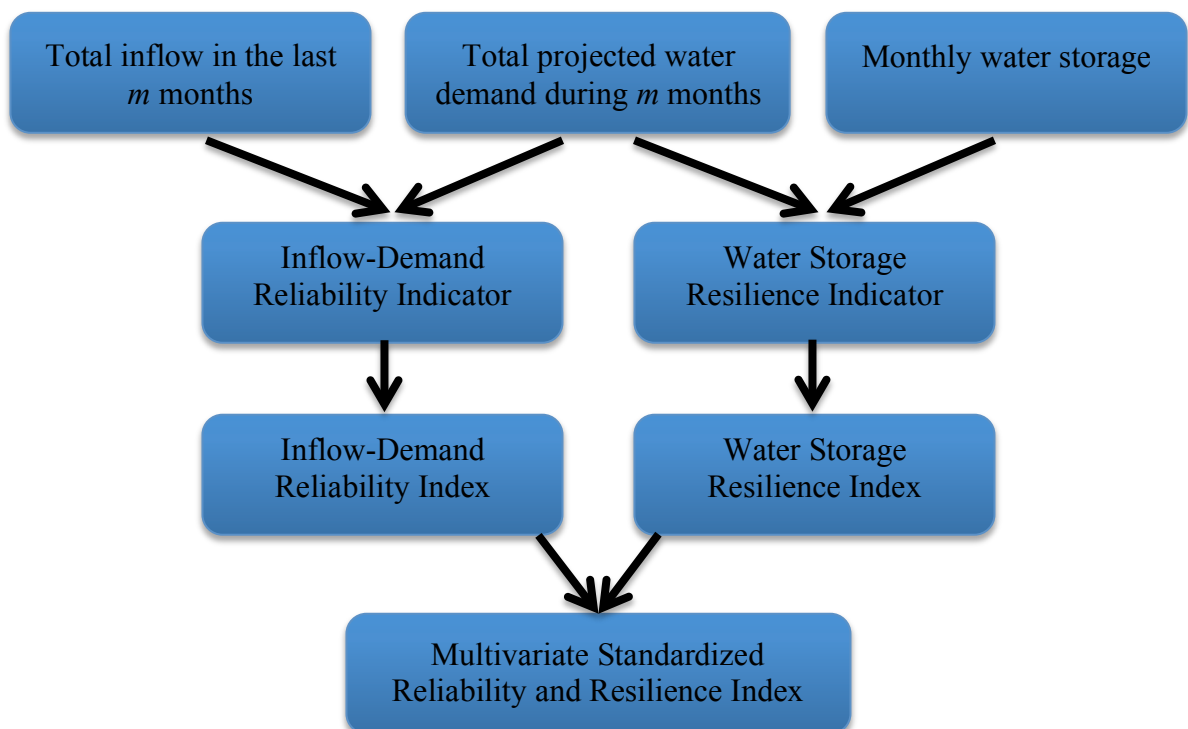


Figure 6.2 Components of the Multivariate Standardized Reliability and Resilience Index (MSRRI)

demand reliability (IDR) indicator is derived by computing the sum of the percent change of inflow with respect to water demand during the projected timeframe:

$$\alpha_t = \frac{\sum_{i=t-m+1}^t Q_{in_i} - Q_{est_t}}{Q_{est_t}}, Q_{est_t} = \begin{cases} \sum_{i=t-12}^{t-13+m} (Q_{out})_i, & \text{if } m = 6 \\ \sum_{i=t-m+1}^t (Q_{out})_i, & \text{if } m = 12 \end{cases} \quad (6.4)$$

where Q_{in_i} is the monthly inflow to the reservoir (i ranges from month 1 to N , which is the sample size), m is the selected timeframe in months (6 months for within-year and 12 for over-year systems), Q_{est_t} is the total estimated water demand during projected timeframe, and time step t ranges from month 13 to N ($t=13, \dots, N$). Here, the first 6 or 12 months (depending on the type of the reservoir system) of the data are used to estimate the demand in the projected timeframe. The total water demand for the projected timeframe (next m months) is estimated based on the same period in the previous year. For this reason, the index can only be estimated starting from the second year of the data ($t=1, \dots, 12$).

This indicator (inflow-demand reliability) corresponds to the “top-down” approach where the available inflow is assessed relative to the water demand. In other words, this indicator shows whether the available water (inflow to the system) is sufficient to satisfy water demand, regardless of the storage in the reservoir. The other indicator introduced here in respect to “bottom-up” methodology is the water storage resilience indicator. This indicator is defined based on monthly inflow, monthly water demand, monthly storage and total water demand during the timeframe. Water storage resilience (WSR) indicator is computed on monthly basis and shows whether the reservoir storage is sufficient to satisfy water demand for the selected time period (m):

$$\beta_t = \frac{S_t + Q_{in_t} - Q_{out_t} - O_{min} - Q_{est_t}}{Q_{est_t}} \quad (6.5)$$

Where, S_t is the reservoir storage at month t , where $t=1,3,\dots,N$, O_{min} is the reservoirs minimum operational storage, Q_{in_t} is the monthly inflow to the reservoir at month t , Q_{est_t} is the total estimated water demand during projected timeframe (either 6 or 12 months) as discussed before, and Q_{out_t} is the monthly water demand at month t . For convenience and easy cross-comparison, the two indicators are standardized using the standard normal distribution. First, the marginal probabilities of both indicators (water storage resilience and inflow-demand reliability) are calculated [Gringorten, 1963].

$$P(x_t) = \frac{l-0.44}{N+0.12} \quad (6.6)$$

where N is the sample size, l denotes the rank of non-zero indicator (α or β) data from the smallest to largest, and $P(x_t)$ is the corresponding empirical probability at month t . The empirical probabilities are then transformed into a standardized index (SI) as:

$$SI(x) = \varphi^{-1}(P(x)) \quad (6.7)$$

$$SI(P(x)) = \begin{cases} \text{if } 0 < P(x) \leq 0.5, + \left(k - \frac{C_0 + C_1 k + C_2 k^2}{1 + d_1 k + d_2 k^2 + d_3 k^3} \right) \text{ and } k = \sqrt{\ln(1/P(x)^2)} \\ \text{if } 0.5 < P(x) \leq 1, - \left(k - \frac{C_0 + C_1 k + C_2 k^2}{1 + d_1 k + d_2 k^2 + d_3 k^3} \right) \text{ and } k = \sqrt{\ln[1/(1 - P(x))^2]} \end{cases} \quad (6.8)$$

where φ is the standard normal distribution function and $P(x)$ is the empirical probability. One can also standardize the empirical probability values by a commonly used approximation (Equation 6.8), in which, $C_0 = 2.515517$, $C_1 = 0.802583$, $C_2 = 0.010328$, $d_1 = 1.432788$, $d_2 = 0.189269$, and $d_3 = 0.001308$ [Abramowitz and Stegun, 1972; Edwards, 1997; Naresh Kumar et

al., 2009; Farahmand and AghaKouchak, 2015]. Substituting α and β with x from Equations 6.6 to 8, leads to standardized indices for inflow-demand and water storage resilience as (hereafter, $SI(\alpha)$ and $SI(\beta)$). The two univariate indicators are then combined using a multivariate framework [Yue et al., 1999]:

$$P_{j_t} = \Pr (SI(\alpha) \leq SI(\alpha_t), SI(\beta) \leq SI(\beta_t)) \quad (6.9)$$

where, P_{j_t} is the joint (multivariate) empirical probability at month t , calculated using the two inflow-demand reliability index $SI(\alpha_t)$ and water storage resilience index $SI(\beta_t)$. Having the two univariate indicators, the joint empirical probability can be derived using the multivariate model of the Gringorten plotting position introduced by [Yue et al., 1999]:

$$P_{j_t}(SI(\alpha_t), SI(\beta_t)) = \frac{l-0.44}{N+0.12} \quad (6.10)$$

where, l is the number of occurrences of the pair $(SI(\alpha_t), SI(\beta_t))$ for $SI(\alpha) \leq SI(\alpha_t)$ and $SI(\beta) \leq SI(\beta_t)$, and N is the sample size. We define the Multivariate Standardized Reliability and Resilience Index (*MSRRI*) by standardizing the joint distribution function of the inflow-demand reliability index and water storage resilience index:

$$MSRRI = \varphi^{-1}(P_j) \quad (6.11)$$

where the joint empirical probability P_j can be standardized using Equation 6.8. The final index (*MSRRI*) is not only a hybrid index (consisting of two fundamentally different indicators) but also covers both common approaches in vulnerability studies (top-down and bottom-up). In all standardized conditions, positive values indicate sufficient water to satisfy demand, while negative values indicate shortage of water relative to the demand. Here, the standardization is

based on a nonparametric approach that does not require any assumption regarding the distribution function of the original data [Farahmand and AghaKouchak, 2015]. Figure 6.2 summarizes the steps for deriving MSRRI based on the two univariate indicators.

6.4 Results and discussion

In the following section, the proposed indicators have been applied to case studies in Melbourne, Australia and California, USA.

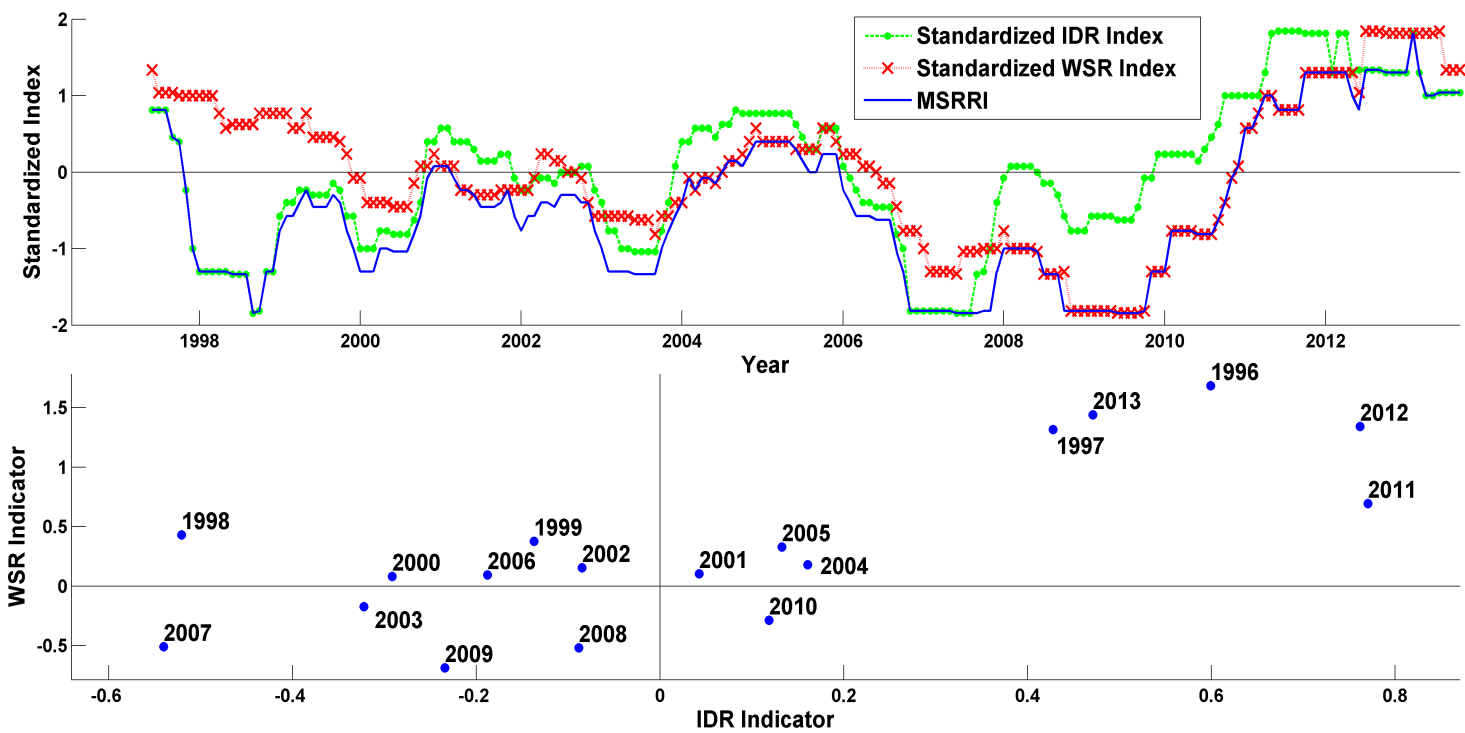


Figure 6.3 Melbourne's major reservoirs; (top panel): Standardized Water Storage Resilience (WSR) Index, Inflow-Demand Reliability (IRD) Index, and Multivariate Standardized Resilience and Reliability Index (MSRRI); (bottom panel) annual Water Storage Resilience Indicator vs. annual Inflow-Demand Reliability Indicator.

The presented method can be used for a group of reservoirs (Figure 6.3) that serve a region or individual reservoirs (Figures 6.4 - 6.6). First, an example application is provided for Melbourne major reservoirs, which consists of Thomson, Maroondah, O'Shannassy, and Upper Yarra

reservoirs. All together, they provide 80% of Melbourne's water demand. This reservoir system is categorized as an over-year system as they take more than one year to fill. Therefore, the timeframe for this system is defined as 12 months. Melbourne's fresh water supply mostly depends on surface runoff and thus, substantially below average precipitation and inflow could lead to water scarcity in the region [Grant *et al.*, 2013].

Figure 6.3 displays the water storage resilience (WSR) and inflow-demand reliability (IDR) indicators as well as MSRRI. As shown, the IDR and WSR indicators behave differently in trend and severity providing information on climatic and reservoir conditions, respectively. The differences between IDR and WSR can explain the differences between meteorological and local reservoir conditions. For example, $IDR < 0$ and $WSR > 0$ indicate onset of a low inflow condition (i.e., hydrologic drought) based on input relative to demand ($IDR < 0$), while there is sufficient storage to satisfy the demand ($WSR > 0$). On the other hand, $IDR > 0$ and $WSR < 0$ indicate above average inflow relative to demand ($IDR > 0$), while still storage is below average and cannot satisfy the demand ($WSR < 0$). MSRRI combines these two types of information and provides information on the overall condition of the system. This information is particularly important when describing recovery from a drought relative to storage and demand. The top panel in Figure 6.3 shows the three indices together for Melbourne's major reservoirs on a monthly scale. We focus on the Melbourne's Millennium Drought [Grant *et al.*, 2013; Van Dijk *et al.*, 2013] started in 1998 and lasted over a decade. As shown in Figure 6.3, IDR shows deficit starting 1998 while WSR slowly changes and shows deficit in 2000. In other words, the input deficit starts in 1998; however, the water supply can satisfy the demand because of the storage until 2000. On the other hand, IDR shows drought recovery at the end of 2009 based on inflow relative to demand, while WSR does not show a full recovery until 2011 when the reservoir storage has recovered. Relying

on a univariate index (either inflow-demand reliability or water storage resilience) cannot clearly reveal the severity of water stress. The MSRRI detects the onset and persistence of the water supply stress throughout the Millennium drought based on both IDR and WSR. In fact, given that MSRRI is based on the joint distribution of IDR and WSR, MSRRI indicates the onset of the stress based on IDR and recovery of the system based on WSR. This behavior of MSRRI provides an assessment of the overall stress on the system including the system resilience.

In the bottom panel of Figure 6.3, the annual IDR indicator is plotted against WSR indicator. This figure shows four possible combinations of IDR and WSR: $IDR > 0$ and $WSR > 0$; $IDR < 0$ and $WSR < 0$; $IDR < 0$ and $WSR > 0$; and $IDR > 0$ and $WSR < 0$ (see Table 6.1).

Table 6.1 Four categories of reservoir conditions based on inflow-demand reliability and reliability indicators

Water storage resilience Indicator	Inflow-demand reliability Indicator	Current water storage can satisfy total demand during next timeframe	Total inflow during previous timeframe can provide total demand during next timeframe
Positive	Positive	Yes	Yes
Positive	Negative	Yes	No
Negative	Positive	No	Yes
Negative	Negative	No	No

For example, a positive value WSR indicator and a negative value of IDR indicator imply that the water storage can satisfy the demand for the next timeframe, while the inflow cannot. In other words, the climatic condition (top-down) is not favorable, but the local reservoirs (bottom-up) can satisfy the demand. In Figure 6.3(bottom), in 1996 both of the indicators are positive (the top right corner of the plot), but they decrease and reach a record low in 2009 indicating significant water scarcity with respect to both inflow and storage. In 2009, the inflow and storage were not sufficient to satisfy the demand for another year. However, significant rainfall in 2010

improved the condition ($IDR > 0$ showing climatic recovery), and by 2011 the system fully recovered with respect to inflow and storage.

As mentioned earlier, the presented method can also be applied to individual reservoirs. In the following, MSRRRI is evaluated against univariate water stress indicators for Shasta Lake, Lake Oroville, and Trinity Lake, which outflow to the Sacramento River, the Feather River, and the Trinity River, respectively. The timeframes are six months for Shasta Lake and Lake Oroville, and 12 months for Trinity Lake. California, like Melbourne region, often experiences prolonged droughts and water stress due to its semi-arid and variable climate, and high population [Gao *et al.*, 2012]. Nearly 41% of California's water demand is provided by surface runoff, and hence reservoirs are very important for water management and providing resilience [Stanton and Fitzgerald, 2011]. California has been in a prolonged drought since 2012 resulting in record low storage and significant water stress on the system [Aghakouchak *et al.*, 2014; AghaKouchak *et al.*, 2014].

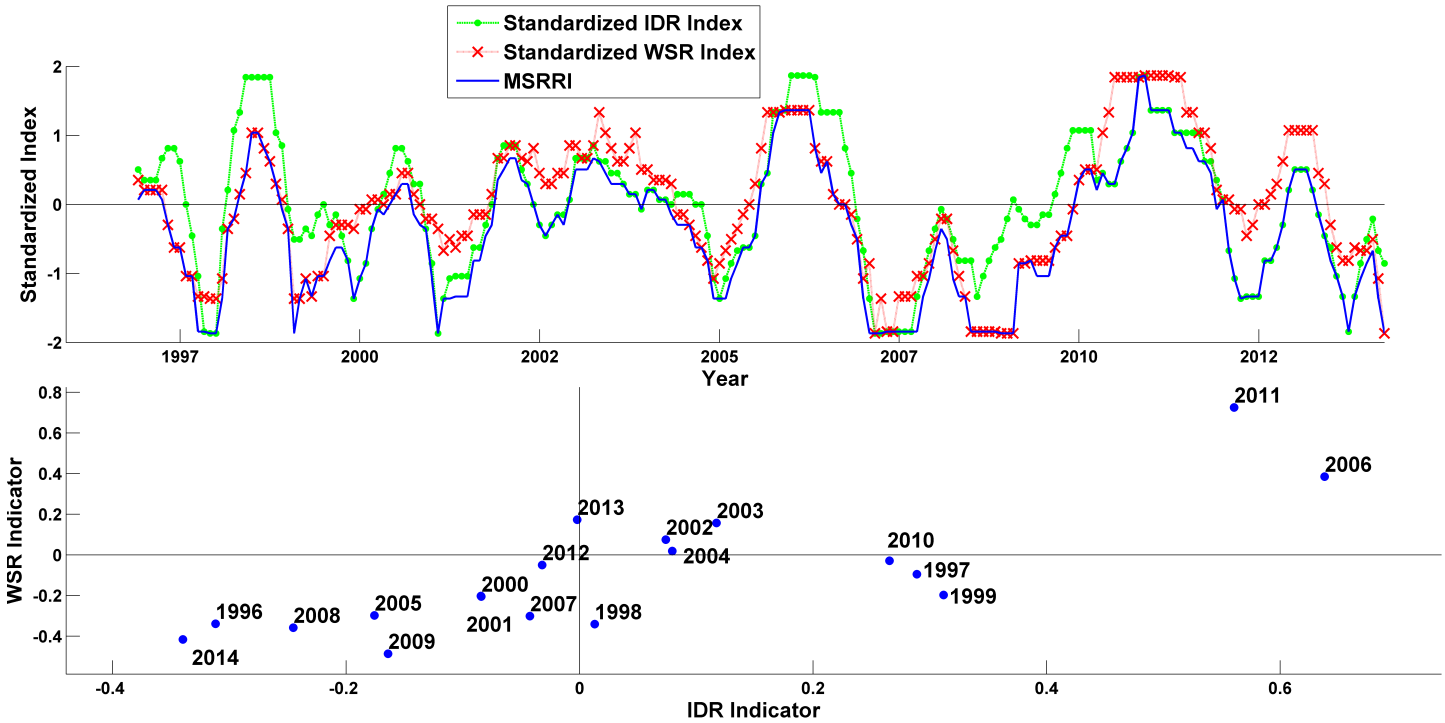


Figure 6.4 Shasta Lake(top panel): Standardized Water Storage Resilience (WSR) Index, Inflow-Demand Reliability (IRD) Index, and Multivariate Standardized Resilience and Reliability Index (MSRRI); (bottom panel) annual Water Storage Resilience Indicator vs. annual Inflow-Demand Reliability Indicator.

Figures 6.4 – 6.6 display IDR, WSR, MSRRI and WSR vs. IDR for Shasta Lake, Lake Oroville, and Trinity Lake, respectively. Shasta Lake univariate and multivariate indicators, plotted in Figure 6.4, show that despite high inflows ($IDR > 0$) during 1997, 1998, 1999, and 2010 as compared to the other years, lack of sufficient storage resulted in $WSR < 0$. In contrast, in 2013, although the water storage was sufficient to satisfy the total water demand, the net inflow to the reservoir was the main driver of water stress in the following year (compare IDR and WSR in Figure 6.4(top) and see Figure 6.4(bottom)).

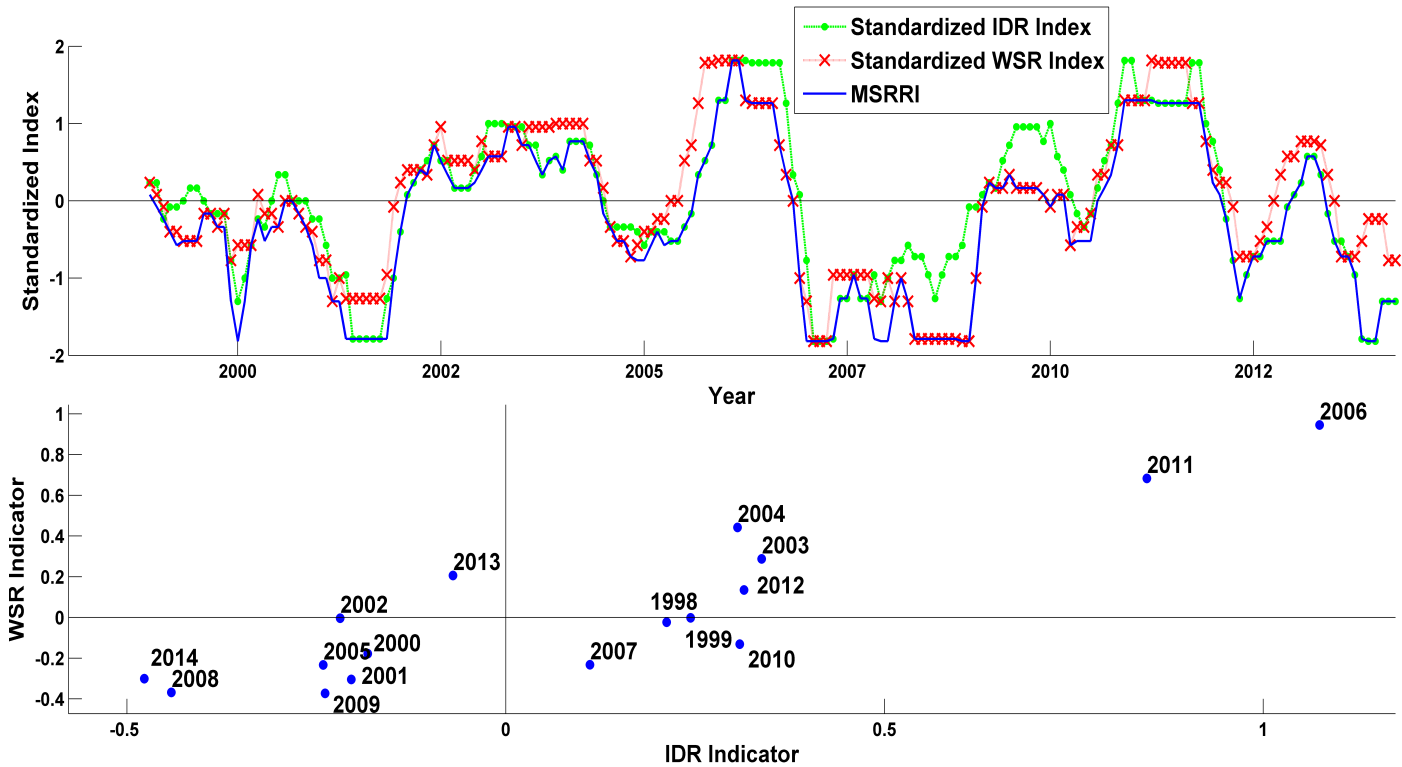


Figure 6.5 Lake Oroville; (top panel): Standardized Water Storage Resilience (WSR) Index, Inflow-Demand Reliability (IRD) Index, and Multivariate Standardized Resilience and Reliability Index (MSRRI); (bottom panel) annual Water Storage Resilience Indicator vs. annual Inflow-Demand Reliability Indicator.

Given that MSRRI is based on the joint distribution function of IDR and WSR, it responds to changes in either or both of the indicators. Figure 6.4 (Top Panel) shows that in 2004 MSRRI begins to decline much earlier than both IDR and WSR indicators, showing an earlier detection of water stress. According to the California Department of Water Resources data, in 2006 Shasta Lake had the highest volume of water storage. In contrast, it experienced the lowest volume of water storage in 2008 [California Data Exchange Center – Reservoirs, 2013]. As shown in Figure 6.4, MSRRI is consistent with Shasta Lake observations. Analysis of Lake Oroville’s (Figure 6.5) reveals that during 2007 and 2010, the study area is net positive (wet condition) with respect to inflow relative to the demand.

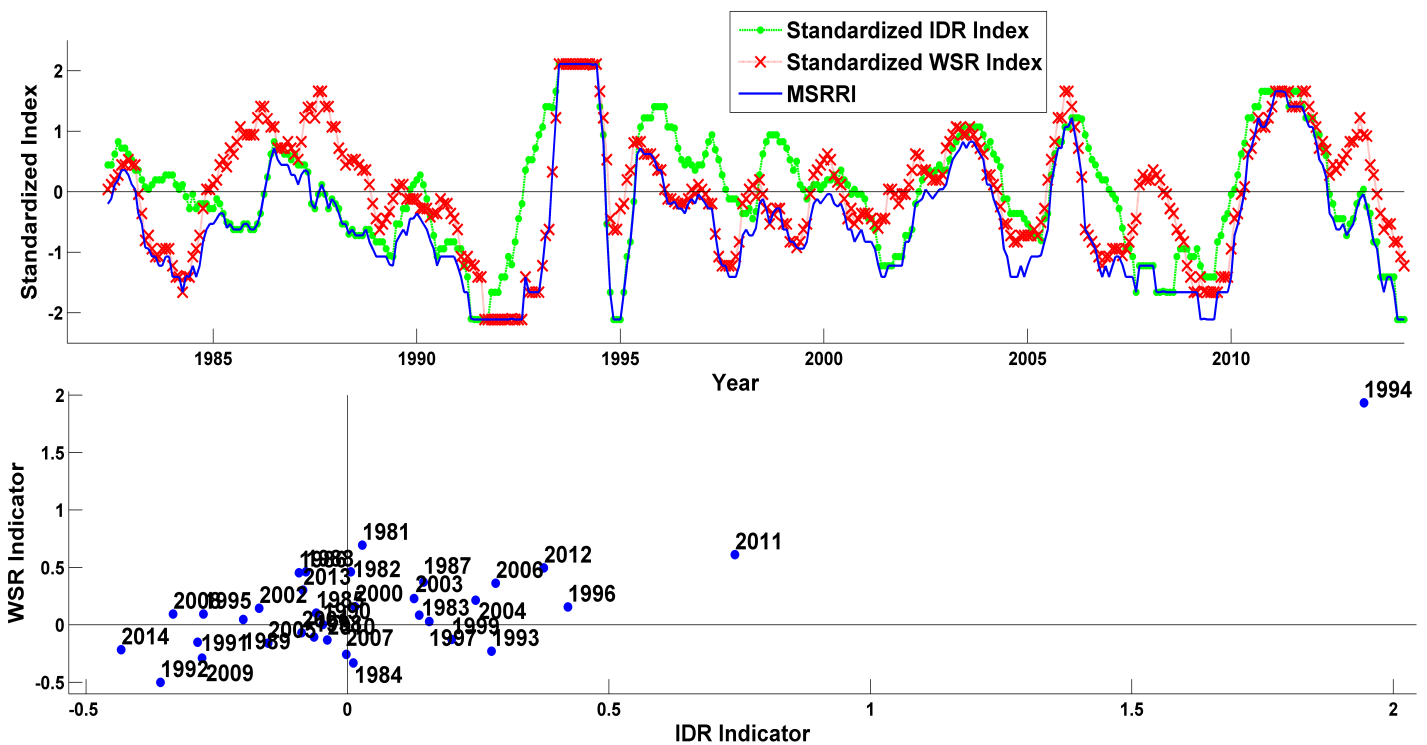


Figure 6.6 Trinity Lake; (top panel): Standardized Water Storage Resilience (WSR) Index, Inflow-Demand Reliability (IRD) Index, and Multivariate Standardized Resilience and Reliability Index (MSRRI); (bottom panel) annual Water Storage Resilience Indicator vs. annual Inflow-Demand Reliability Indicator.

This means above average inflow and precipitation in the region. However, MSSRI indicates that with respect to storage, Lake Oroville has a deficit and the system has not recovered. The storage recovers in 2011; however, between 2012 and 2013 the inflow relative to the demand (IDR) decreases. By the end of 2013, the storage is still sufficient to satisfy the demand. In 2014, with a record low IDR, significant stress is observed in WSR and MSRRI. A similar figure is provided for Trinity Lake from 1980 to 2013 (Figure 6.6). The dry and wet signals based on MSRRI, precipitation and reservoir storage data from the California Department of Water Resources [California Data Exchange Center – Reservoirs, 2013] are consistent (not shown here for brevity). For example, the three reservoirs show positive MSRRI values in 2006 which is consistent with the observed above average precipitation (see Figure 6.7a).

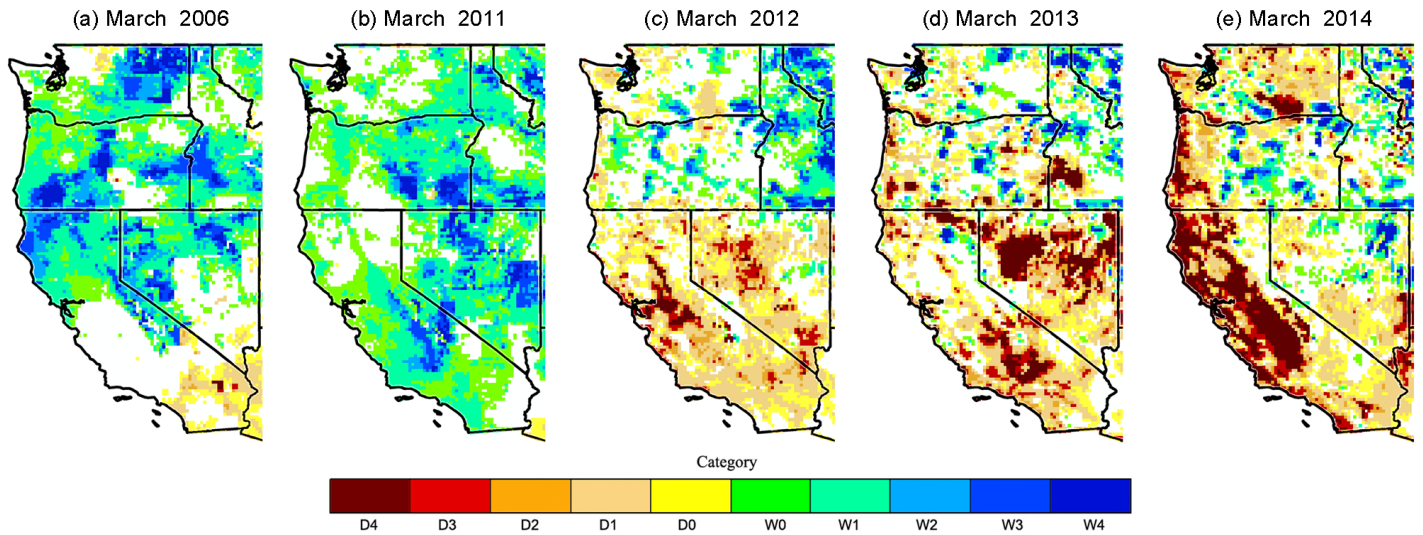


Figure 6.7 Dry and wet conditions in March 2006, 2011, 2012, 2013 and 2014 (based on 6-month standardized precipitation index). D0: Abnormally Dry; D1: Moderate Drought; D2: Severe Drought; D3: Extreme Drought; D4: Exceptional Drought; W0: Abnormally Wet; W1: Moderate Wetness; W2: Severe Wetness; W3: Extreme Wetness; W4: Exceptional Wetness.

Also, the three reservoirs show that the inflow and reservoir conditions were above average in 2011 (see Figure 6.7b). However, three years of consecutive dry conditions (see Figure 6.7c-e) led to critical conditions in 2014 (i.e., record low IDR and near record low WSR).

6.5 Conclusions

Classical performance indices such as resilience, reliability, and vulnerability that have been widely used for quantifying the complex interplay of climate variability and system resilience (here, surface water reservoirs). These indices help understanding water stress based on different factors including large scale meteorological and climatic conditions (top-down) representing climate variability and change. Furthermore, there are indicators that can describe local resilience

of the water resources system (bottom-up) to cope with variability and extreme conditions (i.e., considering the effect of reservoirs). These indicators are often used separately. In this paper, a framework is suggested that integrates both top-down and bottom-up approaches for assessing water stress due to both climatic conditions and local reservoir levels. In other words, the suggested framework considers both large scale variability and the local capacity for coping with extremes. The indicator representing a top-down approach is termed Inflow-Demand Reliability (IDR) indicator (Equation 6.4). The indicator describing the system resilience is named Water Storage Resilience (WSR) indicator (Equation 6.5). Both indicators are described relative to the average demand to consider the human dimension of the water stress. The hybrid framework that combines the two concepts is termed the Multivariate Standardized Reliability and Resilience Index (MSRRI). This model offers a unique approach for estimating the overall water stress including the local capacity to cope with extremes. MSRRI is based on a nonparametric approach that does not require parameter estimation or any a priori assumption on the underlying distribution function of the original data. In this study, MSRRI and individual univariate indicators (IDR and WSR) are used to assess water stress in Melbourne, Australia and California, USA. MSRRI is applied to both individual reservoirs and a group of reservoirs. The results show that MSRRI is superior to univariate indices because it captures both early onset and persistence of water stress over time. MSRRI provides information on not only inflow deficit, but also how long it takes to recover from an extreme dry condition based on reservoir levels. In other words, a positive value of MSRRI indicates that the demand can be satisfied during the timeframe of analysis, regardless of the inflow conditions. Example applications of MSRRI for several extreme wet and dry conditions are provided including the Australian Millennium Drought (1998-2010) and the 2014 California Drought. One interesting feature of MSRRI is that it is

standardized similar to the currently available standardized drought indicators such as the Standardized Precipitation Index (SPI). This means that values of MSRRI can be compared with the commonly used drought indicators. Current drought indicators do not consider the local capacity to cope with dry conditions. MSRRI, on the other hand, can provide complementary information on drought recovery based on reservoir storage and demand that cannot be achieved from the commonly used drought indicators such as SPI.

CHAPTER 7: Summary and Conclusions

The overarching goal of this study was to address the research gaps outlined in Chapter 1. In this dissertation, several novel evaluation metrics are introduced that can be used for evaluation of errors and biases in climate model simulations used as input for climate change assessment studies (Chapter 2). These indicators are then used for evaluating 34 CMIP5 climate model simulations (Chapter 3). Furthermore, a modeling framework is introduced for improving global hydrologic models through a comprehensive multi-objective calibration approach (Chapter 4). Then, an additional layer for including man-made reservoirs is integrated in the above mentioned calibrated global hydrologic model. This additional layer is designed to account for local reservoirs that provide resilience against climatic change and variability (Chapter 5). Finally, having global hydrologic simulations and the local response including the effects of reservoirs, a multivariate framework is introduced for water stress assessment using both climatic information and local reservoir conditions (Chapter 6). In the following, the conclusions for each chapter are summarized:

Chapters 2: Combining categorical and continuous evaluation metrics: toward a volumetric contingency table

In this part of the study, the contingency table categorical metrics are extended to volumetric indices for evaluation of gridded data relative to a reference data set. The suggested volumetric indicators decompose the total volumetric error (bias) into volumetric error terms associated with hit, false, and miss components in simulations. Volume of water associated with the hit, false, and miss are important for assessing water availability under climatic change and variability. The results show

that the volumetric indices provide additional information beyond the commonly used categorical metrics that can be useful in evaluating gridded data sets. Chapter 2 contributes to ongoing metrics development efforts for validation and verification of climate model simulations.

Chapters 3: Evaluation of CMIP5 Precipitation Simulations

The objective of this chapter was to cross-validate CMIP5 historical simulations of precipitation relative to GPCP reference data, quantifying model pattern discrepancies and biases for both entire data distributions and their upper tails. From the results of the Volumetric Hit Index (*VHI*) pattern analysis of the total monthly precipitation amounts, it is found that most CMIP5 simulations are in fairly good agreement with GPCP observations in many areas, but model replication of observed precipitation patterns over deserts and certain sub-continental regions (e.g. northern Eurasia, central Australia) is problematical. The *VHI* of the multi-model ensemble mean and median also are found to be superior to most CMIP5 model simulations overall.

Analyses of total biases (*B*) in CMIP5 simulations reveal that most models overestimate precipitation in regions of steep topography, while underestimating it leeward of the mountains, as well as over many other arid regions. Moreover, while most climate model simulations show low *B* values over Europe, there are considerable inter-model variations in bias over Australia and Amazonia.

At high quantiles ($> 75\%$ and $> 90\%$) of the distribution of monthly precipitation, the Quantile Bias (*QB*) analyses indicate that CMIP5 simulations show more glaring discrepancies in precipitation amounts with respect to the GPCP satellite observations. While continuing to overestimate precipitation in regions of steep topography, the models generally underestimate it in tropical locations such as Amazonia, central Africa, and southern Asia, as well as in broad

swaths of the extra-tropics such as Australia, the arid regions of northern Africa and central Asia, and northern China, Russia, and Canada. At high precipitation quantiles also, the CMIP5 models show substantially reduced agreement with the patterns of the GPCP reference data. Except over North America, Amazonia, and central Africa, most CMIP5 simulations are lacking in predictive skill ($VHI_{75} \sim 0$) for the higher tail of the precipitation distribution. Note, however, that a low VHI_{75} does not necessarily imply the absence of locally simulated precipitation, only that its amount falls below the given reference data's threshold value.

These results thus suggest that, while today's climate model simulations are generally in agreement with satellite-based gauge-adjusted estimates of total monthly precipitation in many areas, they presently are not well suited for simulating upper quantiles of the precipitation distribution. Such distribution errors, which have persisted across CMIP phases, are often characterized by a general tendency for the models to precipitate too frequently in light amounts, but too rarely in the intense downbursts that are occasionally observed. In focusing on the upper tails of the precipitation distributions, the present study reveals such model intensity errors in a particularly stark way.

The persistence of these upper-tail errors in all the evaluated CMIP5 simulations is indicative of the presence of general deficiencies in the models. For instance, these systematic precipitation distribution errors do not seem to be very sensitive to inter-model differences in horizontal resolution (e.g. the MRI-ESM1_esm model, with a 160×320 grid, does not clearly outperform other coarser-resolution models in Table 3.1). Substantive differences in error structure also are not apparent between the “_esm” historical simulations with prescribed CO_2 *emissions* and those with prescribed CO_2 *concentrations*. Thus, it is likely that these systematic precipitation errors are due more to general model shortcomings in representing the dynamics or physics of climatic

phenomena than to inter-model differences in greenhouse forcings or horizontal resolution.

An underestimation or incorrect placement of intense tropical and extra-tropical precipitation is also clearly displayed by the CMIP5 simulations analyzed in the present study (e.g. in Figures 3.2 and 3.4). Such ubiquitous precipitation errors suggest that improvements, not only in model convective parameterizations, but also in their representation of sub-grid scale cloud microphysical processes that regulate droplet auto-conversion, accretion, and through-fall, may be essential for better simulation of the observed global precipitation distributions.

Chapter 4: Improving global hydrologic modeling using a multi-objective calibration framework

One of the limitations of most, if not all, current global hydrological models is that they are either used uncalibrated or only calibrated based on runoff. Using long-term historical satellite-based observations, this chapter proposes a framework for multi-objective calibration and parameter estimation. An existing model was modified by adding a soil moisture layer for calibration with satellite observations. A multi objective calibration approach was implemented for calibration based on annual surface runoff and monthly soil moisture. The results show that before calibration, the original model overestimated soil moisture substantially. This issue has significant impacts on assessing climatic change and variability impacts on the water cycle and may lead to misleading results. Applying the multi-objective calibration approach to the global hydrological model resulted in improved simulation results (i.e., significant reduction in the relative errors of both runoff and soil moisture simulations against the observations).

Chapter 5: Accounting for local resilience in climate change impact assessments

In this chapter, the global modeling framework, discussed in Chapter 4, is applied for climate change impact assessment. An additional layer including man-made reservoirs is included in the model to account for local reservoirs that provide resilience against extremes. The novelty of this global-local modeling framework is to take advantage of models in different scales (global hydrology and local reservoirs). The modeling framework is applied over the Melbourne major reservoirs using CMIP5 climate simulations as input. Furthermore, the study uses 17 different operational water demand scenarios ranging from very optimistic to very unfavorable. These scenarios consider population and industrial growth and involve different consumption behaviors. The Results suggest that for a thorough analysis of climate change impact on water resources, including the effect of local resilience (i.e., reservoirs) and the expected demand in the future are fundamental. Our results showed explicit consideration of water demand scenarios is a major step forward from a more large-scale and common top-down climate change impact assessment approach toward a bottom-up local assessment. The framework allows investigating regions, considering the local capacity to cope with extremes and water consumption behaviors which are ignored in large-scale and top-down climate change studies.

Chapter 6: A hybrid multivariate framework for describing water stress relative to the local capacity to cope with extremes

In this chapter, a framework is suggested that integrates both top-down and bottom-up approaches for assessing water stress due to both climatic conditions and local reservoir levels, respectively. In other words, the suggested framework considers both large scale variability and the local capacity for coping with extremes. The indicator representing a top-down approach is

termed Inflow-Demand Reliability (IDR) indicator. The indicator describing the system resilience is named Water Storage Resilience (WSR) indicator. Both indicators are described relative to the average demand to consider the human dimension of the water stress. The hybrid framework that combines the two concepts is termed the Multivariate Standardized Reliability and Resilience Index (MSRRI). This model offers a unique approach for estimating the overall water stress including the local capacity to cope with extremes. MSRRI is based on a nonparametric approach that does not require parameter estimation or any a priori assumption on the underlying distribution function of the original data. The results show that MSRRI is superior to univariate indices because it captures both early onset and persistence of water stress over time. MSRRI provides information on not only inflow deficit, but also how long it takes to recover from an extreme dry condition based on reservoir levels. One interesting feature of MSRRI is that it is standardized similar to the currently available standardized drought indicators such as the Standardized Precipitation Index (SPI).

REFERENCES

- Abramowitz, M. E., and I. A. . E. Stegun (1972), Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. National Bureau of Standards Applied Mathematics Series 55. Tenth Printing.,
- Adler, R. F. et al. (2003), The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present), *Journal of Hydrometeorology*, 4(6), 1147–1167, doi:10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2.
- AghaKouchak, A., and A. Mehran (2013), Extended contingency table: Performance metrics for satellite observations and climate model simulations, *Water Resources Research*, 49(10), 7144–7149, doi:10.1002/wrcr.20498.
- Aghakouchak, A., and A. Mehran (2013), Extended contingency table: Performance metrics for satellite observations and climate model simulations, *Water Resources Research*, 49(10), 7144–7149.
- AghaKouchak, A., and N. Nakhjiri (2012), A near real-time satellite-based global drought climate data record, *Environmental Research Letters*, 7, 1–8, doi:10.1088/1748-9326/7/4/044037.
- AghaKouchak, A., E. Habib, and A. Bárdossy (2010), Modeling Radar Rainfall Estimation Uncertainties: Random Error Model, *Journal of Hydrologic Engineering*, 15(4), 265–274, doi:10.1061/(ASCE)HE.1943-5584.0000185.
- AghaKouchak, A., A. Behrangi, S. Sorooshian, K. Hsu, and E. Amitai (2011), Evaluation of satellite-retrieved extreme precipitation rates across the central United States, *Journal of Geophysical Research*, 116(D2), D02115, doi:10.1029/2010JD014741.
- Aghakouchak, A., A. Mehran, H. Norouzi, and A. Behrangi (2012), Systematic and random error components in satellite precipitation data sets, *Geophysical Research Letters*, 39(9).
- Aghakouchak, A., D. Feldman, M. J. Stewardson, J.-D. Saphores, S. Grant, and B. Sanders (2014), Australia’s drought: lessons for California., *Science (New York, N.Y.)*, 343(6178), 1430–1, doi:10.1126/science.343.6178.1430.
- AghaKouchak, A., L. Cheng, O. Mazdiyasni, and A. Farahmand (2014), Global warming and changes in risk of concurrent climate extremes: Insights from the 2014 California drought, *Geophysical Research Letters*, 41(24), 8847–8852, doi:10.1002/2014GL062308.
- Alcamo, J., P. Döll, T. Henrichs, F. Kaspar, B. Lehner, T. Rösch, and S. Siebert (2003), Development and testing of the WaterGAP 2 global model of water use and availability, *Hydrological Sciences Journal*, 48(February 2015), 317–337, doi:10.1623/hysj.48.3.317.45290.

- Allen, R. G., L. S. Pereira, D. Raes, and M. Smith (1998), Crop evapotranspiration: guidelines for computing crop water requirements. Irrigation and Drainage Paper 56., , 300.
- Anagnostou, E. N., W. F. Krajewski, D.-J. Seo, and E. R. Johnson (1998), Mean-Field Rainfall Bias Studies for WSR-88D, *Journal of Hydrologic Engineering*, 3(3), 149–159, doi:10.1061/(ASCE)1084-0699(1998)3:3(149).
- Anagnostou, E. N., V. Maggioni, E. I. Nikolopoulos, T. Meskele, F. Hossain, and A. Papadopoulos (2010), Benchmarking High-Resolution Global Satellite Rainfall Products to Radar and Rain-Gauge Rainfall Estimates, *IEEE Transactions on Geoscience and Remote Sensing*, 48(4), 1667–1683, doi:10.1109/TGRS.2009.2034736.
- Arnell, N. W. (1999a), A simple water balance model for the simulation of streamflow over a large geographic domain, *Journal of Hydrology*, 217, 314–335, doi:10.1016/S0022-1694(99)00023-2.
- Arnell, N. W. (1999b), Climate change and global water resources, *Global Environmental Change*, 9(June), doi:10.1016/S0959-3780(99)00017-5.
- Barker, F., R. Faggian, and A. J. Hamilton (2011), A History of Wastewater Irrigation in Melbourne, Australia, *Journal of Water Sustainability*, 1(2), 31–50.
- Barnett, T. P. et al. (2008), Human-induced changes in the hydrology of the western United States, *Science*, 319(5866), 1080–1083.
- Van Beek, L. P. H. (2008), Forcing PCR-GLOBWB with CRU data,
- Van Beek, L. P. H., and M. F. P. Bierkens (2008), *The Global Hydrological Model PCR-GLOBWB: Conceptualization, Parameterization and Verification*.
- Van Beek, L. P. H., Y. Wada, and M. F. P. Bierkens (2011), Global monthly water stress: 1. Water balance and water availability, *Water Resour. Res.*, 47.
- Van Beek, L. P. H., Y. Wada, and M. F. P. Bierkens (2011), Global monthly water stress: 1. Water balance and water availability, *Water Resources Research*, 47(July 2010), doi:10.1029/2010WR009791.
- Behrangi, A., B. Khakbaz, T. C. Jaw, A. AghaKouchak, K. Hsu, and S. Sorooshian (2011), Hydrologic evaluation of satellite precipitation products over a mid-size basin, *Journal of Hydrology*, 397(3-4), 225–237, doi:10.1016/j.jhydrol.2010.11.043.
- Bergström, S., and V. P. Singh (1995), The HBV model., *Computer models of watershed hydrology.*, 443–476.
- Bolvin, D. T., R. F. Adler, G. J. Huffman, E. J. Nelkin, and J. P. Poutiainen (2009), Comparison of GPCP Monthly and Daily Precipitation Estimates with High-Latitude Gauge

- Observations, *Journal of Applied Meteorology and Climatology*, 48(9), 1843–1857, doi:10.1175/2009JAMC2147.1.
- Bony, S. et al. (2006), How Well Do We Understand and Evaluate Climate Change Feedback Processes?, *Journal of Climate*, 19(15), 3445–3482, doi:10.1175/JCLI3819.1.
- Boyle, D. P., H. V. Gupta, and S. Sorooshian (2000), Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods, *Water Resources Research*, 36(12), 3663–3674, doi:10.1029/2000WR900207.
- Ter Braak, C. J. F., and J. a. Vrugt (2008), Differential Evolution Markov Chain with snooker updater and fewer chains, *Statistics and Computing*, 18, 435–446, doi:10.1007/s11222-008-9104-9.
- Brekke, Levi D, Barsugli, J. J. (2013), Uncertainties in Projections of Future Changes in Extremes, in *Extremes in a Changing Climate*, pp. 309–345, Springer.
- Brown, B. G., R. R. Bullock, C. A. David, J. H. Gotway, M. B. Chapman, A. Takacs, E. Gilleland, K. Manning, and J. Mahoney (2004), New verification approaches for convective weather forecasts, in *11th Conf. Aviation, Range, and Aerospace Meteorology*.
- Brown, J. R., C. Jakob, and J. M. Haynes (2010), An Evaluation of Rainfall Frequency and Intensity over the Australian Region in a Global Climate Model, *Journal of Climate*, 23(24), 6504–6525, doi:10.1175/2010JCLI3571.1.
- Catto, J. L., L. C. Shaffrey, and K. I. Hodges (2010), Can Climate Models Capture the Structure of Extratropical Cyclones?, *Journal of Climate*, 23(7), 1621–1635, doi:10.1175/2009JCLI3318.1.
- Catto, J. L., C. Jakob, G. Berry, and N. Nicholls (2012), Relating global precipitation to atmospheric fronts, *Geophysical Research Letters*, 39(10), n/a–n/a, doi:10.1029/2012GL051736.
- Catto, J. L., C. Jakob, and N. Nicholls (2013), A global evaluation of fronts and precipitation in the ACCESS model, *Aust. Meteorol. Oceanogr. Soc. J.*, 63, 191–203.
- Cayan, D. R., E. P. Maurer, M. D. Dettinger, M. Tyree, and K. Hayhoe (2008), Climate change scenarios for the California region, *Climatic Change*, 87(1), S21–S42.
- Cayan, D. R., T. Das, D. W. Pierce, T. P. Barnett, M. Tyree, and A. Gershunov (2010), Future dryness in the southwest US and the hydrology of the early 21st century drought., *Proceedings of the National Academy of Sciences of the United States of America*, 107(50), 21271–21276, doi:10.1073/pnas.0912391107.

- Chadwick, R., I. Boutle, and G. Martin (2013), Spatial Patterns of Precipitation Change in CMIP5: Why the Rich Do Not Get Richer in the Tropics, *Journal of Climate*, 26(11), 3803–3822, doi:10.1175/JCLI-D-12-00543.1.
- CHAMPION, A. J., K. I. HODGES, L. O. BENGTSSON, N. S. KEENLYSIDE, and M. ESCH (2011), Impact of increasing resolution and a warmer climate on extreme weather from Northern Hemisphere extratropical cyclones, *Tellus A*, 63(5), 893–906, doi:10.1111/j.1600-0870.2011.00538.x.
- Chang, F. J., L. Chen, and L. C. Chang (2005), Optimizing the reservoir operating rule curves by genetic algorithms, *Hydrological Processes*, 19(April 2003), 2277–2289, doi:10.1002/hyp.5674.
- Christensen, J. H., F. Boberg, O. B. Christensen, and P. Lucas-Picher (2008), On the need for bias correction of regional climate change projections of temperature and precipitation, *Geophysical Research Letters*, 35(20), L20709, doi:10.1029/2008GL035694.
- Christensen, N., N. Christensen, D. P. Lettenmaier, and D. P. Lettenmaier (2007), A multimodel ensemble approach to assessment of climate change impacts on the hydrology and water resources of the Colorado River basin, *Hydrology and Earth System Sciences*, 11, 1417–1434.
- Crutchfield, S. (2012), *U.S. Drought 2012: Farm and Food Impacts*.
- Dai, A. (2006), Precipitation Characteristics in Eighteen Coupled Climate Models, *Journal of Climate*, 19(18), 4605–4630, doi:10.1175/JCLI3884.1.
- Davis, C. A., B. G. Brown, R. Bullock, and J. Halley-Gotway (2009), The Method for Object-Based Diagnostic Evaluation (MODE) Applied to Numerical Forecasts from the 2005 NSSL/SPC Spring Program, *Weather and Forecasting*, 24(5), 1252–1267, doi:10.1175/2009WAF2222241.1.
- Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan (2002), A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE Transactions on Evolutionary Computation*, 6(2), 182–197, doi:10.1109/4235.996017.
- Dee, D. P. et al. (2011), The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, 137(April), 553–597, doi:10.1002/qj.828.
- Dessai, S., and M. Hulme (2004), Does climate adaptation policy need probabilities?, *Climate Policy*, 4(2), 107–128, doi:10.1080/14693062.2004.9685515.
- Van Dijk, A. I. J. M., H. E. Beck, R. S. Crosbie, R. a M. De Jeu, Y. Y. Liu, G. M. Podger, B. Timbal, and N. R. Viney (2013), The Millennium Drought in southeast Australia (2001-

- 2009): Natural and human causes and implications for water resources, ecosystems, economy, and society, *Water Resources Research*, 49, 1040–1057, doi:10.1002/wrcr.20123.
- Dirmeyer, P. a., a. J. Dolman, and N. Sato (1999), The Pilot Phase of the Global Soil Wetness Project, *Bulletin of the American Meteorological Society*, 80, 851–878, doi:10.1175/1520-0477(1999)080<0851:TPPOTG>2.0.CO;2.
- Döll, P., F. Kaspar, and B. Lehner (2003), A global hydrological model for deriving water availability indicators: Model tuning and validation, *Journal of Hydrology*, 270, 105–134, doi:10.1016/S0022-1694(02)00283-4.
- Dorigo, W. A., K. Scipal, R. M. Parinussa, Y. Y. Liu, W. Wagner, R. A. M. de Jeu, and V. Naeimi (2010), Error characterisation of global active and passive microwave soil moisture datasets, *Hydrology and Earth System Sciences*, 14(12), 2605–2616, doi:10.5194/hess-14-2605-2010.
- Dosio, A., and P. Paruolo (2011), Bias correction of the ENSEMBLES high-resolution climate change projections for use by impact models: Evaluation on the present climate, *Journal of Geophysical Research*, 116(D16), D16106, doi:10.1029/2011JD015934.
- Edwards, D. C. (1997), Characteristics of 20th Century Drought in the United States at Multiple Time Scales, , 174.
- Entekhabi, D., E. Njoku, P. O’Neill, M. Spencer, T. Jackson, J. Entin, E. Im, and K. Kellogg (2008), The Soil Moisture Active/Passive Mission (SMAP), in *IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium*, vol. 3, pp. III – 1–III – 4, IEEE.
- Entekhabi, D., R. H. Reichle, R. D. Koster, and W. T. Crow (2010a), Performance Metrics for Soil Moisture Retrievals and Application Requirements, *Journal of Hydrometeorology*, 11(3), 832–840, doi:10.1175/2010JHM1223.1.
- Entekhabi, D. et al. (2010b), The Soil Moisture Active Passive (SMAP) Mission, *Proceedings of the IEEE*, 98(5), 704–716, doi:10.1109/JPROC.2010.2043918.
- Farahmand, A., and A. AghaKouchak (2015), A Generalized Framework for Deriving Nonparametric Standardized Drought Indicators, *Advances in Water Resources*, doi:10.1016/j.advwatres.2014.11.012.
- Feddema, J., K. Oleson, G. Bonan, L. Mearns, W. Washington, G. Meehl, and D. Nychka (2005), A comparison of a GCM response to historical anthropogenic land cover change and model sensitivity to uncertainty in present-day land cover representations, *Climate Dynamics*, 25(6), 581–609, doi:10.1007/s00382-005-0038-z.
- Fekete, B. M. (2002), High-resolution fields of global runoff combining observed river discharge and simulated water balances, *Global Biogeochemical Cycles*, 16(3), doi:10.1029/1999GB001254.

- Fekete, B. M., C. J. Vörösmarty, and W. Grabs (2002), High-resolution fields of global runoff combining observed river discharge and simulated water balances, *Global Biogeochemical Cycles*, 16(3), 15–1–15–10, doi:10.1029/1999GB001254.
- Fiering, M. B. (1982), Alternative indices of resilience, *Water Resources Research*, 18(1), 33–39, doi:10.1029/WR018i001p00033.
- Gaetani, M., and E. Mohino (2013), Decadal Prediction of the Sahelian Precipitation in CMIP5 Simulations,
- Gao, H., C. Birkett, and D. P. Lettenmaier (2012), Global monitoring of large reservoir storage from satellite remote sensing, *Water Resources Research*, 48(February), 1–12, doi:10.1029/2012WR012063.
- Gebremichael, M., and W. F. Krajewski (2007), Application of copulas to modeling temporal sampling errors in satellite-derived rainfall estimates, *Journal of Hydrologic Engineering*, 12(4), 404–408.
- Ghan, S., X. Bian, A. Hunt, and A. Coleman (2002), The thermodynamic influence of subgrid orography in a global climate model, *Climate Dynamics*, 20(1), 31–44, doi:10.1007/s00382-002-0257-5.
- Gilleland, E. (2013), Testing Competing Precipitation Forecasts Accurately and Efficiently: The Spatial Prediction Comparison Test, *Monthly Weather Review*, 141(1), 340–355, doi:10.1175/MWR-D-12-00155.1.
- Gleckler, P. J., K. E. Taylor, and C. Doutriaux (2008), Performance metrics for climate models, *Journal of Geophysical Research*, 113(D6), D06104, doi:10.1029/2007JD008972.
- Gong, L., E. Widén-Nilsson, S. Halldin, and C. Y. Xu (2009), Large-scale runoff routing with an aggregated network-response function, *Journal of Hydrology*, 368(1-4), 237–250, doi:10.1016/j.jhydrol.2009.02.007.
- Gosling, S. N., and N. W. Arnell (2011), Simulating current global river runoff with a global hydrological model: Model revisions, validation, and sensitivity analysis, *Hydrological Processes*, 25, 1129–1145, doi:10.1002/hyp.7727.
- Gourley, J. J., J. M. Erlingis, Y. Hong, and E. B. Wells (2012), Evaluation of Tools Used for Monitoring and Forecasting Flash Floods in the United States, *Weather and Forecasting*, 27(1), 158–173, doi:10.1175/WAF-D-10-05043.1.
- Grant, S. B. et al. (2012), Taking the “Waste” Out of “Wastewater” for Human Water Security and Ecosystem Sustainability, *Science*, 337, 681–686, doi:10.1126/science.1216852.
- Grant, S. B., T. D. Fletcher, D. Feldman, J. D. Saphores, P. L. M. Cook, M. Stewardson, K. Low, K. Burry, and A. J. Hamilton (2013), Adapting urban water systems to a changing climate:

- Lessons from the millennium drought in southeast Australia, *Environmental Science and Technology*, 47, 10727–10734, doi:10.1021/es400618z.
- Gringorten, I. I. (1963), A plotting rule for extreme probability paper, *Journal of Geophysical Research*, 68(3), 813–814, doi:10.1029/JZ068i003p00813.
- Guilyardi, E., V. Balaji, B. Lawrence, S. Callaghan, C. Deluca, S. Denvil, M. Lautenschlager, M. Morgan, S. Murphy, and K. E. Taylor (2013), Documenting Climate Models and Their Simulations, *Bulletin of the American Meteorological Society*, 94(5), 623–627, doi:10.1175/BAMS-D-11-00035.1.
- Haario, H., E. Saksman, and J. Tamminen (2001), An adaptive metropolis algorithm. Bernoulli 7 223–242, *Mathematical Reviews (MathSciNet): MRI828504 Digital Object Identifier: doi*, 10, 3318737.
- Haddeland, I., T. Skaugen, and D. P. Lettenmaier (2006), Anthropogenic impacts on continental surface water fluxes, *Geophysical Research Letters*, 33(April), 2–5, doi:10.1029/2006GL026047.
- Haddeland, I. et al. (2011), Multimodel estimate of the global terrestrial water balance: setup and first results, , 869–884, doi:10.1175/2011JHM1324.1.
- Haerter, J. O., S. Hagemann, C. Moseley, and C. Piani (2011), Climate model bias correction and the role of timescales, *Hydrology and Earth System Sciences*, 15(3), 1065–1079, doi:10.5194/hess-15-1065-2011.
- Hagemann, S. (2002), Report No . 336 An Improved Land Surface Parameter Dataset for Global and Regional Climate Models by Authors, *Planck Institute for Meteorology Report*, (336), 21.
- Hagemann, S. et al. (2013), Climate change impact on available water resources obtained using multiple global climate and hydrology models, *Earth System Dynamics*, 4, 129–144, doi:10.5194/esd-4-129-2013.
- Haile, A. T., E. Habib, and T. Rientjes (2013), Evaluation of the climate prediction center (CPC) morphing technique (CMORPH) rainfall product on hourly time scales over the source of the Blue Nile River, *Hydrological Processes*, 27(12), 1829–1839, doi:10.1002/hyp.9330.
- Hallegatte, S. (2009), Strategies to adapt to an uncertain climate change, *Global Environmental Change*, 19, 240–247, doi:10.1016/j.gloenvcha.2008.12.003.
- Hanasaki, N., S. Kanae, and T. Oki (2006), A reservoir operation scheme for global river routing models, *Journal of Hydrology*, 327, 22–41, doi:10.1016/j.jhydrol.2005.11.011.
- Hanasaki, N., S. Kanae, T. Oki, K. Masuda, K. Motoya, Y. Shen, and K. Tanaka (2007), An integrated model for the assessment of global water resources – Part 1: Input

meteorological forcing and natural hydrological cycle modules, *Hydrology and Earth System Sciences Discussions*, 4(October), 3535–3582, doi:10.5194/hessd-4-3535-2007.

Hanasaki, N., S. Kanae, T. Oki, K. Masuda, K. Motoya, N. Shirakawa, Y. Shen, and K. Tanaka (2008), An integrated model for the assessment of global water resources – Part 1 : Model description and input meteorological forcing, *Hydrol. Earth Syst. Sci.*, 12, 1007–1025, doi:10.5194/hess-12-1027-2008.

Hao, Z., A. AghaKouchak, and T. J. Phillips (2013), Changes in concurrent monthly precipitation and temperature extremes, *Environmental Research Letters*, 8(3), 034014, doi:10.1088/1748-9326/8/3/034014.

Harberg, R. J. (1997), *Planning and Managing Reliable Urban Water Systems*, American Water Works Association.

Hashimoto, T., J. R. Stedinger, and D. P. Loucks (1982), Reliability, resiliency, and vulnerability criteria for water resource system performance evaluation, *Water Resources Research*, 18(1), 14, doi:10.1029/WR018i001p00014.

Hazen, A. (1914), Storage to be Provided in Impounded Reservoirs for Municipal Water Supply, *Transactions of the American Society of Civil Engineers*, 77, 1539–1640.

Henderson-Sellers, A., A. J. Pitman, P. K. Love, P. Irannejad, and T. H. Chen (1995), The Project for Intercomparison of Land Surface Parameterization Schemes (PILPS): Phases 2 and 3, *Bulletin of the American Meteorological Society*, 76(4), 489–503, doi:10.1175/1520-0477(1995)076<0489:TPFIOL>2.0.CO;2.

Hirota, N., and Y. N. Takayabu (2013), Reproducibility of precipitation distribution over the tropical oceans in CMIP5 multi-climate models compared to CMIP3, *Climate Dynamics*, 41(11-12), 2909–2920, doi:10.1007/s00382-013-1839-0.

Hirpa, F. A., M. Gebremichael, and T. Hopson (2010), Evaluation of High-Resolution Satellite Precipitation Products over Very Complex Terrain in Ethiopia, *Journal of Applied Meteorology and Climatology*, 49(5), 1044–1051, doi:10.1175/2009JAMC2298.1.

Hossain, F., and G. J. Huffman (2008), Investigating Error Metrics for Satellite Rainfall Data at Hydrologically Relevant Scales, *Journal of Hydrometeorology*, 9(3), 563–575, doi:10.1175/2007JHM925.1.

Hsu, K., X. Gao, S. Sorooshian, and H. V. Gupta (1997), Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks, *Journal of Applied Meteorology*, 36(9), 1176–1190, doi:10.1175/1520-0450(1997)036<1176:PEFRSI>2.0.CO;2.

- Huffman, G. J., R. F. Adler, D. T. Bolvin, and G. Gu (2009), Improving the global precipitation record: GPCP Version 2.1, *Geophysical Research Letters*, 36(17), L17808, doi:10.1029/2009GL040000.
- Huntington, T. G. (2006), Evidence for intensification of the global water cycle: Review and synthesis, *Journal of Hydrology*, 319, 83–95, doi:10.1016/j.jhydrol.2005.07.003.
- Hurst, H. E. (1951), Long term storage capacity of reservoirs, , 116, 770–799.
- Ionescu, C., R. J. T. Klein, J. Hinkel, K. S. Kavi Kumar, and R. Klein (2008), Towards a Formal Framework of Vulnerability to Climate Change, *Environmental Modeling & Assessment*, 14(1), 1–16, doi:10.1007/s10666-008-9179-x.
- IPCC (2007), *Climate Change 2007: Impacts, Adaptation, and Vulnerability*, Exit EPA Disclaimer Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change [Parry, Martin L., Canziani, Osvaldo F., Palutikof, Jean P., van der Linden, Paul J., and Hanson, Clair E. (eds.)]. Cambri.
- Jackson, T.J., Entekhabi, D., Njoku, E. (2005), The Hydros Mission and Validation of Soil Moisture Retrievals.
- JAIN, S. K., and P. K. BHUNYA (2008), Reliability, resilience and vulnerability of a multipurpose storage reservoir / Confiance, résilience et vulnérabilité d'un barrage multi-objectifs, *Hydrological Sciences Journal*, 53(2), 434–447, doi:10.1623/hysj.53.2.434.
- Jiang, J. H. et al. (2012), Evaluation of cloud and water vapor simulations in CMIP5 climate models using NASA “A-Train” satellite observations, *Journal of Geophysical Research*, 117(D14), D14105, doi:10.1029/2011JD017237.
- Joetzjer, E., H. Douville, C. Delire, and P. Ciais (2013), Present-day and future Amazonian precipitation in global climate models: CMIP5 versus CMIP3, *Climate Dynamics*, 41(11-12), 2921–2936, doi:10.1007/s00382-012-1644-1.
- John, V. O., and B. J. Soden (2007), Temperature and humidity biases in global climate models and their impact on climate feedbacks, *Geophysical Research Letters*, 34(18), L18704, doi:10.1029/2007GL030429.
- Karl, T. R., J. M. Melillo, and T. C. Peterson (2009), *Global Climate Change Impacts in the United States*.
- Kennedy, J., R. C. Eberhart, and Y. Shi (2001), chapter seven - The Particle Swarm, in *The Morgan Kaufmann Series in Artificial Intelligence*, edited by J. Kennedy and R. C. E. B. T.-S. I. Shi, pp. 287–325, Morgan Kaufmann, San Francisco.
- Kerr, Y. H., P. Waldteufel, J.-P. Wigneron, J. Martinuzzi, J. Font, and M. Berger (2001), Soil moisture retrieval from space: the Soil Moisture and Ocean Salinity (SMOS) mission, *IEEE*


Transactions on Geoscience and Remote Sensing, 39(8), 1729–1735,
doi:10.1109/36.942551.

- Kerr, Y. H. et al. (2012), The SMOS Soil Moisture Retrieval Algorithm, *IEEE Transactions on Geoscience and Remote Sensing*, 50(5), 1384–1403, doi:10.1109/TGRS.2012.2184548.
- KJELDSEN, T. R., and D. ROSBJERG (2004), Choice of reliability, resilience and vulnerability estimators for risk assessments of water resources systems, *Hydrological sciences journal*, 49(5), 755–767.
- Kumar, S., V. Merwade, J. L. Kinter, and D. Niyogi (2013), Evaluation of Temperature and Precipitation Trends and Long-Term Persistence in CMIP5 Twentieth-Century Climate Simulations, *Journal of Climate*, 26(12), 4168–4185, doi:10.1175/JCLI-D-12-00259.1.
- Kundzewicz, Z., L. Mata, N. Arnell, P. Döll, B. Jimenez, K. Miller, T. Oki, Z. Sen, and I. Shiklomanov (2008), The implications of projected climate change for freshwater resources and their management, *Hydrological Sciences Journal*, 53(1), 3 – 10.
- Kundzewicz, Z. W., and J. Kindler (1995), Multiple criteria for evaluation of reliability aspects of water resource systems, *IAHSAISH Publication*, (231), 217–224.
- Lebsock, M., H. Morrison, and A. Gettelman (2013), Microphysical implications of cloud-precipitation covariance derived from satellite remote sensing, *Journal of Geophysical Research: Atmospheres*, 118(12), 6521–6533, doi:10.1002/jgrd.50347.
- Li, H., J. Sheffield, and E. F. Wood (2010), Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching, *Journal of Geophysical Research*, 115(D10), D10101, doi:10.1029/2009JD012882.
- 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 general circulation models, *Journal of Geophysical Research*, 99, 14415, doi:10.1029/94JD00483.
- Liepert, B. G., and M. Previdi (2012), Inter-model variability and biases of the global water cycle in CMIP3 coupled climate models, *Environmental Research Letters*, 7(1), 014006, doi:10.1088/1748-9326/7/1/014006.
- Liu, C., R. P. Allan, and G. J. Huffman (2012), Co-variation of temperature and precipitation in CMIP5 models and satellite observations, *Geophysical Research Letters*, 39(13), n/a–n/a, doi:10.1029/2012GL052093.
- Liu, Y. Y., R. M. Parinussa, W. A. Dorigo, R. A. M. De Jeu, W. Wagner, A. I. J. M. van Dijk, M. F. McCabe, and J. P. Evans (2011), Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals, *Hydrology And Earth System Sciences*, 15(2), 425–436, doi:10.5194/hess-15-425-2011.

- Liu, Z., A. Mehran, T. J. Phillips, and A. AghaKouchak (2014), Seasonal and regional biases in CMIP5 precipitation simulations, *Climate Research*, 60, 35–50, doi:10.3354/cr01221.
- Loew, a., T. Stacke, W. Dorigo, R. De Jeu, and S. Hagemann (2013), Potential and limitations of multidecadal satellite soil moisture observations for selected climate model evaluation studies, *Hydrology and Earth System Sciences*, 17, 3523–3542, doi:10.5194/hess-17-3523-2013.
- Madani, K. (2010), Game theory and water resources, *Journal of Hydrology*, 381(3-4), 225–238, doi:10.1016/j.jhydrol.2009.11.045.
- Mastrandrea, M. D., N. E. Heller, T. L. Root, and S. H. Schneider (2010), Bridging the gap: Linking climate-impacts research with adaptation planning and management, *Climatic Change*, 100, 87–101, doi:10.1007/s10584-010-9827-4.
- Matalas C. Nicholas, F. B. M. (1977), *Climate, Climatic Change, and Water Supply*, The National Academies Press, Washington, DC.
- 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.
- McDonald, R. I., P. Green, D. Balk, B. M. Fekete, C. Revenga, M. Todd, and M. Montgomery (2011), Urban growth, climate change, and freshwater availability., *Proceedings of the National Academy of Sciences of the United States of America*, 108(15), 6312–6317, doi:10.1073/pnas.1011615108.
- McMahon, T. A., A. J. Adeloye, and S.-L. Zhou (2006), Understanding performance measures of reservoirs, *Journal of Hydrology*, 324(1-4), 359–382, doi:10.1016/j.jhydrol.2005.09.030.
- McMahon, T. a., G. G. S. Pegram, R. M. Vogel, and M. C. Peel (2007), Revisiting reservoir storage-yield relationships using a global streamflow database, *Advances in Water Resources*, 30, 1858–1872, doi:10.1016/j.advwatres.2007.02.003.
- Mehran, A., and A. AghaKouchak (2014), Capabilities of satellite precipitation datasets to estimate heavy precipitation rates at different temporal accumulations, *Hydrological Processes*, 28(4), 2262–2270, doi:10.1002/hyp.9779.
- Mehran, A., and A. Aghakouchak (2014), Capabilities of satellite precipitation datasets to estimate heavy precipitation rates at different temporal accumulations, *Hydrological Processes*, 28(4), 2262–2270.
- Mehran, A., A. AghaKouchak, and T. J. Phillips (2014), Evaluation of CMIP5 continental precipitation simulations relative to satellite-based gauge-adjusted observations, *Journal of Geophysical Research: Atmospheres*, 119(4), 1695–1707, doi:10.1002/2013JD021152.

- Mehrotra, R., R. Srikanthan, and A. Sharma (2006), A comparison of three stochastic multi-site precipitation occurrence generators, *Journal of Hydrology*, 331, 280–292.
- Meigh, J. R., and E. L. Tate (2002), The Gwava Model - Development of A Global-scale Methodology To Assess The Combined Impact of Climate and Land Use Changes, *EGS XXVII General Assembly*.
- Merabtene, T., A. Kawamura, K. Jinno, and J. Olsson (2002), Risk assessment for optimal drought management of an integrated water resources system using a genetic algorithm, *Hydrological Processes*, 16(11), 2189–2208, doi:10.1002/hyp.1150.
- Merritt, W. S., Y. Alila, M. Barton, B. Taylor, S. Cohen, and D. Neilsen (2006), Hydrologic response to scenarios of climate change in sub watersheds of the Okanagan basin, British Columbia, *Journal of Hydrology*, 326(1-4), 79–108, doi:10.1016/j.jhydrol.2005.10.025.
- Miller, F. et al. (2010), Resilience and vulnerability : complementary or conflicting concepts?,
- Milly, P. C. D., and a. B. Shmakin (2002), Global Modeling of Land Water and Energy Balances. Part I: The Land Dynamics (LaD) Model, *Journal of Hydrometeorology*, 3(1969), 283–299, doi:10.1175/1525-7541(2002)003<0283:GMOLWA>2.0.CO;2.
- Min, S.-K., D. Simonis, and A. Hense (2007), Probabilistic climate change predictions applying Bayesian model averaging, *Philosophical Transactions of the Royal Society A - Mathematical Physical and Engineering Sciences*, 365(1857), 2103–2116, doi:10.1098/rsta.2007.2070.
- Moise, A. F., and F. P. Delage (2011), New climate model metrics based on object-orientated pattern matching of rainfall, *Journal of Geophysical Research*, 116(D12), D12108, doi:10.1029/2010JD015318.
- Moody, P., and C. Brown (2013), Robustness indicators for evaluation under climate change: Application to the upper Great Lakes, *Water Resources Research*, 49(6), 3576–3588, doi:10.1002/wrcr.20228.
- Moss, R. H. et al. (2010), The next generation of scenarios for climate change research and assessment., *Nature*, 463(7282), 747–756, doi:10.1038/nature08823.
- Moy, W.-S., J. L. Cohon, and C. S. ReVelle (1986), A Programming Model for Analysis of the Reliability, Resilience, and Vulnerability of a Water Supply Reservoir, *Water Resources Research*, 22(4), 489–498, doi:10.1029/WR022i004p00489.
- Müller Schmied, H., S. Eisner, D. Franz, M. Wattenbach, F. T. Portmann, M. Flörke, and P. Döll (2014), Sensitivity of simulated global-scale freshwater fluxes and storages to input data, hydrological model structure, human water use and calibration, *Hydrology and Earth System Sciences Discussions*, 11, 1583–1649, doi:10.5194/hessd-11-1583-2014.

- Nalbantis, I., and D. Koutsoyiannis (1997), A parametric rule for planning and management of multiple-reservoir systems, *Water Resources Research*, 33(9), 2165–2177, doi:10.1029/97WR01034.
- Naresh Kumar, M., C. S. Murthy, M. V. R. Sessa Sai, and P. S. Roy (2009), On the use of Standardized Precipitation Index (SPI) for drought intensity assessment, *Meteorological Applications*, 16(3), 381–389, doi:10.1002/met.136.
- Nijssen, B., G. M. O'Donnell, D. P. Lettenmaier, D. Lohmann, and E. F. Wood (2001), Predicting the discharge of global rivers, *Journal of Climate*, 14, 3307–3323, doi:10.1175/1520-0442(2001)014<3307:PTDOGR>2.0.CO;2.
- Norouzi, H., M. Temimi, W. B. Rossow, C. Pearl, M. Azarderakhsh, and R. Khanbilvardi (2011), The sensitivity of land emissivity estimates from AMSR-E at C and X bands to surface properties, *Hydrology and Earth System Sciences*, 15(11), 3577–3589, doi:10.5194/hess-15-3577-2011.
- Oki, T., Y. Agata, S. Kanae, T. Saruhashi, D. Yang, and K. Musiak (2001), Global assessment of current water resources using total runoff integrating pathways, *Hydrological Sciences Journal*, 46(February 2015), 983–995, doi:10.1080/02626660109492890.
- Oliveira, R., and D. P. Loucks (1997), Operating rules for multireservoir systems, *Water Resources Research*, 33(4), 839, doi:10.1029/96WR03745.
- Palmer, M. a., C. a. Reidy Liermann, C. Nilsson, M. Flörke, J. Alcamo, P. S. Lake, and N. Bond (2008), Climate change and the world's river basins: Anticipating management options, *Frontiers in Ecology and the Environment*, 6, 81–89, doi:10.1890/060148.
- Pfahl, S., and H. Wernli (2012), Quantifying the Relevance of Cyclones for Precipitation Extremes, *Journal of Climate*, 25(19), 6770–6780, doi:10.1175/JCLI-D-11-00705.1.
- Phillips, T. J., and P. J. Gleckler (2006), Evaluation of continental precipitation in 20th century climate simulations: The utility of multimodel statistics, *Water Resources Research*, 42(3), n/a–n/a, doi:10.1029/2005WR004313.
- Pinker, R. T., D. Sun, M.-P. Hung, C. Li, and J. B. Basara (2009), Evaluation of Satellite Estimates of Land Surface Temperature from GOES over the United States, *Journal of Applied Meteorology and Climatology*, 48(1), 167–180, doi:10.1175/2008JAMC1781.1.
- Polcher, J., K. Laval, L. Dümenil, J. Lean, and P. R. Rowntree (1996), Comparing three land surface schemes used in general circulation models, *Journal of Hydrology*, 180, 373–394, doi:10.1016/0022-1694(95)02886-2.
- Preston, B. L., E. J. Yuen, and R. M. Westaway (2011), Putting vulnerability to climate change on the map: a review of approaches, benefits, and risks, *Sustainability Science*, 6(2), 177–202, doi:10.1007/s11625-011-0129-1.

- Qian, Y., S. J. Ghan, and L. R. Leung (2009), Downscaling hydroclimatic changes over the Western US based on CAM subgrid scheme and WRF regional climate simulations, *International Journal of Climatology*, n/a–n/a, doi:10.1002/joc.1928.
- Reclamation (2011), *SECURE Water Act Section 9503(c)*  *Reclamation Climate Change and Water*.
- Reed, S., V. Koren, M. Smith, Z. Zhang, F. Moreda, D. J. Seo, and D. Participants (2004), Overall distributed model intercomparison project results, *Journal of Hydrology*, 298(1-4), 27–60.
- Reichler, T., and J. Kim (2008), Uncertainties in the climate mean state of global observations, reanalyses, and the GFDL climate model, *Journal of Geophysical Research: Atmospheres*, 113(D5), n/a–n/a, doi:10.1029/2007JD009278.
- Rogelj, J., M. Meinshausen, and R. Knutti (2012), Global warming under old and new scenarios using IPCC climate sensitivity range estimates, *Nature Climate Change*, 2(4), 248–253, doi:10.1038/nclimate1385.
- Rost, S., D. Gerten, A. Bondeau, W. Lucht, J. Rohwer, and S. Schaphoff (2008), Agricultural green and blue water consumption and its influence on the global water system, *Water Resources Research*, 44, 1–17, doi:10.1029/2007WR006331.
- Schaller, N., I. Mahlstein, J. Cermak, and R. Knutti (2011), Analyzing precipitation projections: A comparison of different approaches to climate model evaluation, *Journal of Geophysical Research*, 116(D10), D10118, doi:10.1029/2010JD014963.
- Schewe, J. et al. (2014), Multimodel assessment of water scarcity under climate change., *Proceedings of the National Academy of Sciences of the United States of America*, 111(9), 3245–50, doi:10.1073/pnas.1222460110.
- Schlenker, W., W. M. Hanemann, and A. C. Fisher (2007), Water availability, degree days, and the potential impact of climate change on irrigated agriculture in California, *Climatic Change*, 81(2007), 19–38, doi:10.1007/s10584-005-9008-z.
- Seager, R. et al. (2007a), Model projections of an imminent transition to a more arid climate in southwestern North America, *Science*, 316(5828), 1181–1184, doi:10.1126/science.1139601.
- Seager, R. et al. (2007b), Model projections of an imminent transition to a more arid climate in southwestern North America., *Science (New York, N.Y.)*, 316(2007), 1181–1184, doi:10.1126/science.1139601.
- Sheffield, J., and E. F. Wood (2007), Characteristics of global and regional drought, 1950-2000: Analysis of soil moisture data from off-line simulation of the terrestrial hydrologic cycle, *Journal of Geophysical Research: Atmospheres*, 112, 1–21, doi:10.1029/2006JD008288.

- Shiklomanov, I. A., A. I. Shiklomanov, R. B. Lammers, B. J. Peterson, and C. J. Vorosmarty (2000), The Dynamics of River Water Inflow to the Arctic Ocean, in *The Freshwater Budget of the Arctic Ocean SE - 13*, vol. 70, edited by E. Lewis, E. P. Jones, P. Lemke, T. Prowse, and P. Wadhams, pp. 281–296, Springer Netherlands.
- Sillmann, J., V. V. Kharin, X. Zhang, F. W. Zwiers, and D. Bronaugh (2013), Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate, *Journal of Geophysical Research: Atmospheres*, 118(4), 1716–1733, doi:10.1002/jgrd.50203.
- Simonović, S. (1995), *Modelling and management of sustainable basin-scale water resource systems*, International Association of Hydrological Sciences, Wallingford.
- Sivakumar, B. (2011), Global climate change and its impacts on water resources planning and management: assessment and challenges, *Stochastic Environmental Research and Risk Assessment*, 25(4), 583–600, doi:10.1007/s00477-010-0423-y.
- Sorooshian, S., K.-L. Hsu, X. Gao, H. V. Gupta, B. Imam, and D. Braithwaite (2000), Evaluation of PERSIANN System Satellite–Based Estimates of Tropical Rainfall, *Bulletin of the American Meteorological Society*, 81(9), 2035–2046, doi:10.1175/1520-0477(2000)081<2035:EOPSSE>2.3.CO;2.
- Sorooshian, S. et al. (2011), Advanced Concepts on Remote Sensing of Precipitation at Multiple Scales,
- Srinivasan, K., T. R. Neelakantan, P. S. Narayan, and C. Nagarajukumar (1999), Mixed-Integer Programming Model for Reservoir Performance Optimization, *Journal of Water Resources Planning and Management*, 125(5), 298–301, doi:10.1061/(ASCE)0733-9496(1999)125:5(298).
- Stacke, T., and S. Hagemann (2012), Development and evaluation of a global dynamical wetlands extent scheme, *Hydrology and Earth System Sciences*, 16, 2915–2933, doi:10.5194/hess-16-2915-2012.
- Stanton, E., and E. Fitzgerald (2011), California Water Supply and Demand: Technical Report, , (February), 31.
- Steinschneider, S., and C. Brown (2012), Dynamic reservoir management with real-option risk hedging as a robust adaptation to nonstationary climate, *Water Resources Research*, 48(5), n/a–n/a, doi:10.1029/2011WR011540.
- Stephens, G. L., T. L’Ecuyer, R. Forbes, A. Gettleman, J.-C. Golaz, A. Bodas-Salcedo, K. Suzuki, P. Gabriel, and J. Haynes (2010), Dreary state of precipitation in global models, *Journal of Geophysical Research*, 115(D24), D24211, doi:10.1029/2010JD014532.

- Stoll, S., H. J. Hendricks Franssen, M. Butts, and W. Kinzelbach (2011), Analysis of the impact of climate change on groundwater related hydrological fluxes: A multi-model approach including different downscaling methods, *Hydrology and Earth System Sciences*, 15, 21–38, doi:10.5194/hess-15-21-2011.
- Storn, R., and K. Price (1997), Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces, *Journal of global optimization*, 341–359, doi:10.1023/A:1008202821328.
- Stott, P. A., N. P. Gillett, G. C. Hegerl, D. J. Karoly, D. A. Stone, X. Zhang, and F. Zwiers (2010), Detection and attribution of climate change: a regional perspective, *Wiley Interdisciplinary Reviews: Climate Change*, 1(2), 192–211, doi:10.1002/wcc.34.
- Sudler, C. E. (1927), Storage Required for the Regulation of Stream Flow, *Transactions of the American Society of Civil Engineers*, 91(2), 622–660.
- Sun, Y., S. Solomon, A. Dai, and R. W. Portmann (2007), How Often Will It Rain?, *Journal of Climate*, 20(19), 4801–4818, doi:10.1175/JCLI4263.1.
- Tang, B. (1993), Orthogonal Array-Based Latin Hypercubes, *Journal of the American Statistical Association*, 88(424), 1392–1397, doi:10.1080/01621459.1993.10476423.
- Tarroja, B., A. Aghakouchak, R. Sobhani, D. Feldman, S. Jiang, and S. Samuelsen (2014), Science of the Total Environment Evaluating options for Balancing the Water-Electricity Nexus in California : Part 1 – Securing Water Availability, *Science of the Total Environment*, The, 497-498, 697–710, doi:10.1016/j.scitotenv.2014.06.060.
- Taylor, K. E., R. J. Stouffer, and G. a. Meehl (2012), An overview of CMIP5 and the experiment design, *Bulletin of the American Meteorological Society*, 93(april), 485–498, doi:10.1175/BAMS-D-11-00094.1.
- Taylor, R. G. et al. (2013), Ground water and climate change, *Nature Climate Change*, 3(April), 322–329, doi:10.1038/nclimate1744.
- Tian, Y., C. D. Peters-Lidard, J. B. Eylander, R. J. Joyce, G. J. Huffman, R. F. Adler, K. Hsu, F. J. Turk, M. Garcia, and J. Zeng (2009), Component analysis of errors in satellite-based precipitation estimates, *Journal of Geophysical Research*, 114(D24), D24101, doi:10.1029/2009JD011949.
- Trenberth, K. E. (2001), Climate Variability and Global Warming, *Science*, 293 (5527), 48–49, doi:10.1126/science.293.5527.48.
- Trenberth, K. E., P. G. Ambenje, R. Bojariu, D. R. Easterling, a M. G. Klein, D. E. Parker, J. a Renwick, F. Rahimzadeh, M. M. Rusticucci, and B. J. Soden (2007), Introduction Changes in Surface Climate : Temperature Changes in Surface Climate : Precipitation and

- Atmospheric Moisture Changes in the Free Atmosphere, *Climate Change 2007*, (September), 1–4.
- Turk, F. J., P. Arkin, M. R. P. Sapiano, and E. E. Ebert (2008), Evaluating High-Resolution Precipitation Products, *Bulletin of the American Meteorological Society*, 89(12), 1911–1916, doi:10.1175/2008BAMS2652.1.
- Vogel, R. M., and R. A. Bolognese (1995), Storage-Reliability-Resilience-Yield Relations for Over-Year Water Supply Systems, *Water Resources Research*, 31(3), 645–654, doi:10.1029/94WR02972.
- Vorosmarty, C. J. Green, P. Salisbury, J. Lammers, R. B. (2000), Global Water Resources: Vulnerability from Climate Change and Population Growth, *Science*, 289(JULY), 284–288, doi:10.1126/science.289.5477.284.
- Vrugt, J. a, and B. a Robinson (2007), Improved evolutionary optimization from genetically adaptive multimethod search., *Proceedings of the National Academy of Sciences of the United States of America*, 104, 708–711, doi:10.1073/pnas.0610471104.
- Vrugt, J. a., C. J. F. ter Braak, C. G. H. Diks, B. a. Robinson, J. M. Hyman, and D. Higdon (2009a), Accelerating Markov Chain Monte Carlo Simulation by Differential Evolution with Self-Adaptive Randomized Subspace Sampling , *International Journal of Nonlinear Sciences and Numerical Simulation* , 10 (March), 273, doi:10.1515/IJNSNS.2009.10.3.273.
- Vrugt, J. a., B. a. Robinson, and J. M. Hyman (2009b), Self-adaptive multimethod search for global optimization in real-parameter spaces, *IEEE Transactions on Evolutionary Computation*, 13(2), 243–259, doi:10.1109/TEVC.2008.924428.
- Wada, Y., L. P. H. van Beek, D. Viviroli, H. H. Durr, R. Weingartner, and M. F. P. Bierkens (2011a), Global monthly water stress: 2. Water demand and severity of water stress, *Water Resour. Res.*, 47.
- Wada, Y., L. P. H. Van Beek, D. Viviroli, H. H. Drr, R. Weingartner, and M. F. P. Bierkens (2011b), Global monthly water stress: 2. Water demand and severity of water stress, *Water Resources Research*, 47, 1–17, doi:10.1029/2010WR009792.
- Wada, Y., L. P. H. Van Beek, and M. F. P. Bierkens (2012), Nonsustainable groundwater sustaining irrigation: A global assessment, *Water Resources Research*, 48(November 2011), doi:10.1029/2011WR010562.
- Wagner, W., W. Dorigo, R. de Jeu, D. Fernandez, J. Benveniste, E. Haas, and M. Ertl (2012), Fusion of Active and Passive Microwave Observations To Create an Essential Climate Variable Data Record on Soil Moisture, *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, I-7(September), 315–321, doi:10.5194/isprsannals-I-7-315-2012.

- Wanders, N., D. Karssenber, a. De Roo, S. M. De Jong, and M. F. P. Bierkens (2014), The suitability of remotely sensed soil moisture for improving operational flood forecasting, *Hydrology and Earth System Sciences*, 18, 2343–2357, doi:10.5194/hess-18-2343-2014.
- Ward, M. N., C. M. Brown, K. M. Baroang, and Y. H. Kaheil (2013), Reservoir performance and dynamic management under plausible assumptions of future climate over seasons to decades, *Climatic Change*, 118, 307–320, doi:10.1007/s10584-012-0616-0.
- Watanabe, S., S. Kanae, S. Seto, P. J.-F. Yeh, Y. Hirabayashi, and T. Oki (2012), Intercomparison of bias-correction methods for monthly temperature and precipitation simulated by multiple climate models, *Journal of Geophysical Research*, 117(D23), D23114, doi:10.1029/2012JD018192.
- Wehner, M. (2013), *Methods of Projecting Future Changes in Extremes*, Springer Science C Business Media Dordrecht.
- Wehner, M. F., R. L. Smith, G. Bala, and P. Duffy (2009), The effect of horizontal resolution on simulation of very extreme US precipitation events in a global atmosphere model, *Climate Dynamics*, 34(2-3), 241–247, doi:10.1007/s00382-009-0656-y.
- Werth, S., and A. Güntner (2009), Calibration analysis for water storage variability of the global hydrological model WGHM, *Hydrology and Earth System Sciences Discussions*, 6, 4813–4861, doi:10.5194/hessd-6-4813-2009.
- Weverberg, K. Van, A. M. Vogelmann, W. Lin, E. P. Luke, A. Cialella, P. Minnis, M. Khaiyer, E. R. Boer, and M. P. Jensen (2013), The Role of Cloud Microphysics Parameterization in the Simulation of Mesoscale Convective System Clouds and Precipitation in the Tropical Western Pacific,
- Wilcox, E. M., and L. J. Donner (2007), The Frequency of Extreme Rain Events in Satellite Rain-Rate Estimates and an Atmospheric General Circulation Model, *Journal of Climate*, 20(1), 53–69, doi:10.1175/JCLI3987.1.
- Wilks, D. S. (2006), *Statistical Methods in the Atmospheric Sciences*, Academic Press, Burlington, MA (2nd edition).
- Wood, A. W., D. P. Lettenmaier, and R. N. Palmer (1997a), Assessing climate change implications for water resources planning, *Climatic Change*, 37(1), 203–228, doi:10.1023/A:1005380706253.
- Wood, E. F., D. Lettenmaier, X. Liang, B. Nijssen, and S. W. Wetzel (1997b), Hydrological Modeling of Continental-Scale Basins, *Annual Review of Earth and Planetary Sciences*, 25, 279–300, doi:10.1146/annurev.earth.25.1.279.

- Xu, Z., and Z.-L. Yang (2012), An Improved Dynamical Downscaling Method with GCM Bias Corrections and Its Validation with 30 Years of Climate Simulations, *Journal of Climate*, 25(18), 6271–6286, doi:10.1175/JCLI-D-12-00005.1.
- Yapo, P. O., H. V. Gupta, and S. Sorooshian (1998), Multi-objective global optimization for hydrologic models, *Journal of Hydrology*, 204(1-4), 83–97, doi:10.1016/S0022-1694(97)00107-8.
- Yue, S., T. B. M. . Ouarda, B. Bobée, P. Legendre, and P. Bruneau (1999), The Gumbel mixed model for flood frequency analysis, *Journal of Hydrology*, 226(1), 88–100, doi:10.1016/S0022-1694(99)00168-7.
- Ziervogel, G., S. Bharwani, and T. E. Downing (2006), Adapting to climate variability: Pumpkins, people and policy, *Natural Resources Forum*, 30(4), 294–305, doi:10.1111/j.1477-8947.2006.00121.x.