Spring 1-13-2015

Impacts of Climate Change on Hydrologic Processes in the Colorado River Basin

Pablo A. Mendoza
University of Colorado at Boulder, pablocorgan@gmail.com

IMPACTS OF CLIMATE CHANGE ON HYDROLOGIC PROCESSES IN THE COLORADO RIVER BASIN

by

PABLO A. MENDOZA

B.S., Universidad de Chile, 2006

P.E., Universidad de Chile, 2010

M.S., Universidad de Chile, 2010

A thesis submitted to the faculty of the Graduate School of the University of Colorado in partial fulfillment of the requirement for the degree of

Doctor of Philosophy

Department of Civil, Environmental, and Architectural Engineering

2015

This thesis entitled:

Impacts of Climate Change on Hydrologic Processes in the Colorado River Basin

written by Pablo A. Mendoza

has been approved for the

Department of Civil, Environmental, and Architectural Engineering

Balaji Rajagopalan

Martyn P. Clark

Jeffrey R. Arnold

Levi Brekke

Joseph Kasprzyk

Edith Zagona

Date_____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Pablo A. Mendoza (Ph.D., Civil, Environmental, and Architectural Engineering)

Impacts of Climate Change on Hydrologic Processes in the Colorado River Basin

Thesis directed by Professor Balaji Rajagopalan

The clear evidence of human-induced hydroclimatic shifts in the western U.S., and especially within the Colorado River Basin (CRB), has motivated the research community to pursue reliable estimates of future climate change impacts on hydrological processes. Although the large body of literature in this field shows consensus on the future drying of the CRB, there are large uncertainties in the magnitude of hydrologic changes because of differences in experimental approaches. Hence, this work focuses on the effects of methodological choices, in particular those related to hydrologic modeling, on the portrayal of climate change impacts at three catchments located in the headwaters of the CRB.

A commentary on the current development of complex process-based hydrologic models is first presented. It is argued that the relatively poor performance of such models may occur due to restrictions on their ability to refine their portrayal of physical processes, and improving hydrological models requires integrating the strengths of prior knowledge of hydrologic processes with the strengths of data driven inference. An assessment of the effects of hydrologic model choice and parameter calibration on projected hydrologic changes follows. Here, it is demonstrated that the subjective selection of model structures may introduce large uncertainties to hydrologic projections. Based on this, the third part of this study compares the effects of hydrologic model choice and parameter estimation strategies on projected climate change impacts. The main finding here is that the choice of parameter estimation methods can provide similar or larger uncertainties in some hydrologic processes when compared to uncertainties

coming from model choice. The fourth component of this study evaluates the effects of regional climate model (RCM) configuration and forcing scaling on projected hydrologic changes. The results illustrate the implications of RCM configuration on projected changes of the water cycle, and provide an integrated view of the interplay between forcings and hydrologic model structures in the portrayal of climate change impacts. Finally, a commentary on the main results of this study and possible ways to move forward is provided. It is asserted that a greater role of expert knowledge is required to improve model selection and parameter estimation, and also for incorporating non-stationarity in climate change impact studies.

DEDICATION

To my parents, my sister, and my grandmother Lidia.

ACKNOWLEDGMENTS

I am grateful and deeply indebted to my mentors, Martyn Clark and Balaji Rajagopalan, for their infinite support and for believing in me. Without their creativity, constant encouragement, guidance, genuine care and patience, the completion of this amazing journey would have been impossible. Thank you for giving me the unique opportunity to do research with you. I also thank my committee members, Jeffrey Arnold, Levi Brekke, Joseph Kasprzyk, and Edith Zagona, for their interest on my work and for their insightful and helpful comments.

I would like to acknowledge the U.S. Army Corps of Engineers, the U.S. Bureau of Reclamation and the Cooperative Institute of Research and Environmental Sciences (CIRES) for funding this study. I also wish to thank everyone at the Research Applications Laboratory (RAL) of the National Center for Atmospheric Research. In particular, I am extremely grateful for Naoki Mizukami, Mark Raleigh, Andy Newman, Ethan Gutmann, Kyoko Ikeda, Michael Barlage, Roy Rasmussen, and Andy Wood for their feedback and advice during this process.

I acknowledge my former advisors, Ximena Vargas and James McPhee, for introducing me to the fascinating world of hydrological sciences, and for encouraging me to do research in this field. I also thank my past and current colleagues from Universidad de Chile, CU Boulder – especially from the water program at CEAE and the Department of Geography – and from many other groups for their advice, support and especially their friendship. There are far too many names to list here.

Finally, I want to thank all my family and closest friends. I am especially grateful for my parents, Myriam and Gustavo, for their unconditional love and dedication, for all the sacrifices they made throughout the years to ensure my access to a good education, and for teaching me

values to become a good man. I also thank my sister, Evelyn, for inspiring me to dream big. Finally, I acknowledge my grandmother Lidia for her everlasting love, devotion and extremely sharp sense of humor. Wherever you are, this body of work is for you.

CONTENTS

TABLES

FIGURES

CHAPTER 1: Introduction

1.1 Background

Given the strong evidence of ongoing intensification of the water cycle from regional to continental scales (e.g. Huntington 2006), most key questions facing humanity in regards of climate projections for the 21st century are water-related (Lall 2014). According to the Intergovernmental Panel on Climate Change (IPCC 2013), the nature, extent and timing of climate change impacts on water resources are expected to vary regionally, depending on the mitigation and adaptation capabilities of different social and environmental systems.

During the last decade, several studies conducted over the western United States have detected consistent trends in hydroclimatic variables such as temperature, precipitation, snow water equivalent, soil moisture and runoff (e.g. Barnett et al. 2005; Regonda et al. 2005; Mote et al. 2005; Stewart et al. 2004, 2005; Knowles et al. 2006; Hamlet et al. 2007; Cayan et al. 2010). Even more, it has been found that a substantial portion of these trends is human-induced (Barnett et al. 2008; Pierce et al. 2008; Das et al. 2009; Hidalgo et al. 2009), increasing concern about the consequences of anthropogenic warming on water supply over this region, which relies heavily on snow accumulation during winter. Within this large domain, the situation of the Colorado River basin (CRB) – a major basin that spans seven states in the U.S. and two states in Mexico – is critical, as the system is already over-allocated (Bureau of Reclamation 2012) and future warming might increase the probability of supply and demand imbalances (McCabe and Wolock 2007).

The vulnerability of the CRB reservoir system storage to changes in runoff has raised the interest of the research community to better understand the implications of precipitation and temperature changes on hydrologic processes, which are augmented by the semi-arid nature of

the basin (Loaiciga et al. 1996). A growing body of literature on this topic (recently reviewed by Vano et al. 2014) shows an overall consensus about future drying of this region. For example, Figure 1.1 displays projected variations in precipitation minus evapotranspiration (P –E) reported by Seager et al. (2012) for the headwaters of the Colorado River basin under a severe emission scenario (Meinshausen et al. 2011), where most simulations project a decrease in P−E for upcoming two-decade periods, and intensified drying as the 21st century advances. Furthermore, runoff changes in the CRB are expected to be induced by the combined effects of (i) decreased snowpack in winter – resulting from the increased rain to snow ratio –, (ii) earlier runoff peaks due to rising temperature in spring, and (iii) precipitation change in summer (Gao et al. 2011). Hence, a better comprehension of hydrologic sensitivities to climate change in the CRB is critical to develop adaptation strategies for an already stressed water supply system.

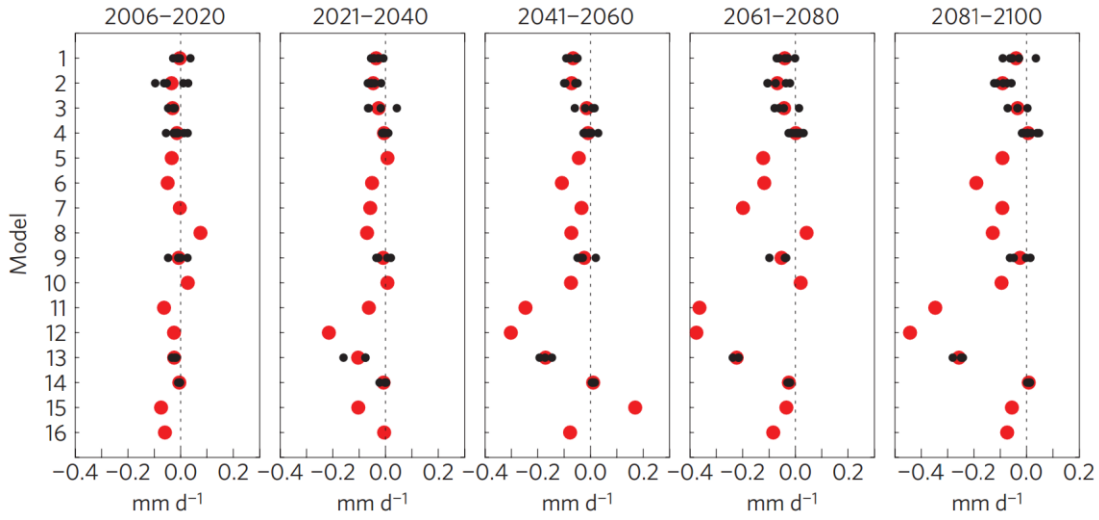

Figure 1.1: The difference in P−E relative to 1951–2000 for two-decade periods of the present century for the Colorado River headwaters region (37°–42° N, 112°–106° W). The vertical axis refers to the individual global climate models. Black dots show the change in P−E for the five two-decade periods for the individual continuous runs with the model. The red dots show the ensemble mean change for each model. For models with single continuous runs only the red dot is plotted. The twenty-first-century projections use the Representative Concentration Pathway (RCP) 8.5 emission scenario (from Seager et al. 2012).

1.2 Motivation

In spite of the general agreement regarding future drying of the CRB, there are considerable discrepancies among studies in terms of the magnitude of projected hydrologic changes. For instance, while Hoerling and Eischeid (2007) reported a 45 % decline in runoff at Lees Ferry by mid-century, Christensen and Lettenmaier (2007) projected a decrease of 6 % for the same period. Likewise, many more examples showing divergent results and conclusions can be found in the literature (e.g. Milly et al. 2005; Christensen et al. 2004; Ray et al. 2008; Hoerling et al. 2009; Rasmussen et al. 2011, 2014; Harding et al. 2012; Vano et al. 2012; Ficklin et al. 2013; Vano and Lettenmaier 2014), suggesting that the large uncertainty in hydrologic projections arise from differences in methodological choices (Vano et al. 2014).

To deal with the above problem, many authors have relied on the 'cascade of uncertainty' paradigm (e.g. Wilby and Dessai 2010), which attempts to quantify the uncertainty at each step of the modeling process, including (Figure 1.2): (i) choice of greenhouse gas emission scenarios, (ii) choice of climate model(s), (iii) choice of climate model initial conditions, (iv) choice of meteorological forcing downscaling methods, (v) choice of hydrological model structures, and (vi) choice of hydrological model parameter sets. Although this approach has demonstrated practical utility for quantifying the overall uncertainty and its main contributors (e.g. Wilby and Harris 2006; Chen et al. 2011), it typically relies on statistical methods to estimate climate change impacts at local scales, giving much less attention to uncertainties derived from hydrological modeling (Bastola et al. 2011). This is problematic because the choice of a hydrologic model structure (Vano et al. 2012) and parameter values (Wilby 2005) –commonly based on pragmatic considerations – have been proven to be very important for hydrologic change assessment. More worrisome, despite outstanding advances in the development of

process-based hydrologic models (e.g. Abbott et al. 1986; Wigmosta et al. 1994; VanderKwaak and Loague 2001; Ivanov et al. 2004; Maxwell and Miller 2005; Rigon et al. 2006; Qu and Duffy 2007; Lawrence et al. 2011; Niu et al. 2011) that have been later used in the assessment of climate change impacts, there is still limited understanding of their shortcomings and relative importance compared to parameter identification strategies, especially in terms of individual hydrologic processes (e.g. evapotranspiration, generation of surface flow and baseflow, snowpack).

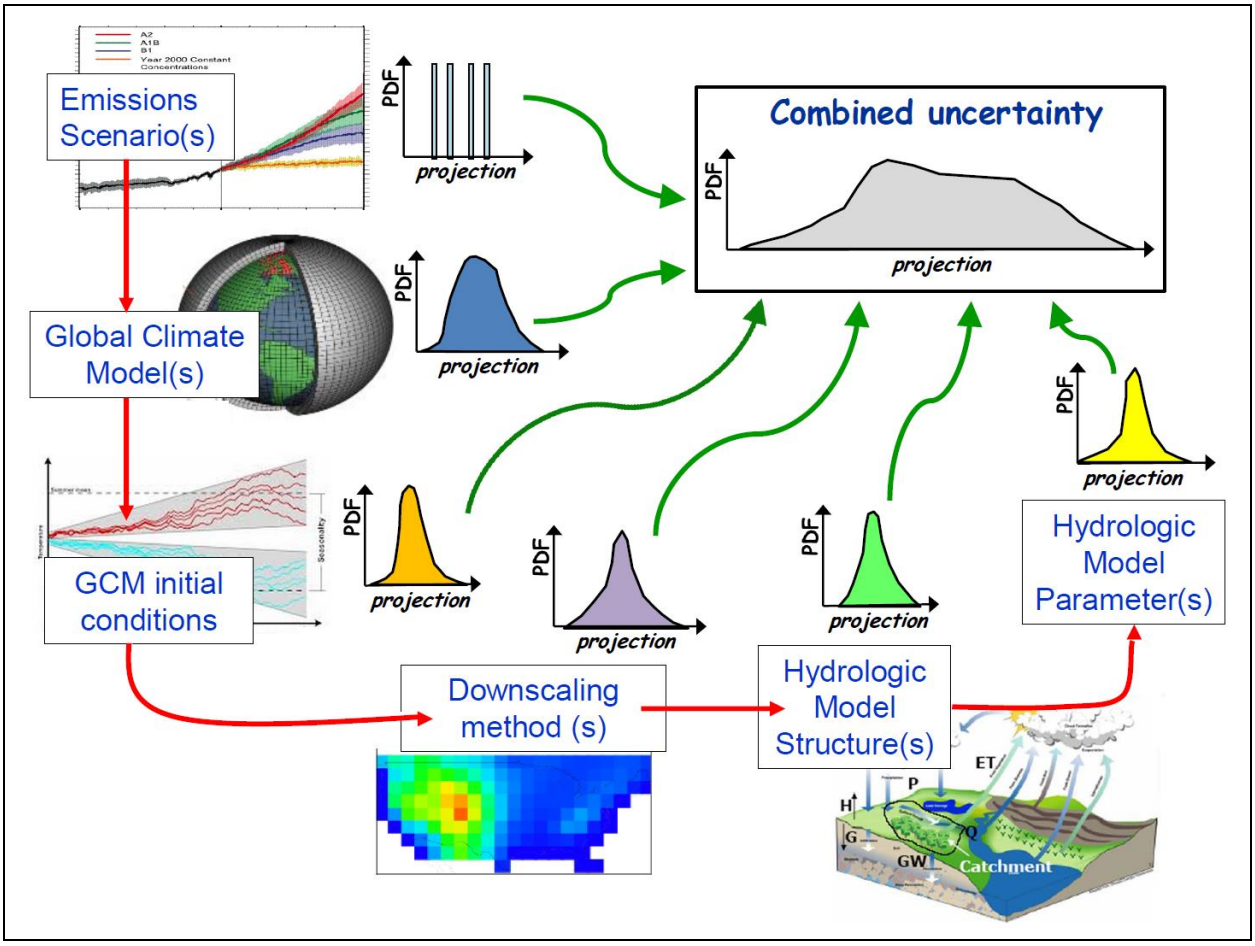

Figure 1.2: Representation of the "cascade of uncertainty" paradigm. The diagram illustrates the key methodological choices for the assessment of change impacts, and how their uncertainties combine to generate a final projected envelope.

Therefore, a better understanding of uncertainties in hydrologic modeling decisions is critical to advance model predictive capabilities under a changing climate. The ultimate goal of this doctoral research is to provide a suite of best practices for the assessment of climate change impacts on hydrologic processes, which will be the result from addressing the following questions:

1. In spite of their complexity and physical realism, distributed process-based models perform similarly to, or only slightly better than, traditional bucket-style rainfall-runoff models (Reed et al. 2004; Smith et al. 2012). What are the main a priori constraints of complex process-based models that currently contribute to their relatively poor performance?
2. What are the effects of model structure selection and parameter calibration on the portrayal of climate change impacts?
3. How does the subjective choice of hydrologic models and parameter identification strategies affect projected changes in water balance and hydrological processes, and how do these effects compare?
4. How do the configuration of regional climate models (RCM) and forcing scaling affect model fidelity and projected hydrologic changes?
5. Moving forward, how can we improve our understanding of hydrologic sensitivity to climate variability and change?

1.3 Outline

Chapter 2 provides a detailed description of the material and methods used in this research. Specifically, a characterization of the case study basins selected, meteorological forcing datasets, hydrologic model structures and signature measures of hydrologic behavior is provided.

Chapter 3 addresses Question 1 in the form of commentary, where it is suggested that hydrologists and land-surface modelers may be unnecessarily constraining the behavioral agility of very complex physics-based models. It is argued that the relatively poor performance of such models can occur due to restrictions on their ability to refine their portrayal of physical processes, in part because of strong a-priori constraints in: (i) the representation of spatial variability and hydrologic connectivity, (ii) the choice of model parameterizations, and (iii) the choice of model parameter values. A specific example of problems associated with strong a-priori constraints on parameters in a land surface model is provided. Finally, some directions for accelerating progress in the development of hydrologic modeling frameworks are proposed and discussed.

Chapter 4 examines the role of hydrologic model selection and parameter calibration on the assessment of climate change impacts to address Question 2. To this end, model performance and projected hydrologic changes obtained through four different hydrologic model structures are compared before (i.e., using default parameters) and after calibration with the Shuffled Complex Evolution algorithm (SCE-UA; Duan et al. 1992, 1993). Hydrologic changes are examined via a climate change scenario where the NCAR Community Climate System Model Version 3 (CCSM3 ; Collins et al. 2006) change signal is used to perturb the boundary conditions of the Weather Research and Forecasting (WRF ; Skamarock et al. 2008) regional climate model configured at 4-km resolution. Inter-model differences in the portrayal of climate change impacts are discussed in terms of annual water balance, monthly changes in specific variables and signature measures of hydrologic behavior.

Chapter 5 addresses Question 3 by examining the relative importance of hydrologic model choice and parameter estimation strategies on projected changes in annual water balance,

monthly simulated processes and signature measures of hydrologic behavior. To this end, hydrologic changes obtained with the model structures described in Chapter 2, whose parameters were obtained using a common calibration strategy, are compared with those coming from parameter sets identified using multiple options for different calibration decisions (objective function, multiple local optima and calibration forcing dataset), using a unique model structure. The results presented in this study prompt the need to incorporate uncertainty in model structure and model parameters to avoid an over-confident portrayal of climate change impacts.

Chapter 6 examines and compares the effects of regional climate model (RCM) configuration and output re-scaling on the portrayal of climate change impacts (Question 4). Specifically, it is assessed how the above decisions affect: (i) historical performance in terms of hydrologic signature measures, and (ii) hydrologic changes due to a climate perturbation, with focus on the annual water balance and catchment processes. Meteorological forcings for current and a modified climate scenario are obtained at three spatial resolutions (4-, 12- and 36-km) from dynamical downscaling with the WRF regional climate model, and hydrologic changes are computed using four hydrologic model structures and two sets of hydrologic model parameters (the same included in Chapter 4). This study provides an integrated portrait of the effects that meteorological forcing datasets, model structures and model parameters have on hydrologic change projections.

Finally, Chapter 7 addresses Question 5 by providing a brief overview of the existing paradigms used in the evaluation of climate change impacts, a summary of the main findings of previous chapters, and some recommendations for moving towards a process-based assessment of the sensitivity of hydrologic processes to climate variability and change.

CHAPTER 2: Material and methods

2.1 Study area

The importance of the CRB for water management and decision making, together with strong evidence of a hydroclimatic shift over the past decades (Miller and Piechota 2008, 2011), has motivated several studies to generate streamflow projections under different future climate scenarios over the area (e.g. Milly et al. 2005; Christensen and Lettenmaier 2007; Hoerling et al. 2009; Bureau of Reclamation 2012). Much of the water for this region comes from the high-elevation area – the Colorado Headwaters – that works as a natural reservoir during the winter, storing precipitation as snowpack and glaciers. Given the strategic relevance of this sub-domain, we select three basins in the Colorado Headwaters Region – Yampa River at Steamboat Springs, East River at Almont and Animas River at Durango – whose location and elevation ranges are shown in Figure 2.1. These catchments have been included in many past climate change studies (e.g. Wilby et al. 1999; Sankarasubramanian and Vogel 2002; Mastin et al. 2011; Milly and Dunne 2011) and, because of their relatively small size compared to the CRB, they offer a unique opportunity to perform extensive analysis involving thousands of model runs (e.g. sensitivity analysis and hydrologic model calibration), to evaluate different approaches in climate change impact assessment, and also to provide detailed understanding of physical processes in the headwaters of the CRB.

Table 2.1 summarizes the main hydroclimatic characteristics of the three basins for which historical data are available, over an 8-year period (Oct/2000 - Sep/2008). Mean basin precipitation ranges between 700 mm/year to 900 mm/year, while mean basin elevation is above 2500 m.a.s.l. Among these basins, the East River at Almont has the largest runoff ratio (0.42), and the Yampa at Steamboat Springs has the lowest runoff ratio (0.32, with the lowest runoff and

precipitation amounts). The land surface of the Yampa and Animas River basins is predominantly covered by deciduous forests (26 % at Yampa and 23 % at Animas) and evergreen forests (37 % at Yampa and 39 % at Animas), while the land surface of the East River basin is mainly covered by evergreen forests (29 %) and grassland/herbaceous (26 %).

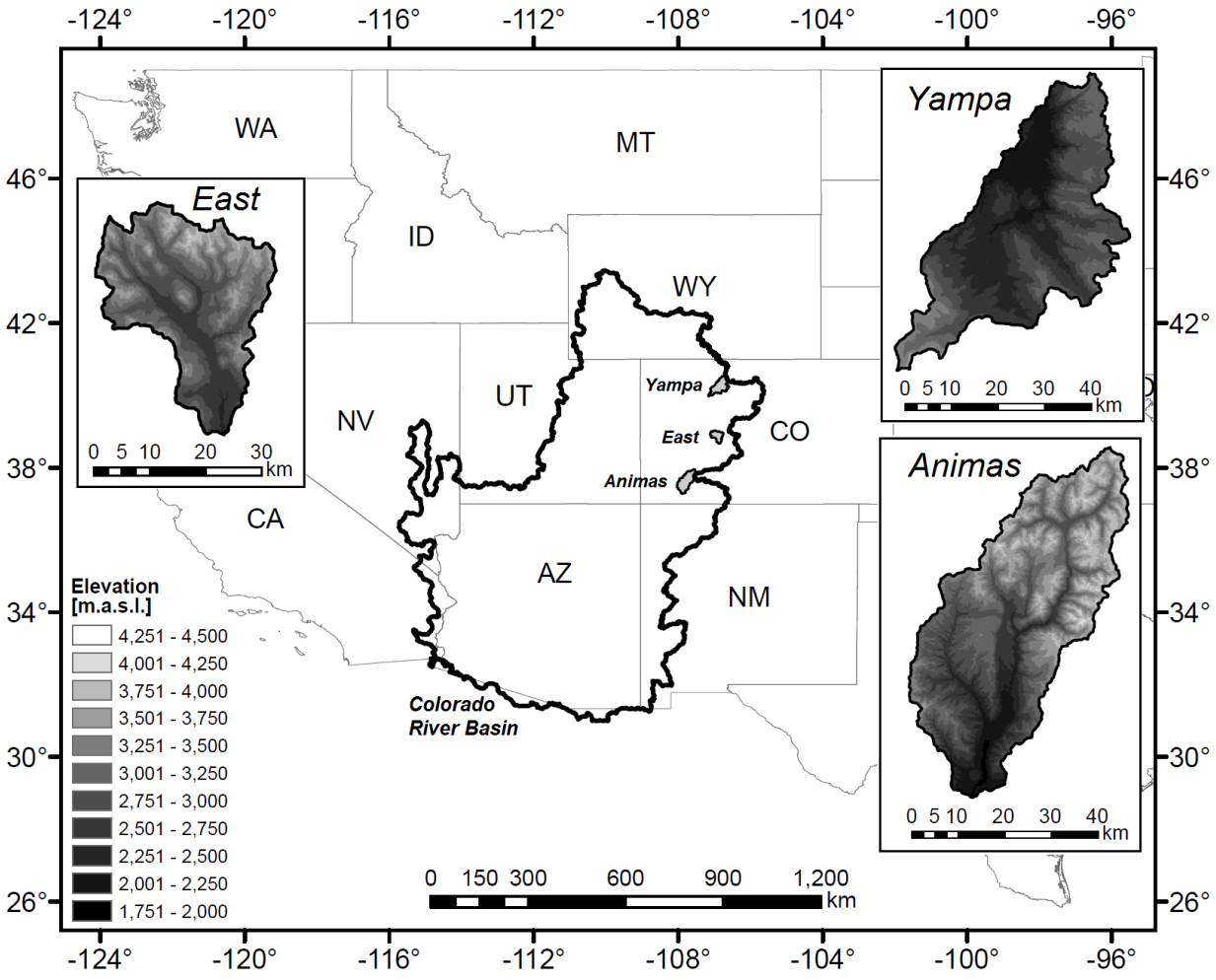

Figure 2.1: Location of the basins of interest.

Table 2.1: Three study watersheds' characteristics. Hydrologic variables correspond to the period Oct/2000-Sep/2008. P, R, PE, RR and DI denote basin-averaged mean annual values of precipitation, runoff, potential evapotranspiration, runoff ratio and dryness index, respectively.

Location	Area (km^2)	Mean basin elevation (m.a.s.l.)	Mean annual runoff (mm/yr)	Mean Precipitation from WRF (mm/yr)	Mean annual PE* (mm/yr)	Mean annual RE (R/P)	Mean annual DI (PE/P)
Yampa at Steamboat Springs	1468	2674	228	717	953	0.32	1.33
East at Almont	748	3127	327	782	757	0.42	0.97
Animas at Durango	1819	3098	365	883	885	0.41	1.00

*PE obtained from PRMS by using a Jensen-Haise formulation (Jensen et al. 1969)

2.2 Meteorological forcings

We use dynamically downscaled climate datasets obtained with the Weather Research and Forecasting (WRF) regional climate model (Skamarock et al. 2008) to force hydrologic simulations and compute hydrologic changes due to a climate perturbation. These datasets consist of historical (control run, CTRL) and pseudo global warming (PGW) outputs at three different horizontal resolutions (4-, 12- and 36-km). Specifications of these WRF simulations are fully described in Rasmussen et al. (2014), but briefly reviewed below. The initial and 3-hourly lateral boundary conditions were taken from the North American Regional Reanalysis (NARR; Mesinger et al. 2006) coarse resolution dataset (~32 km). The model physics options used in that study included the Noah Land Surface Model (Noah-LSM) version 3.2 with upgraded snow physics (Chen and Dudhia 2001; Barlage et al. 2010), the Thompson mixed-phase cloud microphysics scheme (Thompson et al. 2008), the Yonsei University planetary boundary layer (Hong et al. 2006) and the Community Atmosphere Model's (CAM) longwave and shortwave radiation schemes (Collins et al. 2006). Because the use of an horizontal grid spacing of 6 km or less is able to accurately estimate vertical motions driven by topography (Ikeda et al. 2010; Rasmussen et al. 2011), a convective parameterization was included for the 12- and 36-km

simulations – using the Betts-Miller-Janjić scheme (Janjić 1994)– but not for the 4-km simulation (Rasmussen et al. 2014).

The PGW approach (Schär et al. 1996; Hara et al. 2008; Kawase et al. 2009; Rasmussen et al. 2011) consists of adding a mean climate perturbation to the initial and 3-hourly boundary conditions, here taken from NARR. The climate perturbation used was based on expected changes from the NCAR CCSM3 model forced by the A1B scenario. This perturbation is generated by subtracting the current 10-yr (1995-2005) monthly climatology from a future 10-yr (2045-2055) monthly climatology.

Meteorological data from WRF simulations is available for all horizontal resolutions at hourly time steps, for both historical and modified climatic conditions. The output variables and temporal disaggregation used depend on specific hydrologic model requirements (Table 2.2).

2.3 Hydrologic models

To explore the effects of the choice of hydrologic model structure, we choose four hydrologic/land surface models: the US Geological Survey's Precipitation Runoff Modeling System (PRMS; Leavesley et al. 1983; Leavesley and Stannard 1995), the Variable Infiltration Capacity model (VIC; Wood et al. 1992; Liang et al. 1994, 1996) the Noah Land Surface Model (Noah-LSM; Ek 2003; Mitchell et al. 2004) and the Noah Land Surface Model with Multiple Parameterizations (Noah-MP; Niu et al. 2011; Yang et al. 2011). Our choice is based on the fact that the four models cover different degrees of complexity in terms of conceptualization of vegetation, soil and seasonal snowpack (see Figure 2.2 and Table 2.3 for further details), and also have different parameterizations for some hydrologic processes (e.g. different model equations for canopy storage, baseflow, etc.). Additionally, these hydrologic model structures have been used in several research studies (e.g. Wilby et al. 1999; Haddeland et al. 2002; Hay et

al. 2002; Hay and Clark 2003; Christensen and Lettenmaier 2007; Barlage et al. 2010; Yang et al. 2011; Cai et al. 2014). Our experimental design considers a hydrologic model spatial resolution identical to that used in the 4-km WRF simulations performed by Rasmussen et al. (2014), though simulation time steps, forcing variables and land cover data used for a priori parameter estimates vary depending on specific model requirements (see Table 2.2 for further details).

Table 2.2: Summary of data sources and simulation setup used in this study

Model	Vegetation data	Soil data	Forcing variables*	Spatial/temporal discretization
PRMS	USGS 1-km gridded vegetation type and density data (USDA 1992)	State soils geographic (STATSGO) 1-km gridded soils data (USDA 1994)	Daily precipitation; maximum and minimum daily temperature.	4 km and $\Delta t = 24$ h
VIC	UMD 1-km Global Land Cover Classification (Hansen et al. 2000)	State soils geographic (STATSGO) 1-km gridded soils data (USDA 1994)	Precipitation, temperature, shortwave and longwave radiation, wind speed, relative humidity and air pressure.	4 km and $\Delta t = 1$ h
Noah-LSM and Noah-MP	National Land Cover Data Base, 2006 (Fry et al. 2011).	State soils geographic (STATSGO) 1-km gridded soils data (USDA 1994)	Precipitation, temperature, shortwave and longwave radiation, wind speed, relative humidity and air pressure.	4 km and $\Delta t = 1$ h

*Air temperature at 2 m and wind speed at 10 m are used for hydrologic simulations.

In this study we use a single suite of physics options for Noah-MP, including a Ball-Berry type model for canopy stomatal resistance, the Community Land Model (CLM; Oleson et al. 2010) soil stress function to control stomatal resistance, the SIMTOP model for runoff and groundwater (Niu et al. 2005), a Monin-Obukhov similarity theory-based drag coefficient, supercooled liquid water and frozen soil permeability based on Niu and Yang (2006), a two-stream radiation transfer scheme applied only to the vegetated fraction, a snow surface albedo parameterization based on the Canadian Land Surface Scheme (CLASS; Verseghy 1991),

partitioning of precipitation into snowfall and rainfall based on Jordan (1991) and a Noah-type lower boundary of soil temperature. Readers are referred to Niu et al. (2011) for a full description of each model component.

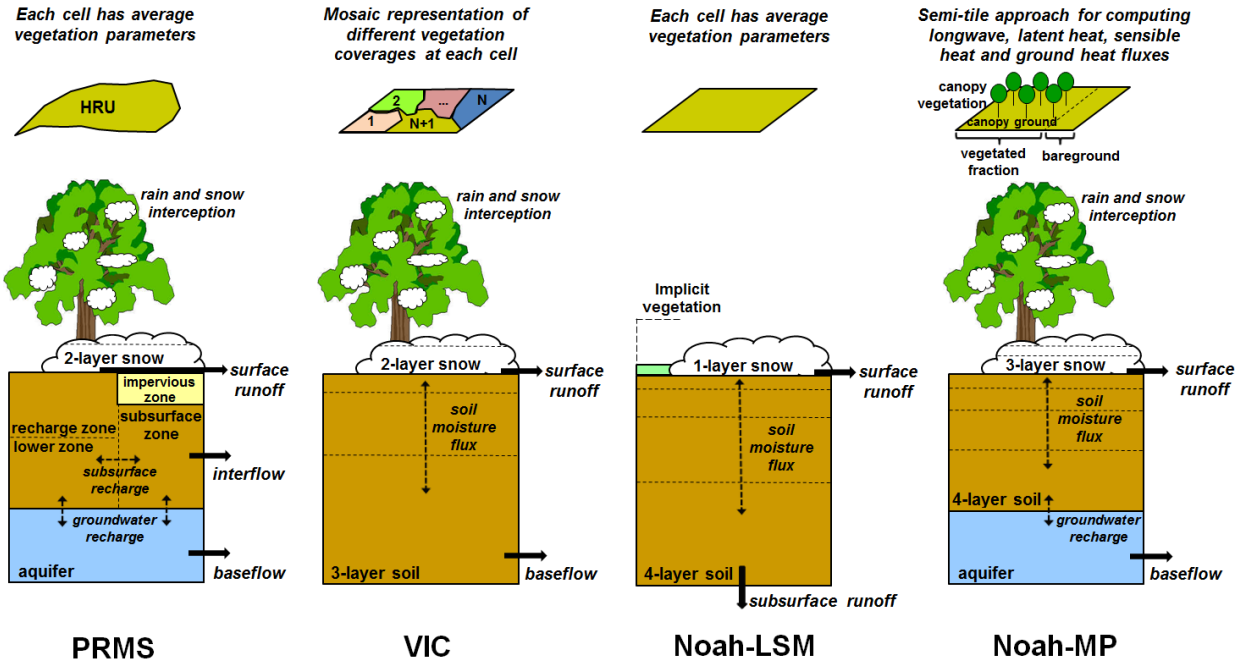

Figure 2.2: Comparison of model architectures used in this study: Precipitation Runoff Modeling System (PRMS), Variable Infiltration Capacity (VIC), Noah Land Surface Model (Noah-LSM) and Noah Land Surface Model with Multiple Parameterization Options (Noah-MP).

Table 2.3: Overview of hydrologic model components used in this study

Model	Snow accumulation and melt	Canopy storage	Moisture in the soil column/surface runoff	Baseflow
PRMS	2-layer energy/mass balance model. Snowpack energy balance is computed every 12 hours.	Precipitation can be intercepted by and evaporated from the plant canopy. Precipitation that is not intercepted by the canopy layer (throughfall) is distributed to the watershed land surface. Interception of precipitation by the plant canopy is computed during a time step as a function of plant-cover density and the storage available on the predominant plant-cover type in each HRU.	Surface runoff and infiltration are computed using a non-linear variable-source-area method allowing for cascade flow.	The groundwater zone is conceptualized as a linear reservoir (ie. baseflow is computed as a linear function of groundwater storage).
VIC	2-layer energy/mass balance model.	Water enters 1-layer canopy reservoir, and can leave as canopy evaporation, transpiration or throughfall. Canopy throughfall occurs when additional precipitation exceeds the storage capacity of the canopy. Different vegetation classes are allowed within a unique grid cell via a 'mosaic' approach, where energy and water balance terms are computed independently for each coverage class (vegetation and bare soil).	An infiltration capacity function is defined. Vertical movement of moisture through soil follows 1-D Richards equation.	Defined as a function of the soil moisture in the third layer (Arno formulation). The function is linear below a soil moisture threshold, and becomes nonlinear above that threshold.
Noah-LSM	1-layer energy/mass balance model that simulates snow accumulation, sublimation, melting and heat exchange at snow-atmosphere and snow-soil interfaces.	One canopy layer, simple canopy resistance. Simple Jarvis type of canopy resistance function, single linearized energy balance equation representing combined ground/vegetation surface, considering seasonal LAI and green vegetation fraction.	Surface runoff is computed as the difference between throughfall and a maximum infiltration rate. Vertical movement of moisture through soil layers follows 1-D Richards equation.	Computed as the product of a scaling factor between 0 and 1 and the hydraulic conductivity of the bottom layer.
Noah-MP	3-layer energy/mass balance model that represents percolation, retention and refreezing of meltwater within the snowpack.	Snow interception includes loading/unloading, melt/refreeze capabilities, and sublimation of canopy-intercepted snow, along with detailed representation of transmission and attenuation of radiation through the canopy, within- and below-canopy turbulence, and different options to represent the biophysical controls on transpiration.	Surface runoff is an exponential function of depth to water table. Vertical movement of moisture through soil layers follows 1-D Richards equation.	Baseflow is parameterized as an exponential decaying function of the water table level (SIMTOP).

2.4 Signature measures of hydrologic behavior

We use six hydrologic signature measures (Yilmaz et al. 2008; Stewart et al. 2005) to quantify model fidelity and projected changes in catchment behavior. These metrics are intended to represent different hydrologic processes from overall precipitation partitioning into ET and runoff to vertical redistribution of excess precipitation via percolation. All the signatures are derived based on daily runoff time series. The notation, short description, mathematical formulation and physical process associated with each signature measure are detailed in Table 2.4. Similar diagnostic evaluation metrics have been used in past studies with multiple purposes, such as model evaluation (e.g. Herbst et al. 2009; Majone et al. 2012; Pfannerstill et al. 2014), catchment classification (e.g. Oudin et al. 2010; Ley et al. 2011; Sawicz et al. 2011; Carrillo et al. 2011), sensitivity analysis (e.g. van Werkhoven et al. 2008; Wagener et al. 2009b), hydrologic model structure identification (e.g. Hartmann et al. 2013; Hrachowitz et al. 2014), analysis of spatial distribution of hydrologic processes (e.g. McMillan et al. 2014) and the choice of realistic model parameter values in terms of process representations (e.g. Pokhrel and Gupta 2009; van Werkhoven et al. 2009; Kollat et al. 2012; Pokhrel et al. 2012).

Table 2.4: Signature measures used to evaluate projected changes in catchment behavior

Notation	Short description	Equation	Hydrologic process
RR	Runoff Ratio	$RR = R / P$	Overall water balance (ET processes).
FMS	FDC Mid-segment Slope	$FMS = \frac{\log(Q_{m1}) - \log(Q_{m2})}{m_1 - m_2}$	Variability, or flashiness, of the flow magnitudes.
FHV	FDC High-segment Volume	$FHV = \sum_{h=1}^{H} Q_h$	Measure of the catchment response to high rainfall/snowmelt events.
FLV	FDC Low-segment Volume	$FLV = \sum_{l=1}^{L} [\log(Q_l) - \log(Q_L)]$	Measure of the long-term baseflow processes
FMM	FDC Median	$FMM = median(\log(\text{FDC}))$	Measure of mid-range flows.
CTR	Center Time of Runoff	$CTR = \frac{\sum_{i=1}^{N} t_i Q_i}{\sum_{i=1}^{N} Q_i}$	Seasonality of runoff.

R: basin-averaged mean annual runoff.
P: basin-averaged mean annual precipitation.

Q_{m1}: flow with exceedance probability of $m_1 = 0.2$.
Q_{m2}: flow with exceedance probability of $m_2 = 0.7$.
$h = 1, 2,...,H$ are the flow indices into the array of flows with exceedance probabilities lower than 0.02.
$l = 1,2,...,L$ is the index into the array of flow values located within the low-flow segment (0.7-1.0 exceedance probabilities), being L the index for minimum flow.
N: total number of days in a water year.

CHAPTER 3: Are we unnecessarily constraining the agility of complex process-based models?

3.1 Introduction

The hydrologic community has made substantial investments in the development of complex physics-based models that provide detailed representations of the dominant physical processes and their interactions (e.g. Abbott et al., 1986; Wigmosta et al., 1994; VanderKwaak and Loague, 2001; Ivanov et al., 2004; Maxwell and Miller, 2005; Rigon et al., 2006; Qu and Duffy, 2007; Lawrence et al., 2011; Niu et al., 2011). In spite of their complexity and physical realism, distributed process-based models perform similarly to, or only slightly better than, traditional bucket-style rainfall-runoff models (e.g., Reed et al., 2004; Smith et al., 2012). In this commentary we discuss some issues that can result in relatively low performance of complex models, illustrate some of these shortcomings through an example application, and make practical recommendations that should lead to improved physics-based model simulations.

3.2 On the need for model agility

Over the last four decades, a number of important issues related to process representation and model performance have been widely discussed (Freeze and Harlan, 1969; Bergstrom, 1991; Blöschl and Sivapalan, 1995; Beven, 2000, 2002, 2006; Sivapalan et al., 2003; Kirchner, 2006; Clark et al., 2008, 2011; Gupta et al., 2008, 2012; Wagener et al., 2009a; Beven and Cloke, 2012). A fundamental challenge is developing models that represent how the spatial variability in hydro-meteorological fields, topography, vegetation and soils combines to produce fluxes of energy and water at catchment, regional and global scales (e.g., Reggiani et al., 1999; Beven, 2002; Kollet et al., 2010). Meeting this challenge requires extensive evaluation and refinement of

model representations of hydrological processes (Beven 2002; Clark et al. 2011a), in particular those related to representing fluxes of water and energy at the spatial scale of the model discretization (Samaniego et al. 2010).

However, a factor that complicates the problem of evaluating and refining the behavior of process-based models is that many of them have fixed representations of spatial variability (e.g., a single spatial resolution and configuration, parameter look-up tables with limited number of soil and vegetation classes), fixed representations of model physics (e.g., a single set of process representations), and fixed (hard-coded) model parameter values. Such strong a-priori constraints arguably reflect over-confidence in the spatio-temporal representation of physics-based equations describing complex systems, which are heterogeneous across different spatial scales and often poorly characterized by direct measurement (Kirchner 2006), resulting in models with insufficient ability to adequately simulate the heterogeneity of biophysical and hydrological processes.

In view of such problems, we believe that hydrologic and land surface modeling systems should be *agile* (i.e., have the capability to adjust model equations and parameters to faithfully represent observed processes), in order to enable testing multiple hypotheses of hydrologic behavior (Clark et al. 2011a). Specifically, modeling frameworks should be agile enough to support at least the following key aspects: (i) the capability to modify the representation of spatial variability and hydrologic connectivity (e.g., support different spatial resolutions, grid cell versus hydrologic response units, mosaic versus semi-tile approach to represent sub-grid heterogeneity), (ii) the capability to modify model parameterizations for individual processes (e.g., different soil stress functions for evapotranspiration, non-linear reservoir versus multiple parallel reservoirs for baseflow), and (iii) the capability to modify model parameter values.

Furthermore, these features should be extensible to facilitate iterative improvements in the representation of complex systems (i.e., model reconfiguration) as new data that might support new hypotheses becomes available (Son and Sivapalan, 2007; Fenicia et al., 2008; Clark et al., 2011).

The need for model agility is increasingly recognized, and many modeling frameworks are now available that facilitate experimenting with competing modeling alternatives (Clark et al. 2011a). For instance, Pomeroy et al. (2007) developed the Cold Regions Hydrologic Model (CRHM) to experiment with different alternative representations of cold region processes; Clark et al. (2008) developed the Framework for Understanding Structural Errors (FUSE) to test different parameterizations of soil hydrology used in traditional bucket-style rainfall-runoff models; Niu et al. (2011) developed the Noah-MP model with the aim to experiment with several model parameterizations of biophysical and hydrological processes used in land-surface models; Essery et al. (2013) developed the Joint UK Land Environment Simulator (JULES) Investigation Model (JIM) to test different options to simulate snow processes. Nevertheless, these modeling frameworks lack an integrated supporting system for experimenting with different representations of spatial variability, a broad range of physics parameterizations (i.e., they are *all* somewhat limited in scope), and different model parameter values. For example, FUSE includes only simple parameterizations of soil hydrology, and is focused on spatially lumped structures; JIM is restricted to snowpack processes; and Noah-MP is limited to a semi-tile grid structure.

We believe that modeling systems addressing at least the three requirements proposed above (capability to modify spatial variability and hydrologic connectivity, capability to modify model parameterizations of individual processes and capability to modify model parameter values) will

provide a robust framework for the assessment of differences among process representations in existing hydrological models, and to accelerate future model development and improvement.

3.3 An example of unnecessary constraints in a complex process-based model: Treating uncertain model parameters as physical constants

Somewhat paradoxically, many physics-based models set uncertain model parameters to fixed values. For example, transfer functions that link measurable properties of the landscape (e.g., clay and sand contents, percent of organic matter) with model parameters (e.g., soil porosity, saturated soil hydraulic conductivity) typically include fixed coefficients. These coefficients should ideally be described by a sampling distribution, as they are commonly obtained through statistical analysis of data samples taken for a given region and spatial domain. Similarly, 'observable' model parameters (e.g., saturated hydraulic conductivity, soil porosity, vegetation height) are defined as single values for each model element. This is problematic because such parameters are difficult to define precisely given large within-element spatial heterogeneity and errors associated with direct and indirect measurement techniques. More worrisome, many of the functional 'free' parameters (e.g., coefficients in conceptual baseflow and surface runoff parameterizations) are hard-coded as spatially constant values. Setting model parameters to fixed values effectively treats them as physical constants, neglecting the large uncertainty in their estimates and the large impact that they have on model predictions.

In this section we provide an example of the impact of fixed model parameters through analysis of the Noah Land Surface Model with Multiple Parameterization Options (Noah-MP; Niu et al., 2011). We discuss the existence of hard-coded model parameters, and demonstrate how hydrological simulations can be very sensitive to their values.

3.3.1 Hard-coded model parameters

Although Noah-MP has look-up tables to define soil and vegetation parameter values for different soil and land cover types, it incorporates several hard-coded parameters for snow and runoff processes. This is a typical problem in many complex physics based models, such as the Variable Infiltration Capacity model (VIC; Wood et al., 1992; Liang et al., 1994, 1996) and the Community Land Model (CLM; Oleson et al., 2010). Many other examples can be found in the literature, including models that have hard-wired constants based on limited experimental data.

As an example, Figure 3.1 displays a section of code wherein Noah-MP developers have commented that several snow and runoff parameters could be treated as "adjustable"; however, adjusting these parameters requires manual alteration of the appropriate lines of code and subsequent recompiling of the model subroutine before a new parameter trial can be conducted. This severely constrains the ability to conduct extensive sensitivity analysis and/or parameter estimation. Similarly, hard-coded parameters can be found in the CLASS snow albedo parameterization (Figure 3.2), where minimum and maximum snow albedo have been set to 0.55 and 0.84 respectively (dimensionless units), and time decay in snow albedo has been set to 0.01 (units of h^{-1}). One would expect these hard-coded parameters to vary regionally and seasonally, and there is no apparent justification for setting the parameter values to globally fixed constants when, in fact, they are subject to large estimation and scaling uncertainties, and therefore more appropriately described by probability density functions.

```
! runoff parameters used for SIMTOP and SIMGM:
  REAL, PARAMETER :: TIMEAN = 10.5   !gridcell mean topgraphic index (global mean)
  REAL, PARAMETER :: FSATMX = 0.38   !maximum surface saturated fraction (global mean)

! adjustable parameters for snow processes

  REAL, PARAMETER :: MEXP     = 1.0 ! 2.50   !melting factor (-)
  REAL, PARAMETER :: Z0SNO    = 0.002  !snow surface roughness length (m) (0.002)
  REAL, PARAMETER :: SNOW_IWC = 0.03   !liquid water holding capacity for snowpack (m3/m3) (0.03)
  REAL, PARAMETER :: SWEMX    = 1.00   !new snow mass to fully cover old snow (mm)
                                       !equivalent to 10mm depth (density = 100 kg/m3)
```

Figure 3.1: Section of the source code of Noah-MP containing some fixed runoff and snow model parameters.

```
! ----------------------- local variables ------------------------------------
  INTEGER :: IB          !waveband class

! ----------------------------------------------------------------------------
! zero albedos for all points

        ALBSND(1: NBAND) = 0.
        ALBSNI(1: NBAND) = 0.

! when cosz > 0

         ALB = 0.55 + (ALBOLD-0.55) * EXP(-0.01*DT/3600.)

! 1 mm fresh snow(SWE) -- 10mm snow depth, assumed the fresh snow density 100kg/m3
! here assume 1cm snow depth will fully cover the old snow

         IF (QSNOW > 0.) then
           ALB = ALB + MIN(QSNOW*DT,SWEMX) * (0.84-ALB)/(SWEMX)
         ENDIF

         ALBSNI(1)= ALB          ! vis diffuse
         ALBSNI(2)= ALB          ! nir diffuse
         ALBSND(1)= ALB          ! vis direct
         ALBSND(2)= ALB          ! nir direct
```

Figure 3.2: Section of the source code of Noah-MP containing the snow albedo CLASS parameterization.

3.3.2 Model performance and parameter sensitivity

In this example, we configure Noah-MP to simulate runoff in three headwater catchments in the Colorado River basin: the Yampa River at Steamboat Springs (1468 km^2), the East River at Almont (748 km^2) and the Animas River at Durango (1819 km^2). The predominant land

surface cover of these basins is deciduous and evergreen forest, while the hydrology is mainly dominated by snow processes. All hydrologic model simulations were forced using hourly reanalysis outputs from the Weather Research and Forecasting (WRF) model (Skamarock et al. 2008), using the 4-km simulations described by Rasmussen et al. (2014).

The period for hydrologic model simulations is October 2000 through September 2008 (hourly time steps), and the first two years are used as a warm-up period (so that analysis is restricted to October 2002 through September 2008). No horizontal routing of surface overland flow, subsurface flow, or channel flow is performed; instead, basin-average runoff is computed as the average of the 1D (vertical) 4-km model grid cells. Finally, model outputs are aggregated to daily time steps to compute evaluation metrics.

Figure 3.3a displays scatter plots with runoff model simulations versus observations (period Oct/2002 – Sep/2008) at the three basins of interest, using default parameter values. The RMSE values and Nash-Sutcliffe efficiencies indicate that model performance is quite poor, especially at the East River and Animas River basins, and that parameter calibration is a necessary step to improve model fidelity. This point seems obvious for the hydrologic community (especially for applied hydrologists relying on traditional bucket-style rainfall-runoff models), where tremendous advances have been achieved in terms of parameter estimation methods (e.g., Duan et al., 1992; Gupta et al., 1998; Yapo et al., 1998; Vrugt et al., 2003a, 2003b, 2006; Pokhrel et al., 2012), sensitivity analysis (e.g. Tang et al., 2007; van Werkhoven et al., 2008; Foglia et al., 2009; Wagener et al., 2009b; Göhler et al., 2013; Rakovec et al., 2014), ensemble simulation and verification (e.g., Carpenter and Georgakakos, 2004; De Lannoy et al., 2006; Pauwels and De Lannoy, 2009) and parameter uncertainty quantification (e.g. Beven and Binley, 1992; Uhlenbrook et al., 1999; Vrugt et al., 2005; Kavetski et al., 2006a, 2006b; Kuczera

et al., 2006; Thyer et al., 2009). However, use of these techniques is less common in the land surface community, where most attention has been focused on improving process parameterizations, typically using fixed parameter values obtained from the literature.

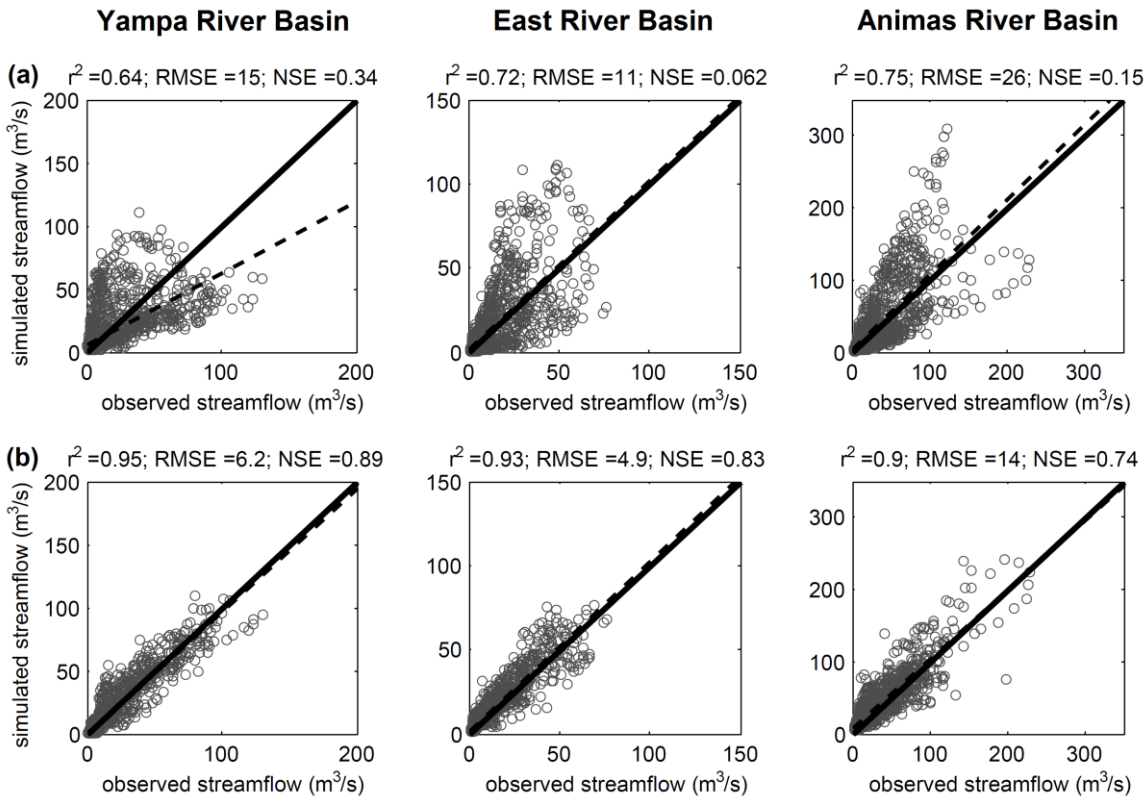

Figure 3.3: Model streamflow simulations versus observations for the period Oct/2002 - Sep/2008 using (a) default parameter values (top row), and (b) calibrated values for six originally hard-coded parameters: *f*, $R_{sb,max}$, λ_m, m_s, α_{min} and κ. (bottom row). The solid line is the 1:1 line, and the dotted line is the linear regression. In all panels, r^2, RMSE and NSE denote coefficient of determination, root mean squared error and Nash-Sutcliffe efficiency, respectively.

With the aim to identify which parameters have the largest impact on model predictions, we use the Distributed Evaluation of Local Sensitivity Analysis (DELSA) method (Rakovec et al. 2014) to evaluate the sensitivity of a suite of metrics (Table 3.1) to variations in the model

parameters (Table 3.2). All parameters listed in Table 3.2 can be considered 'observable' (i.e., a priori values can be specified by direct measurement or using indirect procedures), except the following (i.e., 'free' model parameters): empirical canopy wind parameter (w_{rp}), runoff decay factor (f), baseflow coefficient ($R_{sb,max}$), maximum surface saturated fraction (F_{sat}), exponent used in the curves for the melting season (m_s) and the exponent in snow decay albedo relationship (κ). From these free parameters, five of them were originally hard-coded (f, $R_{sb,max}$, F_{sat}, m_s and κ). It is noteworthy that this DELSA application required modification of the source code in order to 'uncover' all runoff and snow parameters (i.e., increase model agility), whose values were originally hard-coded.

Table 3.1: Objective functions included in DELSA

Name	Description	Equation
RMSE	Root Mean Squared Error	$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(Q_t^{sim} - Q_t^{obs}\right)^2}$
%BiasRR	Percent bias in runoff ratio	$\%BiasRR = \frac{\sum_{t=1}^{N}\left(Q_t^{sim} - Q_t^{obs}\right)}{\sum_{t=1}^{N} Q_t^{obs}} \times 100$
%BiasFMS	Percent bias in FDC midsegment slope	$\%BiasFMS = \frac{\left[\log\left(Q_{m1}^{sim}\right) - \log\left(Q_{m2}^{sim}\right)\right] - \left[\log\left(Q_{m1}^{obs}\right) - \log\left(Q_{m2}^{obs}\right)\right]}{\left[\log\left(Q_{m1}^{obs}\right) - \log\left(Q_{m2}^{obs}\right)\right]} \times 100$
%BiasCTR	Percent bias in runoff seasonality	$\%BiasCTR = \frac{CTR_{flow}^{sim} - CTR_{flow}^{obs}}{CTR_{flow}^{obs}} \times 100$

Q_t^{sim} : simulated flow for time step t

Q_t^{obs} : observed flow for time step t

Q_{m1}^{sim} : simulated flow with exceedance probability of $m_1 = 0.2$

Q_{m1}^{obs} : observed flow with exceedance probability of $m_1 = 0.2$

Q_{m2}^{sim} : simulated flow with exceedance probability of $m_2 = 0.7$

Q_{m2}^{obs} : observed flow with exceedance probability of $m_2 = 0.7$

$CTR^{flow} = \frac{\sum_{i=1}^{365} t_i Q_i}{\sum_{i=1}^{365} Q_i}$: centroid of daily hydrograph for an average water year

Table 3.2: Parameters of Noah-MP considered in this example application. The parameter ranges investigated (columns 5 and 6) were selected based on literature review of the different model components. The explanation of ranges in multipliers (if the parameter is spatially distributed in the basin) or raw values (if the parameter is spatially uniform) is provided in the 'Comment' column, together with the associated references.

Parameter	Description	Units	Distributed*	Range		Comment
				min	max	
	Soil parameters[a]					
b	Clapp-Hornberger b parameter	-	yes	**0.42**	**1.84**	Multipliers obtained from b exponent values in the range 2 - 15 (Cosby et al. 1984).
θ_{sat}	porosity	$m^3\ m^{-3}$	yes	**0.88**	**1.14**	Multipliers obtained from porosity values in the range 0.35 - 0.53 (Cosby et al. 1984), porosity constrained to be larger than field capacity.
Ψ_{sat}	saturated soil matric potential	$m\ m^{-1}$	yes	**0.15**	**2.20**	Multipliers obtained from saturated matric potential ranging from 0.02 to 0.78 (Cosby et al. 1984).
K_{sat}	saturated soil hydraulic conductivity	$m\ s^{-1}$	yes	**0.20**	**9.56**	Multipliers obtained from range 5×10^{-7} – 5×10^{-5} for ksat (Cosby et al. 1984).
k_{qtz}	soil quartz content	-	yes	**0.29**	**1.37**	Multipliers obtained from range 0.1-0.82 for quartz content (Hogue et al. 2005; Rosero et al. 2010).
	Vegetation parameters[a]					
$z_{0,veg}$	momentum roughness length	m	yes	**0.17**	**2.39**	Multipliers obtained from range 0.01-2.6 m (Dorman and Sellers 1989; Xia et al. 2012a).
ρ_L	leaf reflectance	-	yes	**0.90**	**1.10**	Multipliers based on average standard deviations reported by *Asner et al.* [1998].
ρ_S	stem reflectance	-	yes	**0.90**	**1.10**	Multipliers based on average standard deviations reported by *Asner et al.* [1998].
τ_L	leaf transmittance	-	yes	**0.90**	**1.10**	Arbitrary ±10 % multipliers, constrained by variations in leaf reflectance ($\rho + \tau \leq 1$).
τ_S	stem transmittance	-	yes	**0.90**	**1.10**	Arbitrary ±10 % multipliers, constrained by variations in leaf reflectance ($\rho + \tau \leq 1$).
χ_L	leaf/stem orientation index	-	yes	**0.50**	**1.67**	Multipliers defined such that max. absolute orientation index is 0.5 (Prihodko et al. 2008).
w_{rp}	empirical canopy wind parameter	m^{-1}	no	0.18	10	Obtained from *Goudriaan* [1977].
T_{min}	minimum temperature for photosynthesis	K	yes	**1.00**	**1.03**	Multipliers obtained from range 265-281 °K (Sacks et al. 2007).
$V_{max,25}$	maximum rate of carboxylation at 25° C	$\mu mol(CO2)\ m^{-2}\ s^{-1}$	yes	**0.65**	**1.35**	Multipliers obtained from standard deviations reported by (Kattge et al. 2009).
m_{ps}	slope of conductance-to-photosynthesis relationship	-	yes	**0.67**	**1.33**	Multipliers obtained from the slope range 4 - 12 (Sellers et al. 1996; Wolf et al. 2006).
SAI_{fw}	monthly stem area index, one sided (fall/winter)	$m^2\ m^{-2}$	yes	**0.10**	**2.14**	Multipliers obtained from stem area index range 0.01 - 3.0 (Otto et al. 2011).
SAI_{ss}	monthly stem area index, one sided (spring/summer)	$m^2\ m^{-2}$	yes	**0.10**	**1.88**	Multipliers obtained from stem area index range 0.01 - 3.0 (Otto et al. 2011).
LAI_{fw}	monthly leaf area index, one sided (fall/winter)	$m^2\ m^{-2}$	yes	**0.10**	**3.18**	Multipliers obtained from leaf area index range 0.01-7 (Myneni et al. 1997; Dorman and Sellers 1989; Hastie et al. 2002).
LAI_{ss}	monthly leaf area index, one sided (spring/summer)	$m^2\ m^{-2}$	yes	**0.10**	**1.27**	Multipliers obtained from leaf area index range 0.01-7 (Myneni et al. 1997; Dorman and Sellers 1989; Hastie et al. 2002).
	Runoff parameters[b]					
f	runoff decay factor	m^{-1}	no	1.0	10	Based values reported in *Beven* [1997].
$R_{sb,max}$	baseflow coefficient	$mm\ s^{-1}$	no	0.5	8	Based on *Niu et al.* [2005].
λ_m	grid cell mean topographic index	-	no	7.35	13.65	Variations up to 30 % from the default hard-coded value (10.35).
F_{sat}	maximum surface saturated fraction	-	no	0.29	0.46	Based on *Niu et al.* [2005].
	Snow parameters[b]					
m_s	exponent used in the curves for the melting season	-	no	0.5	3	Based on range in *Niu and Yang* [2007].
$z_{0,sno}$	snow surface roughness length	m	no	0.0001	0.01	Based on range suggested by *Marks and Dozier* [1992] and *Reba et al.* [2012].
θ_{lwc}	liquid water holding capacity for snowpack	$m^3\ m^{-3}$	no	0.01	0.08	Based on ranges in *Amorocho and Espildora* [1966] and *Anderson* [1973].
SWE_{new}	new snow mass to fully cover old snow	mm	no	0.5	5	Minimum is 50 % of default value; maximum obtained from *Xia et al.* [2012].
α_{min}	minimum snow albedo	-	no	0.45	0.65	Based on *Aguado* [1985] and *Dirmhirn and Eaton* [1975].
α_{max}	maximum snow albedo	-	no	0.70	0.95	Based on *Aguado* [1985] and *Essery and Etchevers* [2004].
κ	exponent in snow decay albedo relationship	h^{-1}	no	0.001	0.1	Based on *Essery and Etchevers* [2004].

*If the parameter is distributed, its sensitivity is analyzed on the basis of its multipliers. Although description and units refer to actual parameters in Noah-MP, parameter values in bold represent the multiplier values (instead of actual parameters).

[a]Exposed to users

[b]Hard-coded parameters

The results in Figure 3.4 demonstrate very high sensitivity for model parameters that were originally hard-coded. Specifically, RMSE is most sensitive to the monthly leaf area index for spring/summer (LAI_{ss}), the runoff decay factor (f), the exponent used in the snow depletion curves for the melting season (m_s) and the exponent in the snow decay albedo relationship (κ). In the case of the runoff ratio (%BiasRR), the most sensitive parameter is f, followed by the Clapp-Hornberger b parameter, the saturated hydraulic conductivity (K_{sat}), the slope of conductance-to-photosynthesis relationship (m_{ps}) and LAI_{ss}. When looking at variations in flashiness of runoff (%BiasFMS), the most sensitive parameters are f and κ, followed by the Clapp-Hornberger b parameter and the empirical canopy wind parameter (w_{rp}). Finally, the sensitivity in runoff seasonality (%BiasCTR) is mostly explained by variations of m_s, κ and the minimum snow albedo, α_{min} (i.e., snow parameters). The reader can also note that, among free parameters, relative differences in sensitivities between formerly hard-coded parameters and those exposed depend on the metric examined. For instance – when looking at RMSE – f, m_s and κ (originally hard-coded) are the most sensitive free parameters, followed by w_{rp} and $R_{sb,max}$ which have similar sensitivity. Nevertheless, w_{rp} becomes the second most sensitive free parameter (after f) when the objective criterion is %BiasRR. Overall, the most sensitive parameters are those that were formerly hard-coded – this result holds for all basins and all objective functions.

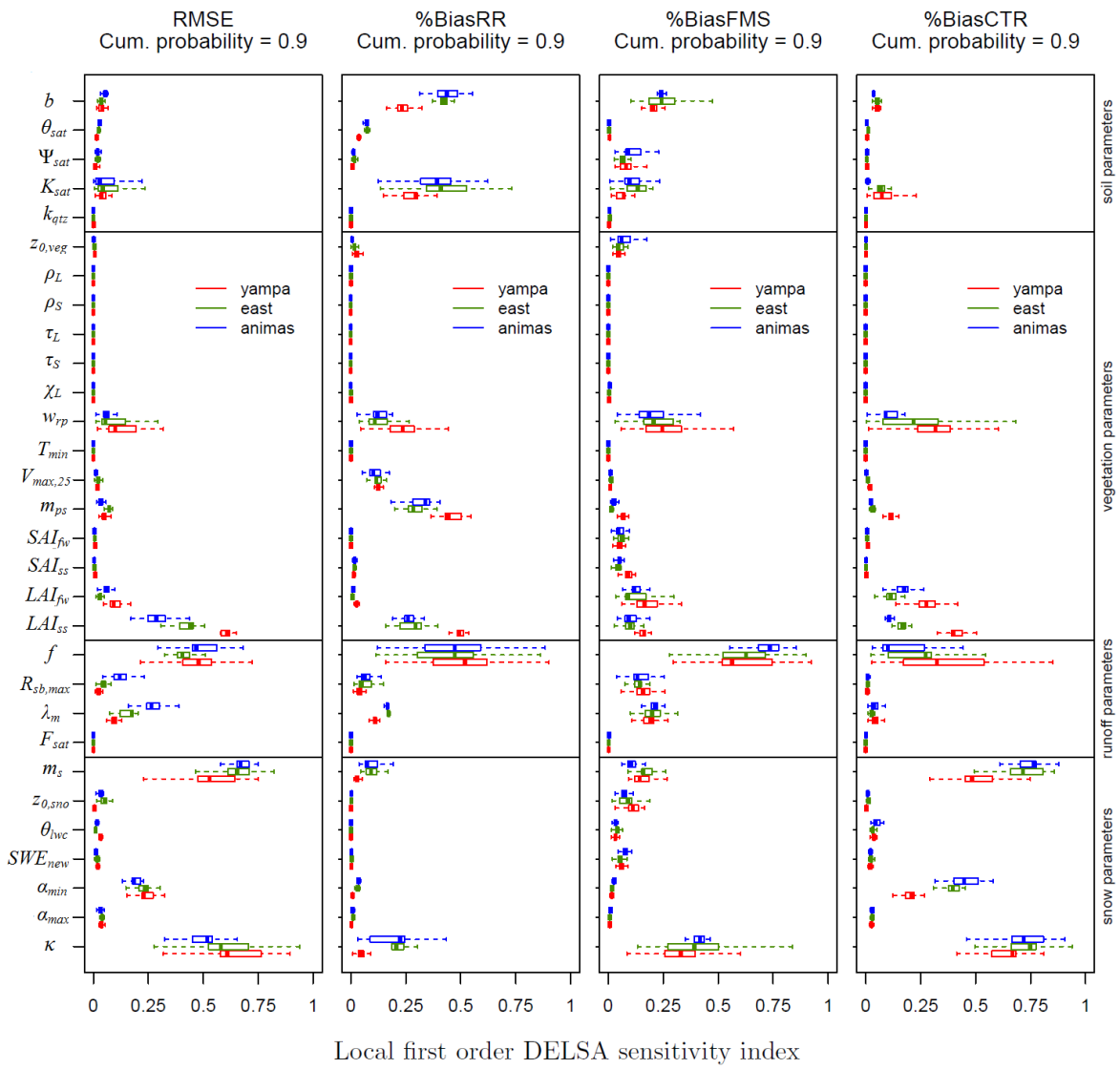

Figure 3.4: 90 % quantiles of the full frequency distribution of local first order sensitivity indices for several objective functions: root mean squared error (RMSE), percent bias in runoff ratio (%BiasRR), percent bias in flashiness of runoff (%BiasFMS) and percent bias in runoff seasonality (%BiasCTR). The uncertainty estimates are obtained by bootstrapping (resampled 1000 times). The vertical bold line in the boxplot is the median, the body of a boxplot shows the interquantile range (Q75–Q25) and the whiskers represent the sample minima and sample maxima. In DELSA, the assessment of parameter sensitivity is based on local gradients of the model performance index with respect to model parameters at multiple points throughout the parameter space. DELSA indices scale between 0 and 1, and larger values are associated with very sensitive parameters.

In summary, the response to high-intensity precipitation events, flashiness of runoff and seasonality are highly sensitive to variations of snow and runoff parameters (all of them originally hard-coded), while soil and vegetation parameters become more relevant when evaluating model behavior in terms of evapotranspiration processes. This suggests that calibration efforts aimed to improve model fidelity should include some of the hard-coded parameters in Table 3.2. To test this idea, we perform a simple calibration experiment aimed to adjust six runoff and snow parameters (f, $R_{sb,max}$, λ_m, m_s, α_{min} and κ) using the Shuffled Complex Evolution (SCE-UA) algorithm (Duan et al., 1992, 1993), by minimizing the root mean squared error between observed and simulated daily streamflow (RMSE) for the period October 1, 2002 to September 30, 2008. These parameters were selected because they showed the largest sensitivities for RMSE among formerly hard-coded parameters. The results displayed in Figure 3.3b clearly demonstrate how the inclusion of these parameters in the calibration process improves model accuracy (e.g., higher NSE and r^2, and lower RMSE).

3.3.3 The physical basis of hard-coded parameters

Complex models represent physical processes at a fine level of granularity (i.e., detail), and, as such, it is possible to impose much stronger a-priori constraints on model behavior in comparison to simpler, conceptual models. Noah-MP explicitly simulates all energy fluxes at the snow surface, as opposed to more parsimonious temperature-index models that represent snowmelt as an empirical function of temperature. Process-based models can therefore simulate accelerated snow melt during rain-on-snow events when turbulent heat fluxes are an important component of the snow-surface energy balance (e.g., Marks et al. 1999), whereas such processes are poorly represented in the temperature-index snow models as the empirical relationships between temperature and snow melt are applied consistently to all snow melt events. While the

temperature-index snow models may have fewer parameters, their values can be difficult to constrain correctly because the empirical functions implicitly represent a wide range of physical processes.

Although the strong physical basis of more complex process-based models provides powerful justification for their widespread use, it is important to recognize that empirical functions are widely used in such models at a much finer level of granularity. For example, consider the albedo decay parameterization from the source code illustrated in Figure 3.2, noting that the albedo decay parameter (κ) is one of the most sensitive model parameters for the criteria examined here (Figure 3.4). The physical processes affecting decreases in snow albedo over time include rounding and growth of the snow grains, deposition of dust on the snow surface, among others. These physical processes are included in some models (e.g., Jordan 1991; Flanner et al. 2007), but in Noah-MP the albedo decay rate is set to be constant over time. The lumping of multiple physical processes into a single albedo decay parameter is hence very similar to the lumping of all snow-surface energy fluxes into a single empirical temperature-melt expression, and there is no physical basis to treat time decay in snow albedo as a fixed constant in both space and time. This is a common problem, since other more flexible albedo parameterizations reported in the literature (e.g., Yang et al. 1997) also lack a proper justification for fixing parameters defining the albedo decay rate.

This specific and compelling example underscores a fundamental issue in process-based modeling: it is important to carefully specify the uncertainty of the different model parameters and process parameterizations (Montanari and Koutsoyiannis 2012), and retain the flexibility to adjust the model parameters to suit different hydroclimatic regimes. Because most physical processes are parameterized to some extent, treating uncertain model parameters as fixed

physical constants can unnecessarily constrain the agility of process-based models and severely limit their applicability to scales and locations for which these "parameters" have not been tuned. Furthermore, the reasons for imposing hard wired parameters may be obvious to the original model developers (e.g., related to the lack of measurements at the spatial scales of application, or the need to impose boundaries for some coefficients), these reasons may be less obvious to future model developers and users. Exposing parameter values to users is, in our opinion, a transparent and informative practice that supports future model development and improvement. Further, it is naive to believe that these hard-coded numerical values (e.g., the value 0.55 in the equation of Figure 3.2) can be denoted by precise values instead of probability density functions, considering that they have been either specified based on order-of-magnitude considerations or estimated via statistical analysis.

Moreover, ignoring the spatial scales for which physically-based equations describing fluxes of water and energy were derived (e.g., Richards' equation) and the spatial scale at which the empirical parameterizations were originally estimated (e.g., the Clapp-Hornberger pedo-transfer-functions or the saturated soil hydraulic conductivity in Table 3.2) will induce large uncertainties due to inappropriate scaling or averaging procedures, which in turn will propagate into model states and fluxes. Hydrologic theory (e.g., Darcy's law) developed at the scale of laboratory experiments (0.01–0.1 m) may be appropriate for predictions at the point scale, but may need to be modified for applications at larger scales (e.g., hillslope, catchment and beyond) due to effects of non-linearities, heterogeneities of landscape properties (e.g., vegetation, soils) and preferential flow of water through the soil matrix (Beven 2002). For instance, although the Clapp-Hornberger *b* parameter (as defined in Table 3.2) appears to be valid for a grid whose area is either 1 m^2 or 100 km^2 (i.e., it is implied to be quasi-scale invariant), it depends on the soil

texture (Clapp and Hornberger 1978), implying that the equations at which this parameter appears should be estimated at a scale for which the soil texture can be assumed quasi-homogeneous but still with some degree of uncertainty. Since a cell of 10 km^2 will contain many kinds of soil types, a scaling procedure should be performed to estimate the "effective" soil saturated hydraulic conductivity that best represents the sub-grid variability of soil within the given cell (Samaniego et al. 2010).

3.4 Where to from here?

Our call for increased agility of process-based models contributes to the debate on the "correct" approach to modeling (e.g. Freeze and Harlan, 1969; Beven, 2002; Reggiani and Schellekens, 2003; Loague and VanderKwaak, 2004). "Physics-based" models reflect a high level of confidence in the spatio-temporal representativity of physics-based equations describing complex systems, encoding very strong a-priori assumptions regarding individual processes (e.g. Abbott et al., 1986; Wigmosta et al., 1994; VanderKwaak and Loague, 2001; Ivanov et al., 2004; Maxwell and Miller, 2005; Rigon et al., 2006; Qu and Duffy, 2007; Lawrence et al., 2011; Niu et al., 2011), which hinder the representation of hydrologic process idiosyncrasies in specific catchments. By contrast, "conceptual" models begin with limited a-priori assumptions and infer knowledge through interpretation of how catchments respond to external forcing (e.g. Burnash et al., 1973; Lindström et al., 1997; Perrin et al., 2003; Fenicia et al., 2011), but are typically highly parameterized and do not explicitly represent many of the dominant physical processes necessary to reasonably simulate hydrological processes under changing hydroclimatic and land use conditions. The relative strengths of the so-called "physics-based" and "conceptual" modeling philosophies are therefore in their respective reliance on prior knowledge and data-based

inference. A key challenge in moving forward is to integrate these strengths to improve model representation of hydrological processes.

Finding a good balance between strong prior knowledge and data-based inference requires stepping back from specific model equations and examining the major decisions in the development of process-based hydrological models: (1) what schemes should we use to represent spatial variability and hydrologic connectivity throughout the model domain; (2) what parameterizations should we use to simulate the fluxes of water and energy at the spatial scale of the model discretization; and (3) what values should we use for the model parameters. When viewed from this perspective, there is no real distinction between physics-based and conceptual models: there is a continuum of modeling approaches (Gupta et al. 2012), with inter-model differences simply defined by decisions on which processes are represented explicitly, the spatial resolution used to simulate them, and the methods used to estimate model parameter values. The fundamental question follows from the key challenge just expressed: *How can we integrate our understanding of environmental physics with the available data to both define the structure of a hydrological model and define suitable values for model parameters?*

In our opinion, improving hydrological models requires developing effective methods to define and discriminate among competing modeling options, including both model structure and model parameters. This involves both (1) increasing the physical realism of traditional rainfall-runoff models and reducing the reliance on traditional model calibration methods that are plagued by compensatory errors and unrealistic hydrologic process simulations; and (2) increasing the agility of physically-motivated modeling systems to better suit local conditions. Modeling advances require explicitly simulating all dominant biophysical and hydrological processes, and focusing attention on detailed process-based evaluation of the suitability of

different methods to represent spatial variability and hydrological connectivity, different scale-appropriate flux parameterizations, and different approaches to estimate model parameter values. Implementing this vision requires effective methods for a controlled and systematic approach to model development and improvement (Clark et al. 2011a; Gupta et al. 2012), obtained by incorporating multiple modeling options into agile physics-based modeling frameworks and by applying a process-based philosophy for model evaluation and diagnosis.

Further, achieving this vision requires reconciling more agile models with the available data in order to identify suitable model structures and model parameter values. A useful solution to this problem can be found in the 'diagnostic approach' for model evaluation, based on confronting information contained in the data with the information provided by models (Gupta et al. 2008), and in the use of probabilistic representation of process parameterization equations (Bulygina and Gupta 2011). The diagnostic approach has proved to be useful for finding optimal parameter sets that provide a more realistic representation of catchment processes (e.g., Pokhrel and Gupta, 2009; van Werkhoven et al., 2009; Kollat et al., 2012; Pokhrel et al., 2012). The combined use of a diagnostic evaluation approach with inverse estimation and data assimilation methods (see Liu and Gupta, 2007 and Gupta et al., 2012 for an overview of techniques) can reduce the dimensionality of the model evaluation problem (e.g., focus on a subset of processes), and facilitate the reconfiguration of agile models (e.g., refinement of model equations, state and parameter updating) using information extracted from new datasets.

3.5 Concluding remarks

In this commentary we argue that the relatively poor performance of very complex physics-based hydrologic models can originate from unnecessary constraints that make it difficult to experiment with different kinds of spatial variability and process parameterizations.

As in the example presented here, it is typical for parameters in complex models to be specified using values reported in the literature, often based on limited data or order-of-magnitude considerations. This practice constrains our abilities to conduct extensive analysis and limits our opportunities to improve model fidelity and characterizing model uncertainty.

In view of this, we encourage an expanded and more comprehensive evaluation of critical modeling assumptions, building on the advances in multiple hypothesis modeling methodologies (e.g. Pomeroy et al., 2007; Clark et al., 2008, 2011; Fenicia et al., 2011; Niu et al., 2011; Essery et al., 2013). Future modeling systems should incorporate the capability to modify representations of spatial variability and hydrologic connectivity, individual process representations, numerical schemes, and couplings with other model components (e.g., atmosphere, sediment transport). Ongoing development of more agile versions of Noah-MP is just one example of active research in this area. Moreover, model reconfiguration capabilities should be able to cater to variable data availability (e.g., more complex model structures and meaningful specification of parameter values as more information is available) and integrate mechanisms for uncertainty quantification and analysis (e.g., ensemble generation, data assimilation, statistical post-processing and visualization). Such capabilities are necessary to facilitate diagnosis of model adequacy problems, refine model representations of natural processes, understand the major sources of uncertainty in model simulations, and identify critical areas for future research.

Finally, future research should also investigate robust physically-based scaling theories that can explain, and hence simulate, the heterogeneity of biophysical and hydrologic processes across multiple spatial scales. Progress in this direction will facilitate improved predictions of water and energy fluxes across different scales and locations, while constituting a necessary step

towards addressing the grand challenge of hyper-resolution large-scale modeling proposed by Wood et al. (2011).

CHAPTER 4: Effects of hydrologic model choice and calibration on the portrayal of climate change impacts

4.1 Introduction

There is now general agreement in the scientific community that the rising levels of carbon dioxide in the atmosphere are modifying historical climate conditions (IPCC 2013). One of the most relevant impacts of future climate change on society is changes in regional water availability for municipal, industrial, mining, irrigation, hydropower generation and other activities (Xu 1999; Brekke et al. 2009; Wagener et al. 2010). This situation is particularly critical for the Colorado River Basin (CRB) due to the susceptibility of runoff variations due to changes in precipitation and temperature, which stem from changes in evapotranspiration processes and snowpack accumulation/melt patterns (Christensen and Lettenmaier 2007). This vulnerability, together with the importance of the CRB for water resources supply for the growing regions of western and southwestern US, has motivated many climate change studies in this area, based on different modeling approaches and, therefore, resulting in a diverse set of conclusions (Milly et al. 2005; Christensen and Lettenmaier 2007; Hoerling and Eischeid 2007; Ray et al. 2008; Hoerling et al. 2009; Rasmussen et al. 2011, 2014; Miller et al. 2011, 2012; Vano et al. 2012, 2014).

The large uncertainty in estimates of hydrologic changes (i.e. changes in hydrologic variables obtained from hydrologic models) due to climate perturbation is not surprising for the hydrologic research community. In recent decades, many sources of uncertainty for quantifying climate change impacts on water resources have been identified (Chen et al. 2011), including: (1) selection of greenhouse gas emission scenarios, (2) choice of climate model(s), (3) specification of climate model initial conditions, (4) choice of meteorological forcing downscaling methods,

(5) selection of hydrological model structures and (6) choice of hydrological model parameter sets. Understanding risks associated with climate change requires estimating the uncertainty at each step of the modeling process (Xu 1999; Bergström et al. 2001; Wilby 2005; Wilby and Harris 2006; Graham et al. 2007; Chen et al. 2011; Vano et al. 2014). Among these elements, the choices of climate model (Murphy et al. 2004) and downscaling methods (Gutmann et al. 2012, 2014) have received significant attention, because recent studies have found that these are the main contributors to overall uncertainty (Wilby and Harris 2006; Chen et al. 2011).

Although a considerable number of past studies focused on the treatment of uncertainty in climate change projections, only a few have focused on hydrologic model structures and parameter uncertainty. For instance, Wilby (2005) explored parameter stability and identifiability using two hydrologic model structures, finding that (i) transferability of model parameters between wet and dry periods depends on the representativeness of the training period, and that (ii) model structure uncertainty on projected streamflow can be comparable to the uncertainty due to choice of emission scenario when the simplest model (low flows period) is considered. Jones et al. (2006) applied three different models in 22 Australian catchments covering a wide range of climates, and demonstrated that runoff variations due to changes in rainfall and evapotranspiration are clearly model dependent. Jiang et al. (2007) compared outputs from six hydrological models for mean annual and monthly changes in hydrologic variables due to perturbations of precipitation and temperature, finding that (i) differences across models depend on the climate scenario, the season and the variable of interest, and (ii) models without thresholds in soil moisture have larger differences in projected changes in soil storage. Poulin et al. (2011) used two different hydrological models to compare the effects of model structure against parameter equifinality on the uncertainty of hydrologic simulations, finding that model

structure uncertainty dominates. More recently, Miller et al. (2012) found that hydrologic model choice has a large effect on the portrayal of climate change impact in the San Juan River basin. Vano et al. (2012) evaluated hydrologic changes due to perturbed climate scenarios using six hydrologic/land surface models in the CRB, demonstrating large inter-model differences in runoff changes due to shifts in precipitation and temperature. Surfleet et al. (2012) compared a large-scale approach, a basin scale-approach and a site-specific approach in the Santiam River Basin (USA), showing that differences in the portrayal of climate change impacts can be attributed to scale and the ability of the models to capture local hydrological processes.

Despite the increasing awareness of the implications of hydrologic model structures on the estimation of climate change impacts on hydrology, the effects of model representation of specific processes (e.g. evapotranspiration, snow accumulation and ablation, percolation) on the overall hydrologic model response still remains unclear. In view of this, the main goal of this chapter is to compare hydrologic changes obtained with different hydrologic model structures in terms of annual water balance, monthly simulated processes (e.g. ET, snowpack, soil moisture) and signature measures of hydrologic behavior (e.g. runoff seasonality, long-term baseflow), for uncalibrated and calibrated model simulations.

4.2 Approach

4.2.1 Meteorological forcings

Meteorological data from WRF simulations is available at hourly time steps and a 4 km-resolution for both historical and PGW conditions during the period October/2000 - September/2008. The variables and temporal disaggregation used depend on specific hydrologic model requirements (further details in Chapter 2). Figure 4.1 includes basin-averaged monthly precipitation and temperature from WRF for current and future climate scenarios over the period

October/2002 – September/2008. Note that PGW simulations reflect increases in precipitation during fall/winter and the beginning of spring, and a decrease in precipitation during summer over all basins. On the other hand, the increase in temperature tends to be uniform throughout the year in all basins. These signals in precipitation and temperature changes are present at each individual water year (not shown here), although monthly precipitation amounts can vary at the basins of interest from year to year.

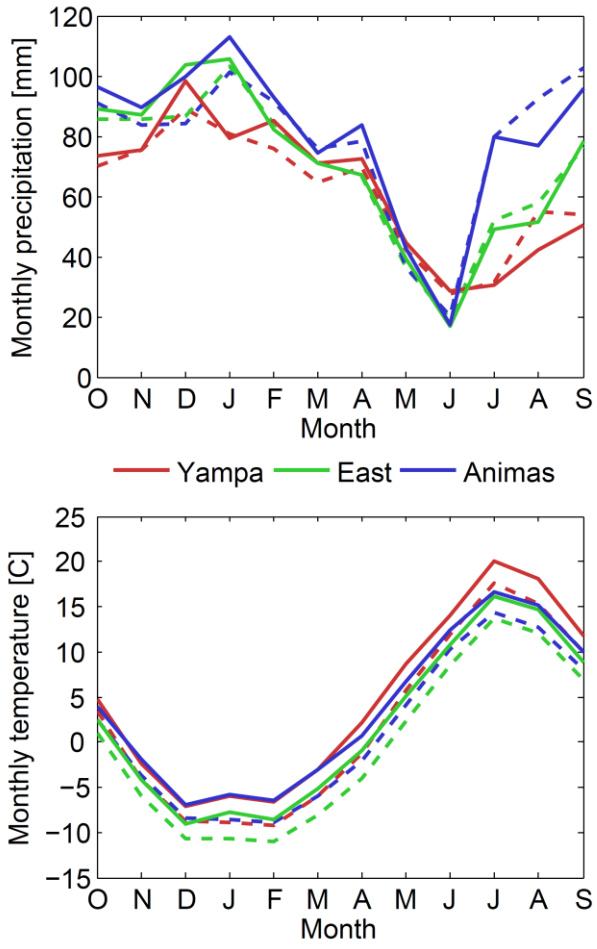

Figure 4.1: Basin-averaged monthly precipitation (top panel) and temperature (bottom panel) values for CTRL (dashed lines) and PGW (solid lines) WRF outputs used in this study (period Oct/2002 - Sep/2008).

The single choice of GCM, emission scenario and the time period over which the climate perturbation was obtained is certainly an important limitation for this study, since they affect the magnitude and direction of climatic shifts. Indeed, Vano et al. (2014) demonstrated the large effects of these decisions on long-term runoff projections over the upper CRB, including results from 19 GCMs and three emission scenarios (A2, A1B and B1) obtained by Seager et al. (2007) and Christensen and Lettenmaier (2007). However, they also noted that higher future greenhouse gas emissions broadly translate to a warmer and, in most cases, drier climate, implying that a general decrease in runoff should be expected in this region. Although high-resolution climate models limit the number of scenarios that can be analyzed, they offer a more realistic representation of climate features that strongly depend on terrain complexity (Rasmussen et al. 2011, 2014), providing better meteorological fields for the assessment of climate change impacts on hydrology.

4.2.2 Hydrologic/land surface models

We choose four hydrologic/land surface models: the US Geological Survey's Precipitation Runoff Modeling System (PRMS; Leavesley et al. 1983; Leavesley and Stannard 1995), the Variable Infiltration Capacity model (VIC; Wood et al. 1992; Liang et al. 1994, 1996) the Noah Land Surface Model (Noah-LSM; Ek 2003; Mitchell et al. 2004) and the Noah Land Surface Model with Multiple Parameterizations (Noah-MP; Niu et al. 2011; Yang et al. 2011). Further details on inter-model differences, information requirements and model setup are included in Chapter 2.

4.2.3 Experimental setup

All model simulations are carried out for the period between October 1, 2000 and September 30, 2008, using the first two years to initialize model states. As done for many past large scale (i.e., continental to global scale) hydrologic modeling experiments (e.g. Mitchell et al. 2004; Gerten et al. 2004; Xia et al. 2012b), we first compute hydrologic changes using default parameter values obtained from the information sources described in Table 2.2. Therefore, a comparison of hydrologic change estimates obtained from uncalibrated (i.e. use of default parameters) and calibrated model simulations will provide a comprehensive assessment of the caveats behind traditional methodologies used for climate change impact evaluation.

We calibrate all the models for all basins with the Shuffled Complex Evolution (SCE-UA) algorithm (Duan et al. 1992, 1993) by minimizing the root mean square error between observed and simulated daily streamflow for the period between October 1, 2002 and September 30, 2008. Given the short length of WRF reanalysis datasets and that our main priority is to analyze hydrologic change signals, we decided to perform calibration and compute hydrologic changes over the entire period Oct/2002 – Sep/2008 instead of splitting it into calibration and validation datasets. In this study, runoff from hydrologic model simulations is obtained as the sum of surface runoff and baseflow, including also interflow if the model is PRMS.

PRMS does not have an explicit river network routing scheme for streamflow; instead, it has a cascade module used to define connections for routing flow from upslope to downslope hydrologic response units and stream segments and among ground-water reservoirs (Markstrom et al. 2008). In VIC, Noah-LSM and Noah-MP, no horizontal routing of surface overland flow, subsurface flow, or channel flow is performed. Instead, basin-average runoff is taken as the average of the 1D (vertical) 4-km model grid cells' runoff. During the calibration process, we

preserve the spatial variability of a priori model parameters (in case they are spatially distributed) through the adjustment of multiplier values that are applied for each parameter within the entire watershed. We adjust only those parameter multipliers identified as the most sensitive after performing a Distributed Evaluation of Local Sensitivity Analysis (DELSA; Rakovec et al. 2014). The reader is referred to Appendix A for a list with the parameters included in the calibration of each model.

Once the calibration process is finished, hydrologic changes are computed for the period Oct/2002 - Sep/2008 by forcing the models with the same meteorological datasets used for uncalibrated simulations.

4.3 Results and discussion

4.3.1 Model performance

Figure 4.2 summarizes model performance for the period October/2002 - September/2008 in terms of mean annual streamflow, monthly streamflow and flow duration curves, for both uncalibrated and calibrated simulations. None of the hydrologic model structures considered in this study are able to reproduce seasonal runoff patterns or flow duration curves using default parameter values (Figure 4.2a). Although this is not surprising and has been widely reported in the literature, many studies that seek to characterize the water balance at the continental scale make use of non-calibrated or semi-calibrated land surface models (e.g. Mitchell et al. 2004; Xia et al. 2012b). Importantly, the inclusion of a 'classic' calibration process based on the minimization of the root mean square error (RMSE) between simulated and observed total runoff still leaves inconsistencies across different model structures (Figure 4.2b). Some models show large errors in mean annual runoff or seasonal runoff patterns even after calibration, and the FDC is not accurately represented by any model, particularly for low-flows.

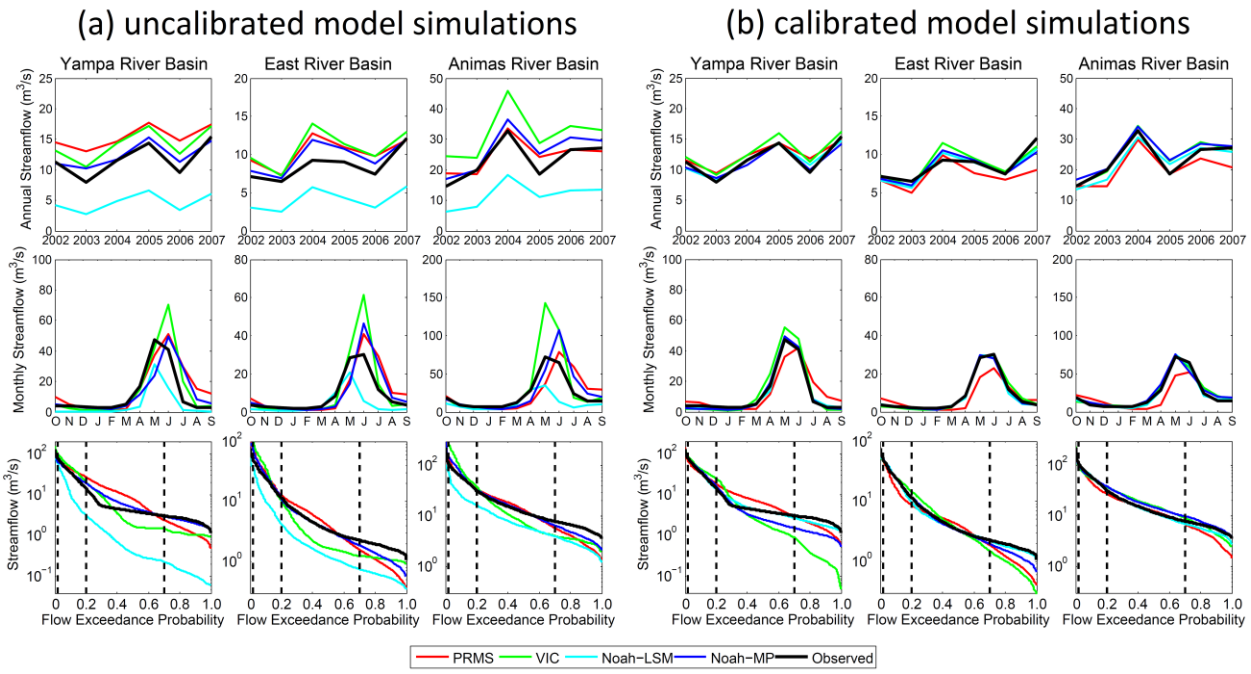

Figure 4.2: Historical streamflow simulation outputs for the period Oct/2002 - Sep/2008 for all basins: mean annual streamflow for all water years (top), mean monthly flows (middle) and flow duration curves (bottom).

In order to assess how much model performance improves functional catchment behavior through a traditional single-objective calibration strategy, we analyze the differences between simulated and observed values of signature measures of hydrologic behavior for both uncalibrated and calibrated model simulations (Figure 4.3). Parameter adjustment clearly improves the simulation of those signatures whose formulations are closer to the objective function used for calibration (in this case RMSE, which gives more relative importance to high flows). Consequently, calibration results in smaller inter-model differences in the runoff ratio (RR), the response to large precipitation events (FHV) and mid-range flows (FMM). On the other hand, inter-model differences in the runoff seasonality (CTR), the flashiness of runoff (FMS) and baseflow processes (FLV) are still pronounced after model calibration. Examples of this are Noah-MP and VIC at Yampa when looking at FMS, or Noah-LSM at East and PRMS at

Animas when evaluating baseflow processes (FLV), where calibration has actually degraded the signature measures. While a different objective function (e.g. based on the log of the flows) might improve other metrics (e.g. FLV), no single metric is likely to capture all catchment behaviors.

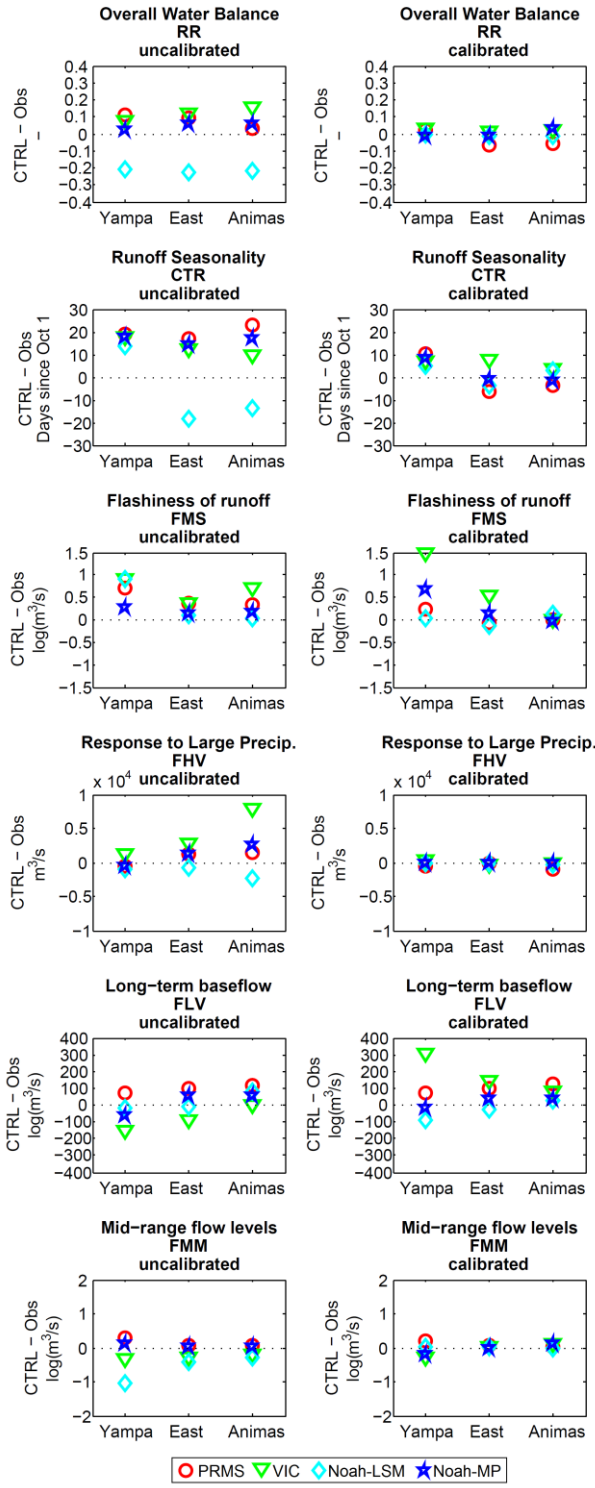

Figure 4.3: Difference between simulated (CTRL) and observed (Obs) signature measures of hydrologic behavior (period Oct/2002 - Sep/2008) obtained from uncalibrated (left panel) and calibrated (right panel) model runs.

4.3.2 Changes in annual water balance

To what extent does parameter calibration decrease the uncertainty in *projected changes* in the overall water balance? To provide an initial answer to this question, we first analyze both uncalibrated and calibrated hydrologic model outputs in the Runoff-ET space for a single climate scenario. In Figure 4.4, the diagonal lines represent basin-averaged mean annual precipitation for current and future climate scenarios over a 6-year average period (Oct/2002 - Sep/2008). The intersection of these lines with the x-axis indicates that all precipitation becomes runoff, while the intersection with the y-axis indicates that the system converts all precipitation into ET. In the same figure, different symbols represent outputs coming from different hydrologic model structures for current climate (unfilled) and future climate (solid). A symbol located exactly on the diagonal lines represents a simulation with negligible changes in storage over the 6-year simulation period, whereas symbols located below the 1:1 line denote increases in storage, and those above denote decreases in storage. Inter-model differences in precipitation partitioning are represented by the distance between different symbols (unfilled or solid), while the distance between a particular symbol (e.g. star for Noah-MP) for current (unfilled) and future (solid) climate scenarios represents the hydrologic change signal.

The results obtained from uncalibrated simulations (Figure 4.4a) indicate that inter-model differences are much larger than the magnitude of hydrologic change signals. Furthermore, all the models have the same hydrologic change signal direction (increase in ET and decrease or negligible change in mean annual runoff) with the exception of Noah-LSM, which projects increases in both runoff and ET (Figure 4.5a). As expected, inter-model differences in runoff (Figure 4.4b) decrease considerably (i.e. less variability along the x-axis) and the direction of hydrologic change signal (Figure 4.5b) is more consistent across models (i.e. less runoff and

more ET for future climate scenario) after calibration, with the exceptions of VIC at Yampa and PRMS at East. Noah-MP stands out from the rest of the models since the direction and magnitude of the signal is not substantially altered after the calibration (compare panels (a) and (b) in Figure 4.5). On the other hand, considerable shifts in projected runoff changes are obtained after calibrating PRMS at East (from -11 mm/year to 6 mm/year), VIC at Yampa (from -7 mm/year to 4 mm/year) and Noah-LSM at all basins (from 8 mm/year to -21 mm/year at Yampa, from 6 mm/year to -21 mm/year at East, and from 12 mm/year to -13 mm/year at Animas). Moreover, an important result from Figure 4.4b is that inter-model differences in precipitation partitioning into runoff and ET are still comparable or even larger than the magnitude of hydrologic change signal, even after model calibration.

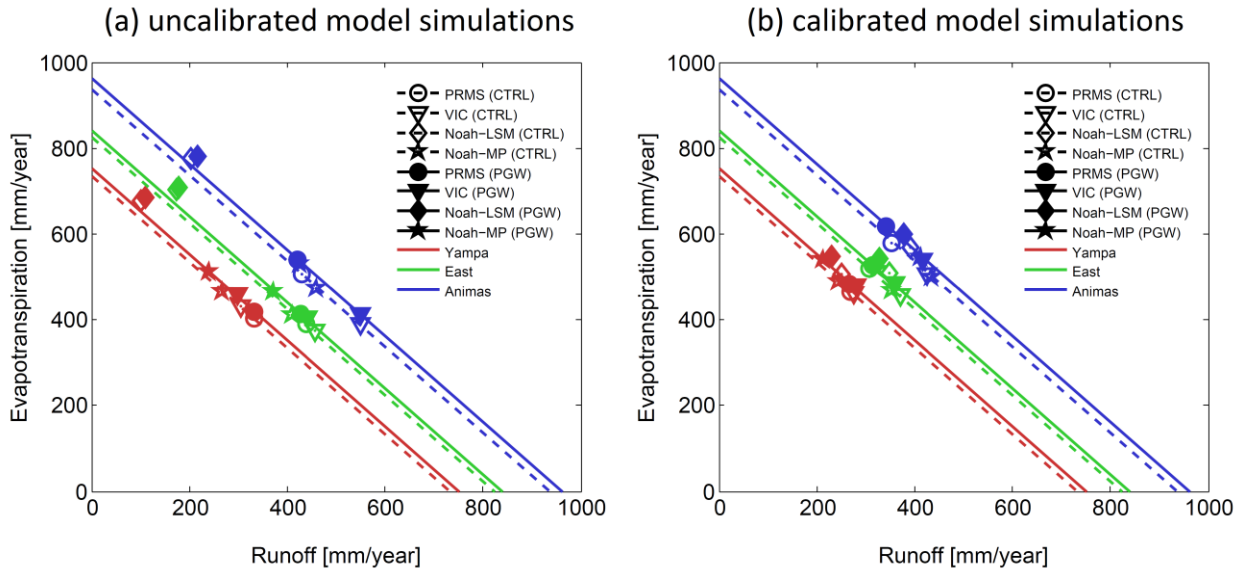

Figure 4.4: Partitioning of current (CTRL) and future (PGW) basin-averaged mean annual precipitation (diagonal, mm/year) into basin-averaged mean annual runoff (x axis, mm/year) and evapotranspiration (y axis, mm/year) across different model structures and basins for the period Oct/2002 - Sep/2008. Results are displayed for (a) uncalibrated model simulations and (b) calibrated model simulations.

Table 4.1: Values of fractional change [(PGW - current climate)/current climate] in basin-averaged total accumulated precipitation, peak SWE, accumulated ET, accumulated surface runoff, accumulated baseflow and accumulated total runoff (sum of surface runoff and baseflow, including interflow if the model is PRMS) averaged for an average water year (Oct/2002-Sep/2008) obtained from both uncalibrated and calibrated model simulations. Also included are the changes in dates of maximum SWE for each basin/model, where the values represent control minus PGW dates of maximum SWE.

Variable	Yampa					East					Animas				
	PRMS	VIC	Noah-LSM	Noah-MP	**Mean**	PRMS	VIC	Noah-LSM	Noah-MP	**Mean**	PRMS	VIC	Noah-LSM	Noah-MP	**Mean**
Total precipitation	0.02	0.02	0.02	0.02	-	0.02	0.02	0.02	0.02	-	0.03	0.03	0.03	0.03	-
Max SWE															
uncalibrated	-0.12	-0.08	-0.12	-0.09	**-0.10**	-0.10	-0.10	-0.14	-0.09	**-0.11**	-0.12	-0.14	-0.14	-0.10	**-0.12**
calibrated	-0.16	-0.10	-0.11	-0.12	**-0.12**	-0.19	-0.04	-0.09	-0.04	**-0.09**	-0.22	-0.06	-0.08	-0.06	**-0.11**
Date of max SWE															
uncalibrated	25	32	7	13	**19.25**	18	13	3	12	**11.50**	12	31	46	31	**30.00**
calibrated	25	2	7	0	**8.50**	25	4	4	6	**9.75**	12	1	5	0	**4.50**
Evapotranspiration															
uncalibrated	0.04	0.07	0.01	0.10	**0.05**	0.07	0.09	0.01	0.13	**0.07**	0.07	0.06	0.01	0.12	**0.06**
calibrated	0.04	0.04	0.08	0.11	**0.07**	0.02	0.06	0.07	0.10	**0.06**	0.07	0.06	0.06	0.09	**0.07**
Baseflow															
uncalibrated	0.03	-0.04	0.08	-0.08	**0.00**	-0.02	-0.05	0.05	-0.04	**-0.02**	-0.03	0.00	0.07	-0.04	**0.00**
calibrated	-0.04	0.01	-0.08	-0.08	**-0.05**	0.01	-0.03	-0.04	-0.05	**-0.03**	-0.11	-0.02	-0.03	-0.04	**-0.05**
Surface runoff															
uncalibrated	-0.01	0.01	0.14	-0.12	**0.01**	-0.03	0.00	-0.01	-0.15	**-0.05**	-0.01	0.01	-0.01	-0.12	**-0.03**
calibrated	0.03	0.02	-0.08	-0.24	**-0.07**	0.03	-0.01	-0.14	-0.14	**-0.07**	0.01	0.00	-0.06	-0.09	**-0.03**
Total runoff															
uncalibrated	0.00	-0.02	0.09	-0.10	**-0.01**	-0.03	-0.04	0.04	-0.09	**-0.03**	-0.02	0.00	0.06	-0.08	**-0.01**
calibrated	0.00	0.01	-0.08	-0.13	**-0.05**	0.02	-0.03	-0.06	-0.09	**-0.04**	-0.03	-0.02	-0.04	-0.05	**-0.03**

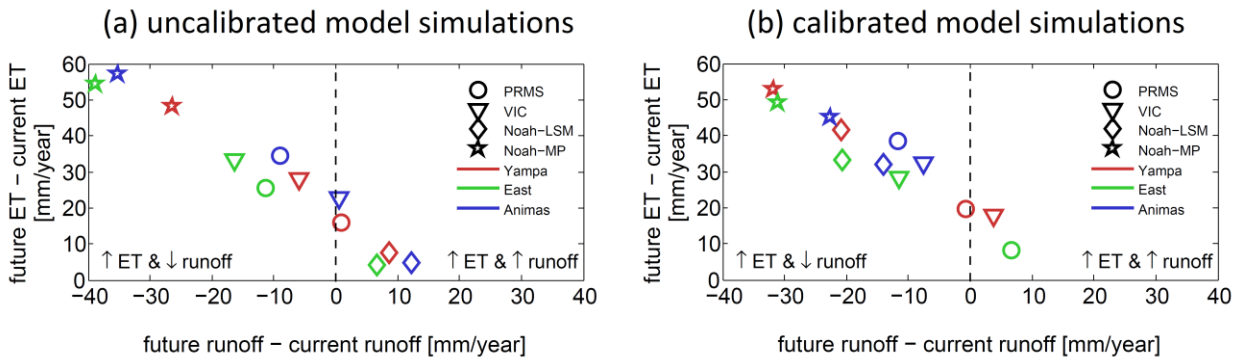

Figure 4.5: Projected changes in basin-averaged mean annual runoff (x axis, mm/year) and evapotranspiration (y axis, mm/year) across different model structures and basins for the period Oct/2002 - Sep/2008. Results are displayed for (a) uncalibrated model simulations and (b) calibrated model simulations.

Table 4.1 summarizes fractional hydrologic changes on an annual basis for both uncalibrated and calibrated model simulations over a 6-year average period (Oct/2002 - Sep/2008). A suite of different variables is included in order to illustrate how model structure selection and parameter calibration may affect the direction and magnitude of projected changes on hydrologic systems. For instance, in the East River basin, the magnitude of fractional changes in maximum SWE increases with PRMS (from -0.10 to -0.19) and decrease with Noah-MP (from -0.09 to -0.04) after the calibration process. Another example is given by baseflow at the Yampa River basin: fractional changes switch from positive to negative values after calibrating PRMS (from 0.03 to -0.04) and Noah-LSM (0.08 to -0.08), but they shift from negative (-0.04) to positive (0.01) values if the model selected is VIC. Similarly, Table 4.1 illustrates the effects of calibration on fractional changes in total runoff (e.g. PRMS at East, VIC at Yampa, Noah-LSM at all basins), capturing (although in different units) the results from Figure 4.5 previously discussed. The key result from Table 4.1 is that the inter-model differences in the hydrologic impacts of the CCSM-WRF climate scenario vary substantially across models (i.e., the

differences in the columns of Table 4.1 for each basin), and the inter-model differences are larger than the mean multi-model change signal for most metrics.

4.3.3 Monthly changes

Figure 4.6 shows mean monthly runoff values obtained from all models for both uncalibrated and calibrated simulations over a 6-year average period (Oct/2002 - Sep/2008). As expected, the use of default parameters (Figure 4.6a) translates into very different catchment responses under current and future climate scenarios, and these differences are also reflected in projected monthly changes (PGW - CTRL). The largest and smallest changes in runoff are obtained from VIC and Noah-LSM, respectively, and the seasonality of these shifts differs substantially across models. For instance, the Noah-LSM simulates increases in runoff during February-April, extending to May for the Yampa River basin, and a decrease during May-June, while Noah-MP generates an increase in runoff during March-May, and a decrease during June-September (Figure 4.6a). Much more consistent results across models are obtained when parameter calibration is performed (Figure 4.6b), and this is reflected in both the magnitude and seasonality of runoff variations. A key question that follows from here is whether inter-model similarities in runoff changes are due to inter-model agreement in changes of other water storages and fluxes.

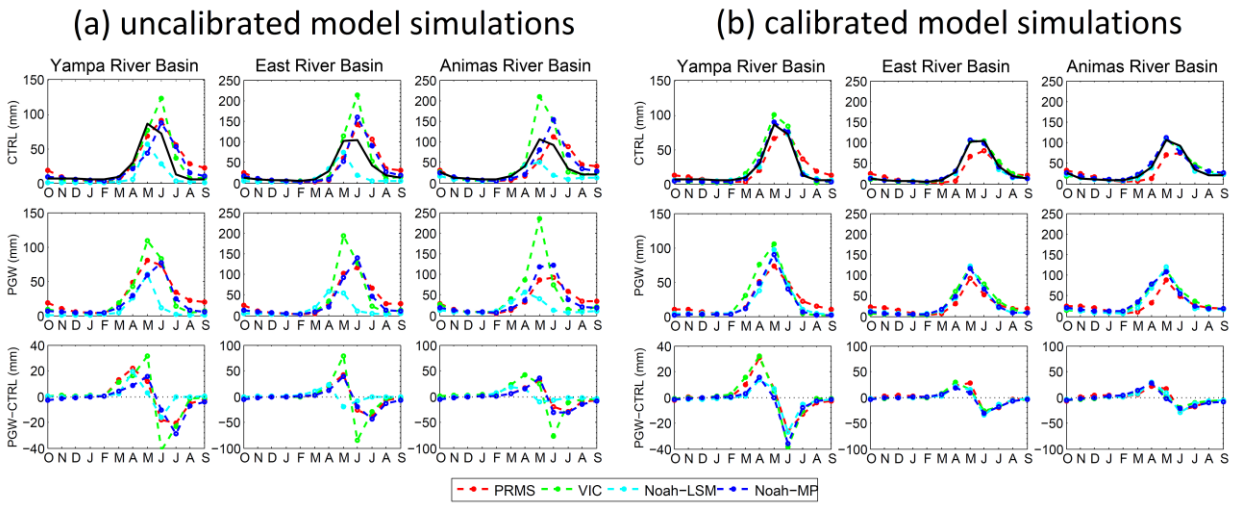

Figure 4.6: Current (CTRL), future (PGW) and changes (PGW - CTRL) in basin-averaged monthly runoff for uncalibrated (left panel) and calibrated (right panel) model simulations over a six-average water year (Oct/2002 - Sep/2008). The back lines in the CTRL panels represent historical observations.

With the aim to explore possible reasons for the (mis)match in projected runoff changes among different model structures, we analyze monthly changes in model states and fluxes obtained from both uncalibrated and calibrated runs (Figure 4.7). The variables included in this analysis are ET, snow water equivalent (SWE), soil moisture, baseflow and surface runoff. To improve consistency in the comparison across models, we consider only the top two soil layers for the computation of soil moisture storage with VIC, Noah-LSM and Noah-MP, and the addition of interflow to surface runoff for PRMS. The panels in Figure 4.7a show large differences in changes for ET, baseflow and surface runoff among models, while more consistent results in terms of seasonal cycles and amplitude are obtained for snowpack (except Noah-LSM) and soil moisture. However, inter-model differences of soil moisture and surface runoff are preserved or emphasized after the calibration process (Figure 4.7b). Furthermore, one can infer from the results displayed in Figure 4.6b and Figure 4.7b that the same runoff changes might be obtained using different hydrologic models due to very different mechanisms; i.e., internal

compensations of model structures and model parameter errors are adjusted through calibration in a way that allows similar responses from different watershed models. The clearest example in this case study is observed in the East River Basin, where monthly changes in runoff are very similar (Figure 4.6b); nevertheless, VIC compensates very large variations in soil moisture with other variables such as ET and baseflow, and PRMS does the same with large variations in ET, SWE and surface runoff.

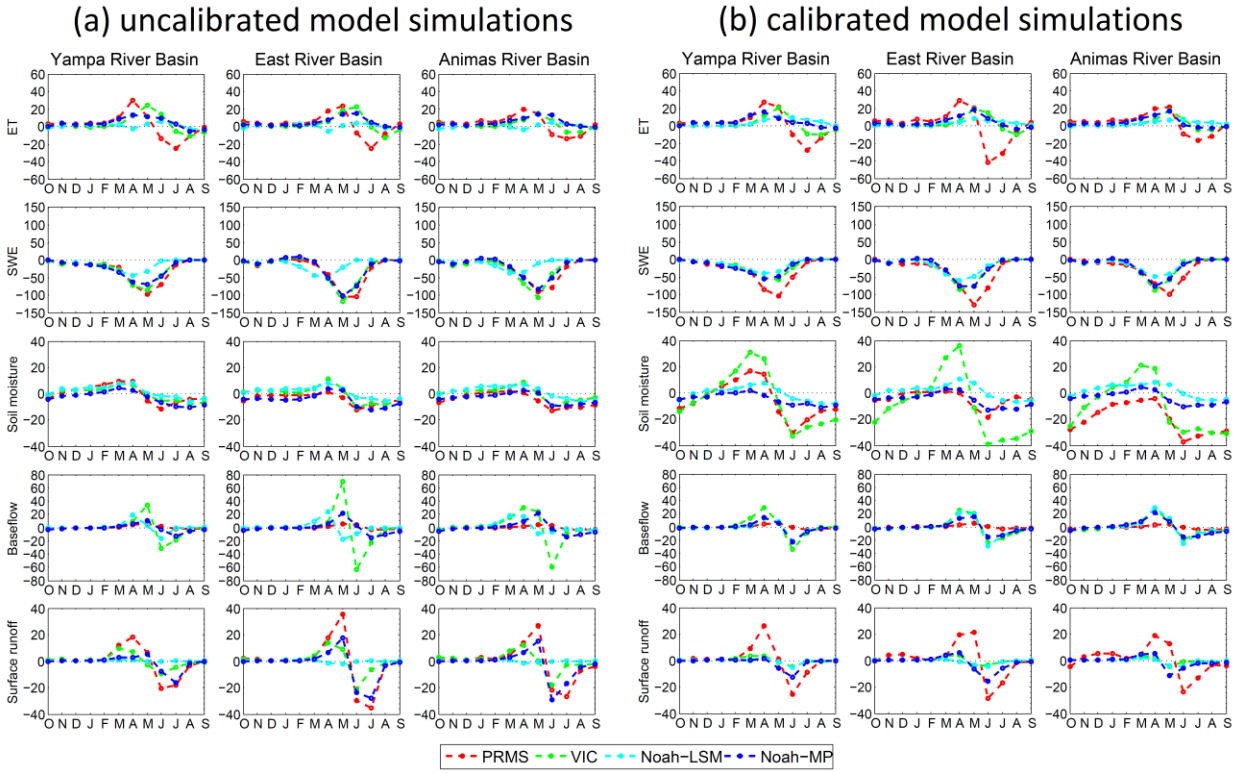

Figure 4.7: Monthly changes (PGW - CTRL) in basin-averaged fluxes and states (mm) for uncalibrated (left panel) and calibrated (right panel) model simulations over a six-average water year (Oct/2002 - Sep/2008).

4.3.4 Projected changes in catchment behavior

Finally, we compare the effects of model choice and parameter adjustment on projected changes in hydrologic signatures. Figure 4.8 illustrates differences between future (PGW) and current (CTRL) signature measures of hydrologic behavior for all (defined in Table 2.4)

models/basins, computed from both uncalibrated (left panel) and calibrated (right panel) model runs. The main result from Figure 4.8 is that calibration helps to decrease the uncertainty associated with model choice in projected changes of those signatures closely related with the objective function selected. Clear examples of this are runoff ratio (RR, except at Yampa), response to large precipitation events (FHV) and mid-range flow levels (FMM). However, the uncertainty due to model structure increases for some signatures and basins (e.g. runoff seasonality – CTR – at Yampa and East, flashiness of runoff – FMS – at Yampa, and baseflow processes – FLV – at Yampa and Animas). Moreover, different hydrologic model structures can provide opposite changes (signal) of some signature metrics even after calibration (e.g. FLV and FMS).

It is interesting to see that for both uncalibrated and calibrated model outputs, the only consistent signal obtained with all models is a negative change in runoff seasonality (CTR), which is directly related with an expected decrease in snowpack under the PGW scenario (i.e. shorter accumulation season and earlier melt season). For the case of calibrated model simulations, a general reduction of high flow volumes (FHV) occurs regardless of the model choice (except PRMS at Yampa). The results in Figure 4.8 illustrate the strong interplay between model structure and model parameters, and suggest the following hypothesis: different calibration approaches may lead to very different answers from those displayed in the right panel of Figure 4.8 or, put differently, that subjective decisions on configuring and calibrating hydrologic models may have unexpected and underappreciated impacts on the portrayal of climate change impacts. Current work is focused on this problem in order to get a better comprehension of uncertainties introduced by model structure selection and different parameter estimation strategies.

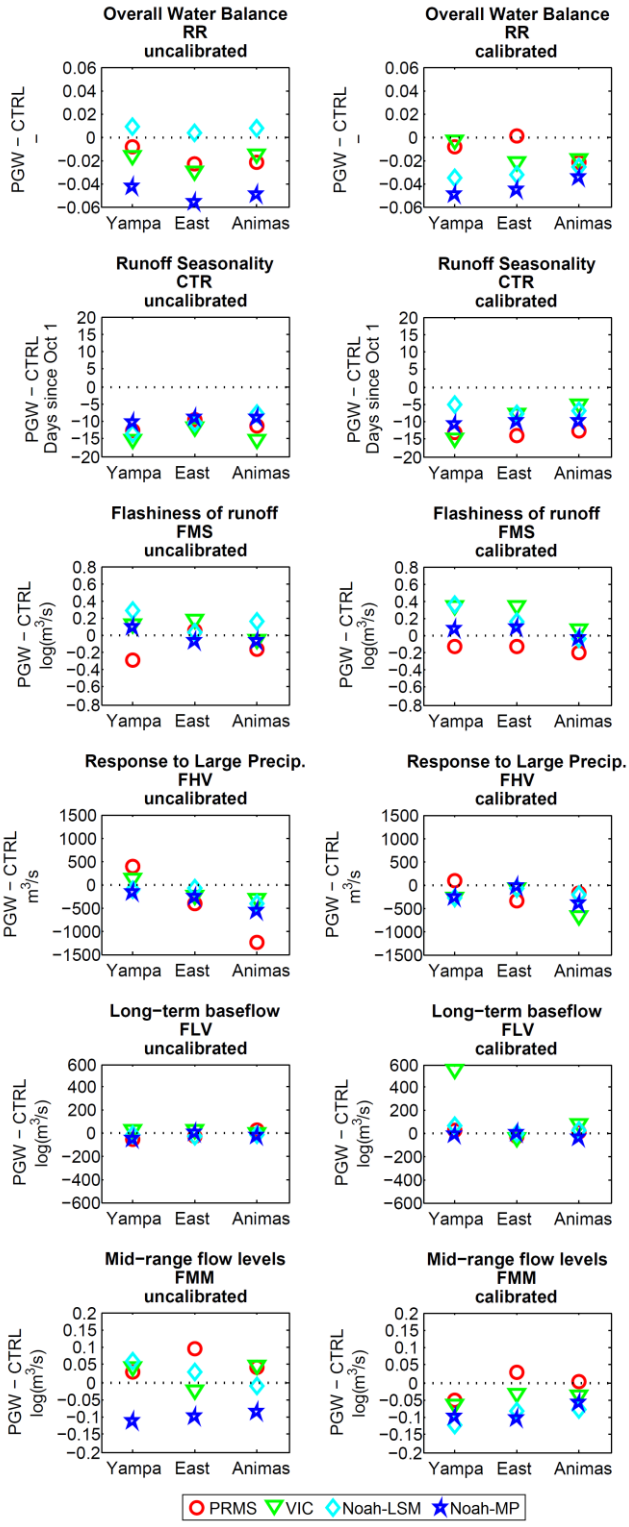

Figure 4.8: Impact of climate change on signature measures of hydrologic behavior for both uncalibrated (left panel) and calibrated (right panel) model simulations over a six-average water year (Oct/2002 - Sep/2008).

4.4 Conclusions

This study aims to improve our understanding of the effects of hydrologic model choice on the portrayal of climate change impacts. Specifically, we assess the effects of model structure selection on: (1) historical performance in terms of hydrologic signature measures, and (2) hydrologic changes due to a climate perturbation, with focus on the overall water balance and catchment processes. Because several efforts aimed to characterize future changes on the hydrology at the continental or global scales have made use of hydrologic/land surface models with little or no calibration, we include in our analysis a comparison between uncalibrated and calibrated model outputs. Our main findings are:

- Inter-model differences in portrayal of climate change impacts are substantial, even after calibration. These differences reflect on projected changes in overall water balance, monthly changes in individual simulated processes and signature measures of hydrologic behavior.
- In this chapter, better values for specific process evaluation metrics (i.e. signature measures) were obtained over the historical period Oct/2002 - Sep/2008 only if their mathematical formulation was close to the root mean square error between simulated and observed runoff (i.e. the calibration objective function).
- Consequently, single-objective calibration procedures constrain inter-model differences in climate change impacts for hydrologic metrics that are closely related to the objective function. In this study, calibration improved inter-model agreement on future projected changes of runoff ratio (except at Yampa River basin), response to large precipitation events and mid-range flow levels. However, inter-model agreement decreased when

evaluating the change of other metrics related with flashiness of runoff and baseflow processes.

- Although traditional calibration methods certainly improve inter-model agreement in projected changes of the overall water balance (i.e. partitioning of precipitation into ET and runoff), inter-model differences in the runoff-ET space are comparable and even larger than the hydrologic change signal for the scenario examined here.
- Single-objective calibration approaches aimed to reduce errors in runoff simulations do not necessarily enhance inter-model agreement in projected changes of some hydrological processes such as ET or snowpack. Moreover, identical changes in runoff might be obtained with different hydrologic model structures for very different reasons, indicating that the calibration process is compensating structural and parameter errors to give us "good" runoff simulations, but not to correctly reproduce catchment processes.

The main conclusion from this study is that subjective decisions in the selection of hydrologic model structures and parameters have large effects on the portrayal of climate change impacts. Moreover, these effects may directly impact adaptation strategies. For instance, the diversity of projected changes in runoff amounts and timing affects reservoir operations such as release schedules and magnitudes (Miller et al. 2012); uncertainty in responses to large precipitation events propagates to flood frequency estimates, which are required for design and safety assessment of infrastructure (Raff et al. 2009); uncertainties in ET projections relate with irrigation demands, and should therefore be considered in agricultural adaptation plans; and the diverse responses obtained in terms of long-term baseflow may impact future drought risk evaluation (Wilby and Harris 2006) and policies related with minimum instream flow requirements (Vano et al. 2014).

The implication of our findings is that previous studies evaluating the impacts of climate change on water resources may be over-confident. Moving forward, it is necessary to have a much more comprehensive assessment of the myriad of uncertainties in climate risk assessments; in particular, to improve characterization of uncertainties in hydrologic modeling applications.

CHAPTER 5: Implications of subjective hydrologic model choice and parameter identification strategies on the assessment of climate change impacts

5.1 Introduction

Significant trends in hydrologic variables such as snowpack, runoff, and intensity of extreme hydrologic events (e.g. floods and droughts) have been observed in the last century (e.g. Mote et al. 2005; Regonda et al. 2005; Stewart et al. 2005; Huntington 2006; Hamlet et al. 2007). Such long-term hydrologic shifts relate to changes in the intensification of water cycles associated with a warming climate (e.g. Trenberth 1999). Because of this, there is increasing awareness in the scientific community of the potential effects that changes in precipitation and temperature will have on water resources (IPCC 2013), and water managers have been confronting the task of projecting future changes in hydrology to better enable long-term planning for water resource management.

Quantitative assessment of climate change impacts on the hydrologic cycle (hereafter referred as *hydrologic change*) has been a remarkable challenge, especially if one considers the full range of uncertainties involved (e.g. Bergström et al. 2001; Graham et al. 2007; Brekke et al. 2009; Ludwig et al. 2009; Najafi et al. 2011; Majone et al. 2012; Steinschneider et al. 2012; Surfleet and Tullos 2013; Vano et al. 2014). One approach for addressing this problem has been the development and application of the 'cascade of uncertainty' paradigm, which attempts to quantify the uncertainty at every step in the modeling process (Chen et al. 2011), including: (1) selection of greenhouse gas emission scenarios, (2) choice of climate model(s), (3) specification of climate model initial conditions, (4) choice of meteorological forcing downscaling methods, (5) selection of hydrological model structures and (6) choice of hydrological model parameter sets. Recent studies have demonstrated that among these decisions the choice of global climate

model (GCM) and downscaling methods are the main contributors to overall uncertainty (e.g. Wilby and Harris 2006; Chen et al. 2011), motivating a detailed examination of different options for these methodological decisions (e.g. see Murphy et al. 2004 for GCMs and Gutmann et al. 2012, 2014 for statistical downscaling).

Although the selection of an appropriate hydrologic model structure and the identifiability of hydrologic parameter values are essential to advance our understanding of catchment processes (Son and Sivapalan 2007), uncertainties associated with these components have traditionally been given less attention in climate change impact studies (Bastola et al. 2011; Vano et al. 2012). The choice of a model structure, which in many cases relies on pragmatic considerations, involves several decisions: (i) what physical processes are included, (ii) which parameterizations are used for individual processes, (iii) how are these processes coupled to obtain the system response, and (iv) what numerical method(s) are used to solve model equations. In addition, parameter identification requires the following decisions: (i) defining a priori values for model parameters (especially relevant for distributed models), (ii) the choice of which model parameters to adjust (if any), and (iii) choices related to which calibration strategy to implement (e.g. calibration algorithm, objective function, training period, and forcing datasets). Of course, choosing differently across these options may affect the portrayal of climate change impacts. For instance, the selection of a specific model structure can lead to missing physical processes or the use of inappropriate parameterizations for the catchment(s) of interest. Similarly, parameter estimation strategies may affect the simulation of hydrologic processes, and can also compensate model structural errors – i.e. different models may provide the same answers for the wrong reasons.

In recent years, the hydrologic community has redirected efforts to better understand the effects of hydrologic modeling approaches on the assessment of climate change impacts. While some studies have looked at the implications of model choice (e.g. Boorman and Sefton 1997; Jones et al. 2006; Jiang et al. 2007; Ludwig et al. 2009; Bae et al. 2011; Miller et al. 2012; Vano et al. 2012; Surfleet et al. 2012), others have focused on the effects of parameter values (e.g. Cameron et al. 1999; Wilby 2005; Steele-Dunne et al. 2008; Surfleet and Tullos 2013). A third group of studies has included the combined effects of both sources of uncertainty, but using different experimental designs. For instance, Kay et al. (2009) addressed structural uncertainty by including two hydrologic models, and parameter uncertainty by calibrating several times using slightly different datasets (i.e. dropping one year at a time), finding that uncertainty due to different GCMs dominates over other sources of uncertainty, and that the relative importance of model structure and parameter values was catchment-dependent. Prudhomme and Davies (2009a, b) analyzed hydrologic modeling uncertainty by including two model structures and near-optimal parameter sets, with the result that GCM uncertainty was larger than the uncertainty from downscaling methods, which is itself larger than hydrologic uncertainty. Bastola et al. (2011) analyzed the role of hydrologic model uncertainty (structure and parameters) using a multi-model approach based on the Generalized Likelihood Uncertainty Estimation (GLUE; Beven and Binley 1992) framework and Bayesian Model Averaging (BMA; Raftery et al. 2005), concluding that hydrologic model uncertainties are 'remarkably high' when compared to other sources. Najafi et al. (2011) included four hydrologic model structures (three lumped and one distributed), representing parameter uncertainty with three objective functions for model calibration, and found that differences in projections depend directly on the choice of model structure. They also suggested that differences in model behavior over different seasons are

explained by differences in model states (e.g. soil moisture). Poulin et al. (2011) used two different hydrological models to compare the effects of model structure against parameter equifinality on the uncertainty of hydrologic simulations, finding that model structure uncertainty dominates. Finally, Mendoza et al. (2015a) examined hydrologic changes due to modified climate using four process-based hydrologic/land-surface models over three Colorado Headwater basins, concluding that while traditional single-objective calibration strategies may lead to similar simulations of streamflow from different models, this can be for very different reasons, leaving substantial uncertainty in how different climate scenarios will affect changes in hydrologic processes that are poorly constrained during model calibration.

In summary, previous contributions have clearly shown that the specific approach used for hydrologic modeling can affect the conclusions of climate change impact assessments. However, there is still limited knowledge about the interplay of hydrologic model choice and parameter identification strategies, especially in terms of individual hydrologic processes (e.g. evapotranspiration, generation of surface flow and baseflow, snowpack). This study builds on the previous work by Mendoza et al. (2015a) to further examine how hydrologic modeling strategies affect quantitative assessment of hydrologic changes. In particular, this chapter addresses the following two questions:

1. How do our hydrologic modeling decisions affect projected changes in the annual water balance (i.e. partitioning of precipitation into evapotranspiration and runoff) and hydrologic processes under a climate-changed future scenario?

2. What is the relative importance of model structure choice and parameter estimation strategies on the evaluation of climate change impacts?

To answer these questions, we assess and compare the implications of four hydrologic modeling decisions on the portrayal of climate change impacts: (i) choice of model structure, (ii) choice of objective function for model calibration, (iii) choice of multiple local optimal parameter sets, and (iv) choice of forcing dataset used for model calibration. The remainder of this chapter is organized as follows: the approach is described in Section 5.2; results and discussion are presented in Section 5.3, and the main conclusions are summarized in Section 5.4.

5.2 Approach

5.2.1 Climate change datasets

Meteorological data from WRF simulations is available at hourly time steps and a 4 km-resolution for both historical and PGW conditions during the period October/2000 - September/2008. The variables and temporal disaggregation used depend on specific hydrologic model requirements (further details in Chapter 2). Figure 4.1 includes basin-averaged monthly precipitation and temperature from WRF for current and future climate scenarios over the period October/2002 – September/2008. Note that PGW simulations reflect increases in precipitation during fall/winter and the beginning of spring, and a decrease in precipitation during summer over all basins. On the other hand, the increase in temperature tends to be uniform throughout the year in all basins. These signals in precipitation and temperature changes are present at each individual water year (not shown here), although monthly precipitation amounts can vary at the basins of interest from year to year.

5.2.2 Hydrologic modeling experiments

Table 5.1 provides a summary with hydrologic modeling experiments and the different options explored for each modeling decision. The following aspects are common for all the experiments:

- Hydrologic model simulations are carried out for the period between October 1, 2000 and September 30, 2008, using the first two years to initialize model states.
- Calibration objective functions and signature measures of hydrologic behavior are computed from daily time series of observed and simulated streamflow.
- Hydrologic changes are computed for the period Oct/2002 - Sep/2008 by forcing the model(s) with WRF datasets for the current and future climate scenarios.

A detailed description of the experimental design is provided in the following sub-sections.

Table 5.1: Summary of hydrologic modeling experiments; options explored in each experiment are in bold.

Experiment	Modeling decision	Options			
		Model structure	Objective function	Optimum	Forcing calibration dataset
a	Model structure	**PRMS, VIC, Noah-LSM and Noah-MP**	RMSE	Global	Rasmussen et al. (2014)
b	Objective function	PRMS	**RMSE, LRMSE, KGE, LKGE**	Global	Rasmussen et al. (2014)
c	Multiple local optima	PRMS	RMSE	**Local**	Rasmussen et al. (2014)
d	Forcing calibration data	PRMS	RMSE	Global	**Maurer et al. (2002), Xia et al. (2012) and Rasmussen et al. (2014)**

5.2.2.1 Model structure

We first focus on the following question: *given a single parameter estimation strategy, what are the effects of model choice on the portrayal of climate change impacts?* To find answers, we include the following hydrologic/land surface models: the US Geological Survey's

Precipitation Runoff Modeling System (PRMS; Leavesley et al. 1983; Leavesley and Stannard 1995), the Variable Infiltration Capacity model (VIC; Wood et al. 1992; Liang et al. 1994, 1996) the Noah Land Surface Model (Noah-LSM; Ek 2003; Mitchell et al. 2004) and the Noah Land Surface Model with Multiple Parameterizations (Noah-MP; Niu et al. 2011; Yang et al. 2011). Further details on inter-model differences, information requirements and model setup are included in Chapter 2.

5.2.2.2 Parameter identification strategy

In order to compare the uncertainty due to multiple model structures with the uncertainty arising from the choice of parameter estimation methods, we turn our attention to the following question: *given a single model structure, what are the effects of parameter identification strategies on the portrayal of climate change impacts?* Hence, we compute hydrologic changes using a fixed model structure with parameter sets obtained from: (i) different objective functions, (ii) multiple local optimal parameter sets, and (iii) different calibration forcing datasets. For these experiments, we use PRMS because of its very low computational cost compared to the rest of the hydrologic/land surface models in Figure 2.2. As with the calibrations performed for the other models, we use the SCE-UA algorithm including the parameters listed in Table A.1. The options considered for each calibration decision are detailed below.

5.2.2.2.1 Objective functions

In order to explore the implications of the choice of objective function on hydrologic changes in a controlled way, calibration simulations are forced with historical verification WRF outputs. The first objective function included here is RMSE, which is a standard metric that emphasizes high flows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(Q_t^{sim} - Q_t^{obs}\right)^2} \qquad \textbf{Equation 5.1}$$

Note that the formulation of RMSE is closely related to the Nash-Sutcliffe efficiency (NSE, introduced by Nash and Sutcliffe 1970), which is one of the most utilized indices in hydrology. Indeed, the minimization of RMSE is equivalent to the maximization of NSE. The second objective function selected is the root mean squared error computed with flows in logarithmic space, which emphasizes low flows:

$$LRMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\log(Q_t^{sim}) - \log(Q_t^{obs})\right)^2} \qquad \textbf{Equation 5.2}$$

The third metric considered here is the Kling-Gupta efficiency (KGE, proposed by Gupta et al. 2009), which is an explicit function of errors in: (i) mean flow values (i.e. flow volumes), (ii) variance of streamflow time series, and (iii) correlation between simulations and observations (i.e. timing and shape of the hydrograph). The Kling-Gupta efficiency is given by:

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2} \qquad \textbf{Equation 5.3}$$

where r is the linear correlation coefficient between simulated and observed streamflow time series, α is a measure of relative variability and β represents the bias between observations and simulations. The mathematical expressions for α and β are:

$$\alpha = \sigma_{sim} / \sigma_{obs} \qquad \textbf{Equation 5.4}$$

$$\beta = \mu_{sim} / \mu_{obs} \qquad \textbf{Equation 5.5}$$

where μ_{sim} (μ_{obs}) and σ_{sim} (σ_{obs}) denote the mean and the standard deviation of simulated (observed) flows.

In Equation 5.3, simulated and observed time series are in raw space. Therefore, the last objective function included in this analysis is the Kling-Gupta efficiency computed with flows in

log space, which is denoted by LKGE hereinafter. Because optimal parameter values found via minimization of RMSE (or NSE) may differ from those obtained by maximizing KGE (Gupta et al. 2009), we might expect that LRMSE or LKGE can lead to different optimal parameter sets.

5.2.2.2.2 Multiple local optima

Typically, the parameter set used for evaluating climate change impacts comes from a global single-objective optimization procedure applied over a historical period. However, it is likely that there are many local optimal regions in the parameter space, each one associated with a different hydrologic behavior in terms of the simulation of individual processes (e.g. snow accumulation and ablation, evapotranspiration, percolation, etc.). Moreover, the corresponding local optimal parameter sets can have a very similar performance in terms of the specific model evaluation metric selected for calibration.

To evaluate the effects of choosing any of these parameter sets on hydrologic change estimates, we identified local optimal regions in the parameter space by examining the evolution of SCE-UA outputs obtained from the minimization of RMSE using historical WRF outputs as meteorological forcings. The SCE-UA outputs contain information of $N_c \times (2 \times N_{par} + 1)$ parameter sets for each loop within the optimization process, being $N_c = 10$ the number of complexes (i.e. groups of parameter sets) and $N_{par} = 8$ the number of parameters included for calibration. Local optimal regions are identified by performing *k*-means cluster analysis (MacQueen 1967) over all parameter sets included in loops where complexes have still not merged. The loop that most clearly shows different parameter regions is selected to extract the local optimal parameter sets, which are those that have associated the minimum RMSE within each cluster. In this analysis, we identified five clusters to better illustrate different hydrologic behaviors coming from similar RMSE values.

5.2.2.2.3 Forcing calibration data

Some studies have given special attention to the implications of climate data selection on the response of hydrologic systems, concluding that existing differences in precipitation, temperature, radiation and other variables may have large effects on hydrologic model simulations (e.g. Wayand et al. 2013; Mizukami et al. 2014) and model calibration (e.g. Elsner et al. 2014). Therefore, our final experiment is aimed to quantify the implications of the forcing dataset used for parameter calibration on the evaluation of climate change impacts. For this experiment, we selected two types of retrospective climate datasets – observation based and re-analysis based climate datasets – that can be used for calibration. We compare hydrologic changes obtained with parameters calibrated using historical runs from Rasmussen et al. (2014) with those obtained with PRMS parameters calibrated with the datasets generated by Maurer et al. (2002) and Xia et al. (2012). Below are brief descriptions of these two additional climate datasets.

The Maurer et al. (2002) forcing dataset consists of four daily variables – precipitation, minimum and maximum temperature (T_{min} and T_{max}), and wind speed – as a gridded product derived (excepting wind speed) from observations and interpolated to the 1/8° grids using a consistent set of stations (Maurer et al. 2002). The dataset covers the conterminous United States and portions of Canada and Mexico from 1950 through the present. Maurer et al. (2002) is used as a forcing input for a great variety of hydrologic studies of climate change impacts (e.g. Steinschneider et al. 2012; Miller et al. 2012).

In the Xia et al. (2012) dataset, the precipitation data is a merged product of the NOAA/Climate Prediction Center (CPC) analysis of daily gauge precipitation, a national mosaic 4-km NOAA/National Weather Service (NWS) Stage II radar, 8-km hourly precipitation

analyses using the NOAA/Climate Prediction Center morphing technique (CMORPH; Joyce et al. 2004), and output from the North American Regional Reanalysis (NARR; Mesinger et al. 2006). CPC PRISM-adjusted gauge-only daily precipitation analyses are temporally disaggregated into hourly fields by multiplying daily CPC precipitation by hourly disaggregation weights derived from the national mosaic of 4-km Stage II radar and 8-km hourly CMORPH precipitation analyses. The mosaic of the Stage II product is interpolated to 1/8° grids. All the rest of the variables are generated from the NARR dataset using spatial and temporal disaggregation with the vertical adjustment methods described by Cosgrove et al. (2003).

Because the horizontal grid spacing in Rasmussen et al. (2014) is finer than that in Maurer et al. (2002) and Xia et al. (2012), we keep the basin grid size in the hydrologic models is as in Rasmussen et al. (2014) when using these two datasets, with meteorological forcings assigned to each basin grid cell with the nearest-neighbor interpolation method. For the remainder of this chapter, the results obtained with parameter sets calibrated with forcing datasets from Maurer et al. (2002), Xia et al. (2012) and Rasmussen et al. (2014) will be referred as M02, X12 and R14, respectively.

5.3 Results and discussion

5.3.1 Changes in annual water balance

We first analyze the effects of hydrologic modeling decisions on projected changes in annual water balance (i.e. partitioning of precipitation into ET and runoff). In Figure 5.1, the diagonal lines represent basin-averaged mean annual precipitation for current and future climate scenarios over a 6-year average period (Oct/2002 - Sep/2008). The intersection of these lines with the x-axis indicates that all precipitation becomes runoff, while the intersection with the y-axis indicates that the system converts all precipitation into ET. Each panel represents a

particular experiment, with different symbols representing outputs from different modeling options for current climate (unfilled symbols, forced with WRF-CTRL) and future climate (filled symbols, forced with WRF-PGW). A symbol located exactly on the diagonal lines represents a simulation with no change in storage over the analysis period (i.e. P = ET + R), while deviations from the 1:1 line reflect a change in basin storage over the analysis period (i.e. increases in storage for symbols located below the 1:1 line, and decreases for symbols located above). Inter-option differences (e.g., objective function choices) in precipitation partitioning are represented by the distance between different symbol types; and the distance between unfilled (current climate) and solid (future climate) symbols having the same type (e.g. star for Noah-MP in Figure 5.1a) represents the hydrologic change signal.

Figure 5.1a shows that given a unique parameter estimation strategy, inter-model differences (i.e. discrepancies between hydrologic models) in the partitioning of precipitation into ET and runoff are larger than climate change impacts. Moreover, inter-option differences are also larger than the magnitude of hydrologic change signal when looking at the choice of objective function (Figure 5.1b) and forcing calibration dataset (Figure 5.1d). In general, the uncertainty introduced by these parameter estimation choices overwhelms the uncertainty provided by model choice in the runoff-ET space. Furthermore, some options for a particular modeling decision expand the uncertainty ranges considerably (e.g. calibration with LKGE at the East River basin – shown in Figure 5.1b – focuses on matching low flows, resulting in very high runoff simulations, and therefore very high runoff ratios for both current and future climate scenarios). Finally, the different local optima in the parameter space (Figure 5.1c) have the smallest effect on the overall water balance because all these parameter sets have a similar

RMSE for runoff simulations over the calibration period, and hence are expected to reproduce similar annual runoff volumes.

Figure 5.2 directly compares projected changes in the partitioning of basin-averaged mean annual precipitation into runoff and ET for all hydrologic modeling experiments. In each panel, the dispersion of symbols with the same color across the ΔRunoff – ΔET space (with Δ representing the difference between future and current climate scenarios) represents the uncertainty introduced by a particular modeling decision at a specific basin. The results in Figure 5.2 reveal that all modeling decisions introduce considerable uncertainty around projected changes in mean annual runoff and mean annual ET. Furthermore, the direction of the signal is only preserved in all basins when considering multiple local optimal parameter sets (Figure 5.2c), although the magnitude of the signal changes considerably at the Animas River basin. Indeed, the points defining local optimal parameters for Animas (not shown) are quite separated, compared to East and Yampa, in the parameter plane defined by the two most sensitive PRMS parameters – snarea_curve (parameter defining the snow/area depletion curves) and pref_flow_den (parameter defining the preferential flow) – which explains a larger dispersion in hydrologic changes. When looking at the effects of model choice (Figure 5.2a), one can note that all models – except VIC at Yampa River basin and PRMS at East River basin – project an increase of ET and decrease of runoff. In the case of objective function (Figure 5.2b), however, the only basin for which a consistent signal direction is obtained with all options is Animas. It can also be noted that all parameter sets obtained with different forcing datasets (Figure 5.2d) project increases in ET and decreases in future runoff, except in the case of R14 for the East River basin.

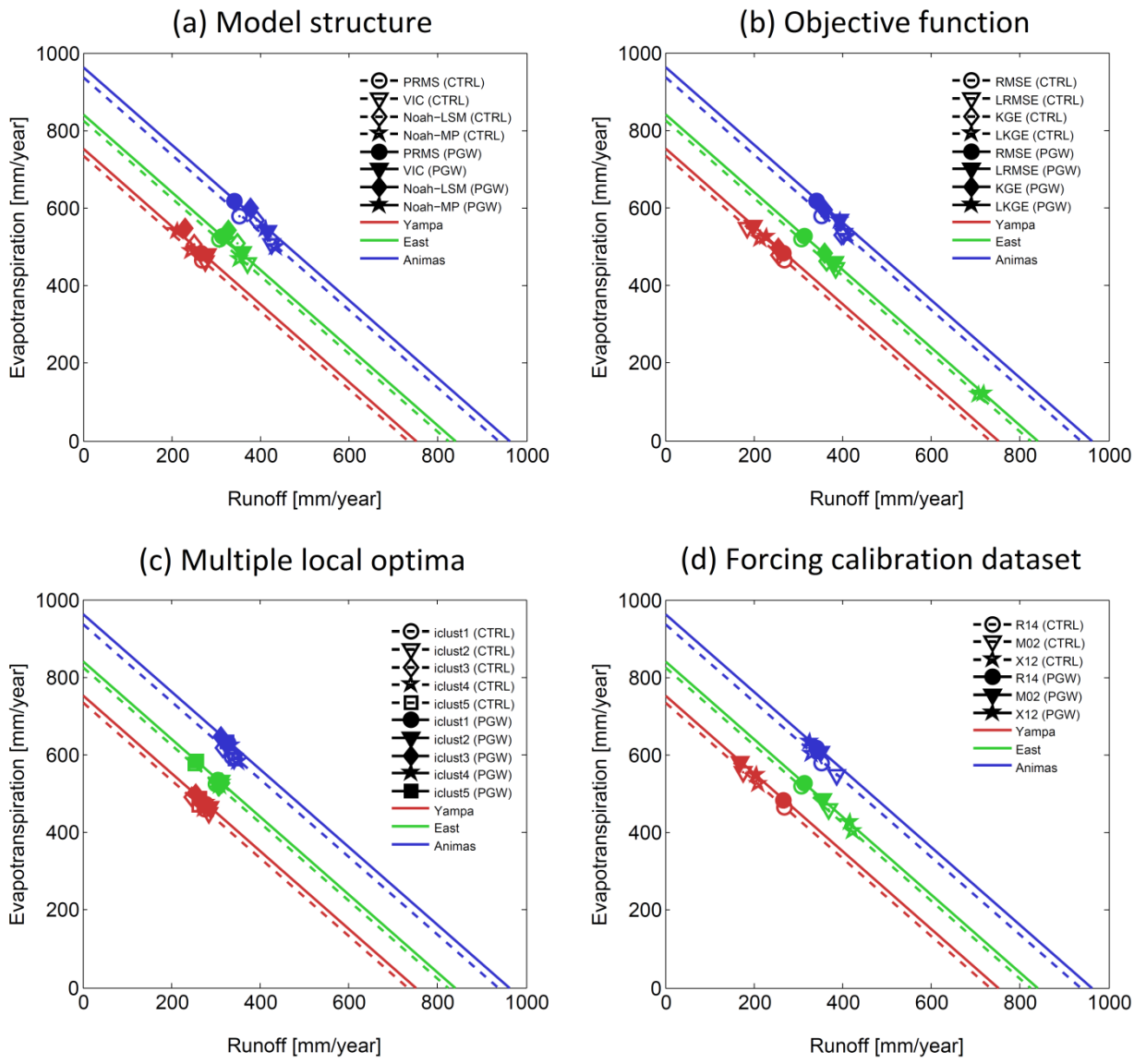

Figure 5.1: Partitioning of current (CTRL) and future (PGW) basin-averaged mean annual precipitation (diagonal, mm/year) into basin-averaged mean annual runoff (x axis, mm/year) and evapotranspiration (y axis, mm/year) across different model structures and basins for the period Oct/2002 - Sep/2008. Results illustrate the uncertainty coming from: (a) choice of model structure, (b) selection of objective function used for calibration, (c) multiple local optimal parameters, and (d) choice of dataset used for calibration.

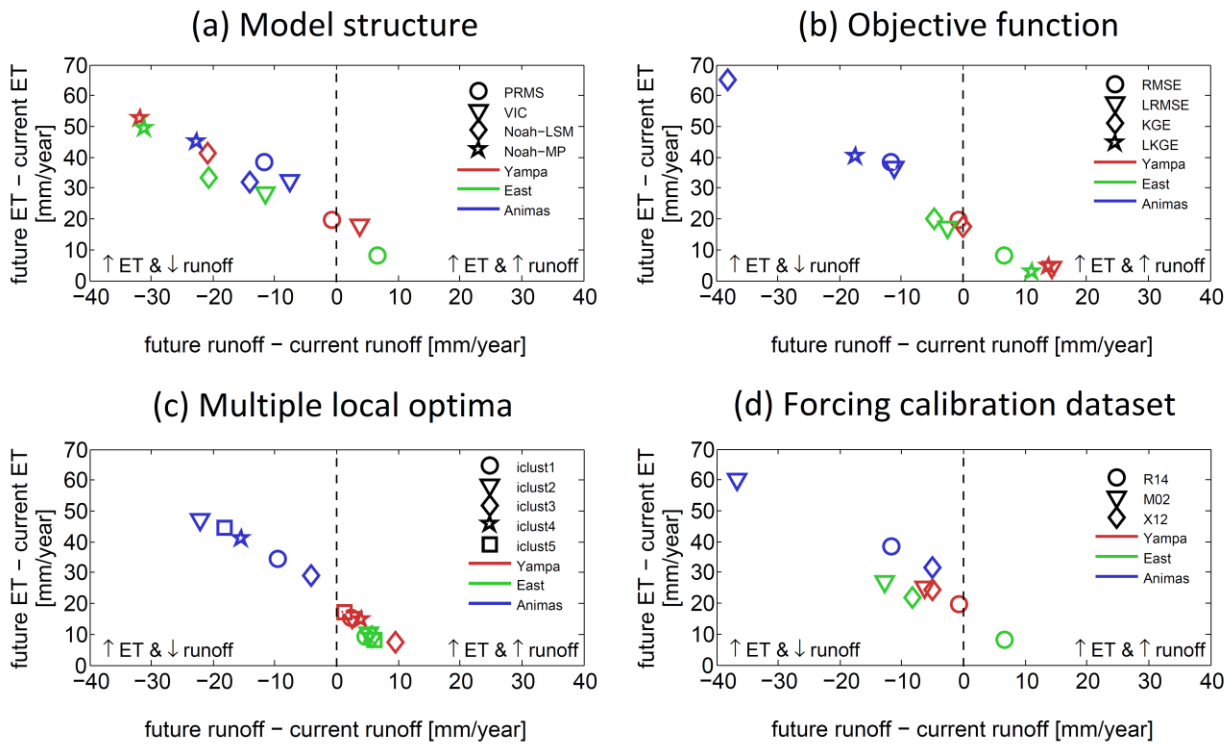

Figure 5.2: Projected changes in basin-averaged mean annual runoff (x axis, mm/year) and evapotranspiration (y axis, mm/year) across different model structures and basins for the period Oct/2002 - Sep/2008. Results illustrate the uncertainty coming from: (a) choice of model structure, (b) selection of objective function used for calibration, (c) multiple local optimal parameters, and (d) choice of dataset used for calibration.

Figure 5.2 also shows that the uncertainty in projected changes introduced by model choice (Figure 5.2a) is larger in the Yampa and East basins than in the Animas; however, parameter estimation decisions introduce more uncertainty in projected changes at Animas than in the Yampa or East basins (compare Figure 5.2b, 5.2c and 5.2d). In other words, the relative effects of modeling decisions on the magnitude of the signal appear to be basin-dependent, which is in agreement with other studies (e.g. Prudhomme and Davies 2009b; Addor et al. 2014). One possible explanation could be differences in hydrologic and physiographic catchment

properties. For instance, the Animas River basin has a much larger response to high precipitation/snowmelt events, which explains why the choice of PRMS parameters– previously selected for their high sensitivity to RMSE (i.e. high flows) – has a larger relative effect on hydrologic changes than the other catchments. In terms of physical properties, the Yampa and East River basins have smaller areas, lower mean slope, smaller elevation range and a flashier response (not shown). However, it is important to note that although previous studies have tried to link the hydrologic response with catchment characteristics using very large samples (e.g. Oudin et al. 2010; Köplin et al. 2012), they have found that still a considerable fraction of basins shows no overlap between physiographic and hydrological similarity.

5.3.2 Monthly changes

This section examines how hydrologic changes due to climate perturbation vary seasonally. Figure 5.3 illustrates the effects of hydrologic modeling decisions on projected changes in basin-averaged monthly runoff. In this plot, the spread in the curves representing different options for a given decision – each in a column in Figure 5.3 – is an indicator of the uncertainty introduced by that specific modeling choice. Figure 5.3 shows that the choices of model structure, objective function and forcing calibration dataset introduce considerable uncertainty, particularly during the melt season, and that all options for modeling decisions are consistent in terms of the seasonality (timing and largest/lower amounts) of projected runoff changes, except the choice of the model forcing dataset (Figure 5.3d). The large uncertainty introduced by forcing calibration dataset is in agreement with previous studies showing how the method used to generate meteorological fields heavily affects hydrologic simulations (Wayand et al. 2013; Elsner et al. 2014; Mizukami et al. 2014). The choice of multiple local optimal

parameter sets (Figure 5.3c) introduces little uncertainty in simulations of monthly runoff changes.

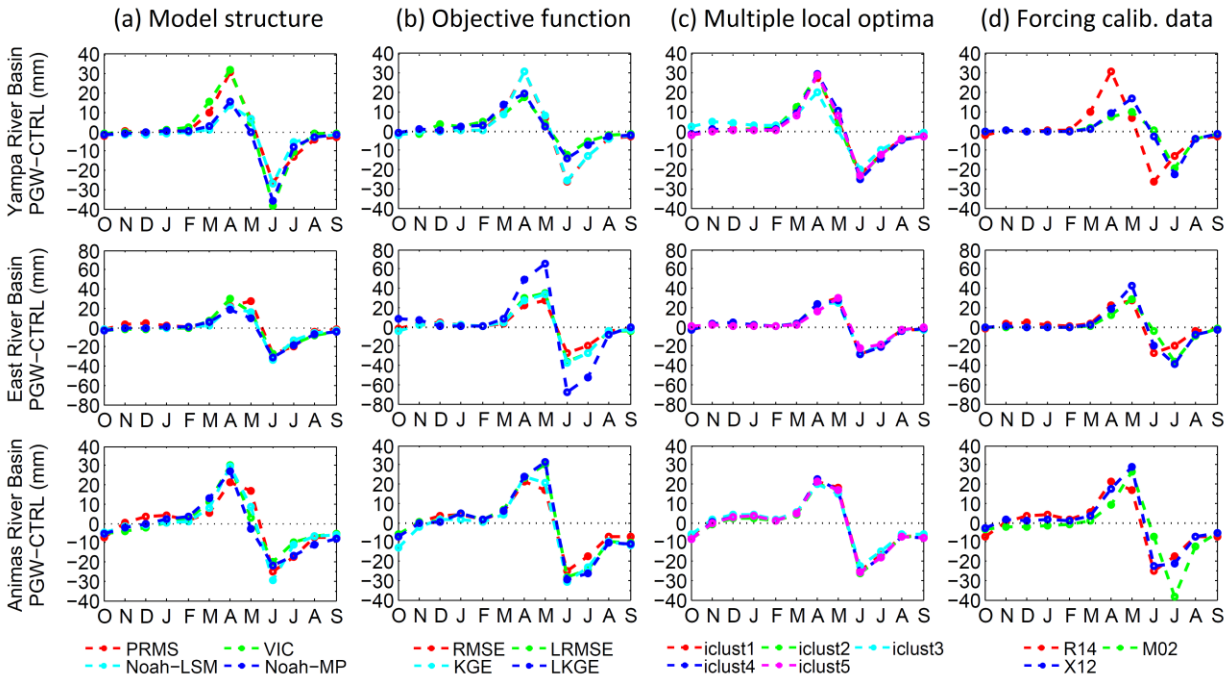

Figure 5.3: Projected changes (PGW - CTRL) in basin-averaged monthly runoff over a six-average water year (Oct/2002 - Sep/2008). Results illustrate the uncertainty coming from: (a) choice of model structure, (b) selection of objective function used for calibration, (c) multiple local optimal parameters, and (d) choice of dataset used for calibration.

A natural question from here is how the effects of modeling decisions compare to each other when examining other hydrologic processes. To address this question, we analyze projected monthly changes in model states and fluxes obtained from different options for each modeling choice. Figures 5.4–5.6 illustrate monthly changes in ET, snow water equivalent (SWE), and soil moisture, respectively, due to a climate perturbation. To improve consistency in the comparison across model structures (experiment (a) in Table 5.1), we consider only the top two soil layers for the computation of soil moisture storage with VIC, Noah-LSM and Noah-MP (see Figure 2.2).

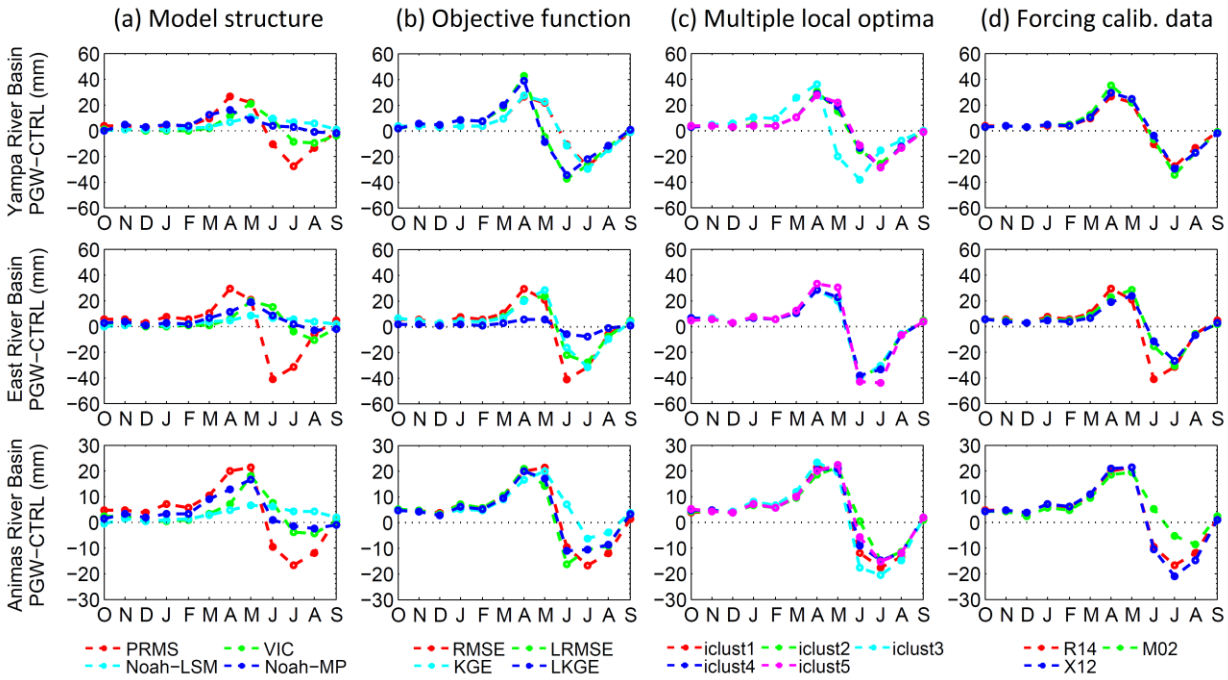

Figure 5.4: Same as in Figure 5.3, but for evapotranspiration (ET).

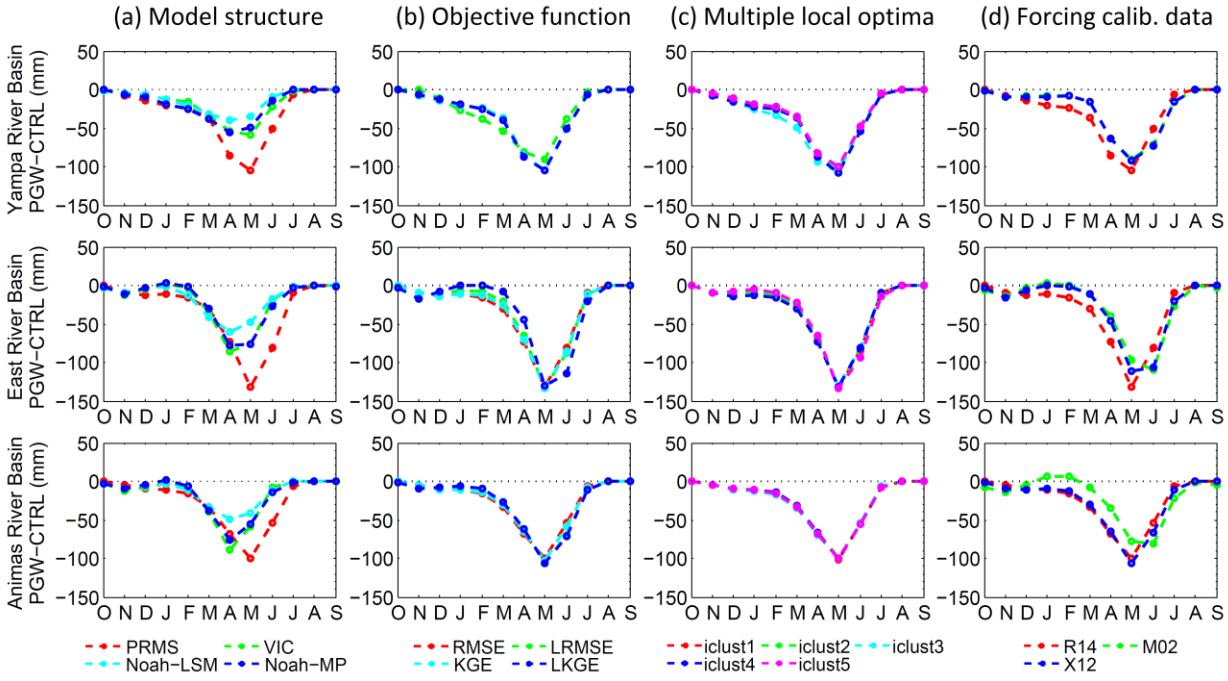

Figure 5.5: Same as in Figure 5.3, but for snow water equivalent (SWE).

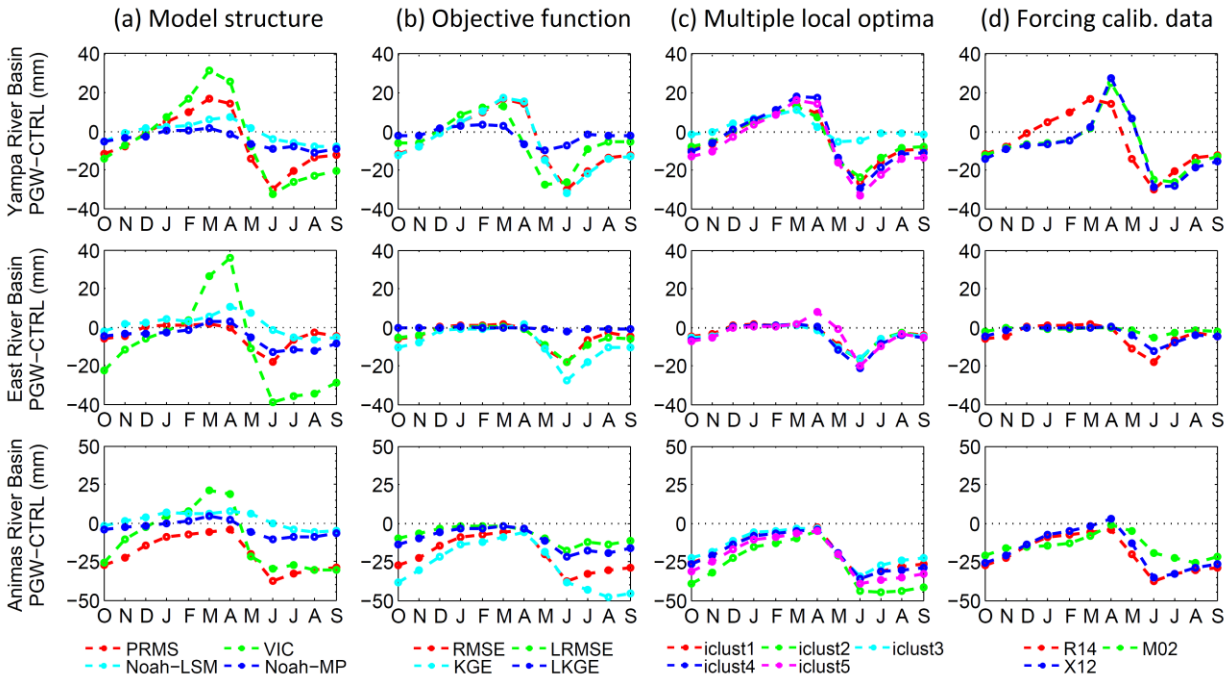

Figure 5.6: Same as in Figure 5.3, but for soil moisture.

Figure 5.4 shows that the choices of model structure – source of discrepancies in terms of canopy conceptualization and potential evapotranspiration (PET) formulations – and objective function – which controls the simulation of the overall water balance (Figure 5.1) – introduce large uncertainties in ET variations at all basins. The selection of multiple local optimal parameter sets mainly affects outputs at the Yampa River basin due to larger discrepancies in local optimal jh_coef values, which control PET (not shown), while forcing calibration data mostly affects the East and Animas River basins, which is consistent with the results displayed in Figure 5.2d. Although Yampa exhibits large difference in monthly runoff changes between R14 and the other two datasets (See Figure 5.3d), all the calibration datasets produce similar ET changes. This might be due to difference in hydroclimatic conditions between Yampa and the other two basins indicated by already higher dryness index in Yampa than the other basins in current climate scenarios (see Table 2.1). There are a number of reasons causing different water

balance changes seen in East and Animas. For example, large difference in ET change between X12 and M02 seen in Animas is likely to be caused by larger inter-forcing differences between M02 and X12 in high elevation especially in shortwave radiation, affecting the simulation of ET processes (Mizukami et al. 2014).

For SWE (Figure 5.5), a larger uncertainty in monthly changes is produced by model choice in all basins when comparing with parameter calibration choices (contrast column one with the others). From the inter-model comparison, PRMS produces the largest reduction of SWE under warming climate among the four models in all basins. Also note that bounds of uncertainty due to any calibration choice do not include other model traces of hydrologic change signal, indicating that, at least for SWE, model choice is more important than parameter estimation strategies explored through PRMS. Nevertheless, it should be noted that parameter sensitivity in PRMS is dominated by and pref_flow_den (i.e. parameter defining the preferential flow, Table A.1), while for the rest of the models snow parameters produced the largest variations in RMSE (not shown), which explains the smaller relative effect of calibration strategies compared to model structure. Additionally, monthly changes are quite consistent among different objective functions and multiple local optimal parameter sets, and the choice of forcing calibration dataset also produces uncertainty in monthly SWE changes, although to a lesser degree in comparison to model choice.

These seasonal characteristics in hydrologic changes are modified when analyzing effects of modeling decisions on soil moisture (Figure 5.6), as all of them produce considerable uncertainty, with the choice of model structure and objective function emerging as the decisions that introduce the most uncertainty in model simulations. In contrast to the SWE results, the choice of calibration options leads to a similar spread in monthly changes to that obtained from

model choice. This is again related to the parameters considered for calibration in PRMS (Table A.1), as pref_flow_den – which is the most sensitive parameter to RMSE – controls the fluxes of water throughout the soil column, specifically between preferential flow and gravity reservoirs.

5.3.3 Projected changes in catchment behavior

Finally, we compare the implications of modeling decisions on projected changes in signature measures of hydrologic behavior (defined in Table 2.4). Figure 5.7 displays differences between future (WRF-PGW) and current (WRF-CTRL) signature measures of hydrologic behavior for all basins, computed from different options for the following hydrologic modeling decisions: (a) choice of model structure, (b) selection of objective function used for calibration, (c) multiple local optimal parameters, and (d) choice of dataset used for calibration.

The results displayed in Figure 5.7 illustrate that the choice of model structure has the largest effect on projected changes of runoff ratio (RR) at Yampa and East River basins, while parameter estimation decisions provide more uncertainty for projected changes in RR at the Animas River basin. In the case of runoff seasonality (CTR), the largest uncertainty in projections comes from the choice of model structure, followed by forcing calibration data. When the analysis focuses on flashiness of runoff (FMS), the largest source of uncertainty is also model structure, followed by the choice of objective function and forcing calibration datasets. The relative importance of modeling decisions on changes in high flow volumes (FHV) is basin-dependent; for instance, while objective function has the largest effect at Yampa, the choice of forcing dataset is the most relevant source of uncertainty at Animas. One can also note that all modeling decisions introduce comparable uncertainties when examining low flows (FLV) and mid-range flow levels (FMM), although the VIC response is an outlier in the set of FLV projections at the Yampa River basin.

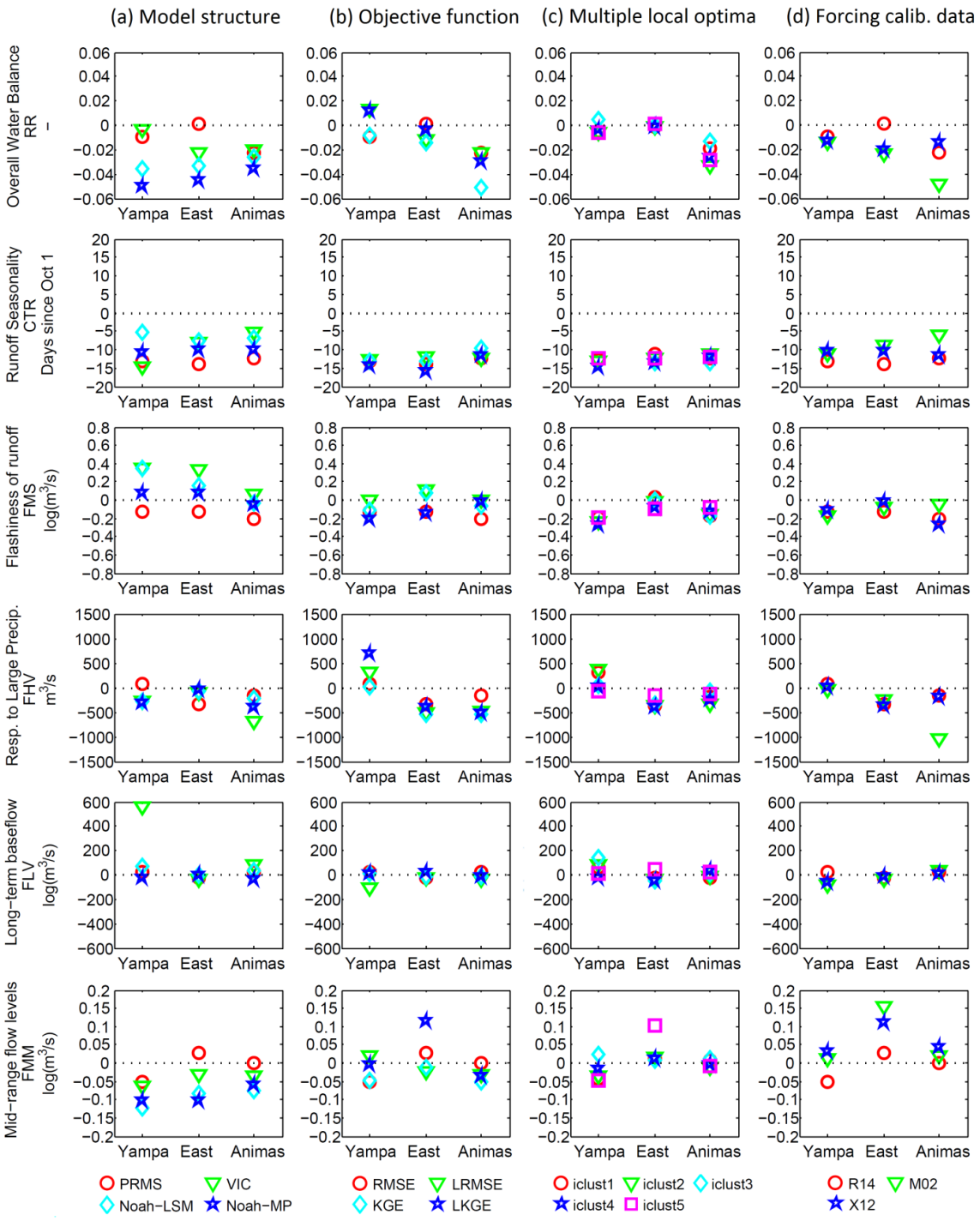

Figure 5.7: Projected changes in signature measures of hydrologic behavior over a six-average water year (Oct/2002 - Sep/2008). Results illustrate the uncertainty coming from: (a) choice of model structure, (b) selection of objective function used for calibration, (c) multiple local optimal parameters, and (d) choice of dataset used for calibration.

Overall, model choice has a larger implication on projected changes in runoff seasonality (CTR) and flashiness of runoff (FMS), while parameter estimation strategy has larger effects on high flow volumes (FHV), which is related with the large first order DELSA sensitivities obtained for the PRMS parameter pref_flow_den. Also, the relative importance of model choice versus parameter estimation decisions on projected changes in runoff ratio (RR) was found to be basin-dependent. Moreover, Figure 5.7 shows that, regardless of the uncertainty introduced, all modeling decisions examined here project earlier runoff volumes (CTR) in all basins. On the other hand, the choice of multiple options for a particular decision may switch the sign in projections for the rest of the signature measures analyzed.

In summary, these results show that uncertainties of climate change impacts on runoff characteristics must be carefully considered when performing particular hydrologic assessments. For example, depending on selection of models and calibration options, projected impacts on floods and droughts can be very different, with potential implications for water resources planning.

5.4 Conclusions

This chapter examines the implications of hydrologic modeling decisions on the portrayal of climate change impacts. Specifically, we analyze the relative effects of model structure choice and parameter estimation strategies on: (1) projected changes in the annual water balance, (2) projected changes in monthly-averaged model states and fluxes, and (3) projected changes in signature measures of catchment behavior. To this end, we compare hydrologic changes coming from four different model structures (PRMS, VIC, Noah and Noah-MP), for which parameters have been obtained using a common calibration strategy, with hydrologic changes coming from a single model structure (PRMS), with parameter sets identified using multiple options for

different calibration decisions (objective function, multiple local optima and calibration forcing dataset). In all experiments, the same current and future climate datasets were used to compute hydrologic changes. For the climate change scenario and basins examined here, the main findings of this study are:

- When comparing the effects of modeling decisions on the partitioning of precipitation into ET and runoff, the uncertainty introduced by the choice of objective function and forcing calibration data can be comparable or larger than that introduced by hydrologic model choice.
- Choices over model structure and parameter estimation strategy have large effects on the direction and magnitude of projected changes in annual water balance, although the direction of the signal was preserved across multiple local optimal parameter sets. Additionally, the relative effects of model structure and parameters were basin-dependent.
- Choices of model structure, objective function, and forcing calibration datasets introduce considerable uncertainty in monthly runoff changes. Moreover, uncertainty coming from parameter identification strategies may overwhelm that provided by model choice (e.g. forcing calibration at Yampa, objective function at East).
- Although model choice has a large effect on monthly changes of hydrologic processes for the climate scenario examined here, the choice of strategy for parameter estimation in a single model structure (PRMS) can introduce similar or larger uncertainties for some processes (e.g. ET at East River basin).
- In our experimental setup, the relative effect of modeling decisions on projected variations in hydrologic signature measures depends on the metric analyzed. While model

structure choice introduced the largest uncertainty in projections of runoff seasonality (CTR) and flashiness of runoff (FMS), parameter estimation decisions were more relevant for changes in high flow volumes (FHV). The relative effects of model choice and parameter estimation methods on projected changes in runoff ratio (RR) were found to be basin-dependent. All modeling decisions analyzed here had an important effect on baseflow (FLV) and mid-range flow levels (FMM).

- Parameter sets with similar performance scores, but located in different regions within the parameter space, may provide an opposite signal when projecting future catchment behavior (e.g. FHV, FLV and FMM at Yampa, FMS and FLV at East, FLV and FMM at Animas).

The main finding of this study is that decisions over configuring and calibrating hydrologic models can have pronounced effects on the portrayal of climate change impacts. Moreover, uncertainties associated with calibration decisions may overwhelm those coming from model choice, indicating that careful attention should be paid to parameter estimation strategies regardless of the hydrologic model structure selected. Finally, the results obtained here suggest that parameter uncertainty should be properly addressed in climate change studies, even if multiple hydrologic model structures are included, in order to avoid an over-confident portrayal of climate change impacts.

CHAPTER 6: Effects of regional climate model configuration and forcing scaling on projected hydrologic changes

6.1 Introduction

The strong evidence of ongoing hydroclimatic shifts in the Western United States (Barnett et al. 2005; Regonda et al. 2005; Mote et al. 2005; Stewart et al. 2005; Knowles et al. 2006; Hamlet et al. 2007) and their potential impacts on water availability – which relies heavily on snowpack accumulation during winter – remains a subject of active debate in the research community. Within this region, the Colorado River basin (CRB) constitutes one of the major water sources for consumption, irrigation and hydropower, among other uses, draining parts of seven states and Mexico, and covering the needs of over 30 million people. Given its strategic relevance, several studies have been conducted with the aim to quantify the effects that changes in precipitation and temperature might have on the hydrology of this area (e.g. Milly et al. 2005; Christensen and Lettenmaier 2007; Hoerling and Eischeid 2007; Ray et al. 2008; Hoerling et al. 2009; Rasmussen et al. 2011, 2014; Miller et al. 2011, 2012; Vano et al. 2012, 2014). Although there is general agreement on the broad picture (i.e. future drying due to the transition to a more arid climate), the wide range of magnitudes in hydrologic projections reflect the tremendous effects that methodological choices (e.g. emission scenario, global climate model, downscaling techniques, hydrologic models) may have on the portrayal of climate change impacts, making it very difficult for water managers the planning and implementation of adaptation strategies for design and safety of infrastructure (Raff et al. 2009), reservoir operation (Miller et al. 2012) or policies for instream flow requirements (Vano et al. 2014).

Despite the fact that the literature is quite rich in examples reflecting the increasing awareness for uncertainty characterization and quantification in the assessment of climate change

impacts (Bergström et al. 2001; Wilby and Harris 2006; Graham et al. 2007; Brekke et al. 2009; Ludwig et al. 2009; Chen et al. 2011; Najafi et al. 2011; Majone et al. 2012; Steinschneider et al. 2012; Surfleet and Tullos 2013; Vano et al. 2014), uncertainties associated with hydrologic modeling have traditionally been given less attention (Bastola et al. 2011). Among these, initial conditions (typically specified or estimated from observations), inputs (i.e. meteorological forcings), model structure and parameters have been highlighted as major error sources (Zehe and Blöschl 2004; Liu and Gupta 2007). In recent years, there has been a proliferation of climate change impact studies aimed to understand the role that the choice of hydrologic model structures (e.g. Boorman and Sefton 1997; Jones et al. 2006; Jiang et al. 2007; Ludwig et al. 2009; Bae et al. 2011; Miller et al. 2012; Vano et al. 2012; Surfleet et al. 2012), hydrologic parameter values (e.g. Cameron et al. 1999; Wilby 2005; Steele-Dunne et al. 2008; Surfleet and Tullos 2013), or the combination that both (e.g. Kay et al. 2009; Prudhomme and Davies 2009a,b; Bastola et al. 2011; Najafi et al. 2011; Poulin et al. 2011; Mendoza et al. 2015a) play on the overall uncertainty in hydrologic projections under modified climate scenarios.

Nevertheless, the interaction between meteorological input data, hydrologic model structures and parameters in hydrologic change estimates still remains unclear. Indeed, most studies looking at the effects of forcing data *configuration* (e.g. in-situ observations, empirical methods, numerical simulations or combinations of them) on hydrologic model simulations have been conducted on a retrospective basis (Flerchinger et al. 2009; Materia et al. 2010; Haddeland et al. 2012; Mo et al. 2012; Wagner et al. 2012; Bohn et al. 2013; Wayand et al. 2013; Elsner et al. 2014; Mizukami et al. 2014; Raleigh et al. 2014), with general agreement about the primary roles that precipitation, temperature and radiation estimates play in simulating hydrologic processes over mountainous regions. A second group of studies has put detail on the effects that

spatial resolution of meteorological forcings may have on hydrologic simulations (e.g. Yang et al. 2003; Liang et al. 2004; Shrestha et al. 2006, 2007; Vano et al. 2014), which is particularly critical when driving distributed hydrologic models. For instance, Shrestha et al. (2006, 2007) tested input forcing datasets with different spatial resolutions on a distributed hydrological model with fixed grid size and a fixed distribution of parameter values in space, concluding that a finer resolution in meteorological fields leads to better model performance. This conclusion seems to be particularly relevant for hydrologic modeling in high mountain regions (Hoerling et al. 2009), where the spatial resolution at which meteorological data is available may have enormous effects on the simulation of snowpack processes (Haddeland et al. 2012).

In this line, recent studies conducted in the Colorado Headwaters Region (Ikeda et al. 2010; Rasmussen et al. 2011, 2014) have explored the effects of forcing configuration using a dynamical downscaling technique. Specifically, they showed that the use of horizontal grid spacing of 6 km or less in a regional climate model (RCM) allows the accurate estimation of vertical motions driven by topography without the need to include a convective parameterization scheme, improving the representation of seasonal snowfall and snowpack. Moreover, Rasmussen et al. (2011, 2014) demonstrated the utility of the pseudo global warming (PGW) approach (Schär et al. 1996; Hara et al. 2008; Kawase et al. 2009) for quantifying climate change impacts on hydrology via coupled high-resolution land-atmosphere simulations. More recently, Mendoza et al. (2015a) examined hydrologic changes due to modified climate with four process-based hydrologic/land-surface models over three Colorado Headwater basins, using the 4-km WRF outputs obtained by Rasmussen et al. (2014) to force hydrologic simulations under current and modified climatic conditions.

This chapter builds on the previous work by Rasmussen et al. (2014) and Mendoza et al. (2015a) to examine how the configuration of forcing datasets generated from RCM simulations and the re-scaling of RCM outputs affect hydrologic change estimates. In particular, these effects are analyzed for: (i) historical simulation of signature measures of hydrologic behavior (e.g. runoff ratio, seasonality, log-term baseflow), and (ii) projected hydrologic change in terms of annual water balance and hydrologic signature measures. In order to characterize the interplay between forcing datasets and model structures, we include in our analyses four different hydrologic/land surface models (PRMS, VIC, Noah and Noah-MP). The remaining of this chapter is organized as follows: the approach is described in Section 6.2; results and discussion are presented in Section 6.3, and the main conclusions are summarized in Section 6.4.

6.2 Approach

6.2.1 Meteorological forcings

Meteorological data from WRF simulations is available at hourly time steps and three different configurations (4-, 12-, and 36-km resolutions), for both historical and PGW conditions during the period October/2000 - September/2008. The variables and temporal disaggregation used depend on specific hydrologic model requirements (further details in Chapter 2).

In this chapter, WRF outputs are used to compare the effects that three different RCM configurations (specifically, grid size and convective parameterization) may have on hydrologic simulations performed using different hydrologic model structures. To compare the effects of WRF configuration – referred as *experiment 1* – with those from re-scaling high-resolution outputs, we create two additional forcing datasets by re-gridding WRF outputs obtained at 4-km to the 12- and 36-km grid cells used by Rasmussen et al. (2014) – referred as *experiment 2*. This is done in two steps: (1) identification of all the 4-km grid points contained in WRF grid cells at

12- and 36-km resolutions, and (2) computation of the new forcing data by spatially averaging the 4-km points contained in the 12- and 36-km grid cells.

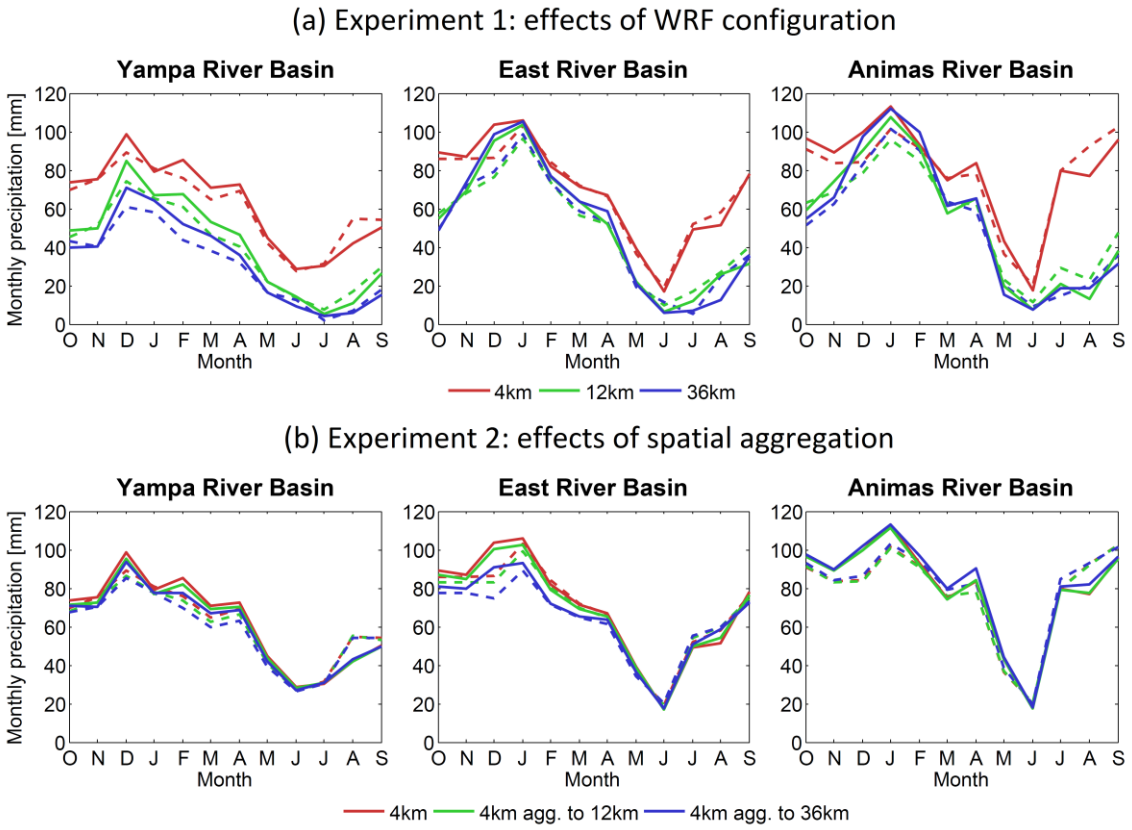

Figure 6.1: Basin-averaged monthly precipitation values for current (CTRL, dashed lines) and future (PGW, solid lines) WRF outputs used in (a) experiment 1 (effects of WRF configuration) and (b) experiment 2 (effects of spatial aggregation), for period Oct/2002 - Sep/2008.

Figure 6.1 includes basin-averaged monthly precipitation values computed from: (a) WRF outputs obtained by Rasmussen et al. (2014) with three different configurations (4-, 12- and 36-km), and (b) 4-km WRF outputs, re-scaled to 12- and 36-km horizontal resolutions. These results correspond to the period October/2002 – September/2008, for current and future climate scenarios. It can be noted that discrepancies in WRF configuration (i.e. differences in model grid size and convective parameterization) affect precipitation amounts heavily (Figure

6.1a). Indeed, 4-km WRF simulations generate much more precipitation in all basins, especially during summer, followed by 12- and 36-km. Moreover, PGW simulations project a decrease in precipitation over fall/winter at all basins regardless of the configuration used, and a general increase during summer months. On the other hand, scaling effects (Figure 6.1b) on monthly precipitation amounts are minor compared to those coming from WRF configuration. At the Yampa and East River basins, smaller monthly precipitation amounts are obtained during October-March for coarser gridded datasets (i.e. 12- and 36-km), and 36-km datasets provide more precipitation during February-April at the Animas River basin.

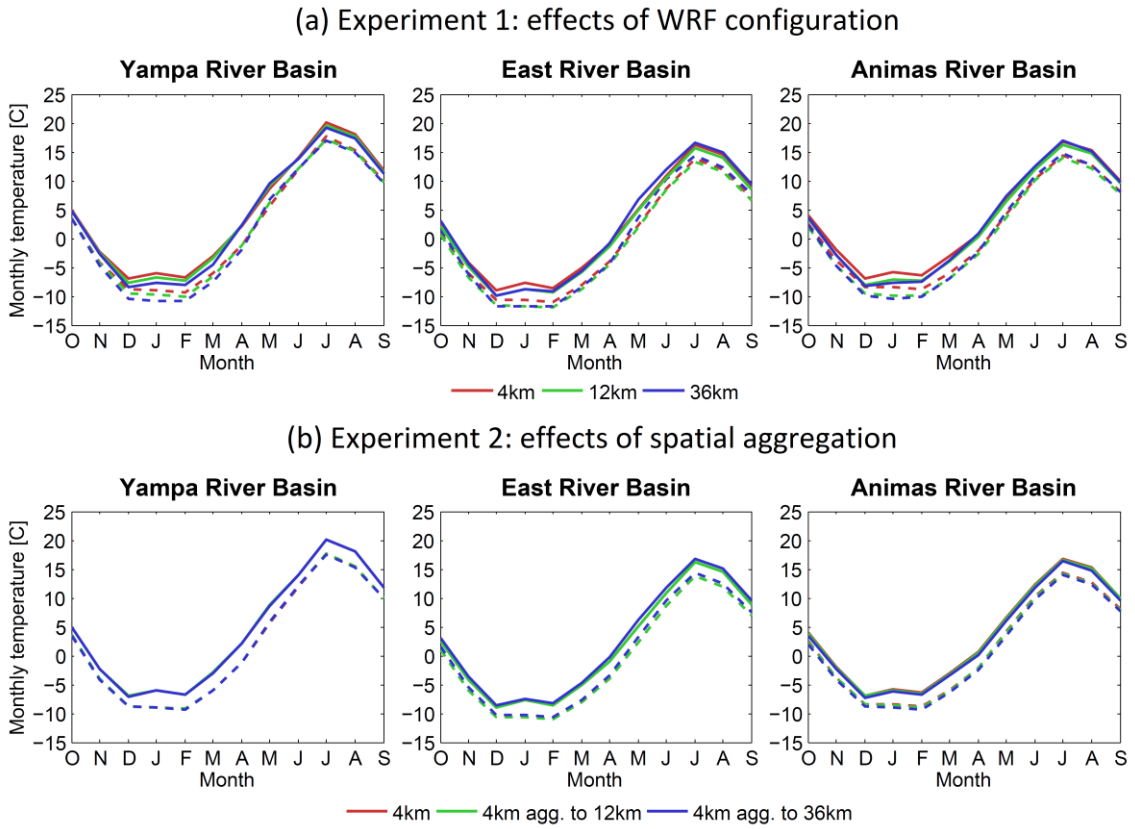

Figure 6.2: Same as in Figure 6.1, but for basin-averaged monthly temperature.

Figure 6.2 displays basin-averaged monthly temperatures computed for experiments 1 (Figure 6.2a) and 2 (Figure 6.2b). There are not substantial differences across different WRF configurations, although 4-km WRF simulations provide higher temperature for both CTRL and

PGW scenarios during December-March. Finally, it can be noted that scaling impacts on basin-averaged monthly temperatures are minimal.

6.2.2 Hydrologic/land surface models

To explore the interaction between meteorological datasets obtained from WRF and the selection of hydrologic model structure, we choose four hydrologic/land surface models: the US Geological Survey's Precipitation Runoff Modeling System (PRMS; Leavesley et al. 1983; Leavesley and Stannard 1995), the Variable Infiltration Capacity model (VIC; Wood et al. 1992; Liang et al. 1994, 1996) the Noah Land Surface Model (Noah-LSM; Ek 2003; Mitchell et al. 2004) and the Noah Land Surface Model with Multiple Parameterizations (Noah-MP; Niu et al. 2011; Yang et al. 2011). Further details on inter-model differences, information requirements and model setup are included in Chapter 2.

6.2.3 Experimental design

In order to assess the effects of WRF configuration and spatial scaling on the portrayal of climate change impacts, we force hydrologic model simulations under historical (CTRL) and modified climate (PGW) scenarios for the following cases:

- Experiment 1: hydrologic model simulations are conducted using 4-km, 12-km and 36-km WRF outputs obtained by Rasmussen et al. (2014).
- Experiment 2: hydrologic model simulations are conducted using 4-km WRF outputs, and two additional datasets obtained from re-scaling 4-km outputs to the original 12- and 36-km WRF grid cells used by Rasmussen et al. (2014).

All model simulations are carried out for the period between October 1, 2000 and September 30, 2008, using the first two years as warming up to initialize model states. We

compute hydrologic changes using the same parameter sets obtained by Mendoza et al. (2015a), i.e. calibrated by minimizing the root mean squared error (RMSE) between observed and simulated daily streamflow (period October 1, 2002 to September 30, 2008) with the Shuffled Complex Evolution (SCE-UA) algorithm (Duan et al., 1992, 1993), using 4-km resolution WRF historical datasets.

The purpose of these experiments is to examine the implications of forcing datasets developed at different spatial resolutions. Therefore, we decide to fix the spatial resolution of hydrological models (grid size) to 4-km – identical to the grid used in the 4-km WRF simulations performed by Rasmussen et al. (2014) – to isolate effects of forcing configuration. Hence, when hydrologic model simulations are forced with 12- and 36-km meteorological datasets, meteorological variables are distributed to hydrologic model grid cells using a nearest neighbor interpolation method as in Shrestha et al. (2006). Hydrologic changes are computed for the period Oct/2002 - Sep/2008 by forcing all hydrologic models with the same current and future WRF datasets.

6.3 Results and discussion

6.3.1 Model performance

We first analyze how hydrologic model performance is affected by WRF configuration (Figure 6.3a) and forcing re-scaling (Figure 6.3b) over the period October/2002 - September/2008. Hence, model fidelity (i.e. accuracy in process representation) is assessed by computing differences between simulated (control, CTRL) and observed (Obs) values of signature measures of hydrologic behavior. The suite of metrics include the runoff ratio (RR), center of time of runoff (CTR), flashiness of runoff (FMS), and low flow volumes (FLV). In Figure 6.3, each evaluation metric is displayed in a different row, hydrologic model structures

are represented by different symbols, and different colors depict different WRF configurations (Figure 6.3a) or spatial forcing resolutions (Figure 6.3b). Therefore, differences between symbols holding different colors indicate the magnitudes of effects of RCM configuration and/or scaling on hydrologic model fidelity.

The results from experiment 1 (Figure 6.3a) show that decreased precipitation amounts obtained with coarser resolutions (12- and 36-km) in WRF (see Figure 6.1a) translate into a degradation in simulated runoff ratios (RR) at all basins. When looking at the center of time of runoff (CTR), however, there is no dependence between the selected WRF configuration (i.e. coarser resolution) and performance, although the use of 12- and 36-km WRF outputs clearly introduce more uncertainty in the multi-model ensemble when comparing to 4-km WRF simulations. The latter effect can also be observed when looking at model fidelity in terms of flashiness of runoff (FMS) and long-term baseflow (FLV). Note that the FLV result obtained from VIC and 36-km WRF for the Yampa River basin (i.e. blue triangle) has been omitted in Figure 6.3a to allow a better visualization and comparison between experiments 1 and 2 (since CTRL – Obs > 2000 log(m^3/s)). With the exception of runoff ratio (RR), enhancement/degradation in performance obtained from switching WRF configurations depends on model/basin.

As shown in Figure 6.3b, the effects of forcing scaling on model fidelity exhibit similar pattern to those from WRF configuration but generally at a much smaller degree for runoff ratio (RR), flashiness of runoff (FMS), and baseflow processes (FLV) throughout all models and basins. However, the representation of runoff seasonality (CTR) is still considerably affected in some models/basins (e.g. Noah-LSM and Noah-MP at Yampa and East, PRMS at East). This might be attributed to the smoothing effect of spatial aggregation on high precipitation/snowmelt

events simulated by 4-km WRF simulations, affecting the historical simulation of high daily runoff events and therefore the computation of center of timing (CTR).

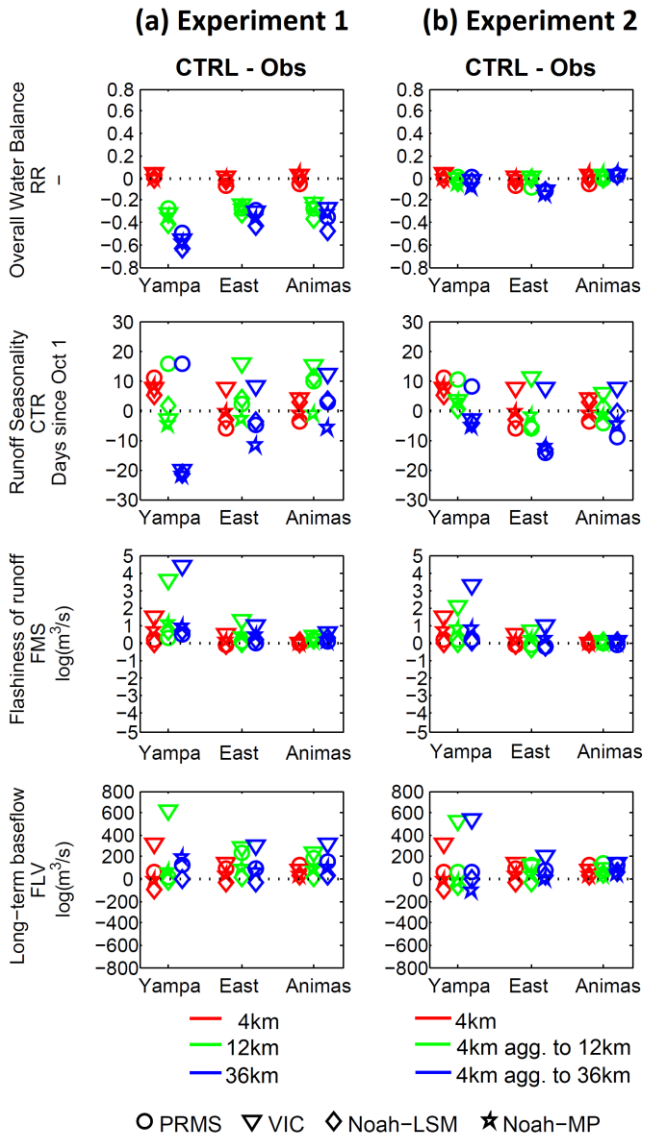

Figure 6.3: Difference between simulated (CTRL) and observed (Obs) signature measures of hydrologic behavior (period Oct/2002 - Sep/2008) obtained from various hydrologic model structures (i.e. different symbols) and forcing datasets (i.e. different colors). Results are displayed for (a) experiment 1 (effects of WRF configuration) and (b) experiment 2 (effects of spatial aggregation).

Figure 6.4 illustrates how hydrologic signature measures obtained from coarse resolution datasets (12- and 36-km) compare with those computed using high-resolution WRF outputs under current climate. The results obtained in experiments 1 and 2 are displayed in Figures 6.4a and 6.4b, respectively, where each column represents a specific signature measure, and each row contains signature measures computed with 12- and 36-km grid spacing, versus metrics obtained with 4-km WRF datasets (i.e. baseline dataset). Results from experiment 1 (Figure 6.4a) demonstrate that while the use of 12- and 36-km WRF outputs has large effects on the simulation of runoff ratio (RR), the implications on other metrics depend on the model structure and/or the basin analyzed. For example, 12-km WRF datasets translate into increases in the center of time of runoff (CTR) when using PRMS, and decreases in the same metric if the model is Noah-MP. Similar effects are observed if hydrologic simulations are forced with 36-km WRF outputs. The effects of WRF configuration on flashiness of runoff (FMS) and long-term baseflow (FLV) tend to be basin-dependent for each model with the exception of VIC, for which larger signature values are simulated in all basins – especially Yampa – when using 12- and 36-km WRF outputs.

The results displayed in Figure 6.4b show that the effects of re-scaling forcing datasets are less pronounced than those from WRF configuration, but still important for some signature measures. Overall, re-gridding 4-km WRF outputs to 12- and 36-km horizontal resolutions generates a reduction in simulated runoff ratio (RR) and center of time of runoff (CTR). Scaling effects translate into very small decreases in flashiness of runoff (FMS) across models/basins with a few exceptions (Yampa when forcing VIC with 12-km and 36-km datasets, and East when forcing VIC with 36-km WRF). Finally, scaling effects on long-term baseflow (FLV) are more pronounced when the forcing grid size is 36-km, translating into increases in FLV with VIC and Noah-LSM, and decreases of the same metric with PRMS and Noah-MP at all basins.

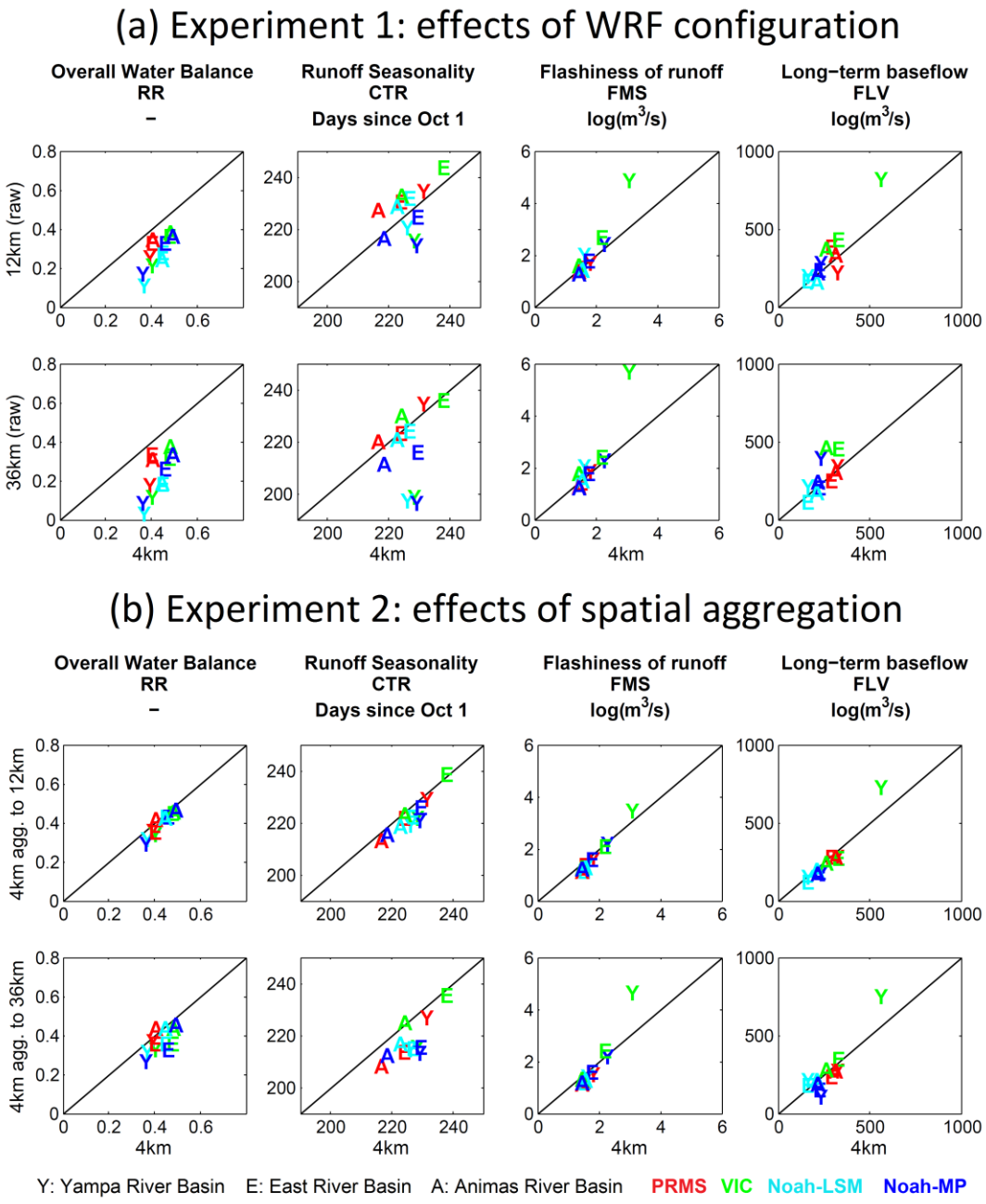

Figure 6.4: Impact of (a) WRF configuration (experiment 1) and (b) spatial aggregation of WRF 4-km resolution datasets on simulated hydrologic signature measures. Each column contains results for a specific metric, while different rows contain outputs from 12-km and 36-km (y axis) versus model outputs using WRF datasets with 4 km horizontal grid space (x axis). In each panel, different letters represent basins and different colors depict results from various hydrologic models (see legend for details).

6.3.2 Changes in annual water balance

In this section, we analyze and compare the effects of WRF configuration (experiment 1) and spatial forcing aggregation (experiment 2) on the partitioning of precipitation into ET and runoff under current and future climate scenarios. In each panel of Figure 6.5, the diagonal lines represent basin-averaged mean annual precipitation for current and future climate scenarios over a 6-year average period (Oct/2002 - Sep/2008). The intersection of these lines with the x-axis indicates that all precipitation becomes runoff, while the intersection with the y-axis indicates that the system converts all precipitation into ET. In each panel, different symbols depict outputs coming from different hydrologic model structures for current climate (unfilled) and future climate (solid) and color differentiate the spatial resolutions. A symbol located exactly on the diagonal lines represents a simulation with negligible changes in storage over the 6-year simulation period (i.e. P = ET + R), whereas symbols located below the 1:1 line denote increases in storage, and those above denote decreases in storage. Inter-model differences in precipitation partitioning are represented by the distance between different symbols (unfilled or solid), while the distance between a particular symbol (e.g. star for Noah-MP) for current (unfilled) and future (solid) climate scenarios represents the hydrologic change signal. It should be noted that uncertainty arising from model choice is represented by the dispersion of symbols holding the same color around the precipitation (diagonal) line.

The results obtained from experiment 1 (Figure 6.5a) show that 4-km WRF simulations generate the largest precipitation amounts under current (CTRL) and future (PGW) climate scenarios at all basins, followed by 12- and 36-km WRF outputs. Inter-model differences in the partitioning of precipitation into runoff and ET are still considerable after calibration (Mendoza et al. 2015a) and tend to increase when hydrologic simulations are forced with 12- and 36-km

WRF datasets. On the other hand, the effects of forcing re-scaling (Figure 6.5b) on basin-averaged annual precipitation are reduced compared to those of WRF configuration. Interestingly, forcing re-scaling can increase inter-model differences (i.e. larger dispersion of symbols when 4-km WRF are re-gridded to 12- and 36-km grid spacing) in precipitation partitioning under current and future climate (e.g. Yampa River basin).

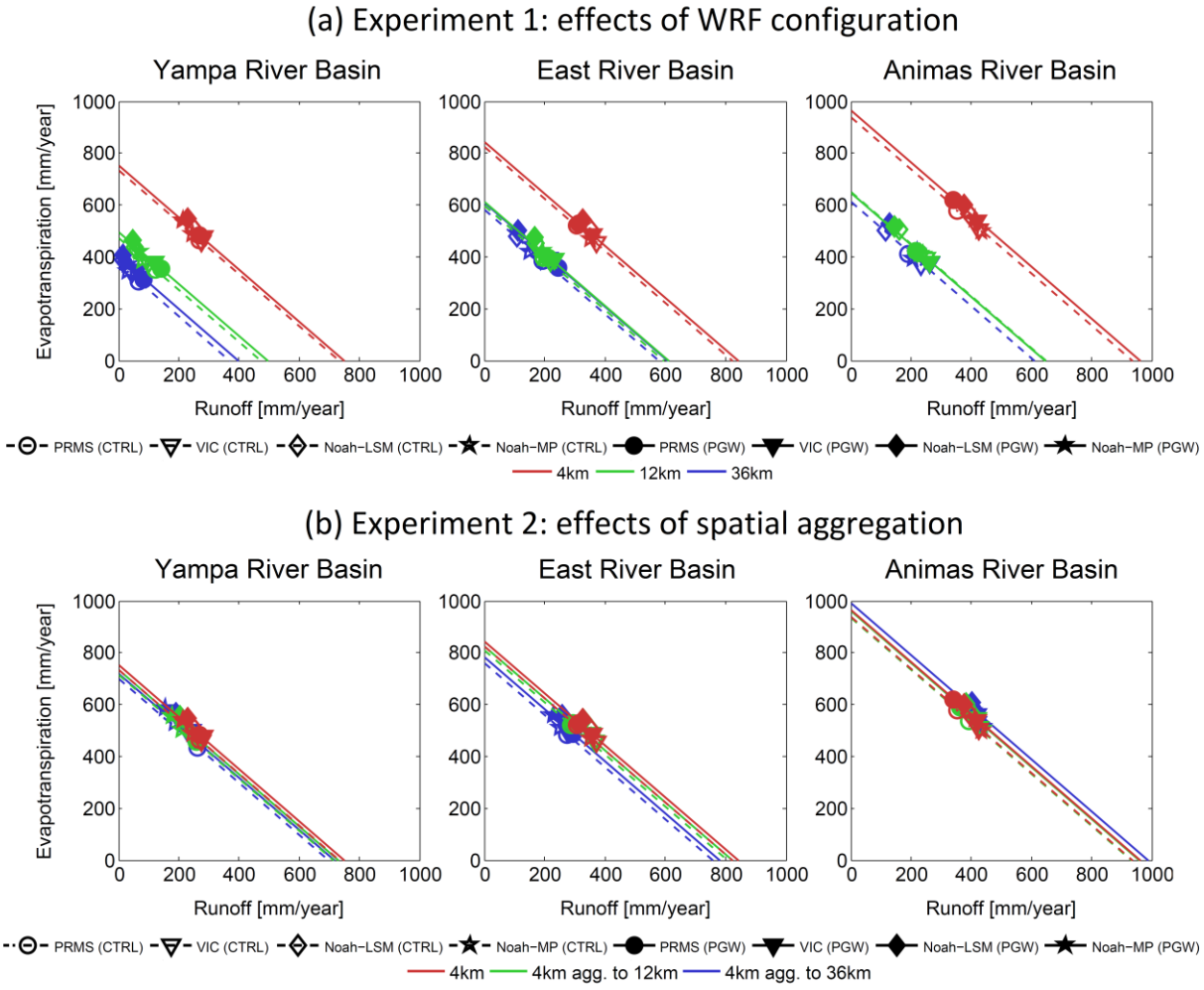

Figure 6.5: Partitioning of current (CTRL) and future (PGW) basin-averaged mean annual precipitation (diagonal, mm/year) into basin-averaged mean annual runoff (x axis, mm/year) and evapotranspiration (y axis, mm/year) obtained from various model structures (i.e. different symbols) and forcing datasets (i.e. different colors) for the period Oct/2002 - Sep/2008. Results are displayed for (a) experiment 1 (effects of WRF configuration) and (b) experiment 2 (effects of spatial forcing aggregation).

How do WRF configuration and forcing re-scaling affect the portrayal of climate change impacts in annual water balance? To answer this question, we compute projected changes in basin-averaged mean annual runoff and ET (Figure 6.6) for each basin (displayed in different columns). In each panel, the dispersion of the same symbol (e.g. triangle for VIC) holding different colors across the ΔRunoff – ΔET space (with Δ representing the difference between future and current climate scenarios) represents the uncertainty introduced by the choice of WRF configuration (Figure 6.6a) or spatial resolution in forcing datasets (Figure 6.6b). Similarly, the dispersion of different symbols holding the same color (e.g. red for 4-km WRF datasets) illustrates the uncertainty associated with hydrologic model choice.

Figure 6.6a reveals that the choice of WRF configuration has large effects on hydrologic changes projected through different hydrologic model structures. These effects are reflected in the *magnitude* (i.e. distance from each symbol to the point ΔRunoff = 0 mm/yr, ΔET = 0 mm/yr) and *direction* (i.e. quadrant in which symbols are located, indicating increase/decrease of runoff and ET) of projected changes in mean annual runoff and mean annual ET obtained with each hydrologic model. Moreover, the dispersion provided by different WRF configurations through a single model structure may be comparable or even larger than that obtained from multiple model structures forced with a unique WRF dataset (e.g. PRMS and Noah-MP simulations at East with calibrated hydrologic model simulations forced with 12-km WRF datasets).

The comparison between Figures 6.6a and 6.6b highlights that the effects of forcing scaling on projected changes in the annual water balance are smaller than those coming from WRF configuration. Indeed, the direction of hydrologic change is mostly preserved when forcing the same hydrologic model with re-scaled datasets (Figure 6.6b). In opposition to the results obtained from experiment 1, the uncertainty provided by model choice forced by any re-scaled

dataset is much larger than the uncertainty provided by the choice of dataset throughout a unique hydrologic model structure. Additionally, re-scaling forcing inputs to coarser resolutions can enhance inter-model differences (i.e. dispersion of symbols holding the same color) in hydrologic change.

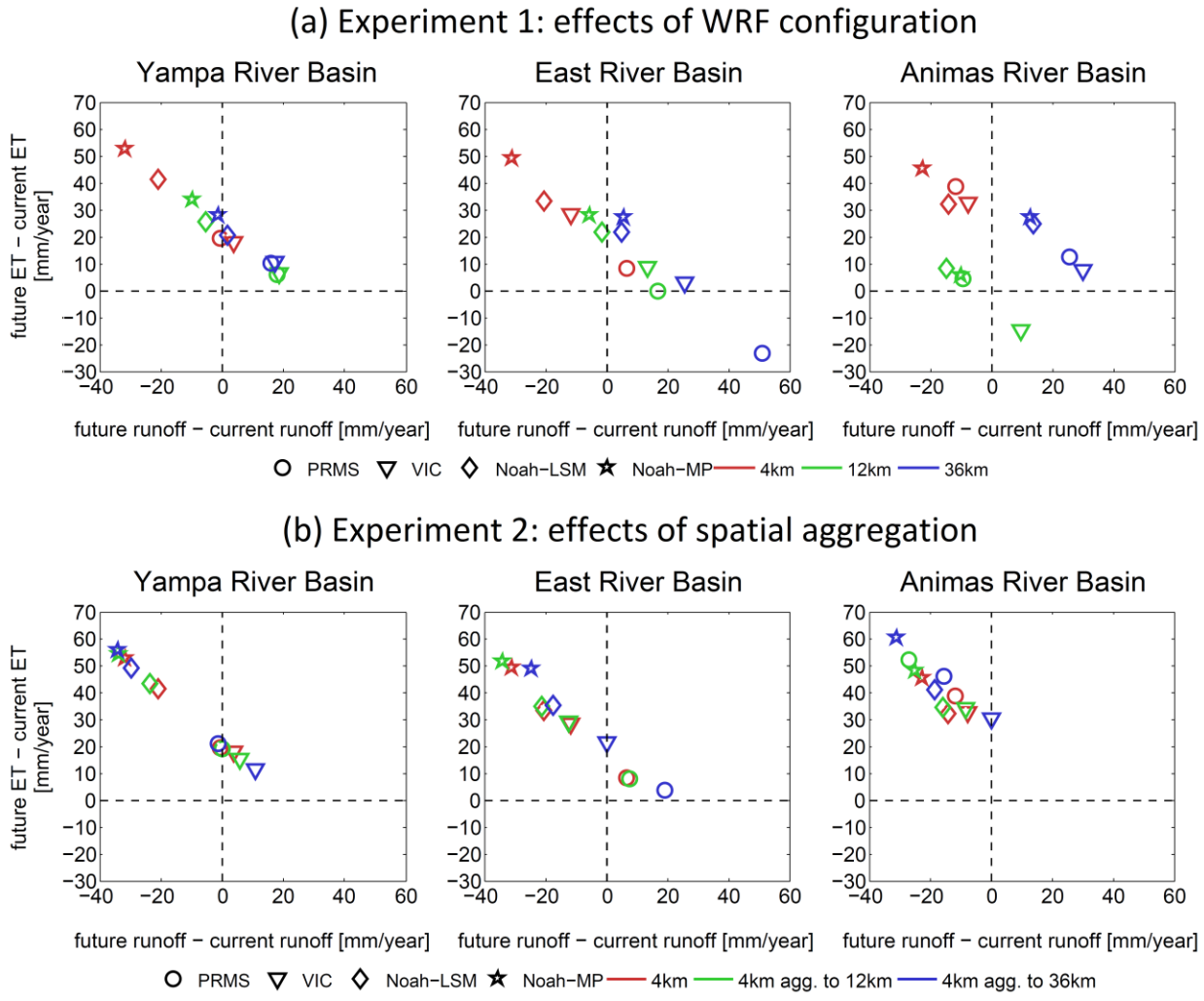

Figure 6.6: Projected changes in basin-averaged mean annual runoff (x axis, mm/year) and evapotranspiration (y axis, mm/year) obtained from various model structures (i.e. different symbols) and forcing datasets (i.e. different colors) for the period Oct/2002 - Sep/2008. Results are displayed for (a) experiment 1 (effects of WRF configuration) and (b) experiment 2 (effects of spatial forcing aggregation).

6.3.3 Projected changes in catchment behavior

Finally, we compare the effects of WRF configuration and forcing re-scaling on projected changes in hydrologic signature measures across multiple model structures (Figure 6.7). The results from experiment 1 (Figure 6.7a) show that the only consistent projection obtained with all WRF configurations is a decrease in the center of time of runoff (CTR), or earlier annual peak flow, under the future climate scenario, although the magnitudes obtained can be different depending on the hydrologic model selected. On the other hand, the use of coarser grid sizes (i.e. 12- and 36-km) and a convective parameterization in WRF generally translate into increased projected changes (i.e. PGW – CTRL) in runoff ratio (RR), which sometimes produce a switch in the sign (i.e. from negative to positive values) of projected variations. When looking at flashiness of runoff (FMS), the effects of WRF configuration on projected changes obtained with a specific hydrologic model structure depend on the basin analyzed. For example, the choice of WRF configuration has little effects on projected changes in FMS obtained with Noah-LSM at Yampa, Noah-MP at East and PRMS at Animas, but large implications in projections for the rest of models/basins. The results for low flow volumes (FLV) show that WRF configuration mostly affects the direction and magnitude of projections obtained with VIC, especially at the Yampa River basin.

According to Figure 6.7b, while the effects of forcing re-scaling on projected changes in runoff ratio (RR) are smaller than those from WRF configuration (Figure 6.7a), they can still switch the sign (e.g. VIC simulations at the Yampa River basin) and magnitude of projections. Although forcing scaling has very minor effects on changes in runoff seasonality (CTR) across all models, it might affect both the magnitude and direction of projections in flashiness of runoff, or FMS (e.g. VIC and Noah-MP at the Yampa and East River basins). Finally, scaling effects on

projected changes in low flow volumes (FLV) are generally smaller than those from WRF configuration, and are mostly reflected in VIC simulations.

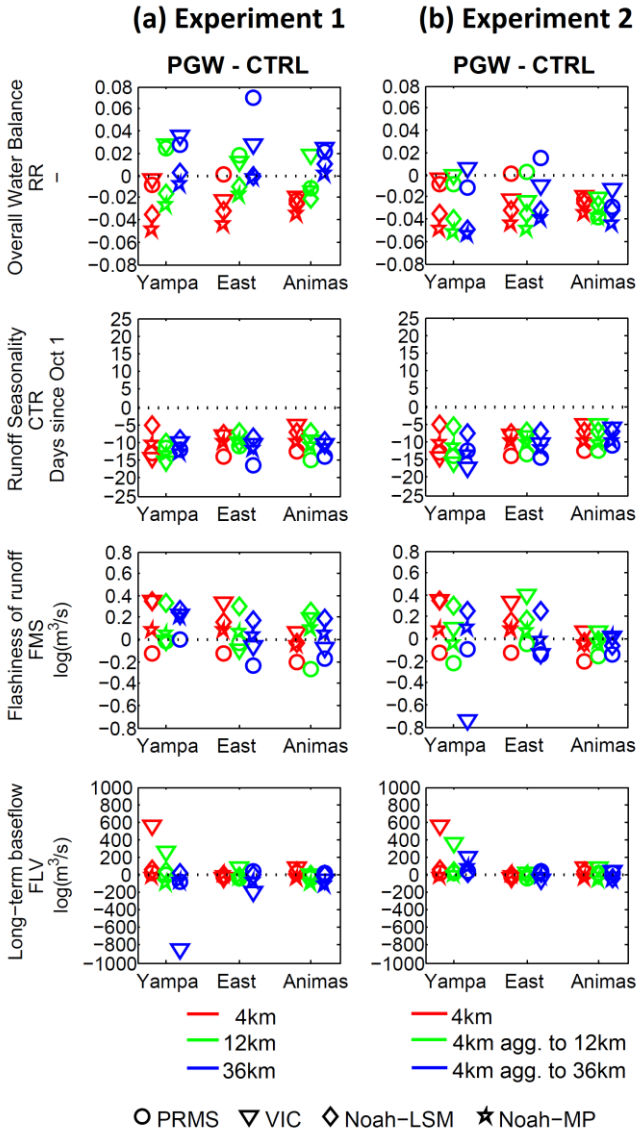

Figure 6.7: Difference between future (PGW) and current (CTRL) simulated signature measures of hydrologic behavior, obtained from various hydrologic model structures (i.e. different symbols) and forcing datasets (i.e. different colors) over a six-average water year (Oct/2002 - Sep/2008). Results are displayed for (a) experiment 1 (effects of WRF configuration) and (b) experiment 2 (effects of spatial forcing aggregation).

6.4 Conclusions

We investigated the implications that regional climate model (RCM) configuration and output re-scaling may have on the portrayal of climate change impacts. Specifically, we assessed the effects of the above decisions on: (i) historical performance in terms of hydrologic signature measures, and (ii) hydrologic changes due to a climate perturbation, with focus on the annual water balance and catchment processes. The analyses were conducted in three catchments located in the headwaters of the Colorado River basin. To explore the interplay between forcing effects and hydrologic model choice, we include four model structures, whose parameters were calibrated against observed runoff using 4-km WRF historical datasets (Mendoza et al. 2015a).

As illustrated by Rasmussen et al. (2014), the choice of WRF configuration (i.e. model grid size and inclusion of convective parameterization) has large effects on simulated precipitation amounts. Specifically, the use of 12- and 36-km grid spacing and a convective parameterization results in the underestimation of basin-averaged annual precipitation totals with respect to 4-km WRF simulations. Therefore, it was found that WRF configuration has larger effects on the historical simulation of hydrologic signature measures in comparison to those provided by forcing re-scaling. However, re-scaling effects on runoff seasonality (CTR) are still considerable due to the attenuation of local high precipitation/snowmelt events.

The water balance analysis revealed that WRF configuration has tremendous effects on the portrayal of hydrologic change at an annual basis (i.e. variations in mean annual runoff and ET), regardless of the hydrologic model structure selected. Moreover, the effects of WRF configuration on hydrologic change may overwhelm the uncertainty from model choice, which surpasses the uncertainty from re-scaled forcings. It was also found that re-scaling forcing

datasets to coarser resolutions may augment inter-model differences in precipitation partitioning and projected changes in runoff and ET.

Forcing scaling effects on projected changes in hydrologic signature measures were found to be generally smaller than those coming from WRF configuration. However, the use of coarser forcing resolutions may translate into a switch in the sign of changes projected by a particular hydrologic model structure (e.g. runoff ratio, flashiness of runoff). Even more, it was found that scaling effects may surpass those associated with WRF configuration when projecting variations in hydrologic behavior (e.g. flashiness of runoff).

The main conclusion from this study is that RCM configuration has enormous implications on the magnitude and direction of hydrologic change signal. Moreover, the results presented here illustrate the strong interplay between meteorological forcings and hydrologic model structures. In order to avoid an over-confident portrayal of climate change impacts, future studies should incorporate an integrated characterization and quantification of the different sources of uncertainty in hydrologic modeling, with particular emphasis on inputs, model structure and parameters.

CHAPTER 7: Towards a process-based assessment of the sensitivity of water resources to climate variability and change

7.1 Introduction

Understanding and predicting the water cycle at multiple spatial and temporal scales has been a key challenge for humanity. We have seen how relatively recent, but clear evidence of ongoing shifts in hydroclimatic variables (e.g. Barnett et al. 2005, 2008; Regonda et al. 2005; Mote et al. 2005; Stewart et al. 2004, 2005; Knowles et al. 2006; Hamlet et al. 2007; Cayan et al. 2010; Pierce et al. 2008; Das et al. 2009; Hidalgo et al. 2009) has increased the awareness of the impact of climate variability and change on water resources, especially considering that an important fraction of the observed trends are human induced.

Given the above problem, the literature has seen a plethora of studies aimed to generate reliable projections of changes in hydrology at the catchment, regional and global scales due to future shifts in precipitation and temperature. Nevertheless, the little agreement in results – particularly runoff – from different studies conducted in the same domain (Vano et al. 2014; Mendoza et al. 2015a) suggests that uncertainties in projections arise from the generally subjective nature of methodological choices. In this chapter we discuss the existing paradigms used in the assessment of climate change impacts, illustrate how different hydrologic modeling decisions can generate an over-confident portrayal of climate change impacts, and make some recommendations to improve our understanding of the sensitivity of the water cycle to climate variability and change.

7.2 Paradigms for the assessment of climate change impacts

There are two main complementary perspectives on climate risk assessment for adaptation (Wilby et al. 2009; Wilby and Dessai 2010): (i) the 'top-down' (or scenario-led) approach – also known as the 'uncertainty cascade' –, and (ii) the 'bottom-up' (or vulnerability-based) approach. The 'top-down' approach relies heavily on the generation of scenarios from different methodological choices (e.g. greenhouse gas emission scenarios, global climate models, downscaling methods, hydrologic models); therefore, implementing this paradigm involves quantifying the uncertainty at each step of the modeling process, resulting in a final envelope for the projected variable(s) of interest (e.g. Wilby and Harris 2006; Chen et al. 2011; Addor et al. 2014). On the other hand, 'bottom-up' methods focus on identifying and reducing the vulnerability to known climate variability, and hence do not necessarily require climate change scenarios (e.g. Lempert et al. 2004; Naylor et al. 2007). In view of the possible mismatch between the outputs generated by 'top-down' frameworks and the needs of decision-makers (e.g. Hallegatte 2009; Romsdahl and Pyke 2009), some authors have argued for combining the strengths of the scenario-led and vulnerability-based approaches, by using climate change projections at a later stage to inform the assessment of vulnerability (e.g. Prudhomme et al. 2010; Wilby and Dessai 2010; Brown et al. 2012).

From the paradigms listed above, the 'top-down' approach has been the primary method in the assessment of climate change impacts on hydrological processes, demonstrating to be particularly useful for identifying the main sources of uncertainty (e.g. Chen et al. 2011; Addor et al. 2014). However, this approach has two main problems: (1) limited practicality (Chen et al. 2011), and (2) little attention to hydrologic process understanding (Bastola et al. 2011). In our opinion, it cannot be expected from every climate change impact study to conduct thousands of

model simulations – as the 'uncertainty cascade' implicitly suggests – to cover the full range of possible uncertainties. Further, the 'top-down' approach addresses the problem of hydrologic model uncertainties from a statistical perspective – i.e. include as many different model structures and parameter sets as possible – rather than a process-based perspective.

In view of such problems, we believe that adopting a *process-based paradigm* is crucial to improve understanding of hydrologic sensitivity to climate variability and change. Implementing this vision requires accepting that some climate impacts (e.g. increases in temperature) can be predicted – also referred as *hard facts* (Blöschl and Montanari 2010) – while using expert knowledge to quantify the effects of hydrologic modeling decisions on individual simulated processes (e.g. runoff efficiency, snowmelt seasonality, soil moisture). This will help to understand *why* certain changes for a hydrological variable are projected, rather than focusing on the magnitude of the changes (Blöschl and Montanari 2010).

7.3 An example of the effects of hydrologic modeling approaches on the portrayal of climate change impacts

In this section, we summarize the main results obtained in chapters 4, 5 and 6. These results are presented in a way that allows visualizing the implications that different hydrologic modeling decisions have on projected changes in the overall water balance, i.e. changes in mean annual runoff and mean annual evapotranspiration (ET).

7.3.1 Effects of hydrologic model choice and calibration

Figure 7.1 shows the standard deviation of projected changes in mean annual runoff and mean annual ET, computed from the multi-model ensemble using default parameter values (red bars) and calibrated parameter values (blue bars). Model calibration was conducted with the

Shuffled Complex Evolution (SCE-UA) algorithm (Duan et al. 1992, 1993) by minimizing the root mean squared error (RMSE) between observed and simulated daily streamflow for the period between October 1, 2002 and September 30, 2008. These results demonstrate that a better match between simulated and observed runoff – achieved through traditional single-objective calibration – does not necessarily improve inter-model agreement in projected changes of the annual water balance. Indeed, the Animas River is the only basin where calibration makes a substantial contribution to decrease uncertainty in projections (Mendoza et al. 2015a). Further, it was obtained that different models can yield to the same results in projected runoff changes, but due to a combination of very different reasons (e.g. different changes in ET, soil moisture).

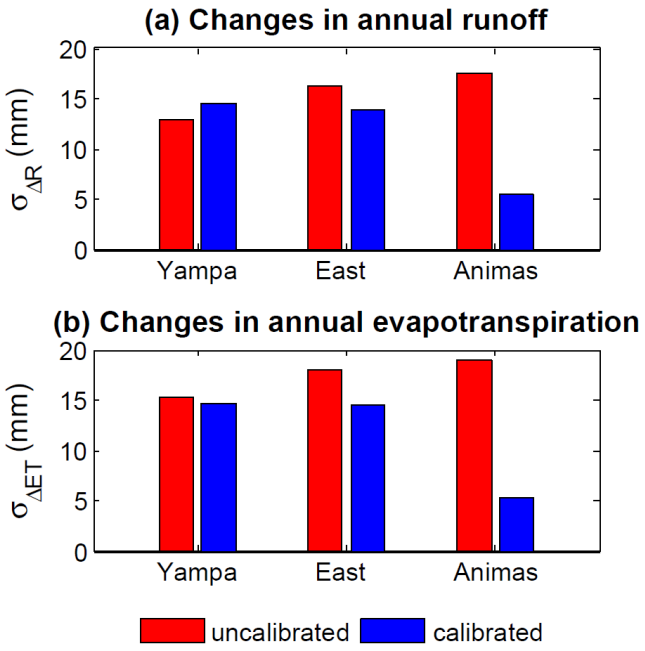

Figure 7.1: Standard deviation of projected changes obtained with four different hydrologic models using uncalibrated (red) and calibrated (blue) model parameter values. Results are displayed for mean annual runoff (top) and mean annual ET (bottom) for the period October/2002 – September/2008.

7.3.2 Comparing the effects of subjective model choice and parameter estimation strategies

In view of the above results, further experiments were conducted in order to compare the effects of model choice with those from parameter estimation strategies (Mendoza et al. 2015c). Figure 7.2 displays the standard deviations of projected hydrologic changes (as in Figure 7.1) coming from four different hydrologic models calibrated with the same strategy (red bars), and those computed with a single model structure (PRMS) using parameter values obtained with four different objective functions (blue bars), five local optimal parameter sets (green bars), and three different forcing datasets (cyan bars).

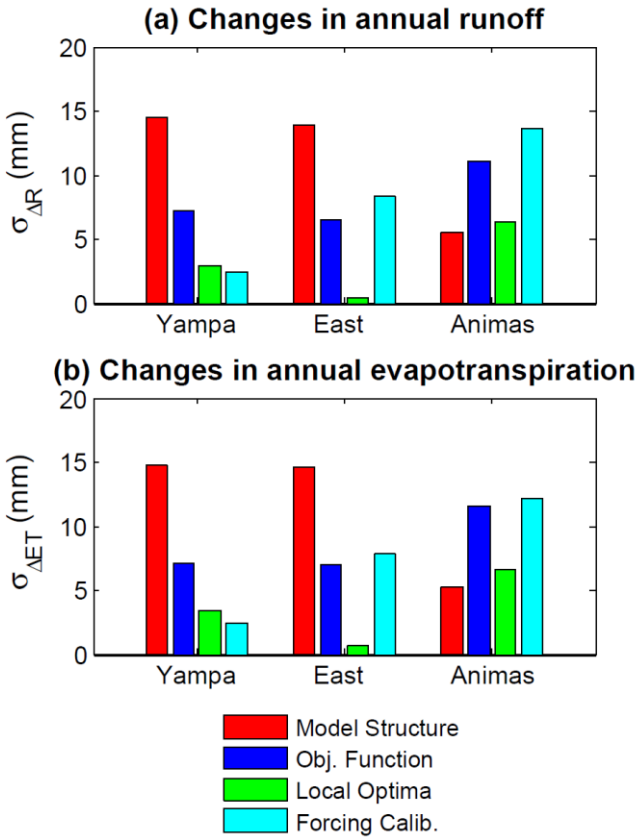

Figure 7.2: Standard deviation of projected changes in mean annual runoff (top) and mean annual ET (bottom) for the period October/2002 – September/2008, obtained from different options associated with four methodological choices: model structure (red), objective function used in model calibration (blue), multiple local optimal parameter sets (green), and forcing dataset used in model calibration (cyan).

The results in Figure 7.2 show that uncertainty from calibration strategy may overwhelm the uncertainty inherent to hydrologic model choice. Even more, it was found that parameter estimation methods may introduce significant uncertainty in projected changes of hydrologic behavior.

7.3.3 Effects of regional climate model configuration and forcing scaling

Finally, numerical experiments were conducted to quantify the effects of regional climate model (RCM) configuration (i.e. horizontal grid spacing and parameterization), and forcing re-scaling on projected hydrologic changes obtained with four model structures (Mendoza et al. 2015d). In the first experiment, hydrologic changes were computed using 4-km, 12-km and 36-km WRF outputs obtained by Rasmussen et al. (2014), where convective parameterization was activated only for 12-km and 36-km simulations (experiment 1). In the second experiment, hydrologic changes were computed with 4-km WRF outputs, and two additional datasets obtained from re-scaling 4-km outputs to the original 12- and 36-km WRF grid cells (experiment 2) used by Rasmussen et al. (2014). Figure 7.3 shows the standard deviation of projected changes in mean annual runoff and mean annual ET, computed from the multi-forcing ensembles in experiments 1 (plain bars) and 2 (hatched bars). The reader can note that WRF configuration has much larger effects than forcing re-scaling on projected changes in the overall water balance.

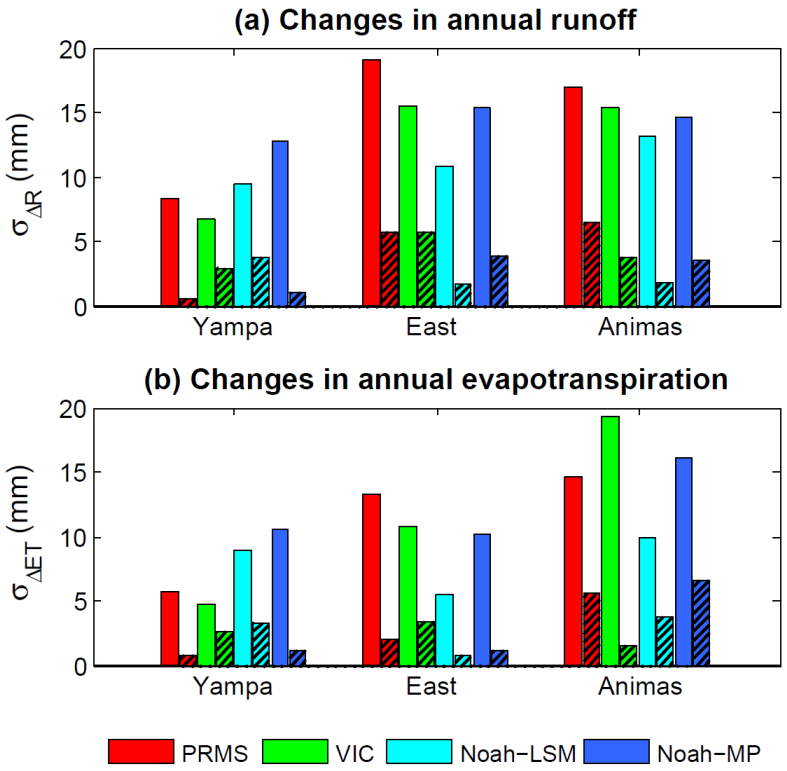

Figure 7.3: Standard deviation of projected changes in mean annual runoff (top) and mean annual ET (bottom) for the period October/2002 – September/2008, obtained from different forcing datasets across four hydrologic model structures. Plain bars represent the uncertainty from WRF configuration, and hatched bars represent the uncertainty from forcing scaling. Parameter values used here were obtained by calibrating each model against observed daily runoff, forcing simulations with 4-km WRF outputs.

Figure 7.4 illustrates the uncertainty in projected changes of mean annual runoff and ET, obtained from multi-model ensembles for each forcing dataset in experiments 1 (plain bars) and 2 (hatched bars). No clear relation was found between uncertainty in projections and WRF configuration across basins. Nevertheless, Figure 7.4 shows that uncertainty in projections increases as coarser forcing resolutions (hatched bars) are used to compute hydrologic changes. Also, the comparison of Figures 7.3 and 7.4 for East and Animas reveals that the method used to

generate forcing datasets can introduce larger uncertainty than model choice in hydrologic change projections.

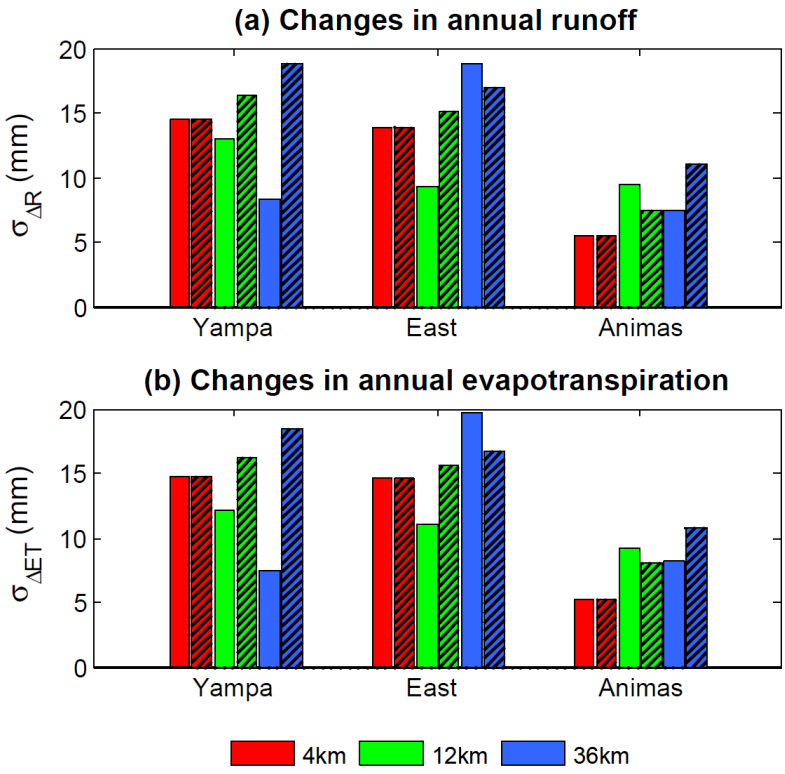

Figure 7.4: Standard deviation of projected changes in mean annual runoff (top) and mean annual ET (bottom) for the period October/2002 – September/2008, obtained from different models across three forcing datasets. Plain bars represent the uncertainty from WRF configuration (i.e. horizontal grid spacing and convective parameterization), and hatched bars represent the uncertainty from forcing scaling. Parameter values used here were obtained by calibrating each model against observed daily runoff, forcing simulations with 4-km WRF outputs.

7.4 Avoiding an over-confident prediction of hydrologic changes

7.4.1 What have we learnt?

Past efforts have proposed several paths to improve confidence in projected hydrologic changes. For example, Blöschl and Montanari (2010) pointed that a critical first step is to

"improve our knowledge of the connections among climate, weather and hydrology *under the current climate*" (italics added). Other general recommendations include the separation and understanding of the different sources of uncertainty (e.g. Harding et al. 2012; Steinschneider et al. 2012), accounting for non-stationary interactions between a changing climate and landscape properties (e.g. Milly et al. 2008; Miller et al. 2011; Surfleet and Tullos 2013), understanding the relative role of climatic variability and land cover change on hydrologic processes at multiple scales (Blöschl et al. 2007), and a better comprehension of climate change impacts on water management planning and decision making (Miller et al. 2012).

Some authors have put detailed emphasis on the improvement of hydrologic modeling decisions. For example, Steinschneider et al. (2012) recommended the incorporation of input (i.e. meteorological forcing) and response data uncertainties in statistical frameworks that quantify hydrologic change uncertainties. Poulin et al. (2011) and Velázquez et al. (2013) recommended the use of hydrological models with different levels of complexity in the assessment of climate change impacts. In this line, Seiller et al. (2012) concluded that use of a multi-model ensemble offers "better climate transposability", i.e. better model accuracy using parameter values calibrated for very different hydroclimatic conditions. Nevertheless, there is also awareness that hydrologic model structures should be carefully selected depending on the study purposes, hydroclimatic regime and basin characteristics (Bae et al. 2011; Najafi et al. 2011). In terms of hydrologic model parameters, Addor et al. (2014) suggests that multi-objective calibration may improve inter-model agreement when using multiple hydrologic model structures. Other studies have found a strong dependence between parameter values and climatic conditions (e.g. Rosero et al. 2010; Vaze et al. 2010; Merz et al. 2011; Coron et al. 2012; Herman et al. 2013). To address this problem, Singh et al. (2011) proposed a trading-space-to-time technique to generate

probabilistic streamflow predictions under different climatic scenarios. More recently, Westra et al. (2014) proposed the inclusion of non-stationarity by allowing hydrologic model parameters to vary in time as a function of selected covariates (e.g. stratify by season, rising and falling hydrograph limbs, etc.).

7.4.2 What have we missed?

Although we agree on the usefulness of multi-model approaches and multi-objective calibration to account for hydrologic model uncertainty, we believe that two key elements should be considered to advance towards a process-based assessment of climate change impacts on hydrology: (1) a diagnostic evaluation approach (Gupta et al. 2008) to build feasible multi-model ensembles and conduct model calibration/evaluation, and (2) *agile* modeling frameworks (Mendoza et al. 2015b) that allow model deconstruction and the incorporation of non-stationarity in model parameters. The inclusion and interaction between the above elements is crucial to understand and attribute hydrologic responses to specific modeling decisions (Clark et al. 2011a).

For example, in the case study presented in section 7.3 we selected four hydrologic models on an ad-hoc basis, rather than from the careful examination of available data to build realistic hypotheses of hydrologic behavior (e.g. McMillan et al. 2011; Clark et al. 2011b). This is problematic because the arbitrary addition of hydrologic models that include unfeasible process representations for the catchment(s) of interest may not only result in the over-estimation of structural uncertainty, but also in practicality issues related with unnecessary extra computational costs. Moreover, the subjective choice of a suite of model structures – developed under very different modeling philosophies – makes it very difficult to isolate the set of decisions as part of the model building process, or remove 'undesirable' modeling approaches (e.g.

potential evapotranspiration formulations, mosaic approach for sub-grid variability, linear reservoir for baseflow generation).

Regarding parametric uncertainty, two main lessons can be extracted from the example provided in section 7.3: (1) single-objective calibration does not necessarily help to reduce inter-model differences in projected hydrologic changes, and (2) uncertainties arising from parameter estimation strategies may overwhelm structural uncertainties. Hence, future studies should incorporate a diagnostic model evaluation approach to guide the parameter search towards *hydrologically consistent* regions (Martinez and Gupta 2011) – i.e. reproduce signature measures of specific catchment processes (e.g. evapotranspiration, high flow volumes, runoff seasonality, long-term baseflow). This could be achieved by constraining solutions or within a multi-objective calibration framework that includes signatures in the formulation of objective functions. Further, parametric uncertainty should be accounted for each model/ensemble member in order to reduce inter-model differences in projected changes of catchment behavior (e.g. Najafi et al. 2011; Velázquez et al. 2013). In our opinion, the availability of more agile hydrologic modeling frameworks (i.e. having the capability to modify spatial variability and hydrologic connectivity, model parameterizations and model parameter values) is crucial for facilitating experimentation towards more reliable estimates of projected hydrologic changes.

Finally, the example presented in section 7.3 illustrates the large effects of RCM configuration on projected hydrologic changes. In view of this, some authors have highlighted the need for further research linking atmospheric and land components of the water cycle using multiple RCMs (e.g. Gao et al. 2011). In our opinion, a strong a priori constraint of regional climate modeling frameworks like WRF is the high computational cost of simulations, making it very difficult to conduct extensive analyses with multiple parameterizations and spatial

configurations. Extending the principle of agility to weather models and moving towards intermediate complexity would facilitate the analyses of multiple scenarios and forcing uncertainty quantification via ensemble simulations.

7.5 Final perspectives

A fundamental problem in the assessment of climate change impacts on hydrologic processes is the large uncertainty associated with each modeling component (e.g. emission scenarios, global climate model, hydrological model, etc.). In our opinion, the widely used 'top-down' approach has two main caveats: (1) limited practicality, and (2) little attention to hydrologic process understanding. We argue that moving towards a process-based paradigm that accounts for a greater role of expert knowledge in decision-making, is crucial to improve understanding of hydrologic sensitivity to climate variability and change. While this approach requires the assumption that some climate impacts are predictable (i.e. *hard facts*), it allows detailed analyses of the implications that hydrologic modeling decisions have on the simulation of individual processes (i.e. focus on the *reasons* of projected ranges, rather than the magnitude of projections).

To obtain more reliable hydrologic predictions under a changing climate, many authors have recommended the use of multi-model ensembles, multi-objective calibration and non-stationary relations between landscape properties and climatic conditions. We believe that integrating a diagnostic evaluation approach with more agile modeling frameworks can facilitate the above tasks through the identification of plausible structures (i.e. set of model equations) and the exclusion of undesired process representations via model deconstruction. Moreover, agile modeling frameworks should facilitate the incorporation of non-stationary parameters and extensive sensitivity analyses and calibration towards hydrologically consistent parameter sets.

Finally, the extension of agility principles to weather models and the simplification of physics representations might facilitate extensive experimentation under different climate change scenarios and forcing uncertainty estimates via ensemble simulations.

Bibliography

Abbott, M., J. Bathurst, J. Cunge, P. O'Connell, and J. Rasmussen, 1986: An introduction to the European Hydrological System - Systeme Hydrologique Europeen, "SHE", 2: Structure of a physically-based, distributed modelling system. *J. Hydrol.*, **87**, 61–77.

Addor, N., O. Rössler, N. Köplin, M. Huss, R. Weingartner, and J. Seibert, 2014: Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments. *Water Resour. Res.*, **50**, 7541–7562, doi:10.1002/2014WR015549.

Aguado, E., 1985: Radiation Balances of Melting Snow Covers at an Open Site in the Central Sierra Nevada, California. *Water Resour. Res.*, **21**, 1649–1654, doi:10.1029/WR021i011p01649.

Amorocho, J., and B. Espildora, 1966: Mathematical Simulation of the snow melting processes. *Dep. Water Sci. Eng. Univ. California, Davis, W.S.E. Pap. 3001.*,.

Anderson, E., 1973: National Weather Service River Forecast system - snow accumulation and ablation model. *NOAA Tech. Memo. NWS HYDRO-17*, 217.

Asner, G. P., C. A. Wessman, D. S. Schimel, and S. Archer, 1998: Variability in Leaf and Litter Optical Properties : Implications for BRDF Model Inversions Using AVHRR , MODIS , and MISR. *Remote Sens. Environ.*, **257**, 243–257.

Bae, D.-H., I.-W. Jung, and D. P. Lettenmaier, 2011: Hydrologic uncertainties in climate change from IPCC AR4 GCM simulations of the Chungju Basin, Korea. *J. Hydrol.*, **401**, 90–105, doi:10.1016/j.jhydrol.2011.02.012.

Barlage, M., and Coauthors, 2010: Noah land surface model modifications to improve snowpack prediction in the Colorado Rocky Mountains. *J. Geophys. Res.*, **115**, D22101, doi:10.1029/2009JD013470.

Barnett, T. P., J. C. Adam, and D. P. Lettenmaier, 2005: Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, **438**, 303–309, doi:10.1038/nature04141.

Barnett, T. P., and Coauthors, 2008: Human-induced changes in the hydrology of the western United States. *Science (80-.).*, **319**, 1080–1083, doi:10.1126/science.1152538.

Bastola, S., C. Murphy, and J. Sweeney, 2011: The role of hydrological modelling uncertainties in climate change impact assessments of Irish river catchments. *Adv. Water Resour.*, **34**, 562–576, doi:10.1016/j.advwatres.2011.01.008.

Bergstrom, S., 1991: Principles and Confidence in Hydrological Modelling. *Nord. Hydrol.*, **22**, 123–136.

Bergström, S., B. Carlsson, M. Gardelin, G. Lindstrom, A. Petterson, and M. Rummukainen, 2001: Climate change impacts on runoff in Sweden-assessments by global climate models, dynamical downscaling and hydrological modelling. *Clim. Res.*, **16**, 101–112.

Beven, K., 1997: TOPMODEL: a critique. *Hydrol. Process.*, **11**, 1069–1085.

Beven, K., 2000: Uniqueness of place and process representations in hydrological modelling. *Hydrol. Earth Syst. Sci.*, **4**, 203–213.

Beven, K., 2002: Towards an alternative blueprint for a physically based digitally simulated hydrologic response modelling system. *Hydrol. Process.*, **16**, 189–206, doi:10.1002/hyp.343.

Beven, K., 2006: Searching for the Holy Grail of scientific hydrology: Qt=(S, R, Δt)A as closure. *Hydrol. Earth Syst. Sci.*, **10**, 609–618, doi:10.5194/hess-10-609-2006.

Beven, K., and A. Binley, 1992: The future of distributed models: model calibration and uncertainty prediction. *Hydrol. Process.*, **6**, 279–298.

Beven, K. J., and H. L. Cloke, 2012: Comment on "Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water" by Eric F. Wood et al. *Water Resour. Res.*, **48**, W01801, doi:10.1029/2011WR010982.

Blöschl, G., and M. Sivapalan, 1995: Scale issues in hydrological modelling: a review. *Hydrol. Process.*, **9**, 251–290.

Blöschl, G., and A. Montanari, 2010: Climate change impacts - □"throwing the dice? *Hydrol. Process.*, **24**, 374–381, doi:10.1002/hyp.7574.

——, and Coauthors, 2007: At what scales do climate variability and land cover change impact on flooding and low flows? *Hydrol. Process.*, **21**, 1241–1247, doi:10.1002/hyp.6669.

Bohn, T. J., B. Livneh, J. W. Oyler, S. W. Running, B. Nijssen, and D. P. Lettenmaier, 2013: Global evaluation of MTCLIM and related algorithms for forcing of ecological and hydrological models. *Agric. For. Meteorol.*, **176**, 38–49, doi:10.1016/j.agrformet.2013.03.003.

Boorman, D., and C. Sefton, 1997: Recognising the uncertainty in the quantification of the effects of climate change on hydrological response. *Clim. Change*, **35**, 415–434.

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

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

Bulygina, N., and H. Gupta, 2011: Correcting the mathematical structure of a hydrological model via Bayesian data assimilation. *Water Resour. Res.*, **47**, W05514, doi:10.1029/2010WR009614.

Bureau of Reclamation, 2012: *Colorado River Basin Water Supply and Demand Study: Technical Report B – Water Supply Assessment*.

Burnash, R., R. Ferral, and R. McGuire, 1973: *A generalized streamflow simulation system - Conceptual modeling for digital computers*. Sacramento, California,.

Cai, X., Z.-L. Yang, C. H. David, G.-Y. Niu, and M. Rodell, 2014: Hydrological evaluation of the Noah-MP land surface model for the Mississippi River Basin. *J. Geophys. Res. Atmos.*, **119**, 23–38, doi:10.1002/2013JD020792.

Cameron, D., K. Beven, and P. Naden, 1999: Flood frequency estimation by continuous simulation under climate change (with uncertainty). *Hydrol. Earth Syst. Sci.*, **4**, 393–405.

Carpenter, T. M., and K. P. Georgakakos, 2004: Impacts of parametric and radar rainfall uncertainty on the ensemble streamflow simulations of a distributed hydrologic model. *J. Hydrol.*, **298**, 202–221, doi:10.1016/j.jhydrol.2004.03.036.

Carrillo, G., P. A. Troch, M. Sivapalan, T. Wagener, C. Harman, and K. Sawicz, 2011: Catchment classification: hydrological analysis of catchment behavior through process-based modeling along a climate gradient. *Hydrol. Earth Syst. Sci.*, **15**, 3411–3430, doi:10.5194/hess-15-3411-2011.

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. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 21271–21276, doi:10.1073/pnas.0912391107.

Chen, F., and J. Dudhia, 2001: Coupling an Advanced Land Surface–Hydrology Model with the Penn State–NCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity. *Mon. Weather Rev.*, **129**, 569–585, doi:10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2.

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

Christensen, N., A. Wood, N. Voisin, D. P. Lettenmaier, and R. N. Palmer, 2004: The effects of climate change on the hydrology and water resources of the Colorado River basin. *Clim. Change*, **62**, 337–363.

Christensen, N. S., 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. *Hydrol. Earth Syst. Sci.*, **11**, 1417–1434, doi:10.5194/hess-11-1417-2007.

Clapp, R., and G. Hornberger, 1978: Empirical equations for some soil hydraulic properties. *Water Resour. Res.*, **14**, 601–604.

Clark, M. P., A. G. Slater, D. E. Rupp, R. A. Woods, J. A. Vrugt, H. V. Gupta, T. Wagener, and L. E. Hay, 2008: Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models. *Water Resour. Res.*, **44**, W00B02, doi:10.1029/2007WR006735.

——, D. Kavetski, and F. Fenicia, 2011a: Pursuing the method of multiple working hypotheses for hydrological modeling. *Water Resour. Res.*, **47**, 1–16, doi:10.1029/2010WR009827.

——, H. K. McMillan, D. B. G. Collins, D. Kavetski, and R. A. Woods, 2011b: Hydrological field data from a modeller's perspective: Part 2: process-based evaluation of model hypotheses. *Hydrol. Process.*, **25**, 523–543, doi:10.1002/hyp.7902.

Collins, W. D., and Coauthors, 2006: The Community Climate System Model Version 3 (CCSM3). *J. Clim.*, **19**, 2122–2143, doi:10.1175/JCLI3761.1.

Coron, L., V. Andréassian, C. Perrin, J. Lerat, J. Vaze, M. Bourqui, and F. Hendrickx, 2012: Crash testing hydrological models in contrasted climate conditions: An experiment on 216 Australian catchments. *Water Resour. Res.*, **48**, doi:10.1029/2011WR011721.

Cosby, B. J., G. M. Hornberger, R. B. Clapp, and T. R. Ginn, 1984: A Statistical Exploration of the Relationships of Soil Moisture Characteristics to the Physical Properties of Soils. *Water Resour. Res.*, **20**, 682–690, doi:10.1029/WR020i006p00682.

Cosgrove, B. A., and Coauthors, 2003: Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. *J. Geophys. Res.*, **108**, 8842, doi:10.1029/2002JD003118.

Das, T., and Coauthors, 2009: Structure and Detectability of Trends in Hydrological Measures over the Western United States. *J. Hydrometeorol.*, **10**, 871–892, doi:10.1175/2009JHM1095.1.

Dirmhirn, I., and F. Eaton, 1975: Some Characteristics of the Albedo of Snow. *J. Appl. Meteorol.*, **14**, 375–379.

Dorman, J., and P. Sellers, 1989: A global climatology of albedo, roughness length and stomatal resistance for atmospheric general circulation models as represented by the simple biosphere model (SiB). *J. Appl. Meteorol.*, **28**, 833–855.

Duan, Q., S. Sorooshian, and V. Gupta, 1992: Effective and Efficient Global Optimization for Conceptual Rainfal-Runoff Models. *Water Resour. Res.*, **28**, 1015–1031.

Duan, Q. Y., V. K. Gupta, and S. Sorooshian, 1993: Shuffled complex evolution approach for effective and efficient global minimization. *J. Optim. Theory Appl.*, **76**, 501–521, doi:10.1007/BF00939380.

Ek, M. B., 2003: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *J. Geophys. Res.*, **108**, 1–16, doi:10.1029/2002JD003296.

Elsner, M. M., S. Gangopadhyay, T. Pruitt, L. D. Brekke, N. Mizukami, and M. P. Clark, 2014: How Does the Choice of Distributed Meteorological Data Affect Hydrologic Model Calibration and Streamflow Simulations? *J. Hydrometeorol.*, **15**, 1384–1403, doi:10.1175/JHM-D-13-083.1.

Essery, R., and P. Etchevers, 2004: Parameter sensitivity in simulations of snowmelt. *J. Geophys. Res.*, **109**, D20111, doi:10.1029/2004JD005036.

——, S. Morin, Y. Lejeune, and C. B Ménard, 2013: A comparison of 1701 snow models using observations from an alpine site. *Adv. Water Resour.*, **55**, 131–148, doi:10.1016/j.advwatres.2012.07.013.

Fenicia, F., J. J. McDonnell, and H. H. G. Savenije, 2008: Learning from model improvement: On the contribution of complementary data to process understanding. *Water Resour. Res.*, **44**, W06419, doi:10.1029/2007WR006386.

——, D. Kavetski, and H. H. G. Savenije, 2011: Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development. *Water Resour. Res.*, **47**, W11510, doi:10.1029/2010WR010174.

Ficklin, D. L., I. T. Stewart, and E. P. Maurer, 2013: Climate change impacts on streamflow and subbasin-scale hydrology in the Upper Colorado River Basin. *PLoS One*, **8**, e71297, doi:10.1371/journal.pone.0071297.

Flanner, M. G., C. S. Zender, J. T. Randerson, and P. J. Rasch, 2007: Present-day climate forcing and response from black carbon in snow. *J. Geophys. Res.*, **112**, D11202, doi:10.1029/2006JD008003.

Flerchinger, G. N., W. Xaio, D. Marks, T. J. Sauer, and Q. Yu, 2009: Comparison of algorithms for incoming atmospheric long-wave radiation. *Water Resour. Res.*, **45**, W03423, doi:10.1029/2008WR007394.

Foglia, L., M. C. Hill, S. W. Mehl, and P. Burlando, 2009: Sensitivity analysis, calibration, and testing of a distributed hydrological model using error-based weighting and one objective function. *Water Resour. Res.*, **45**, W06427, doi:10.1029/2008WR007255.

Freeze, R. A., and R. Harlan, 1969: Blueprint for a physically-based, digitally simulated hydrologic response model. *J. Hydrol.*, **9**, 237–258.

Fry, J., and Coauthors, 2011: Completion of the 2006 National Land Cover Database for the Conterminous United States. *Photogramm. Eng. Remote Sens.*, **77**, 858–864.

Gao, Y., J. a. Vano, C. Zhu, and D. P. Lettenmaier, 2011: Evaluating climate change over the Colorado River basin using regional climate models. *J. Geophys. Res.*, **116**, D13104, doi:10.1029/2010JD015278.

Gerten, D., S. Schaphoff, U. Haberlandt, W. Lucht, and S. Sitch, 2004: Terrestrial vegetation and water balance—hydrological evaluation of a dynamic global vegetation model. *J. Hydrol.*, **286**, 249–270, doi:10.1016/j.jhydrol.2003.09.029.

Göhler, M., J. Mai, and M. Cuntz, 2013: Use of eigendecomposition in a parameter sensitivity analysis of the Community Land Model. *J. Geophys. Res. Biogeosciences*, **118**, 904–921, doi:10.1002/jgrg.20072.

Goudriaan, J., 1977: *Crop micrometeorology: a simulation study*. Pudoc, Center for Agricultural Publishing and Documentation,.

Graham, L. P., J. Andréasson, and B. Carlsson, 2007: Assessing climate change impacts on hydrology from an ensemble of regional climate models, model scales and linking methods – a case study on the Lule River basin. *Clim. Change*, **81**, 293–307, doi:10.1007/s10584-006-9215-2.

Gupta, H. V., S. Sorooshian, and P. O. Yapo, 1998: Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resour. Res.*, **34**, 751–763, doi:10.1029/97WR03495.

Gupta, H. V., T. Wagener, and Y. Liu, 2008: Reconciling theory with observations : elements of a diagnostic approach to model evaluation. *Hydrol. Process.*, **22**, 3802–3813, doi:10.1002/hyp.

——, H. Kling, K. K. Yilmaz, and G. F. Martinez, 2009: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *J. Hydrol.*, **377**, 80–91, doi:10.1016/j.jhydrol.2009.08.003.

——, M. P. Clark, J. A. Vrugt, G. Abramowitz, and M. Ye, 2012: Towards a comprehensive assessment of model structural adequacy. *Water Resour. Res.*, **48**, W08301, doi:10.1029/2011WR011044.

Gutmann, E., T. Pruitt, M. P. Clark, L. Brekke, J. R. Arnold, D. A. Raff, and R. M. Rasmussen, 2014: An intercomparison of statistical downscaling methods used for water resource assessments in the United States. *Water Resour. Res.*, **50**, doi:10.1002/2014WR015559.

Gutmann, E. D., R. M. Rasmussen, C. Liu, K. Ikeda, D. J. Gochis, M. P. Clark, J. Dudhia, and G. Thompson, 2012: A Comparison of Statistical and Dynamical Downscaling of Winter Precipitation over Complex Terrain. *J. Clim.*, **25**, 262–281, doi:10.1175/2011JCLI4109.1.

Haddeland, I., B. V. Matheussen, and D. P. Lettenmaier, 2002: Influence of spatial resolution on simulated streamflow in a macroscale hydrologic model. *Water Resour. Res.*, **38**, doi:10.1029/2001WR000854.

Haddeland, I., J. Heinke, F. Voß, S. Eisner, C. Chen, S. Hagemann, and F. Ludwig, 2012: Effects of climate model radiation, humidity and wind estimates on hydrological simulations. *Hydrol. Earth Syst. Sci.*, **16**, 305–318, doi:10.5194/hess-16-305-2012.

Hallegatte, S., 2009: Strategies to adapt to an uncertain climate change. *Glob. Environ. Chang.*, **19**, 240–247, doi:10.1016/j.gloenvcha.2008.12.003.

Hamlet, A. F., P. W. Mote, M. P. Clark, and D. P. Lettenmaier, 2007: Twentieth-Century Trends in Runoff, Evapotranspiration, and Soil Moisture in the Western United States. *J. Clim.*, **20**, 1468–1486, doi:10.1175/JCLI4051.1.

Hansen, M. C., R. S. DeFries, J. R. Townshend, and R. Sohlberg, 2000: Global land cover classification at 1 km spatial resolution using a classification tree approach. *Int. J. Remote Sens.*, **21**, 1331–1364.

Hara, M., T. Yoshikane, H. Kawase, and F. Kimura, 2008: Estimation of the Impact of Global Warming on Snow Depth in Japan by the Pseudo-Global-Warming Method. *Hydrol. Res. Lett.*, **2**, 61–64, doi:10.3178/hrl.2.61.

Harding, B. L., a. W. Wood, and J. R. Prairie, 2012: The implications of climate change scenario selection for future streamflow projection in the Upper Colorado River Basin. *Hydrol. Earth Syst. Sci.*, **16**, 3989–4007, doi:10.5194/hess-16-3989-2012.

Hartmann, A., T. Wagener, A. Rimmer, J. Lange, H. Brielmann, and M. Weiler, 2013: Testing the realism of model structures to identify karst system processes using water quality and quantity signatures. *Water Resour. Res.*, **49**, 3345–3358, doi:10.1002/wrcr.20229.

Hastie, T., R. Tibshirani, and J. Friedman, 2002: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Second edi. Springer,.

Hay, L. E., and M. P. Clark, 2003: Use of statistically and dynamically downscaled atmospheric model output for hydrologic simulations in three mountainous basins in the western United States. *J. Hydrol.*, **282**, 56–75, doi:10.1016/S0022-1694(03)00252-X.

Hay, L. E., M. P. Clark, R. Wilby, W. J. Gutowski Jr, G. H. Leavesley, Z. Pan, R. W. Arritt, and E. S. Takle, 2002: Use of Regional Climate Model Output for Hydrologic Simulations. *J. Hydrometeorol.*, **3**, 571–590.

Herbst, M., H. V. Gupta, and M. C. Casper, 2009: Mapping model behaviour using Self-Organizing Maps. *Hydrol. Earth Syst. Sci.*, **13**, 395–409.

Herman, J. D., P. M. Reed, and T. Wagener, 2013: Time-varying sensitivity analysis clarifies the effects of watershed model formulation on model behavior. *Water Resour. Res.*, **49**, 1400–1414, doi:10.1002/wrcr.20124.

Hidalgo, H. G., and Coauthors, 2009: Detection and Attribution of Streamflow Timing Changes to Climate Change in the Western United States. *J. Clim.*, **22**, 3838–3855, doi:10.1175/2009JCLI2470.1.

Hoerling, M., and J. Eischeid, 2007: Past peak water in the Southwest. *Southwest Hydrol.*, **6**, 18–19.

Hoerling, M., D. Lettenmaier, D. Cayan, and B. Udall, 2009: Reconciling projections of Colorado River streamflow. *Southwest Hydrol.*, 20–22.

Hogue, T. S., L. Bastidas, H. Gupta, S. Sorooshian, K. Mitchell, and W. Emmerich, 2005: Evaluation and transferability of the Noah land surface model in semiarid environments. *J. Hydrometeorol.*, **6**, 68–84.

Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes. *Mon. Weather Rev.*, **134**, 2318–2341, doi:10.1175/MWR3199.1.

Hrachowitz, M., and Coauthors, 2014: Process consistency in models: The importance of system signatures, expert knowledge, and process complexity. *Water Resour. Res.*, **50**, doi:10.1002/2014WR015484.

Huntington, T. G., 2006: Evidence for intensification of the global water cycle: Review and synthesis. *J. Hydrol.*, **319**, 83–95, doi:10.1016/j.jhydrol.2005.07.003.

Ikeda, K., and Coauthors, 2010: Simulation of seasonal snowfall over Colorado. *Atmos. Res.*, **97**, 462–477, doi:10.1016/j.atmosres.2010.04.010.

IPCC, 2013: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. T.F. Stocker et al., Eds. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA,.

Ivanov, V. Y., E. R. Vivoni, R. L. Bras, and D. Entekhabi, 2004: Preserving high-resolution surface and rainfall data in operational-scale basin hydrology: a fully-distributed physically-based approach. *J. Hydrol.*, **298**, 80–111, doi:10.1016/j.jhydrol.2004.03.041.

Janjić, Z., 1994: The Step-Mountain Eta Coordinate Model: Further Developments of the Convection, Viscous Sublayer, and Turbulence Closure Schemes. *Mon. Weather Rev.*, **122**, 927–945.

Jensen, M., D. Rob, and C. Franzoy, 1969: Scheduling irrigations using climate-crop-soil data. *National Conference on Water Resources Engineering of the American Society of Civil Engineers*, New Orleans, 20.

Jiang, T., Y. D. Chen, C. Xu, X. Chen, X. Chen, and V. P. Singh, 2007: Comparison of hydrological impacts of climate change simulated by six hydrological models in the Dongjiang Basin, South China. *J. Hydrol.*, **336**, 316–333, doi:10.1016/j.jhydrol.2007.01.010.

Jones, R. N., F. H. S. Chiew, W. C. Boughton, and L. Zhang, 2006: Estimating the sensitivity of mean annual runoff to climate change using selected hydrological models. *Adv. Water Resour.*, **29**, 1419–1429, doi:10.1016/j.advwatres.2005.11.001.

Jordan, R., 1991: A One-Dimensional Temperature Model for a Snow Cover. *Cold Reg. Res. Eng. Lab, U.S. Army Corps Eng. Hanover, N.H.*, Spec. Rept. 91–16.

Joyce, R., J. Janowiak, P. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeorol.*, **5**, 487–503.

Kattge, J., W. Knorr, T. Raddatz, and C. Wirth, 2009: Quantifying photosynthetic capacity and its relationship to leaf nitrogen content for global-scale terrestrial biosphere models. *Glob. Chang. Biol.*, **15**, 976–991, doi:10.1111/j.1365-2486.2008.01744.x.

Kavetski, D., G. Kuczera, and S. W. Franks, 2006a: Calibration of conceptual hydrological models revisited: 1. Overcoming numerical artefacts. *J. Hydrol.*, **320**, 173–186, doi:10.1016/j.jhydrol.2005.07.012.

——, ——, and ——, 2006b: Calibration of conceptual hydrological models revisited: 2. Improving optimisation and analysis. *J. Hydrol.*, **320**, 187–201, doi:10.1016/j.jhydrol.2005.07.013.

Kawase, H., T. Yoshikane, M. Hara, F. Kimura, T. Yasunari, B. Ailikun, H. Ueda, and T. Inoue, 2009: Intermodel variability of future changes in the Baiu rainband estimated by the pseudo global warming downscaling method. *J. Geophys. Res.*, **114**, D24110, doi:10.1029/2009JD011803.

Kay, A. L., H. N. Davies, V. A. Bell, and R. G. Jones, 2009: Comparison of uncertainty sources for climate change impacts: flood frequency in England. *Clim. Change*, **92**, 41–63, doi:10.1007/s10584-008-9471-4.

Kirchner, J. W., 2006: Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resour. Res.*, **42**, doi:10.1029/2005WR004362.

Knowles, N., M. D. Dettinger, and D. R. Cayan, 2006: Trends in Snowfall versus Rainfall in the Western United States. *J. Clim.*, **19**, 4545–4559, doi:10.1175/JCLI3850.1.

Kollat, J. B., P. M. Reed, and T. Wagener, 2012: When are multiobjective calibration trade-offs in hydrologic models meaningful? *Water Resour. Res.*, **48**, 1–19, doi:10.1029/2011WR011534.

Kollet, S. J., R. M. Maxwell, C. S. Woodward, S. Smith, J. Vanderborght, H. Vereecken, and C. Simmer, 2010: Proof of concept of regional scale hydrologic simulations at hydrologic resolution utilizing massively parallel computer resources. *Water Resour. Res.*, **46**, W04201, doi:10.1029/2009WR008730.

Köplin, N., B. Schädler, D. Viviroli, and R. Weingartner, 2012: Relating climate change signals and physiographic catchment properties to clustered hydrological response types. *Hydrol. Earth Syst. Sci.*, **16**, 2267–2283, doi:10.5194/hess-16-2267-2012.

Kuczera, G., D. Kavetski, S. Franks, and M. Thyer, 2006: Towards a Bayesian total error analysis of conceptual rainfall-runoff models: Characterising model error using storm-dependent parameters. *J. Hydrol.*, **331**, 161–177, doi:10.1016/j.jhydrol.2006.05.010.

Lall, U., 2014: Debates-The future of hydrological sciences: A (common) path forward? One water. One world. Many climes. Many souls. *Water Resour. Res.*, **50**, 5335–5341, doi:10.1002/2014WR015402.

De Lannoy, G. J. M., P. R. Houser, V. R. N. Pauwels, and N. E. C. Verhoest, 2006: Assessment of model uncertainty for soil moisture through ensemble verification. *J. Geophys. Res.*, **111**, D10101, doi:10.1029/2005JD006367.

Lawrence, D. M., and Coauthors, 2011: Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model. *J. Adv. Model. Earth Syst.*, **3**, M03001, doi:10.1029/2011MS000045.

Leavesley, G. H., and L. G. Stannard, 1995: The precipitation-runoff modeling system-PRMS. *Computer Models of Watershed Hydrology*, V.P. Singh, Ed., Water Resources Publications, Highlands Ranch, CO, 281–310.

Leavesley, G. H., R. W. Lichty, B. M. Troutman, and L. G. Saindon, 1983: *Precipitation-Runoff Modeling System: User's Manual*.

Lempert, R. J., N. Nakicenovic, D. Sarewitz, and M. Schlesinger, 2004: Characterizing climate-change uncertainties for decision-makers. *Clim. Change*, **65**, 1–9.

Ley, R., M. C. Casper, H. Hellebrand, and R. Merz, 2011: Catchment classification by runoff behaviour with self-organizing maps (SOM). *Hydrol. Earth Syst. Sci.*, **15**, 2947–2962, doi:10.5194/hess-15-2947-2011.

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. *J. Geophys. Res.*, **99**, 14,415.14.428.

——, E. F. Wood, and D. P. Lettenmaier, 1996: Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification. *Glob. Planet. Change*, **13**, 195–206, doi:10.1016/0921-8181(95)00046-1.

——, J. Guo, and L. R. Leung, 2004: Assessment of the effects of spatial resolutions on daily water flux simulations. *J. Hydrol.*, **298**, 287–310, doi:10.1016/j.jhydrol.2003.07.007.

Lindström, G., B. Johansson, M. Persson, M. Gardelin, and S. Bergström, 1997: Development and test of the distributed HBV-96 hydrological model. *J. Hydrol.*, **201**, 272–288, doi:10.1016/S0022-1694(97)00041-3.

Liu, Y., and H. V. Gupta, 2007: Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resour. Res.*, **43**, W07401, doi:10.1029/2006WR005756.

Loague, K., and J. E. VanderKwaak, 2004: Physics-based hydrologic response simulation: platinum bridge, 1958 Edsel, or useful tool. *Hydrol. Process.*, **18**, 2949–2956, doi:10.1002/hyp.5737.

Loaiciga, H. A., J. B. Valdes, R. Vogel, J. Garvey, and H. Schwarz, 1996: Global warming and the hydrologic cycle. *J. Hydrol.*, **174**, 83–127.

Ludwig, R., and Coauthors, 2009: The role of hydrological model complexity and uncertainty in climate change impact assessment. *Adv. Geosci.*, **21**, 63–71.

MacQueen, J., 1967: Some methods for classification and analysis of multivariate observations. *Proceedings of 5th Berkeley Symposium on Mathematics, Statistics and Probability*, Berkeley, University of California, 281–297.

Majone, B., C. I. Bovolo, A. Bellin, S. Blenkinsop, and H. J. Fowler, 2012: Modeling the impacts of future climate change on water resources for the Gállego river basin (Spain). *Water Resour. Res.*, **48**, 1–18, doi:10.1029/2011WR010985.

Marks, D., and J. Dozier, 1992: Climate and Energy Exchange at the Snow Surface in the Alpine Region of the Sierra Nevada 2. Snow Cover Energy Balance. *Water Resour. Res.*, **28**, 3043–3054.

——, J. Domingo, D. Susong, T. Link, and D. Garen, 1999: A spatially distributed energy balance snowmelt model for application in mountain basins. *Hydrol. Process.*, **13**, 1935–1959.

Markstrom, S. L., R. G. Niswonger, R. S. Regan, D. E. Prudic, and P. M. Barlow, 2008: *GSFLOW — Coupled Ground-Water and Surface-Water Flow Model Based on the Integration of the Precipitation-Runoff Modeling System (PRMS) and the Modular Ground-Water Flow Model (MODFLOW-2005).*

Martinez, G. F., and H. V. Gupta, 2011: Hydrologic consistency as a basis for assessing complexity of monthly water balance models for the continental United States. *Water Resour. Res.*, **47**, doi:10.1029/2011WR011229.

Mastin, M. C., K. J. Chase, and R. W. Dudley, 2011: Changes in Spring Snowpack for Selected Basins in the United States for Different Climate-Change Scenarios. *Earth Interact.*, **15**, 1–18, doi:10.1175/2010EI368.1.

Materia, S., P. A. Dirmeyer, Z. Guo, A. Alessandri, and A. Navarra, 2010: The Sensitivity of Simulated River Discharge to Land Surface Representation and Meteorological Forcings. *J. Hydrometeorol.*, **11**, 334–351, doi:10.1175/2009JHM1162.1.

Maurer, E., A. W. Wood, and D. P. Lettenmaier, 2002: A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States*. *J. Clim.*, **15**, 3237–3251.

Maxwell, R., and N. Miller, 2005: Development of a coupled land surface and groundwater model. *J. Hydrometeorol.*, **6**, 233–247.

McCabe, G. J., and D. M. Wolock, 2007: Warming may create substantial water supply shortages in the Colorado River basin. *Geophys. Res. Lett.*, **34**, L22708, doi:10.1029/2007GL031764.

McMillan, H., M. Gueguen, E. Grimon, R. Woods, M. Clark, and D. E. Rupp, 2014: Spatial variability of hydrological processes and model structure diagnostics in a 50 km 2 catchment. *Hydrol. Process.*, **28**, 4896–4913, doi:10.1002/hyp.9988.

McMillan, H. K., M. P. Clark, W. B. Bowden, M. Duncan, and R. A. Woods, 2011: Hydrological field data from a modeller's perspective: Part 1. Diagnostic tests for model structure. *Hydrol. Process.*, **25**, 511–522, doi:10.1002/hyp.7841.

Meinshausen, M., and Coauthors, 2011: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Clim. Change*, **109**, 213–241, doi:10.1007/s10584-011-0156-z.

Mendoza, P. A., and Coauthors, 2015a: Effects of hydrologic model choice and calibration on the portrayal of climate change impacts. *J. Hydrometeorol.*, accepted.

——, M. P. Clark, M. Barlage, B. Rajagopalan, L. Samaniego, G. Abramowitz, and H. Gupta, 2015b: Are we unnecessarily constraining the agility of complex process-based models? *Water Resour. Res.*, **50**, doi:10.1002/2014WR015820.

——, ——, N. Mizukami, E. D. Gutmann, J. R. Arnold, L. D. Brekke, and B. Rajagopalan, 2015c: Implications of subjective hydrologic model choice and parameter identification strategies on the assessment of climate change impacts. *Hydrol. Process.*, submitted.

——, ——, ——, K. Ikeda, E. D. Gutmann, L. D. Brekke, J. R. Arnold, and B. Rajagopalan, 2015d: Effects of regional climate model configuration and forcing scaling on projected hydrologic changes. *J. Hydrol.*, in preparation.

Merz, R., J. Parajka, and G. Blöschl, 2011: Time stability of catchment model parameters: Implications for climate impact analyses. *Water Resour. Res.*, **47**, n/a–n/a, doi:10.1029/2010WR009505.

Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Am. Meteorol. Soc.*, **87**, 343–360, doi:10.1175/BAMS-87-3-343.

Miller, W., R. A. Butler, T. Piechota, J. Prairie, K. Grantz, and G. DeRosa, 2012: Water Management Decisions using Multiple Hydrologic Models Within the San Juan River Basin Under Changing Climate Conditions. *J. Water Resour. Plan. Manag.*, **138**, 412–420, doi:10.1061/(ASCE)WR.1943-5452.0000237.

Miller, W. P., and T. C. Piechota, 2008: Regional Analysis of Trend and Step Changes Observed in Hydroclimatic Variables around the Colorado River Basin. *J. Hydrometeorol.*, **9**, 1020–1034, doi:10.1175/2008JHM988.1.

——, and ——, 2011: Trends in Western U.S. Snowpack and Related Upper Colorado River Basin Streamflow. *J. Am. Water Resour. Assoc.*, **47**, 1197–1210, doi:10.1111/j.1752-1688.2011.00565.x.

Miller, W. P., T. C. Piechota, S. Gangopadhyay, and T. Pruitt, 2011: Development of streamflow projections under changing climate conditions over Colorado River basin headwaters. *Hydrol. Earth Syst. Sci.*, **15**, 2145–2164, doi:10.5194/hess-15-2145-2011.

Milly, P. C. D., and K. A. Dunne, 2011: On the Hydrologic Adjustment of Climate-Model Projections: The Potential Pitfall of Potential Evapotranspiration. *Earth Interact.*, **15**, 1–14, doi:10.1175/2010EI363.1.

Milly, P. C. D., K. A. Dunne, and A. V Vecchia, 2005: Global pattern of trends in streamflow and water availability in a changing climate. *Nature*, **438**, 347–350, doi:10.1038/nature04312.

——, J. Betancourt, M. Falkenmark, R. M. Hirsch, Z. W. Kundzewicz, D. P. Lettenmaier, and R. J. Stouffer, 2008: Stationarity is dead: whither water management? *Science (80-.).*, **319**, 573–574, doi:10.1126/science.1151915.

Mitchell, K. E., and Coauthors, 2004: The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *J. Geophys. Res.*, **109**, D07S90, doi:10.1029/2003JD003823.

Mizukami, N., M. P. Clark, A. G. Slater, L. D. Brekke, M. M. Elsner, J. R. Arnold, and S. Gangopadhyay, 2014: Hydrologic Implications of Different Large-Scale Meteorological Model Forcing Datasets in Mountainous Regions. *J. Hydrometeorol.*, **15**, 474–488, doi:10.1175/JHM-D-13-036.1.

Mo, K. C., L.-C. Chen, S. Shukla, T. J. Bohn, and D. P. Lettenmaier, 2012: Uncertainties in North American Land Data Assimilation Systems over the Contiguous United States. *J. Hydrometeorol.*, **13**, 996–1009, doi:10.1175/JHM-D-11-0132.1.

Montanari, A., and D. Koutsoyiannis, 2012: A blueprint for process-based modeling of uncertain hydrological systems. *Water Resour. Res.*, **48**, W09555, doi:10.1029/2011WR011412.

Mote, P. W., A. F. Hamlet, M. P. Clark, and D. P. Lettenmaier, 2005: Declining Mountain Snowpack in Western North America. *Bull. Am. Meteorol. Soc.*, **86**, 39–49, doi:10.1175/BAMS-86-1-39.

Murphy, J., D. Sexton, D. Barnett, G. Jones, M. Webb, M. Collins, and D. Stainforth, 2004: Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature*, **430**, 768–772, doi:10.1038/nature02770.1.

Myneni, R. B., R. Ramakrishna, R. Nemani, and S. W. Running, 1997: Estimation of global leaf area index and absorbed par using radiative transfer models. *IEEE Trans. Geosci. Remote Sens.*, **35**, 1380–1393, doi:10.1109/36.649788.

Najafi, M. R., H. Moradkhani, and I. W. Jung, 2011: Assessing the uncertainties of hydrologic model selection in climate change impact studies. *Hydrol. Process.*, **25**, 2814–2826, doi:10.1002/hyp.8043.

Nash, J., and J. Sutcliffe, 1970: River flow forecasting through conceptual models part I - A discussion of principles. *J. Hydrol.*, **10**, 282–290.

Naylor, R. L., D. S. Battisti, D. J. Vimont, W. P. Falcon, and M. B. Burke, 2007: Assessing risks of climate variability and climate change for Indonesian rice agriculture. *Proc. Natl. Acad. Sci.*, **104**, 7752–7757.

Niu, G., and Z. Yang, 2006: Effects of frozen soil on snowmelt runoff and soil water storage at a continental scale. *J. Hydrometeorol.*, 937–953.

Niu, G.-Y., and Z.-L. Yang, 2007: An observation-based formulation of snow cover fraction and its evaluation over large North American river basins. *J. Geophys. Res.*, **112**, 1–14, doi:10.1029/2007JD008674.

——, ——, R. E. Dickinson, and L. E. Gulden, 2005: A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models. *J. Geophys. Res.*, **110**, D21106, doi:10.1029/2005JD006111.

——, and Coauthors, 2011: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *J. Geophys. Res.*, **116**, 1–19, doi:10.1029/2010JD015139.

Oleson, K. W., and Coauthors, 2010: *Technical Description of version 4.0 of the Community Land Model (CLM)*. Boulder, Colorado, USA,.

Otto, J., T. Raddatz, and M. Claussen, 2011: Strength of forest-albedo feedback in mid-Holocene climate simulations. *Clim. Past*, **7**, 1027–1039, doi:10.5194/cp-7-1027-2011.

Oudin, L., A. Kay, V. Andréassian, and C. Perrin, 2010: Are seemingly physically similar catchments truly hydrologically similar? *Water Resour. Res.*, **46**, 1–15, doi:10.1029/2009WR008887.

Pauwels, V. R. N., and G. J. M. De Lannoy, 2009: Ensemble-based assimilation of discharge into rainfall-runoff models: A comparison of approaches to mapping observational information to state space. *Water Resour. Res.*, **45**, W08428, doi:10.1029/2008WR007590.

Perrin, C., C. Michel, and V. Andréassian, 2003: Improvement of a parsimonious model for streamflow simulation. *J. Hydrol.*, **279**, 275–289, doi:10.1016/S0022-1694(03)00225-7.

Pfannerstill, M., B. Guse, and N. Fohrer, 2014: Smart low flow signature metrics for an improved overall performance evaluation of hydrological models. *J. Hydrol.*, **510**, 447–458, doi:10.1016/j.jhydrol.2013.12.044.

Pierce, D. W., and Coauthors, 2008: Attribution of Declining Western U.S. Snowpack to Human Effects. *J. Clim.*, **21**, 6425–6444, doi:10.1175/2008JCLI2405.1.

Pokhrel, P., and H. V. Gupta, 2009: Regularized calibration of a distributed hydrological model using available information about watershed properties and signature measures. *New Approaches to Hydrological Prediction in Data-sparse Regions (Proc. of Symposium HS.2 at the Joint IAHS and IAH Convention, Hyderabad, India, September 2009)*, 20–25.

——, K. K. Yilmaz, and H. V. Gupta, 2012: Multiple-criteria calibration of a distributed watershed model using spatial regularization and response signatures. *J. Hydrol.*, **418-419**, 49–60, doi:10.1016/j.jhydrol.2008.12.004.

Pomeroy, J. W., D. M. Gray, T. Brown, N. R. Hedstrom, W. L. Quinton, R. J. Granger, and S. K. Carey, 2007: The cold regions hydrological model : a platform for basing process representation and model structure on physical evidence. *Hydrol. Process.*, **21**, 2650–2667, doi:10.1002/hyp.

Poulin, A., F. Brissette, R. Leconte, R. Arsenault, and J.-S. Malo, 2011: Uncertainty of hydrological modelling in climate change impact studies in a Canadian, snow-dominated river basin. *J. Hydrol.*, **409**, 626–636, doi:10.1016/j.jhydrol.2011.08.057.

Prihodko, L., A. S. Denning, N. P. Hanan, I. Baker, and K. Davis, 2008: Sensitivity, uncertainty and time dependence of parameters in a complex land surface model. *Agric. For. Meteorol.*, **148**, 268–287, doi:10.1016/j.agrformet.2007.08.006.

Prudhomme, C., and H. Davies, 2009a: Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK. Part 1: baseline climate. *Clim. Change*, **93**, 177–195, doi:10.1007/s10584-008-9464-3.

——, and ——, 2009b: Assessing uncertainties in climate change impact analyses on the river flow regimes in the UK. Part 2: future climate. *Clim. Change*, **93**, 197–222, doi:10.1007/s10584-008-9461-6.

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

Qu, Y., and C. J. Duffy, 2007: A semidiscrete finite volume formulation for multiprocess watershed simulation. *Water Resour. Res.*, **43**, doi:10.1029/2006WR005752.

Raff, D., T. Pruitt, and L. Brekke, 2009: A framework for assessing flood frequency based on climate projection information. *Hydrol. Earth Syst. Sci.*, **13**, 2119–2136.

Raftery, A. E., T. Gneiting, F. Balabdaoui, and M. Polakowski, 2005: Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *Mon. Weather Rev.*, **133**, 1155–1174, doi:10.1175/MWR2906.1.

Rakovec, O., M. C. Hill, M. P. Clark, A. H. Weerts, A. J. Teuling, and R. Uijlenhoet, 2014: Distributed Evaluation of Local Sensitivity Analysis (DELSA), with application to hydrologic models. *Water Resour. Res.*, **50**, 1–18, doi:10.1002/2013WR014063.

Raleigh, M. S., J. D. Lundquist, and M. P. Clark, 2014: Exploring the impact of forcing error characteristics on physically based snow simulations within a global sensitivity analysis framework. *Hydrol. Earth Syst. Sci.*, **submitted**.

Rasmussen, R., and Coauthors, 2011: High-Resolution Coupled Climate Runoff Simulations of Seasonal Snowfall over Colorado: A Process Study of Current and Warmer Climate. *J. Clim.*, **24**, 3015–3048, doi:10.1175/2010JCLI3985.1.

——, and Coauthors, 2014: Climate Change Impacts on the Water Balance of the Colorado Headwaters: High-Resolution Regional Climate Model Simulations. *J. Hydrometeorol.*, **15**, 1091–1116, doi:10.1175/JHM-D-13-0118.1.

Ray, A., J. Barsugli, K. Averyt, K. Wolter, M. Hoerling, N. Doesken, B. Udall, and R. Webb, 2008: *Climate change in Colorado: a synthesis to support water resources management and adaptation.*

Reba, M. L., D. Marks, T. E. Link, J. Pomeroy, and A. Winstral, 2014: Sensitivity of model parameterizations for simulated latent heat flux at the snow surface for complex mountain sites. *Hydrol. Process.*, **28**, 868–881, doi:10.1002/hyp.9619.

Reed, S., V. Koren, M. Smith, Z. Zhang, F. Moreda, D.-J. Seo, and A. DMIP Participants, 2004: Overall distributed model intercomparison project results. *J. Hydrol.*, **298**, 27–60, doi:10.1016/j.jhydrol.2004.03.031.

Reggiani, P., and J. Schellekens, 2003: Modelling of hydrological responses: the representative elementary watershed approach as an alternative blueprint for watershed modelling. *Hydrol. Process.*, **17**, 3785–3789, doi:10.1002/hyp.5167.

——, S. M. Hassanizadeh, M. Sivapalan, and W. G. Gray, 1999: A unifying framework for watershed thermodynamics: constitutive relationships. *Adv. Water Resour.*, **23**, 15–39, doi:10.1016/S0309-1708(99)00005-6.

Regonda, S. K., B. Rajagopalan, M. Clark, and J. Pitlick, 2005: Seasonal Cycle Shifts in Hydroclimatology over the Western United States. *J. Clim.*, **18**, 372–384, doi:10.1175/JCLI-3272.1.

Rigon, R., G. Bertoldi, and T. M. Over, 2006: GEOtop: A Distributed Hydrological Model with Coupled Water and Energy Budgets. *J. Hydrometeorol.*, **7**, 371–388, doi:10.1175/JHM497.1.

Romsdahl, R. J., and C. R. Pyke, 2009: What does decision support mean to the climate change research community? *Clim. Change*, **95**, 1–10, doi:10.1007/s10584-008-9538-2.

Rosero, E., Z.-L. Yang, T. Wagener, L. E. Gulden, S. Yatheendradas, and G.-Y. Niu, 2010: Quantifying parameter sensitivity, interaction, and transferability in hydrologically enhanced versions of the Noah land surface model over transition zones during the warm season. *J. Geophys. Res.*, **115**, 1–21, doi:10.1029/2009JD012035.

Sacks, W. J., D. S. Schimel, and R. K. Monson, 2007: Coupling between carbon cycling and climate in a high-elevation, subalpine forest: a model-data fusion analysis. *Oecologia*, **151**, 54–68, doi:10.1007/s00442-006-0565-2.

Samaniego, L., R. Kumar, and S. Attinger, 2010: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resour. Res.*, **46**, W05523, doi:10.1029/2008WR007327.

Sankarasubramanian, A., and R. M. Vogel, 2002: Annual hydroclimatology of the United States. *Water Resour. Res.*, **38**, 19–1–19–12, doi:10.1029/2001WR000619.

Sawicz, K., T. Wagener, M. Sivapalan, P. a. Troch, and G. Carrillo, 2011: Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA. *Hydrol. Earth Syst. Sci.*, **15**, 2895–2911, doi:10.5194/hess-15-2895-2011.

Schär, C., C. Frei, D. Lüthi, and H. C. Davies, 1996: Surrogate climate-change scenarios for regional climate models. *Geophys. Res. Lett.*, **23**, 669–672, doi:10.1029/96GL00265.

Seager, R., and Coauthors, 2007: Model projections of an imminent transition to a more arid climate in southwestern North America. *Science (80-.).*, **316**, 1181–1184, doi:10.1126/science.1139601.

——, M. Ting, C. Li, N. Naik, B. Cook, J. Nakamura, and H. Liu, 2012: Projections of declining surface-water availability for the southwestern United States. *Nat. Clim. Chang.*, **3**, 482–486, doi:10.1038/nclimate1787.

Seiller, G., F. Anctil, and C. Perrin, 2012: Multimodel evaluation of twenty lumped hydrological models under contrasted climate conditions. *Hydrol. Earth Syst. Sci.*, **16**, 1171–1189, doi:10.5194/hess-16-1171-2012.

Sellers, P., D. Randall, and G. Collatz, 1996: A revised land surface parameterization (SiB2) for atmospheric GCMs. Part I: Model formulation. *J. Clim.*, **9**, 676–705.

Shrestha, R., Y. Tachikawa, and K. Takara, 2006: Input data resolution analysis for distributed hydrological modeling. *J. Hydrol.*, **319**, 36–50, doi:10.1016/j.jhydrol.2005.04.025.

——, ——, and ——, 2007: Selection of scale for distributed hydrological modelling in ungauged basins. *Predictions in Ungauged Basins: PUB Kick-off (Proceedings of the PUB Kick-off meeting held in Brasilia, 20-22 November 2002)*, 290–297.

Singh, R., T. Wagener, K. van Werkhoven, M. E. Mann, and R. Crane, 2011: A trading-space-for-time approach to probabilistic continuous streamflow predictions in a changing climate – accounting for changing watershed behavior. *Hydrol. Earth Syst. Sci.*, **15**, 3591–3603, doi:10.5194/hess-15-3591-2011.

Sivapalan, M., G. Blöschl, L. Zhang, and R. Vertessy, 2003: Downward approach to hydrological prediction. *Hydrol. Process.*, **17**, 2101–2111, doi:10.1002/hyp.1425.

Skamarock, W. C., and Coauthors, 2008: *A Description of the Advanced Research WRF Version 3*. Boulder, Colorado, USA,.

Smith, M. B., and Coauthors, 2012: Results of the DMIP 2 Oklahoma experiments. *J. Hydrol.*, **418-419**, 17–48, doi:10.1016/j.jhydrol.2011.08.056.

Son, K., and M. Sivapalan, 2007: Improving model structure and reducing parameter uncertainty in conceptual water balance models through the use of auxiliary data. *Water Resour. Res.*, **43**, W01415, doi:10.1029/2006WR005032.

Steele-Dunne, S., P. Lynch, R. McGrath, T. Semmler, S. Wang, J. Hanafin, and P. Nolan, 2008: The impacts of climate change on hydrology in Ireland. *J. Hydrol.*, **356**, 28–45, doi:10.1016/j.jhydrol.2008.03.025.

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

Stewart, I. T., D. R. Cayan, and M. D. Dettinger, 2004: Changes in Snowmelt Runoff Timing in Western North America under a "Business as Usual" Climate Change Scenario. *Clim. Change*, **62**, 217–232.

Stewart, I. T., D. R. Cayan, and M. D. Dettinger, 2005: Changes toward earlier streamflow timing across western North America. *J. Clim.*, **18**, 1136–1155.

Surfleet, C. G., and D. Tullos, 2013: Uncertainty in hydrologic modelling for estimating hydrologic response due to climate change (Santiam River, Oregon). *Hydrol. Process.*, **27**, 3560–3576, doi:10.1002/hyp.9485.

——, D. Tullos, H. Chang, and I.-W. Jung, 2012: Selection of hydrologic modeling approaches for climate change assessment: A comparison of model scale and structures. *J. Hydrol.*, **464-465**, 233–248, doi:10.1016/j.jhydrol.2012.07.012.

Tang, Y., P. Reed, T. Wagener, and K. van Werkhoven, 2007: Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation. *Hydrol. Earth Syst. Sci.*, **11**, 793–817, doi:10.5194/hess-11-793-2007.

Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization. *Mon. Weather Rev.*, **136**, 5095–5115, doi:10.1175/2008MWR2387.1.

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

Trenberth, K., 1999: Conceptual framework for changes of extremes of the hydrological cycle with climate change. *Weather Clim. Extrem.*, **42**, 327–339.

Uhlenbrook, S., J. Seibert, C. Leibundgut, and A. Rodhe, 1999: Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure. *Hydrol. Sci. J.*, **44**, 779–797, doi:10.1080/02626669909492273.

USDA (U.S. Department of Agriculture), 1992: *Forest Land Distribution Data for the United States*.

——, 1994: *State Soil Geographic (STATSGO) Database – Data Use Information*. Misc. Pub.

VanderKwaak, J., and K. Loague, 2001: Hydrologic-response simulations for the R-5 catchment with a comprehensive physics-based model. *Water Resour. Res.*, **37**, 999–1013.

Vano, J. A., and D. P. Lettenmaier, 2014: A sensitivity-based approach to evaluating future changes in Colorado River discharge. *Clim. Change*, **122**, 621–634, doi:10.1007/s10584-013-1023-x.

——, T. Das, and D. P. Lettenmaier, 2012: Hydrologic Sensitivities of Colorado River Runoff to Changes in Precipitation and Temperature. *J. Hydrometeorol.*, **13**, 932–949, doi:10.1175/JHM-D-11-069.1.

——, and Coauthors, 2014: Understanding Uncertainties in Future Colorado River Streamflow. *Bull. Am. Meteorol. Soc.*, **95**, 59–78, doi:10.1175/BAMS-D-12-00228.1.

Vaze, J., D. A. Post, F. H. S. Chiew, J.-M. Perraud, N. R. Viney, and J. Teng, 2010: Climate non-stationarity – Validity of calibrated rainfall–runoff models for use in climate change studies. *J. Hydrol.*, **394**, 447–457, doi:10.1016/j.jhydrol.2010.09.018.

Velázquez, J. A., and Coauthors, 2013: An ensemble approach to assess hydrological models' contribution to uncertainties in the analysis of climate change impact on water resources. *Hydrol. Earth Syst. Sci.*, **17**, 565–578, doi:10.5194/hess-17-565-2013.

Verseghy, D., 1991: CLASS—A Canadian land surface scheme for GCMs. I. Soil model. *Int. J. Climatol.*, **11**, 111–133.

Vrugt, J. A., H. V. Gupta, L. A. Bastidas, W. Bouten, and S. Sorooshian, 2003a: Effective and efficient algorithm for multiobjective optimization of hydrologic models. *Water Resour. Res.*, **39**, 1214, doi:10.1029/2002WR001746.

——, ——, W. Bouten, and S. Sorooshian, 2003b: A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resour. Res.*, **39**, 1201, doi:10.1029/2002WR001642.

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

——, M. P. Clark, C. G. H. Diks, Q. Duan, and B. A. Robinson, 2006: Multi-objective calibration of forecast ensembles using Bayesian model averaging. *Geophys. Res. Lett.*, **33**, L19817, doi:10.1029/2006GL027126.

Wagener, T., P. Reed, K. van Werkhoven, Y. Tang, and Z. Zhang, 2009a: Advances in the identification and evaluation of complex environmental systems models. *J. Hydroinformatics*, **11**, 266–281, doi:10.2166/hydro.2009.040.

——, K. van Werkhoven, P. Reed, and Y. Tang, 2009b: Multiobjective sensitivity analysis to understand the information content in streamflow observations for distributed watershed modeling. *Water Resour. Res.*, **45**, 1–5, doi:10.1029/2008WR007347.

——, and Coauthors, 2010: The future of hydrology: An evolving science for a changing world. *Water Resour. Res.*, **46**, 1–10, doi:10.1029/2009WR008906. http://www.agu.org/pubs/crossref/2010/2009WR008906.shtml (Accessed August 6, 2012).

Wagner, P. D., P. Fiener, F. Wilken, S. Kumar, and K. Schneider, 2012: Comparison and evaluation of spatial interpolation schemes for daily rainfall in data scarce regions. *J. Hydrol.*, **464-465**, 388–400, doi:10.1016/j.jhydrol.2012.07.026.

Wayand, N. E., A. F. Hamlet, M. Hughes, S. I. Feld, and J. D. Lundquist, 2013: Intercomparison of Meteorological Forcing Data from Empirical and Mesoscale Model Sources in the North Fork American River Basin in Northern Sierra Nevada, California. *J. Hydrometeorol.*, **14**, 677–699, doi:10.1175/JHM-D-12-0102.1.

Van Werkhoven, K., T. Wagener, P. Reed, and Y. Tang, 2008: Characterization of watershed model behavior across a hydroclimatic gradient. *Water Resour. Res.*, **44**, 1–16, doi:10.1029/2007WR006271.

——, ——, ——, and ——, 2009: Sensitivity-guided reduction of parametric dimensionality for multi-objective calibration of watershed models. *Adv. Water Resour.*, **32**, 1154–1169, doi:10.1016/j.advwatres.2009.03.002.

Westra, S., M. Thyer, M. Leonard, D. Kavetski, and M. Lambert, 2014: A strategy for diagnosing and interpreting hydrological model nonstationarity. *Water Resour. Res.*, **50**, 5090–5113, doi:10.1002/2013WR014719.

Wigmosta, M., L. Vail, and D. Lettenmaier, 1994: A distributed hydrology-vegetation model for complex terrain. *Water Resour. Res.*, **30**, 1665–1679.

Wilby, R. L., 2005: Uncertainty in water resource model parameters used for climate change impact assessment. *Hydrol. Process.*, **19**, 3201–3219, doi:10.1002/hyp.5819.

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

Wilby, R. L., and S. Dessai, 2010: Robust adaptation to climate change. *Weather*, **65**, 180–185, doi:10.1002/wea.543.

Wilby, R. L., L. E. Hay, and G. H. Leavesley, 1999: A comparison of downscaled and raw GCM output: implications for climate change scenarios in the San Juan River basin, Colorado. *J. Hydrol.*, **225**, 67–91, doi:10.1016/S0022-1694(99)00136-5.

Wilby, R. L., J. Troni, Y. Biot, L. Tedd, B. C. Hewitson, D. M. Smith, and R. T. Sutton, 2009: A review of climate risk information for adaptation and development planning. *Int. J. Climatol.*, **29**, 1193–1215, doi:10.1002/joc.1839.

Wolf, A., K. Akshalov, N. Saliendra, D. A. Johnson, and E. A. Laca, 2006: Inverse estimation of Vc max , leaf area index, and the Ball-Berry parameter from carbon and energy fluxes. *J. Geophys. Res.*, **111**, D08S08, doi:10.1029/2005JD005927.

Wood, E. F., D. P. Lettenmaier, and V. G. Zartarian, 1992: A Land-Surface Hydrology Parameterization With Subgrid Variability for General Circulation Models. *J. Geophys. Res.*, **97**, 2717–2728, doi:10.1029/91JD01786.

——, and Coauthors, 2011: Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water. *Water Resour. Res.*, **47**, W05301, doi:10.1029/2010WR010090.

Xia, Y., and Coauthors, 2012a: Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *J. Geophys. Res.*, **117**, D03109, doi:10.1029/2011JD016048.

——, and Coauthors, 2012b: Continental-scale water and energy flux analysis and validation for North American Land Data Assimilation System project phase 2 (NLDAS-2): 2. Validation of model-simulated streamflow. *J. Geophys. Res.*, **117**, D03110, doi:10.1029/2011JD016051.

Xu, C., 1999: From GCMs to river flow: a review of downscaling methods and hydrologic modelling approaches. *Prog. Phys. Geogr.*, **23**, 229–249, doi:10.1177/030913339902300204.

Yang, D., T. Koike, and H. Tanizawa, 2003: Effect of precipitation spatial distribution on the hydrological response in the upper Tone River of Japan. *Int. Assoc. Hydrol. Sci. Publ.*, 194–202.

Yang, Z.-L., R. E. Dickinson, A. Robock, and K. Y. Vinnikov, 1997: Validation of the Snow Submodel of the Biosphere–Atmosphere Transfer Scheme with Russian Snow Cover and Meteorological Observational Data. *J. Clim.*, **10**, 353–373, doi:10.1175/1520-0442(1997)010<0353:VOTSSO>2.0.CO;2.

——, and Coauthors, 2011: The community Noah land surface model with multiparameterization options (Noah-MP): 2. Evaluation over global river basins. *J. Geophys. Res.*, **116**, 1–16, doi:10.1029/2010JD015140.

Yapo, P. O., H. V. Gupta, and S. Sorooshian, 1998: Multi-objective global optimization for hydrologic models. *J. Hydrol.*, **204**, 83–97, doi:10.1016/S0022-1694(97)00107-8.

Yilmaz, K. K., H. V. Gupta, and T. Wagener, 2008: A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model. *Water Resour. Res.*, **44**, 1–18, doi:10.1029/2007WR006716.

Zehe, E., and G. Blöschl, 2004: Predictability of hydrologic response at the plot and catchment scales: Role of initial conditions. *Water Resour. Res.*, **40**, W10202, doi:10.1029/2003WR002869.

APPENDIX A: Parameters included in hydrologic model calibration

The model parameters included in the calibration process (chapters 4, 5 and 6) are selected based on background sensitivity analysis performed for each hydrologic/land surface model. In this study, we use the Distributed Evaluation of Local Sensitivity Analysis (DELSA; Rakovec et al. 2014) method to quantify parameter sensitivity, using the root mean squared error between observed and simulated streamflow as the objective function. In DELSA, the assessment of parameter sensitivity is based on local gradients of the model performance index with respect to model parameters at multiple points throughout the parameter space. A number of soil, vegetation, runoff and snow parameters were considered in DELSA for each model: 17 for PRMS, 34 for VIC, 17 for Noah-LSM and 30 for Noah-MP.

Based on the sensitivity analysis results, the numbers of parameters calibrated are 8 for PRMS, 9 for VIC, 11 for Noah-LSM and 14 for Noah-MP. These parameters are listed in Tables A1-A4.

Table A.1: Summary of PRMS parameters considered for calibration.

Parameter	Description	Units	Distributed*	Calibration Range	
				min	max
jh_coef**	monthly Jensen-Haise air temperature coefficient	F	no	0.36	2.86
fastcoef_lin	linear flow routing coefficient for fast interflow	d^{-1}	no	0	10
fastcoef_sq	non-linear flow routing coefficient for fast interflow	$in^{-1}\ d^{-1}$	no	0	1.25
pref_flow_den	decimal fraction of the soil zone available for preferential flow	-	no	0	5
soil_moist_max	maximum volume of water per unit area in the capillary reservoir	in	yes	**0**	**2.87**
snarea_curve	snow area-depletion curve values	-	yes	**0**	**1**
tmax_allsnow	monthly maximum air temperature at which precipitation is all snow for the HRU	F	no	-10	40
tmax_allrain	monthly minimum air temperature at an HRU that results in all precipitation during a day being rain	F	no	0	90

*If the parameter is distributed in space, its calibration is performed on the basis of multipliers. Although description and units refer to actual parameters, the values in bold represent the multiplier values (instead of actual parameter values).

**The range is provided for a multiplier applied to each monthly value of jh_coef.

Table A.2: Summary of VIC parameters considered for calibration.

Parameter	Description	Units	Distributed*	Calibration Range	
				min	max
binfilt	variable infiltration curve parameter	-	no	0.001	0.4
Ds	fraction of Dsmax where non-linear baseflow begins	-	no	10^{-5}	1
Dsmax	maximum velocity of baseflow	mm d^{-1}	yes	**0.01**	**2**
Ws	fraction of maximum soil moisture where non-linear baseflow occurs	-	no	9×10^{-4}	1
depth2	thickness of soil layer 2	m	yes	**0.5**	**6**
depth3	thickness of soil layer 3	m	yes	**0.5**	**6**
newalb	new snow albedo	-	no	0.7	0.99
albaa	base in snow albedo function (accumulation)	-	no	0.88	0.99
albtha	base in snow albedo function (melt)	-	no	0.66	0.98

*If the parameter is distributed, its calibration is performed on the basis of multipliers. Although description and units refer to actual parameters, the values in bold represent the multiplier values (instead of actual parameter values).

Table A.3: Summary of Noah-LSM parameters considered for calibration.

Parameter	Description	Units	Distributed*	Calibration Range	
				min	max
maxsmc	soil porosity	m^3 m^{-3}	yes	**0.88**	**1.18**
satdk	saturated soil hydraulic conductivity	m s^{-1}	yes	**0.41**	**1.39**
quartz	soil quartz content	-	yes	**0.29**	**1.37**
refdk	used with refkdt to compute runoff parameter kdt	-	no	2×10^{-8}	2×10^{-4}
fxexp	bare soil evaporation exponent	-	no	0.2	4
refkdt	surface runoff parameter	-	no	0.1	10
czil	Zilintikevich parameter	-	no	0.05	8
cmcmax	maximum canopy water capacity used in canopy evaporation	m	no	5×10^{-5}	2
rsmax	maximum stomatal resistance	m^{-1} s	no	2	10
lvcoef	Livneh coefficient for adjusting snow albedo	-	no	0	1
slope	linear coefficient used to compute subsurface runoff	-	no	0.2	1

*If the parameter is distributed, its calibration is performed on the basis of multipliers. Although description and units refer to actual parameters, the values in bold represent the multiplier values (instead of actual parameter values).

Table A.4: Summary of Noah-MP parameters considered for calibration.

Parameter	Description	Units	Distributed*	Calibration Range	
				min	max
maxsmc	soil porosity	$m^3\ m^{-3}$	yes	**0.88**	**1.18**
wind_rp	empirical canopy wind parameter	m^{-1}	yes	**0.7**	**1.3**
slope_ps	slope of conductance-to-photosynthesis relationship	-	yes	**0.7**	**1.3**
laimss	monthly leaf area index, one sided (Spring/Summer)	-	yes	**0.7**	**1.3**
fff	runoff decay factor	m^{-1}	no	1	8
rsbmx	baseflow coefficient	$mm\ s^{-1}$	no	0.5	8
timean	grid cell mean topographic index	-	no	7.35	13.65
mexp	exponent used in the curves for the melting season	-	no	0.5	3
z0sno	snow surface roughness length	m	no	0.0002	0.02
snow_iwc	liquid water holding capacity for snowpack	$m^3\ m^{-3}$	no	0.02	0.06
swemx	new snow mass to fully cover old snow	mm	no	0.1	20
albmin	minimum snow albedo	-	no	0.44	0.66
albmax	maximum snow albedo	-	no	0.68	1
albdecay	exponent in snow decay albedo relationship	h^{-1}	no	0.001	0.1

*If the parameter is distributed, its calibration is performed on the basis of multipliers. Although description and units refer to actual parameters, the values in bold represent the multiplier values (instead of actual parameter values).