Implications of the Methodological Choices for Hydrologic Portrayals of Climate Change over the Contiguous United States: Statistically Downscaled Forcing Data and Hydrologic Models

NAOKI MIZUKAMI,* MARTYN P. CLARK,* ETHAN D. GUTMANN,* PABLO A. MENDOZA,* ANDREW J. NEWMAN,* BART NIJSSEN,+ BEN LIVNEH,#,++ LAUREN E. HAY,@ JEFFREY R. ARNOLD,& AND LEVI D. BREKKE**

* *National Center for Atmospheric Research, Boulder, Colorado*
+ *University of Washington, Seattle, Washington*
# *Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, Colorado*
@ *U.S. Geological Survey, Denver, Colorado*
& *U.S. Army Corps of Engineers, Seattle, Washington*
** *U.S. Bureau of Reclamation, Denver, Colorado*
++ *Department of Civil, Environmental, and Architectural Engineering, University of Colorado Boulder, Boulder, Colorado*

(Manuscript received 1 October 2014, in final form 16 June 2015)

ABSTRACT

Continental-domain assessments of climate change impacts on water resources typically rely on statistically downscaled climate model outputs to force hydrologic models at a finer spatial resolution. This study examines the effects of four statistical downscaling methods [bias-corrected constructed analog (BCCA), bias-corrected spatial disaggregation applied at daily (BCSDd) and monthly scales (BCSDm), and asynchronous regression (AR)] on retrospective hydrologic simulations using three hydrologic models with their default parameters (the Community Land Model, version 4.0; the Variable Infiltration Capacity model, version 4.1.2; and the Precipitation–Runoff Modeling System, version 3.0.4) over the contiguous United States (CONUS). Biases of hydrologic simulations forced by statistically downscaled climate data relative to the simulation with observation-based gridded data are presented. Each statistical downscaling method produces different meteorological portrayals including precipitation amount, wet-day frequency, and the energy input (i.e., shortwave radiation), and their interplay affects estimations of precipitation partitioning between evapotranspiration and runoff, extreme runoff, and hydrologic states (i.e., snow and soil moisture). The analyses show that BCCA underestimates annual precipitation by as much as −250 mm, leading to unreasonable hydrologic portrayals over the CONUS for all models. Although the other three statistical downscaling methods produce a comparable precipitation bias ranging from −10 to 8 mm across the CONUS, BCSDd severely overestimates the wet-day fraction by up to 0.25, leading to different precipitation partitioning compared to the simulations with other downscaled data. Overall, the choice of downscaling method contributes to less spread in runoff estimates (by a factor of 1.5–3) than the choice of hydrologic model with use of the default parameters if BCCA is excluded.

1. Motivation and scope

A number of observational studies have reported changes in the timing and magnitude of seasonal streamflow patterns across the contiguous United States (CONUS) that may be due to a changing climate (e.g., Regonda et al. 2005; Stewart et al. 2005; Déry et al. 2009; Luce and Holden 2009; Fritze et al. 2011; Sagarika et al. 2014). Responding to the observed hydroclimatic changes, water resource managers are faced with the challenge to predict changes in streamflow and other hydrologic components for long-term water planning. Many hydrologic projections have been made using uncoupled hydrologic simulations forced by downscaled climate model outputs (e.g., Christensen et al. 2004; Dettinger et al. 2004; Christensen and Lettenmaier 2007; Hayhoe et al. 2007; Cayan et al. 2010; Elsner et al. 2010; U.S. Bureau of Reclamation 2011; Bennett et al. 2012; Sharma and Babel 2013).

Uncoupled, offline, or stand-alone hydrologic simulations, typically used to produce the hydrologic

Corresponding author address: Naoki Mizukami, NCAR, P.O. Box 3000, Boulder, CO 80307.
E-mail: mizukami@ucar.edu

DOI: 10.1175/JHM-D-14-0187.1

projections, refer to simulations in which the hydrologic model is not interactively coupled to the atmospheric model. In the course of implementing the uncoupled simulations, modelers confront several methodological choices. In addition to the choice of the global climate models (GCMs) and emission scenarios, the spatial downscaling method used to process the GCM outputs and the hydrologic model configuration (e.g., model structures and model parameters) are important sources of uncertainty in the estimated hydrologic projections (Mendoza et al. 2015a,b; Vano et al. 2014). This study focuses on the latter two methodological choices.

The main objectives of this study are to present 1) how different downscaling methods propagate into different hydrologic portrayals, such as the partitioning between evapotranspiration (ET) and runoff, hydrologic states (snowpack and soil moisture), and runoff characteristics (e.g., high and low runoff); and 2) how sensitivity of hydrologic simulation to choice of downscaling method differs among hydrologic models with varying model complexity. The scope of this study is toward continental-domain hydrologic assessments. Therefore, this paper focuses on statistical downscaling (SD) techniques that have been used for the continental-scale assessments of hydrologic projection. Furthermore, this study includes four statistical downscaling methods used by Gutmann et al. (2014), who evaluated differences in the downscaled precipitation data, and three hydrologic models to evaluate interplays between two methodological choices.

The paper is organized as follows. To provide the background of our study, section 2 describes methodological choices of spatial downscaling of climate model output and the hydrologic models commonly used for large-scale hydrologic assessments. Section 3 describes the meteorological forcing data and hydrologic models used for this study. Section 4 presents intercomparison of statistically downscaled climate data and then examines how their differences in meteorological characteristics affect hydrologic simulations with three hydrologic models. Comparison of hydrologic simulations include the partitioning of runoff and ET, hydrologic states (snow water equivalent and soil moisture), and runoff extremes. Finally, section 5 provides conclusions.

2. Background

a. Spatial downscaling approaches

Hydrologic processes are sensitive to local climate. The effects of local climate variability such as topography are not captured by GCM coarse resolutions (typically >50 km). Therefore, spatial downscaling is a necessary step for hydrologic modeling that uses GCM outputs. There are two main types of downscaling methods: 1) dynamical downscaling where the GCM output is used as boundary and initial conditions to drive a regional climate model at a finer spatial resolution and 2) SD where statistical relations are developed between the local climate variables and GCM output and the relations are then subsequently used to obtain the local meteorological variables from the GCM output. Between dynamical and statistical downscaling techniques, the SD method is an attractive choice for large-scale hydrologic assessments because of its computational efficiency. Also, the SD method corrects for the bias contained in the surface meteorological fields from GCMs, which is difficult to remove in dynamical downscaling (Nicholas and Battisti 2012).

The computational efficiency of SD is crucial for large-scale modeling projects because multiple GCMs with several emission scenarios require repetitive downscaling of hundreds of GCM simulations. One such practical application demonstrating the capability to statistically downscale a large ensemble of climate outputs at a continental scale is detailed in Brekke et al. (2014), a work developed by a collaboration of federal and nonfederal entities, including the U.S. Bureau of Reclamation, U.S. Army Corps of Engineers, and others (Maurer et al. 2014). The partners initially produced and served a 112-member ensemble of phase 3 of the Coupled Model Intercomparison Project (CMIP3; Meehl et al. 2005) climate projections statistically downscaled to 1/8° over the CONUS (Maurer et al. 2007), subsequently translated into hydrologic projections over the western United States (U.S. Bureau of Reclamation 2011). Similar efforts were then applied to phase 5 of the Coupled Model Intercomparison Project (CMIP5) climate projections (Taylor et al. 2012), producing a 234-member ensemble of statistically downscaled climate projections with a 97-member subset translated into hydrology over the CONUS (Maurer et al. 2014; Brekke et al. 2014).

Although many sophisticated SD techniques have been developed, ranging from weather typing and neural networks to weather generators [see Wilby et al. (2004) and Maraun et al. (2010) for further details], the following rather simple techniques are widely used by the water resources community for hydrologic assessments (Maurer et al. 2010; U.S. Bureau of Reclamation 2011; Hanson et al. 2012; Miller et al. 2013; Hay et al. 2014):

- bias-corrected spatial disaggregation (BCSD; Wood et al. 2004) is a method typically applied to monthly values of precipitation and temperature, after which temporal disaggregation is performed to generate

TABLE 1. Statistical downscaling techniques.

SD methods	Descriptions of the procedures	Reference
BCCA	Perform quantile mapping of raw climate model output to spatially aggregated observation-based grid outputs for bias correction, then on a given day select 30 historical analog days based on spatial similarity from coarsened observed grids to develop a linear model that produces the coarse model grid. Finally, apply the developed linear model to finer-resolution historical observation grids to produce downscaled dataset.	Maurer et al. (2010)
BCSDm	Perform a quantile mapping of raw climate model output to spatially aggregated observation-based grid outputs for bias correction and then use a linear interpolation for spatial disaggregation of bias-corrected climate model output. This process is performed with monthly values. Temporal disaggregation is performed by scaling downscaled monthly values to match historical monthly values (month is randomly selected from the historical period).	Wood et al. (2004)
BCSDd	The same as BCSDm except using daily time step and no temporal disaggregation.	Thrasher et al. (2012)
AR	Perform a bilinear interpolation of raw climate model output to finer resolution and then develop a linear regression between ordered observations and spatially disaggregated climate model output for each month.	Dettinger et al. (2004); Stoner et al. (2013)

daily time series (hereafter BCSDm). The BCSD method has also been applied directly to daily GCM output (BCSDd; Thrasher et al. 2012);

- bias-corrected constructed analog (BCCA; Maurer et al. 2010); and
- asynchronous regression (AR; Dettinger et al. 2004; Stoner et al. 2013).

These methods broadly consist of bias correction and spatial disaggregation processes, and Table 1 describes procedures of each SD method in detail. While some communities might refer to these methods as bias correction techniques, the water resources community, for whom this is most relevant, has traditionally referred to them as SD methods; as such, we retain that naming scheme for the remainder of this paper.

Gutmann et al. (2014) evaluated the four SD techniques documented in Table 1 in terms of how they reproduce historical precipitation amounts and key hydrologically relevant attributes, including wet-day fraction, dry (wet) spell length, and precipitation extremes. They showed that each SD method produces different portrayals of the hydrologically relevant attributes as well as precipitation amounts. Accurate portrayals of such hydrologically relevant attributes of downscaled climate data are important when other types of meteorological forcing variables, such as radiation fluxes and humidity, need to be estimated for process-based hydrologic models. To date, the SD techniques reviewed here are typically applied to precipitation and temperature, but not to other meteorological variables required for hydrologic models, though some more sophisticated techniques have been used to downscale all the surface meteorological variables at a daily step (Wilby et al. 1998). A common method to obtain downscaled radiative and humidity data from precipitation and temperature data is to use an empirical algorithm such as the Mountain Microclimate Simulation Model (MTCLIM; Hungerford et al. 1989; Kimball et al. 1997; Thornton and Running 1999; Bohn et al. 2013) in conjunction with the hydrologically relevant attributes derived from downscaled temperature and precipitation. Although precipitation is the primary driver of the hydrological cycle, several studies have illustrated the impact of other climate forcing data, including radiative fluxes, humidity, and wind speed on the hydrologic simulations (e.g., Nasonova et al. 2011; Haddeland et al. 2012; Feld et al. 2013; Pierce et al. 2013; Wayand et al. 2013; Mizukami et al. 2014). Among these variables, errors in shortwave radiation may have the largest impact on hydrologic simulations. Mizukami et al. (2014) showed that large differences between two shortwave radiation estimates—one from climate model reanalysis data and one estimated using MTCLIM with observed temperature and precipitation data—can be the main source of difference in runoff estimates (timing and magnitude) in snowmelt-dominated areas.

The sensitivity of hydrologic simulations to downscaling methods has been examined by many studies for specific basins or regions (e.g., Wilby et al. 2000; Crane et al. 2002; Hay and Clark 2003; Wood et al. 2004; Maurer et al. 2010; Teutschbein et al. 2011). Most of the studies present a comparison between dynamical and statistical approaches for hydrologic model applications. For example, Wood et al. (2004) illustrated that hydrologic responses obtained with BCSDm were comparable to the results from use of dynamic downscaling techniques over the Colorado River basin and the Pacific Northwest. Other studies (e.g., Wilby et al. 2000; Hay and Clark 2003) also confirmed that simpler statistical methods are comparable to computationally expensive dynamical methods in terms of runoff simulations. Given numerous statistical methods developed in recent years, more studies have turned attention to the impact

of SD choice on hydrologic simulations (e.g., Maurer et al. 2010; Teutschbein et al. 2011). However, those studies on SD method impact on hydrologic simulations are based on the use of a single hydrologic model.

b. Hydrologic modeling

The choice of hydrologic model can have an impact on the uncertainty of hydrologic assessments. The uncertainty resulting from the choice of hydrologic model originates mainly from differences in hydrologic model structures (i.e., a set of governing equations representing hydrologic processes, spatial representations of the snow and/or soil column, and land cover) and model parameters (Clark et al. 2015a,b). More recent climate change impact studies use multiple hydrologic models to generate ensemble simulations and evaluate the impact of model choice on hydrologic estimates (e.g., Bastola et al. 2011; Najafi et al. 2011; Poulin et al. 2011; Miller et al. 2012; Surfleet et al. 2012; Vano et al. 2012; Exbrayat et al. 2014; Mendoza et al. 2015a,b). For example, Najafi et al. (2011) performed multiple hydrologic simulations with four hydrologic models of varying complexity forced by a combination of eight GCMs and two emission scenarios in a catchment in Oregon, showing overall larger hydrologic model uncertainty during the low-flow period. These studies highlight the need for multimodel approaches when projecting water availability and uncertainty under climate change. Vano et al. (2012) compared five hydrologic or land surface models with their default parameters to assess the sensitivity of runoff to changes in temperature and precipitation over the Colorado River basin. They found runoff sensitivity to precipitation change, as expressed by elasticity (a ratio of percent change in annual runoff to percent change in annual precipitation), ranges approximately from 2 to 6 at Lees Ferry among the models examined. They illustrated not only differences in the modeled water balance but also hydrologic sensitivity to climate change due to the model choice. Mendoza et al. (2015a) further examined how interplay between the choice of hydrologic model and their model parameters affects assessment of climate change impact on runoff using four hydrologic models over three Colorado headwater basins. The intermodel differences in the estimated runoff changes can be reduced if each model is calibrated compared to the use of the default model parameters, but still remain larger than climate change signals (i.e., runoff changes due to climate change).

3. Forcing data and hydrologic models

This section describes procedures of hydrologic modeling, including climate forcing data or SD datasets, selected hydrologic models, their model parameters, and spinup method to determine initial hydrologic states.

a. Climate datasets

We use four instances of daily precipitation P and maximum and minimum air temperature (T_{max} and T_{min}, respectively) statistically downscaled to a 1/8° grid (~12-km spatial resolution). The SD techniques considered are BCCA, BCSDd, BCSDm, and AR, which were also evaluated by Gutmann et al. (2014). Coarse-resolution climate data from the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalyses (Kalnay et al. 1996; ~1.9° Gaussian grid equivalent to ~210 km) was downscaled to 1/8° resolution over the CONUS. The observation-based gridded climate dataset also at 1/8° resolution developed by Maurer et al. (2002, hereafter M02) was used to train each downscaling method for the period 1979–99 and then to generate downscaled precipitation and temperature data for the same period. M02 has been used for many studies on hydrologic impact of climate change (e.g., Christensen et al. 2004; Hayhoe et al. 2004; Maurer 2007; Brekke et al. 2014). Four SD datasets and the M02 dataset were used as forcing datasets for the hydrologic models described in section 3b.

For each SD dataset and M02, meteorological forcing variables in addition to precipitation and air temperature (i.e., specific humidity, surface pressure, short- and longwave radiation, and wind) required for the offline hydrologic simulations are generated as follows. Derivations of daily short- and longwave radiation and humidity use MTCLIM, version 4.3 (Thornton et al. 2000), in which theoretical daily extraterrestrial insolation is reduced by daily transmittance estimated empirically with diurnal temperature range (DTR) and humidity, and then with a further 25% reduction applied on a day with precipitation (Thornton and Running 1999). Dewpoint temperature is also empirically estimated as a function of T_{min} and P (Kimball et al. 1997). Since daily shortwave radiation and humidity are interrelated in the algorithm (i.e., shortwave radiation is a function of vapor pressure, and vapor pressure is a function of shortwave radiation), both variables were iteratively derived in the MTCLIM algorithm. Estimation of longwave radiation uses an empirical equation developed by Idso (1981) that uses air temperature and vapor pressure to estimate atmospheric emissivity. Surface pressure was computed by adjusting the standard sea level temperature and pressure (288 K and 1013.25 kPa, respectively) with a hydrostatic equation at gridbox elevation. Wind speed for all the datasets is taken from the 2.5° × 2.5° NCEP–NCAR

reanalyses (Kalnay et al. 1996), which is linearly interpolated to the 1/8° grid.

Since model simulations are performed at hourly time steps, temporal disaggregation was made for each daily variable. Daily precipitation values are uniformly distributed throughout the day, and temperature is disaggregated using spline interpolation of the daily maximum and minimum values. Daily total shortwave radiation is disaggregated using the temporal pattern of solar zenith angle for each calendar day. Daily dewpoint is linearly disaggregated between successive days and then converted to specific humidity with subdaily temperature estimates. Surface pressure and wind speed were assumed to be constant throughout the day.

Errors in the phase of precipitation (rain vs snow) can have a large impact on hydrologic simulations, especially over mountain ranges in maritime climate (e.g., Sierra Nevada) in the transient snow zone (Mizukami et al. 2013). For this study, the precipitation type was determined based on air temperature for all the models. All the models use consistent threshold temperatures for snow (0°C), below which all the precipitation falls as snow, and for rain (2°C), above which all the precipitation is rain. These threshold temperature values are within observed precipitation type transition temperatures observed in past studies (e.g., Dai 2008). Between two temperature thresholds, a linear interpolation is used to estimate the percentages of each precipitation type.

b. *Hydrologic model simulations*

Hydrologic simulations were performed over the CONUS using three hydrologic models: 1) the Community Land Model, version 4.0 (CLM4; Oleson et al. 2010); 2) the Variable Infiltration Capacity model, version 4.1.2 (VIC), with energy balance mode (Liang et al. 1994, 1996); and 3) the Precipitation–Runoff Modeling System, version 3.0.4 (PRMS; Leavesley et al. 1983; Leavesley and Stannard 1995). All the models explicitly simulate moisture fluxes (base flow, surface flow, canopy and soil evaporation, and transpiration) and states (soil moisture and snowpack), which are evaluated in this study.

CLM is a land component of the Community Earth System Model (CESM; Hurrell et al. 2013), which has contributed to various global climate change studies such as CMIP5 (Taylor et al. 2012). As a part of CESM, CLM is typically coupled with atmospheric and ocean models to simulate land–atmosphere–ocean interactions, but this study uses offline CLM simulation to focus only on land water and energy processes driven by the meteorological forcings (Lawrence et al. 2011). VIC and PRMS are typically run in an offline mode (i.e., uncoupled to a climate model). VIC has been extensively used for many hydrologic climate assessments over the CONUS, such as climate change (e.g., Christensen and Lettenmaier 2007; U.S. Bureau of Reclamation 2011; Brekke et al. 2014) and drought monitoring (Sheffield et al. 2004, 2012). PRMS has been used to assess the hydrologic response of different emission scenarios for the twenty-first century in selected basins from different hydroclimatic regions across the United States (Hay et al. 2011). Some examples of hydroclimate studies with PRMS include the effects of scale on hydrology (Battaglin et al. 2011), trends in snowfall and groundwater recharge (Bjerklie et al. 2011), impacts on the growing season (Christiansen et al. 2011), effects of baseline conditions (Koczot et al. 2011), changes in spring snowpack (Mastin et al. 2011), statistical comparisons of watershed-scale response (Risley et al. 2011), hydrologic effects of urbanization and climate change (Viger et al. 2011), and impacts on the 1.5-yr flood flow (Walker et al. 2011). For this study, these three models were run with identical spatial resolution (i.e., 1/8° resolution) at hourly time steps (except PRMS, which uses a daily time step).

Brief descriptions of the hydrologic process representations used in each model are given in Table 2. Also, Fig. 1 depicts differences in models' spatial representation of vegetation, snowpack, and subsurface. For example, CLM and VIC use layer structures (15 soil layers and unconfined aquifer for CLM and 3 layers for VIC) while PRMS uses conceptual reservoirs for subsurface soil structure. The horizontal vegetation representation is similar between VIC and CLM and vegetation root depth is prescribed per vegetation tile for both models, but PRMS uses average vegetation representation and thus vegetation parameters.

To evaluate the hydrologic simulations affected by bias of precipitation and temperature solely due to SD techniques, model simulations are conducted for the 19-yr period from the beginning of water year (WY) 1980 through the end of WY 1999 (the water year starts on 1 October and ends on 30 September of the following year). This model simulation period is the same as the training period used for SD methods. Gutmann et al. (2014) showed that the NCEP–NCAR data contain large changes in precipitation between the periods of WYs 1980–99 (SD calibration period) and WYs 2000–08 (SD validation period) over the eastern part of the CONUS, which are not seen in M02 data, potentially dominating precipitation error patterns over the SD validation period.

1) MODEL SPINUP

The model spinup is an important part of the modeling procedure to obtain equilibrium hydrologic states (soil moisture storage, etc.). The models were initialized

TABLE 2. Process components of hydrologic models.

Hydrologic models	Snow accumulation and ablation	Canopy water	ET	Surface runoff and infiltration	Percolation	Base flow	References
CLM	Five-snowpack-layer energy and mass balance (Anderson 1976; Jordan 1991)	Interception as an exponential function of LAI and stem area index, throughfall, and dripping due to excess canopy water capacity, and canopy evaporation	Monin–Obukhov similarity theory	Direct flow from saturation and infiltration excess (exceeding max infiltration capacity) at top soil layer (Niu et al. 2005)	1D Richards equation	Exponential function of water table (Niu et al. 2005)	Oleson et al. (2010)
VIC	Two-snowpack-layer energy and mass balance (Andreadis et al. 2009)	Interception (max canopy hold capacity is 0.2 times LAI), throughfall, and canopy evaporation	Potential ET using Penman–Monteith and further scaled by resistance (aerodynamic and canopy)	Direct runoff from saturation and infiltration excess at top two soil layers. Infiltration is given by an infiltration capacity function (Liang et al. 1996)	1D Richards equation	Linear function below a soil moisture threshold and nonlinear above that threshold at the bottom soil layer (Arno model)	Liang et al. (1994, 1996)
PRMS	Two-snowpack-layer energy and mass balance	Interception (max canopy hold capacity depends on vegetation type), throughfall, and canopy evaporation	Potential ET using Jensen–Haise	Hortonian flow due to infiltration excess and interflow due to saturation excess from preferential-flow reservoir	Use of max soil moisture (threshold) for reservoirs to move soil moisture between reservoirs	A linear function of groundwater storage (gravity reservoir)	Leavesley et al. (1983); Leavesley and Stannard (1995)

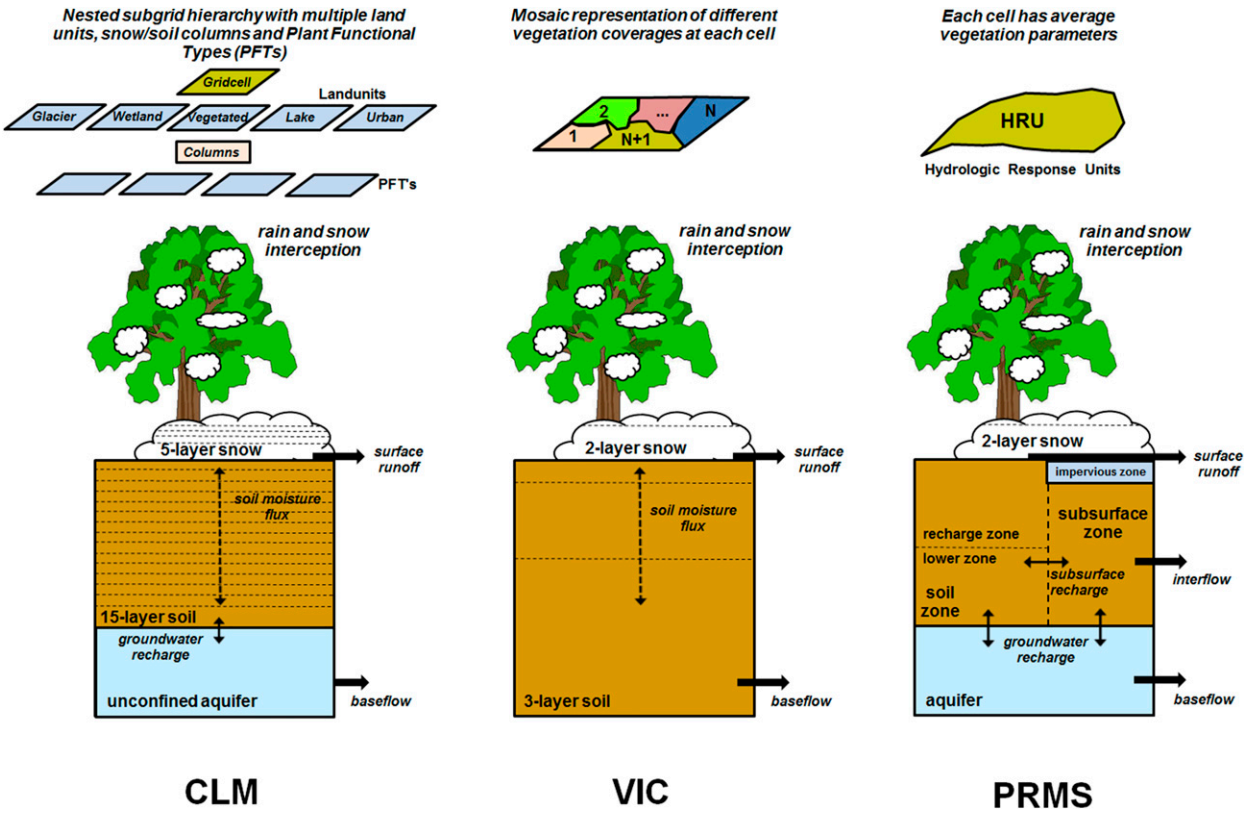

FIG. 1. Comparison of process conceptualizations across models: CLM, VIC, and PRMS.

with each SD dataset by cycling through the first 10-yr period (WYs 1980–90) five times with the respective SD datasets. The model spinup time is dependent on models (Cosgrove et al. 2003). For example, CLM has the unconfined aquifer and requires the longest spinup time to reach equilibrium for water-table depth and water storage in an unconfined aquifer, particularly in dry areas such as the southwestern United States. To reduce the spinup time for CLM, we developed an empirical linear relation between annual mean precipitation (WYs 1979–2008) and soil moisture states (water table and soil moisture) after 145 year simulations over the UCO region (see Fig. 2 for explanation of region acronyms) with the M02 dataset (five cycles of WYs 1979–2008) to obtain the initial estimates of soil moisture states prior to the CLM spinup simulations. The analysis excludes the first water year to further preclude spinup effects.

2) MODEL PARAMETERS

This study used the default model parameter fields (CLM and PRMS) and partially calibrated parameters (VIC) obtained from previous modeling efforts as described in the following subsections. Calibration of the model parameters was not performed in this study despite its potential reduction of intermodel spread of hydrologic simulations as demonstrated by Mendoza et al. (2015a). There were two primary reasons why the models were not calibrated:

1) There are no well-established methods to calibrate the parameters fields consistently across a large scale (i.e., CONUS). Basin-by-basin calibration done by the previous large-scale studies (Hay et al. 2011; Livneh and Lettenmaier 2012; Sheffield et al. 2012; Xia et al. 2012; Brekke et al. 2014) faces unresolved issues on spatial transferability of the parameters calibrated for selected basins to ungauged basins.
2) A traditional calibration optimizes the parameters given calibration forcing data; therefore, calibrated parameters with particular forcing data are not necessarily optimal for other forcing data with different SD methods. This was illustrated by Elsner et al. (2014).

Moreover, this study replicates current practices of continental-scale hydrologic assessments such as recent CMIP3 and CMIP5 hydrologic projection assessments by the U.S. Bureau of Reclamation (U.S. Bureau of Reclamation 2011; Brekke et al. 2014), who used the best available model parameter fields. In both CMIP

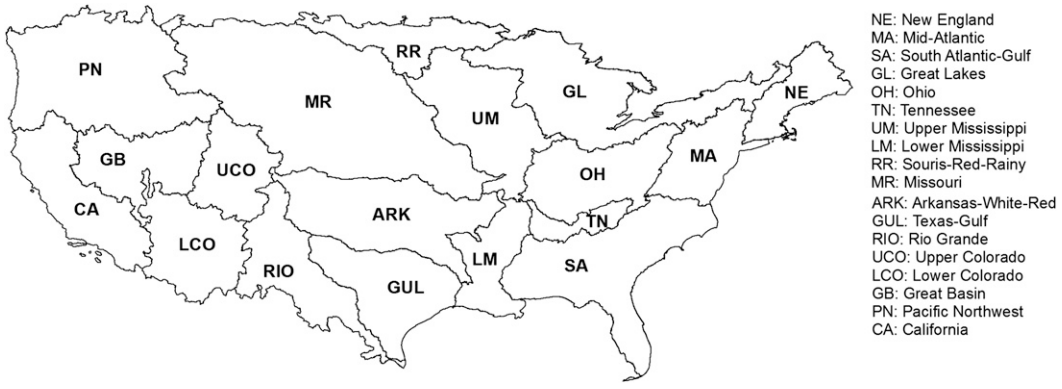

FIG. 2. Map of CONUS with 18 HUC regions.

hydrologic projection assessments with VIC, the calibrations of many VIC soil parameters, particularly the parameters related to conceptual hydrologic process representation, were performed for selected basins across the CONUS (with a focus on the western United States); therefore, it is noted that there is discontinuity of the spatial parameter fields at basin boundaries.

(i) CLM

CLM characterizes the land surface within a grid box by fractions of five different land covers (vegetation, wetland, lake, glacier, and urban), and vegetated cover was further split into 16 possible vegetation and/or forest types, referred to as plant functional types (PFTs; see Fig. 1). Vegetation structures [e.g., leaf area index (LAI), stem area index, and canopy top and bottom heights], canopy optical properties (e.g., reflectance and transmittance), and aerodynamic properties (e.g., roughness length and displacement height) are predefined for each PFT (Oleson et al. 2010). Water and energy fluxes are computed for each PFT and the fluxes from and to a grid box are computed by taking the area-weighted average over all PFTs. The PFT distribution used for this study is based on a global 0.5° PFT grid developed by Lawrence and Chase (2007). Oleson et al. (2010) provides complete details on land surface, subsurface, and snow hydrology parameterizations.

(ii) VIC

A single set of soil parameters related to hydraulic and thermal properties were uniquely assigned to each grid box, while the representation of land cover uses multiple tiles inside a grid box to characterize heterogeneity of land cover; therefore, each tile has unique vegetation parameters such as LAI. VIC soil and vegetation parameters were provided by phase 2 of the North American Land Data Assimilation (NLDAS-2; Xia et al. 2012). As in the CMIP5 hydrologic projection assessments, the VIC parameters used in NLDAS-2 were calibrated for selected river basins inconsistently across the CONUS domain. However, the VIC soil parameter set is one of the latest CONUS-wide fields available for the CONUS-domain VIC simulations.

(iii) PRMS

The PRMS simulations used the spatially distributed parameters that were recently developed by the U.S. Geological Survey (USGS), namely, the USGS Geospatial Fabric (Viger 2014). This geospatial dataset includes the default parameters for each USGS Hydrologic Response Unit (HRU) over the CONUS. The USGS HRU parameters were area-weighted averaged to 1/8° grid box.

4. Results and discussion

This section presents comparisons of the meteorological forcing data, including four SD data and M02, and hydrologic simulations produced with three models forced by each climate dataset. In total, there are 15 sets of the hydrologic simulations from the combinations of three models and five forcing datasets (four downscaled forcing datasets and M02). In this section, a simulation based on a particular combination of the model and forcing data is indicated as model forcing; for example, CLM-M02 indicates a simulation based on CLM forced with M02 data. For clarity, the following terminologies are also used to indicate a particular intercomparison. Intermodel difference refers to the difference in the hydrologic simulations due to different models but forced by the same dataset. Interforcing difference indicates comparison of the hydrologic simulations among different climate datasets. The term bias indicates difference in the forcing variables and associated simulations between SD methods and M02. Last, Fig. 2 shows a

CONUS map with hydrologic unit codes (HUCs) regions, which are referred to throughout this section.

a. Comparison of statistically downscaled forcing data

Although all the SD methods used in this paper perform bias correction of coarse-resolution climate data (i.e., NCEP–NCAR reanalyses) in a similar manner, as described in Table 1, unique downscaling processes for each SD method produce different biases in the downscaled precipitation data. Figure 3 displays annual total precipitation and annual mean daily temperature (i.e., $T_{\min}$ and $T_{\max}$) of M02 (Fig. 3a) and biases of the downscaled dataset for each SD method (Figs. 3b–e). As shown in Fig. 3, BCCA-downscaled precipitation severely underestimates the mean annual precipitation across the CONUS by a large amount (mean bias over the CONUS $= -250\,\mathrm{mm\,yr^{-1}}$). One of the reasons for the large dry bias is that the method tends to smooth out and decrease large precipitation events (Gutmann et al. 2014; Pierce et al. 2014). BCCA-downscaled precipitation is a linear combination of fine-resolution precipitation fields on 30 historical analog days selected based on the domainwide root-mean-square differences (Table 1). Therefore, a coarse climate model precipitation field that has a heavy precipitation in one area but is clear in another area can be matched to analog days that have low precipitation in both areas (Pierce et al. 2014). As a result, selected analog days tend to have smoother spatial patterns, contributing to a decrease of larger precipitation events. The dry bias of BCCA appears to increase as the domain size increases based on two studies (Maurer et al. 2010; Gutmann et al. 2014). BCSDd, BCSDm, and AR have a much smaller CONUS average bias of 8, 3, and $-10\,\mathrm{mm\,yr^{-1}}$, respectively. For daily temperature data, both $T_{\min}$ and $T_{\max}$ biases are less than 0.1°C for all the SD methods over the CONUS except BCSDm, which produced bias in $T_{\min}$ and $T_{\max}$ as large as 0.7° and −0.5°C, respectively. Although BCCA produces a severe low bias in precipitation, the bias in BCCA-downscaled temperature is comparable to the other SD methods. One of the reasons for this reasonable bias pattern in BCCA temperature would be that temperature fields generally have much smoother spatial patterns than precipitation.

We further examine the effects of different SD methods on key hydrologically relevant attributes derived from the downscaled precipitation and temperature data. These attributes include wet-day fraction, defined as fraction of days with precipitation greater than a certain threshold, and DTR. Bias in wet-day fraction impacts precipitation intensity (Gutmann et al. 2014), which further affects infiltration and canopy interception of precipitation. Both attributes also directly influence the other forcing variables (i.e., humidity and downwelling radiation) through MTCLIM. Figure 4 illustrates how those two hydrologically relevant attributes affect shortwave radiation estimates. BCCA and BCSDd produce a much higher wet-day fraction compared to the M02 dataset across the CONUS. Both SD methods produced over +0.25 of mean bias over the CONUS. BCSDm and AR produce less bias than BCCA and BCSDd (within ±0.1 bias across the CONUS). DTR biases in the SD temperature data are shown in Fig. 4 (middle). Biases of $T_{\min}$ and $T_{\max}$ (see Fig. 3, middle and right) directly affect DTR biases. For example, BCCA and BCSDd produced consistent spatial patterns in biases of $T_{\min}$ and $T_{\max}$ across the CONUS, resulting in small DTR bias. However, BCSDm produced a relatively large underestimate over the CONUS (CONUS mean bias: −0.3°C) due to warm bias of $T_{\min}$ and cool bias of $T_{\max}$. AR produced small positive bias in DTR, especially over the MR region (see the region in Fig. 2) where cool bias of $T_{\min}$ and warm bias of $T_{\max}$ are seen.

The annual shortwave radiation bias is shown in Fig. 4 (right). The biases in wet-day fraction and DTR help explain the shortwave radiation error. In this case, the signs of both biases in wet-day fraction and DTR enhance the error of shortwave radiation for all the SD datasets. Clearly, severe overestimation of wet-day fraction in BCCA and BCSDd cause an underestimation of shortwave radiation over the CONUS, excluding the Southwest. On the other hand, negative and positive biases in shortwave radiation in BCSDm and AR, respectively, result from both biases of wet-day fraction and DTR.

We further explore which attribute (wet-day fraction or DTR) affects shortwave radiation estimates more. In MTCLIM [see Thornton et al. (2000) for more details on the algorithms], daily shortwave radiation is computed as the product of three daily variables: theoretical top-of-the atmosphere radiation R_{pot}, cloud-free transmittance $\mathrm{Tr}_{t,\max}$, and cloudy-sky transmittance $\mathrm{Tr}_{f,\max}$. Of those three variables, $\mathrm{Tr}_{f,\max}$ is a function of DTR and wet-day fraction, and $\mathrm{Tr}_{t,\max}$ is weakly related to DTR. The daily cloudy-sky transmittance is given by

$$\mathrm{Tr}_{f,\max} = \begin{cases} 1.0 - 0.9\exp[-B(\mathrm{DTR}^{1.5})] & P = 0 \quad (\mathrm{mm}) \\ \{1.0 - 0.9\exp[-B(\mathrm{DTR}^{1.5})]\}0.75 & P > 0 \quad (\mathrm{mm}) \end{cases}, \tag{1}$$

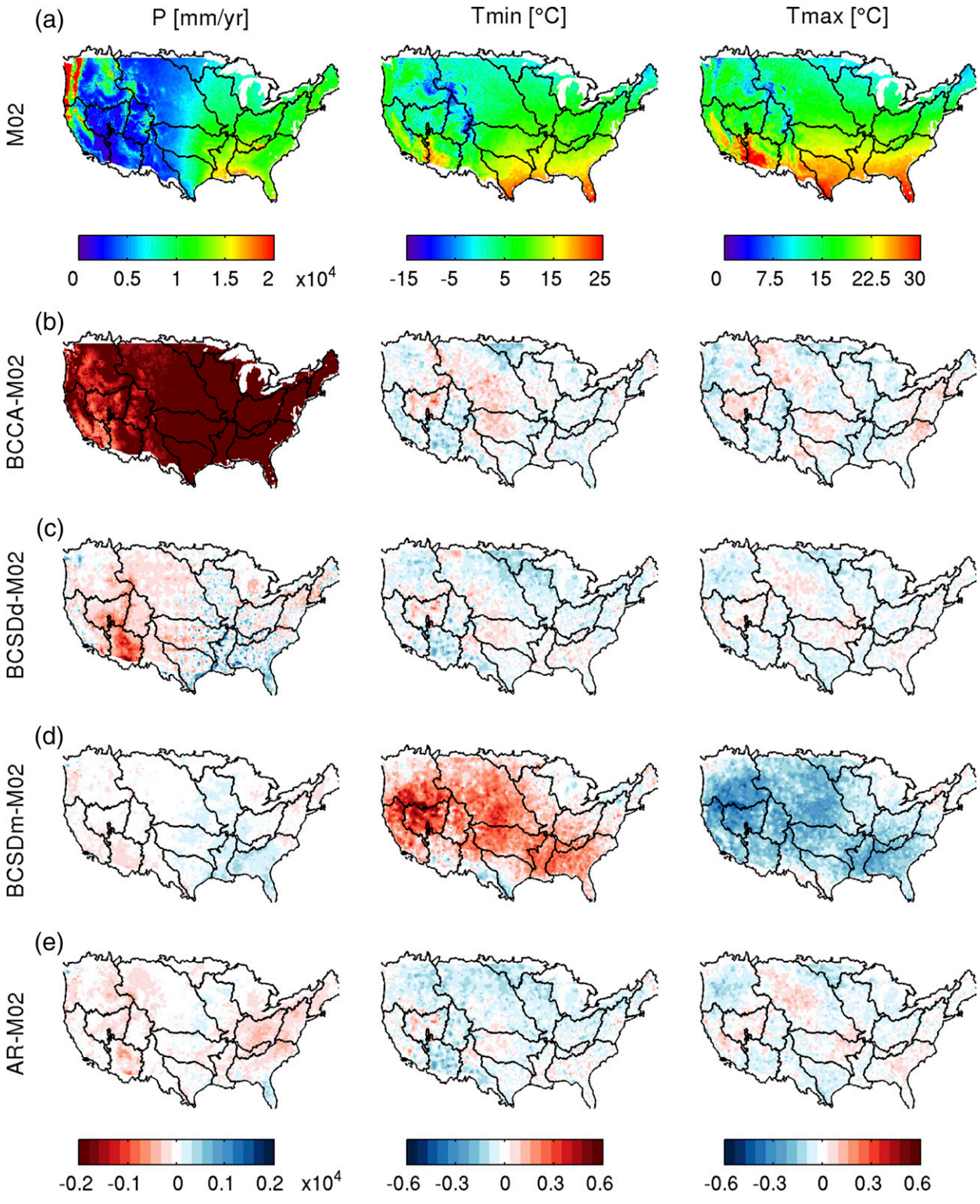

FIG. 3. (a) Annual mean values of P, T_{min}, and T_{max} from M02 and (b)–(e) bias for the four SD methods relative to M02.

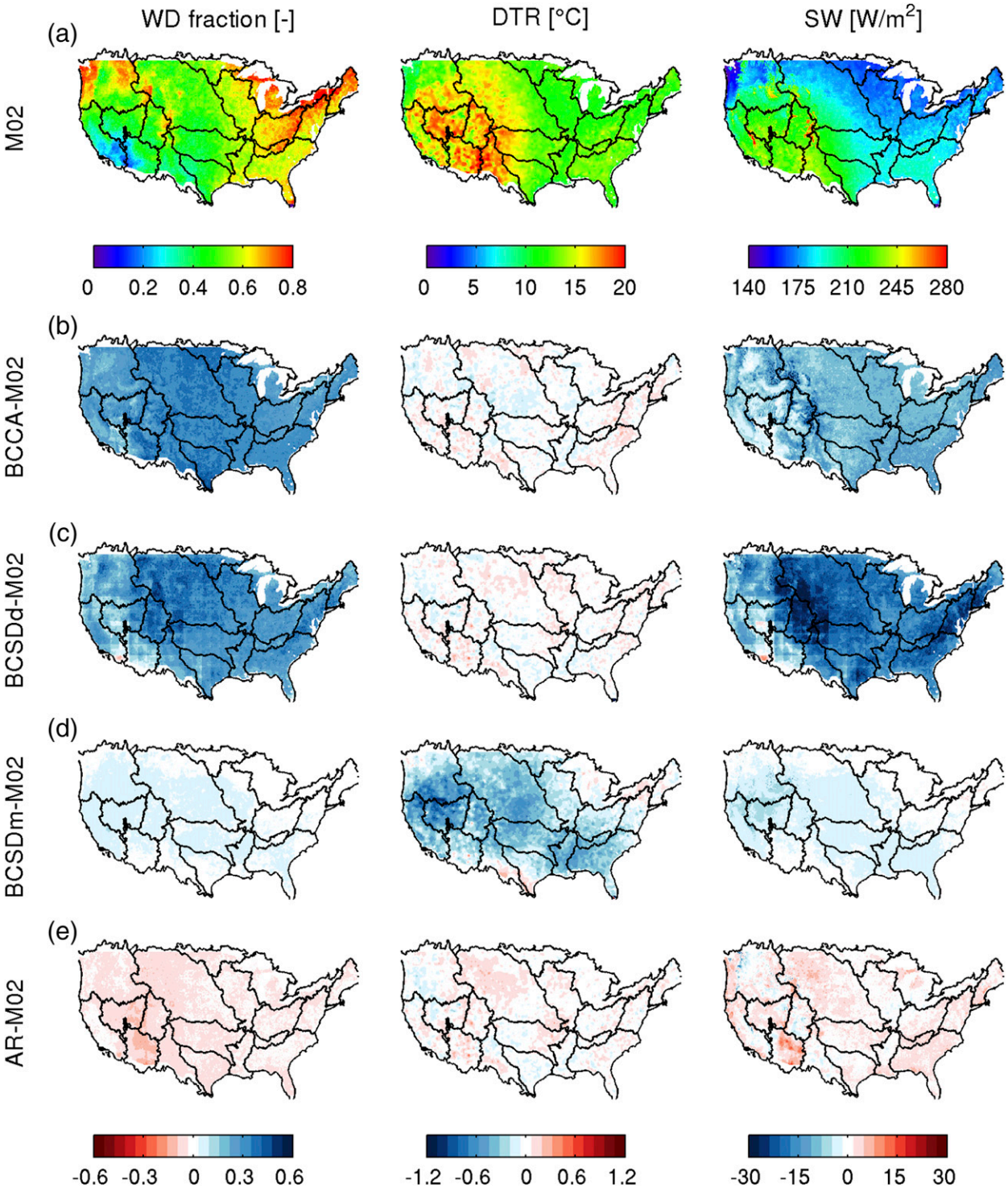

FIG. 4. (a) Annual mean values of wet-day (WD) fraction, DTR, and shortwave radiation (SW) from M02 and (b)–(e) bias for the four SD methods relative to M02. Red (blue) in WD fraction and DTR indicates their biases contribute to positive (negative) bias in SW.

where B is a parameter with a function of 30-day trailing average of DTR. Figure 5 shows $\mathrm{Tr}_{t,\max}$ as a function of DTR indicated by black solid lines, along with the sensitivity to errors in DTR (colored dashed lines) for wet ($P > 0$ mm) and not wet ($P = 0$ mm) days. The higher the DTR value is, the less sensitive $\mathrm{Tr}_{t,\max}$ becomes. Therefore, accuracy in shortwave radiation estimated with MTCLIM is less affected by DTR bias in areas with

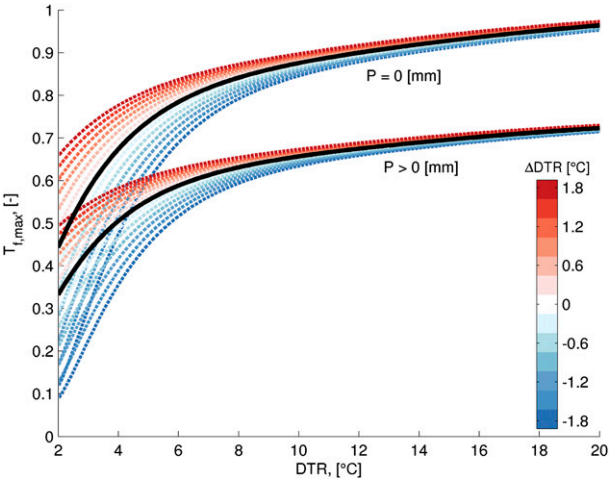

FIG. 5. Cloudy-sky transmittance as a function of DTR (thick black line) and its sensitivity to DTR bias indicated by colored dashed lines for precipitation and nonprecipitation days.

the highest DTR (e.g., southwestern United States; see Fig. 4), and consequently, the error in the wet-day fraction is of greater importance. These results indicate that SD methods require accuracy of precipitation frequency as well as the amount if the downscaled climate data are input into the process-based hydrologic models that compute full energy balance, because errors in wet-day fraction propagate into shortwave radiation estimates with MTCLIM, which can greatly impact snowmelt rates and ET (Mizukami et al. 2014).

b. *Comparison of hydrologic simulations*

This section discusses the effect of model and SD choices on the annual water balance (precipitation partitioning between ET and runoff), characteristics of extreme runoff, and dynamics of hydrologic states (i.e., snowpack and soil moisture). Finally, intermodel and inter-SD forcing differences in the simulations are quantified in terms of annual runoff to provide insight into which choices are likely to have more impact on the hydrologic simulations.

1) PRECIPITATION PARTITIONING INTO ET AND RUNOFF

Biases in SD forcing data via the four SD methods are expected to uniquely propagate into biases in the hydrologic simulations. Figure 6 shows the mean annual runoff and ET estimated with the three models forced by M02 (Fig. 6a), as well as the difference (i.e., bias) compared to M02 for each of the SD datasets (Figs. 6b–e). The spatial patterns in runoff bias are generally in opposite sign of ET bias pattern for each simulation, except for the simulations forced by BCCA data. Large underestimations of ET and runoff in BCCA simulations (note the different color scale than the other SDs) are predominantly caused by the large underestimation of BCCA precipitation. The greatest underestimation of ET is seen in the Great Plains (MR, ARK, GUL, and LM). The greatest runoff underestimation is seen over the eastern United States and in the mountainous western coastal region (e.g., Sierra Nevada and Cascade ranges of CA and PN), where ET is generally energy limited.

More obvious intermodel differences appear for the simulations forced by BCSDd. Even with considerable reduction of shortwave radiation in BCSDd in many regions (Fig. 4) in the midwestern plains (i.e., UM, OH, TN, and LM), VIC and PRMS overestimate ET while CLM does not. Moreover, the magnitude of the ET overestimations from VIC and PRMS is larger than precipitation overestimations over those regions. This occurs because VIC and PRMS produce large positive bias in canopy evaporation (not shown). Both VIC and PRMS store more of the rainfall as canopy interception, which evaporates more freely than water that enters the soil column. Key hydrologically relevant attributes that play a role in canopy evaporation here are both biases of wet-day fraction and precipitation amount. BCSDd results in more precipitation days or high wet-day fraction with relatively small precipitation bias, leading to a lower precipitation intensity [see extreme event section in Gutmann et al. (2014)]. A larger fraction of the less intense precipitation is effectively intercepted by the canopy, leading to increased canopy evaporation. Although the wet day fraction in BCCA is as high as that in BCSDd and both SD methods produce low intensity precipitation (i.e., drizzle), less precipitation is intercepted by the canopy because of a much larger underestimation of precipitation amount, resulting in less canopy evaporation. Both VIC and PRMS produce less transpiration because of less water that infiltrates into the soil as well as the reduction of shortwave radiation. For CLM, the increase of canopy ET does not compensate for the reduction of transpiration; therefore, total ET is reduced. This interesting intermodel difference in ET bias for BCSDd is seen only over the midwestern regions because of denser forest cover than the interior regions such as UCO, LCO, and GB.

For the simulations with BCSDm and AR datasets, which produce less bias of wet-day fraction than BCSDd, it appears that ET bias is caused by biases of precipitation amount or shortwave radiation. When models are forced with BCSDm, they overestimate ET in the eastern United States where slight overestimation of annual precipitation is seen. On the other hand, small

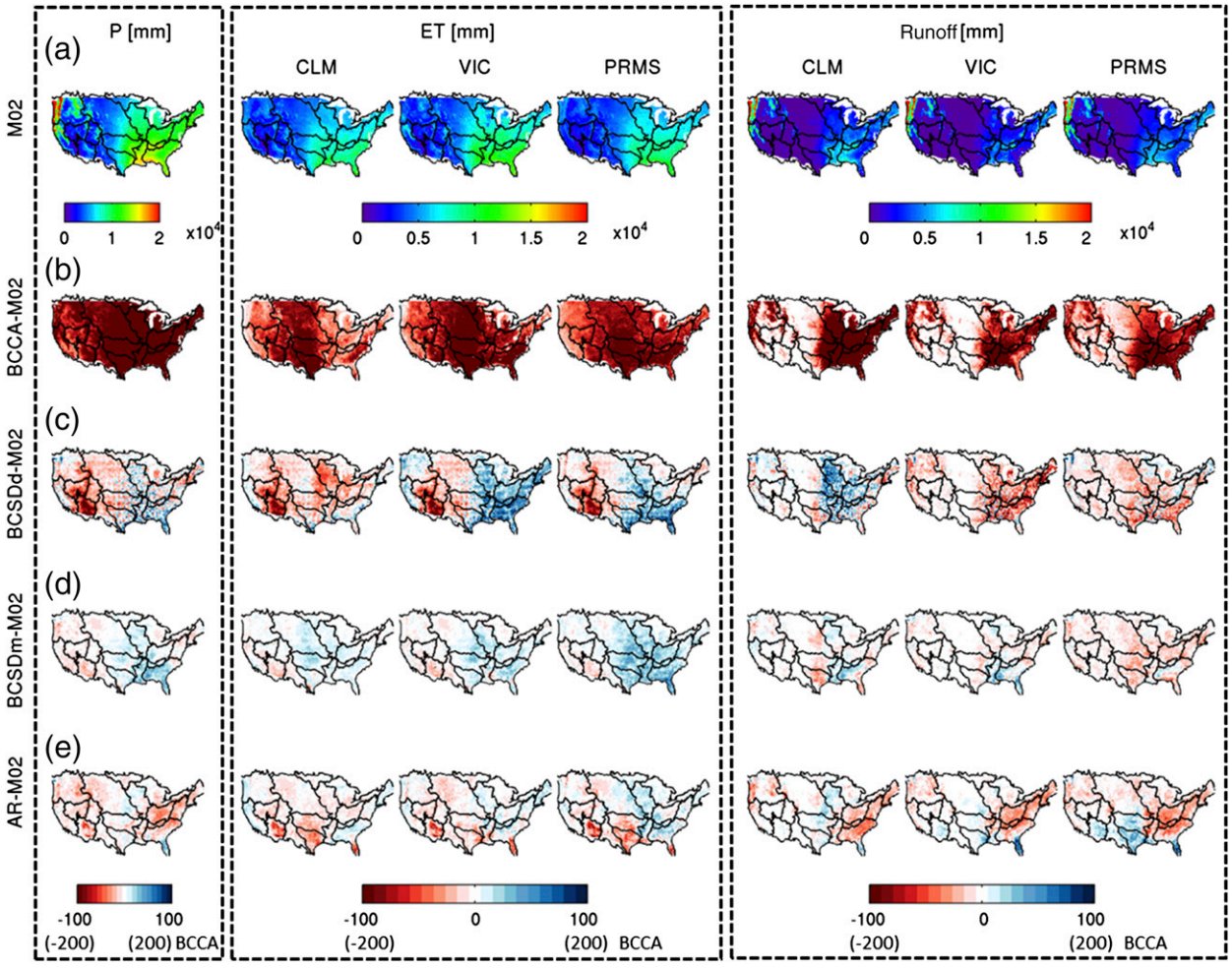

FIG. 6. (a) Values of downscaled annual *P* and simulated water balance (ET and runoff) for three models forced by M02 and (b)–(e) bias for three models forced by the four SD methods relative to M02. Note that the biases range between −200 and 200 mm for BCCA and −100 and 100 mm for the other SDs.

ET increases in the simulations forced by AR, as seen in the eastern United States, are due to a slight positive error in shortwave radiation.

Figure 7 shows the spatial distributions of intermodel differences in the mean annual runoff for M02 and each SD forcing dataset. The spatial patterns of intermodel difference in ET mirrors the runoff difference pattern (not shown here). The spatial patterns of intermodel differences in annual water balance are fairly consistent across all the forcing data but BCCA. Two striking features are seen in Fig. 7. First, larger and more uniform intermodel differences are seen in the eastern part of CONUS than the western part, where large differences are concentrated in high-elevation areas. Second, it is interesting to note that over the midwestern regions (SA, OH, TN, UM, LM, and eastern parts of MR and AR), the sign of the CLM–PRMS difference for BCCA was opposite to those for the other forcings while CLM–VIC and VIC–PRMS comparisons are consistent with all the forcing data. The order of increasing magnitude in the annual runoff simulations is PRMS > CLM > VIC over the midwestern regions when the models were forced by BCCA but CLM > PRMS > VIC for the other forcing data. Note BCCA has the largest underestimation of precipitation compared to the other forcing data (Fig. 3), leading to the lowest runoff with the use of any models. However, the reduction of simulated runoff from CLM-BCCA relative to CLM with the other forcing data was much larger than VIC and PRMS as shown in interforcing difference (see runoff in Fig. 6b). This is why CLM-BCCA produced less runoff than PRMS-BCCA over the midwestern regions.

2) EXTREME RUNOFF

Estimations of extreme runoff values are crucial for many water management problems. For instance, high runoff

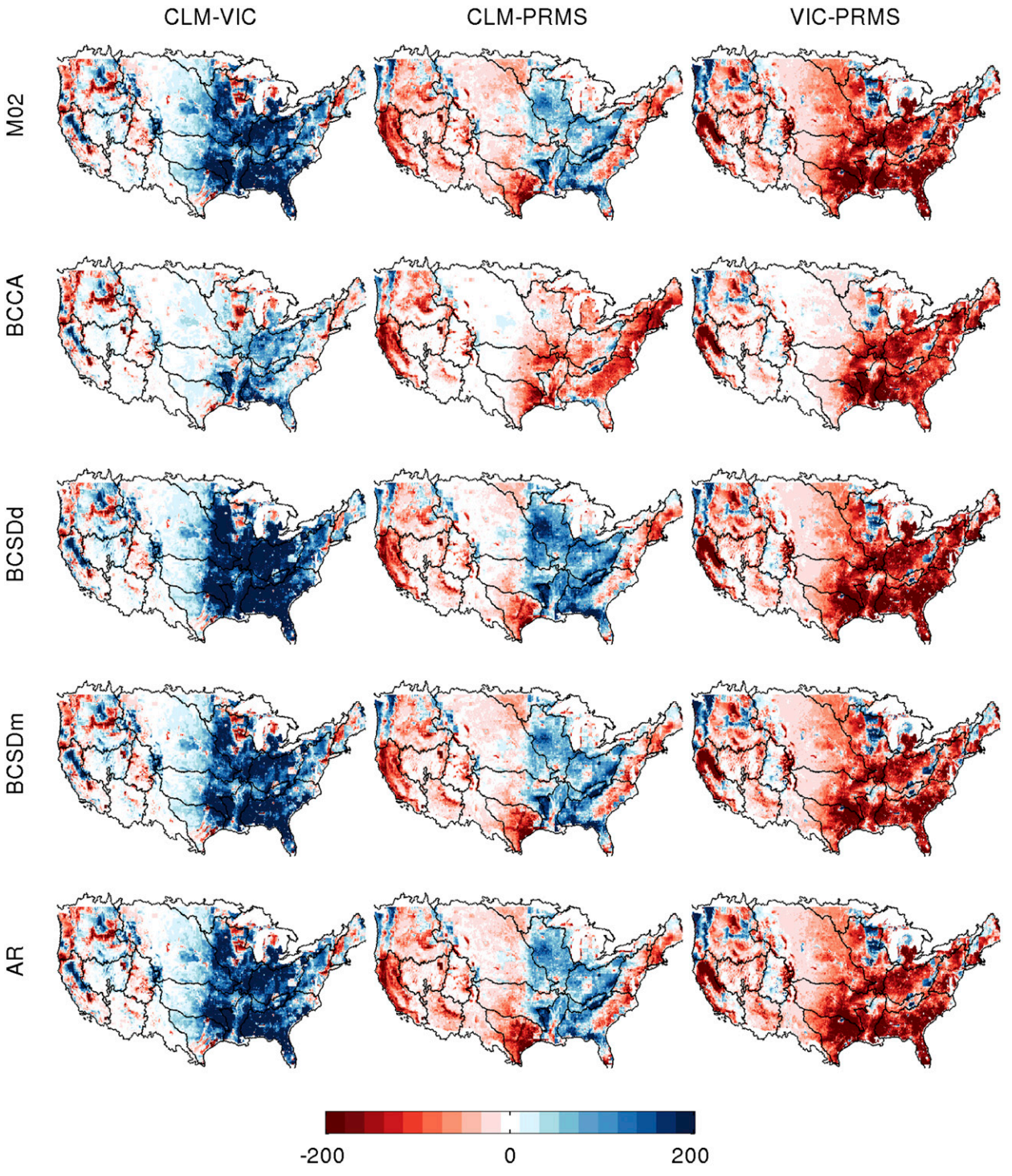

FIG. 7. Spatial distribution of intermodel differences in mean annual runoff for the four SD methods and M02.

estimates are associated with flood events, whereas low runoff estimates are associated with hydrological drought. Of importance is that the changes in extreme runoff may not be necessarily the same as the changes in total annual runoff presented in the previous section.

For high runoff analysis, 20-yr peak runoff (RO_{20yr}) was estimated at each grid cell for all the model simulations. The RO_{20yr} was computed by fitting a log-Pearson type 3 distribution to a 19-yr record of annual peak daily values and then selecting the runoff value at

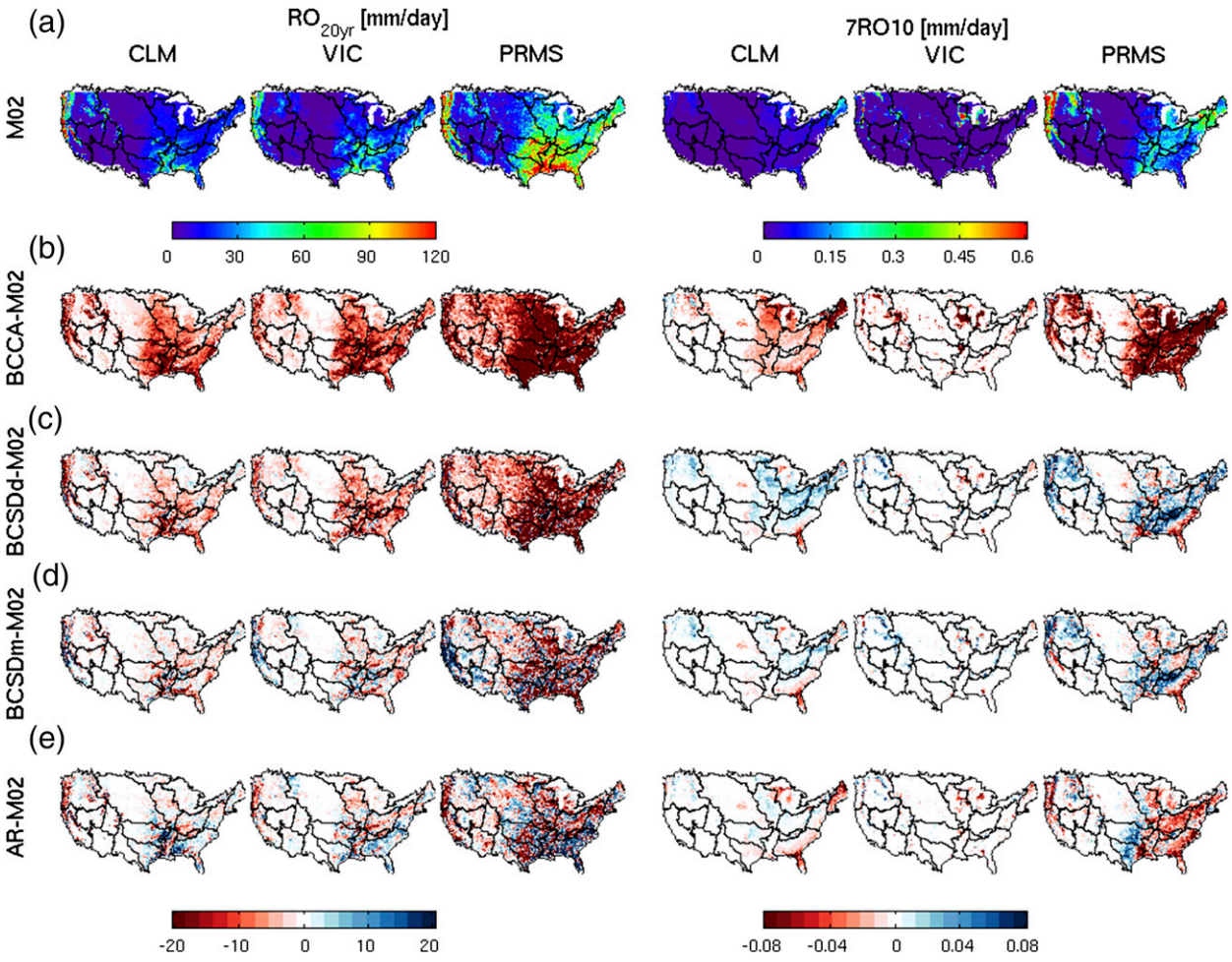

FIG. 8. (a) Values of RO_{20yr} and 7RO10 for three models forced by M02 and (b)–(e) bias for three models forced by the four SD methods relative to M02.

5% exceedance probability. Low runoff analysis used the 7-day, 10-yr annual low runoff (7RO10) since this is widely used in low-flow frequency analysis in the United States. To estimate 7RO10, the minimum values of 7-day mean runoff are first computed for each WY, and then the runoff value at 90% exceedance probability was estimated from log-Pearson type 3 distribution fitted to 19-yr records of the annual lowest runoff values.

Differences in biases of extreme flows due to different SD methods are illustrated in Fig. 8 as spatial distributions and in Fig. 9 as scatterplots. First, for both high- and low-flow estimates, CLM and VIC produce less bias in the interior part of the CONUS (e.g., MR, AR, RIO, UCO, LCO, and GB) than the other parts of the CONUS for all the SD forcings. For the high-flow estimates, spatial bias patterns are similar between CLM and VIC for each SD forcing data (Fig. 8), in which underestimation is concentrated in southern regions (eastern ARK, LM, and TN). However, CLM produces larger scatter in RO_{20yr}, indicating that larger bias exists for each SD dataset than VIC (Fig. 9). On the other hand, PRMS produced even larger biases (Fig. 9) that extend across the CONUS (Fig. 8). For low-flow bias, there are noticeable intermodel differences in spatial bias patterns. Larger biases are seen in the eastern part of CONUS and in the PN region, particularly from PRMS.

Figure 10 shows intermodel comparisons of high and low runoff estimated from the model simulations forced by M02 dataset. Figure 10 also includes comparison of daily mean runoff estimates. Contrary to the mean runoff comparison, there are very large intermodel differences in high and low runoff, indicating quite a large difference in temporal runoff patterns or runoff distribution at a daily step among the models. This could be due to differences in the soil hydrology parameterizations (i.e., percolation, infiltration, and runoff generation) among the models, and parameter

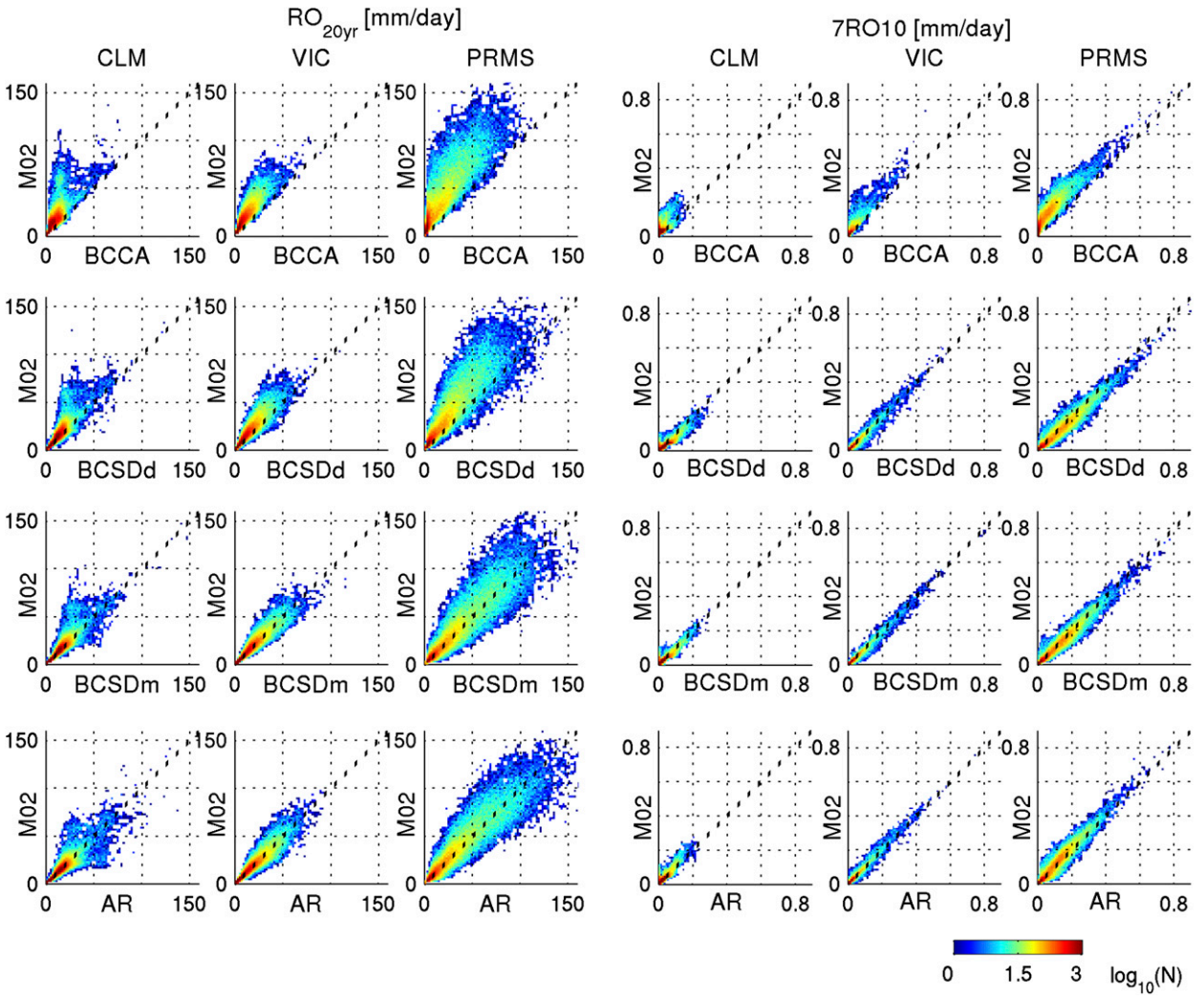

FIG. 9. Scatterplot of extreme runoff [(left) RO_{20yr} and (right) 7RO10] between four SD forcings and M02. Color bar indicates number of occurrences in a 2D bin (bin width is 2 mm for RO_{20yr} and 0.01 mm for 7RO10).

values are also likely to play a role in temporal flow distribution, which was indicated by Nijssen et al. (2003). Little correlation of low runoff estimates among three models implies little agreement of base flow among the three models. High extremes are better correlated than low extremes, despite large systematic intermodel differences. The possible explanations of such completely different intermodel correspondence between high and low runoff are given as follows. High runoff is more dependent on precipitation and snowmelt, which are more directly dictated by the meteorological forcings common among the hydrologic models. On the other hand, low runoff, which originates from base flow, depends on soil and vegetation processes, which are differently described by the model parameterizations and associated parameter values.

3) HYDROLOGIC STATES

Examination of the temporal dynamics of some hydrologic state variables, such as snow water equivalent (SWE) and soil moisture, can help diagnose differences in runoff generation as a result of different methodological choices. This section presents the differences in the selected key attributes of seasonal snowpack and soil moisture dynamics.

(i) Seasonal snowpack

Errors in simulated snow accumulation and ablation can affect seasonal runoff and ET patterns, particularly over the western United States. Climatological seasonal snow patterns were examined with respect to two metrics: the peak amount of SWE and the centroid day of SWE time series for a water year (SWE_{CT}) to infer

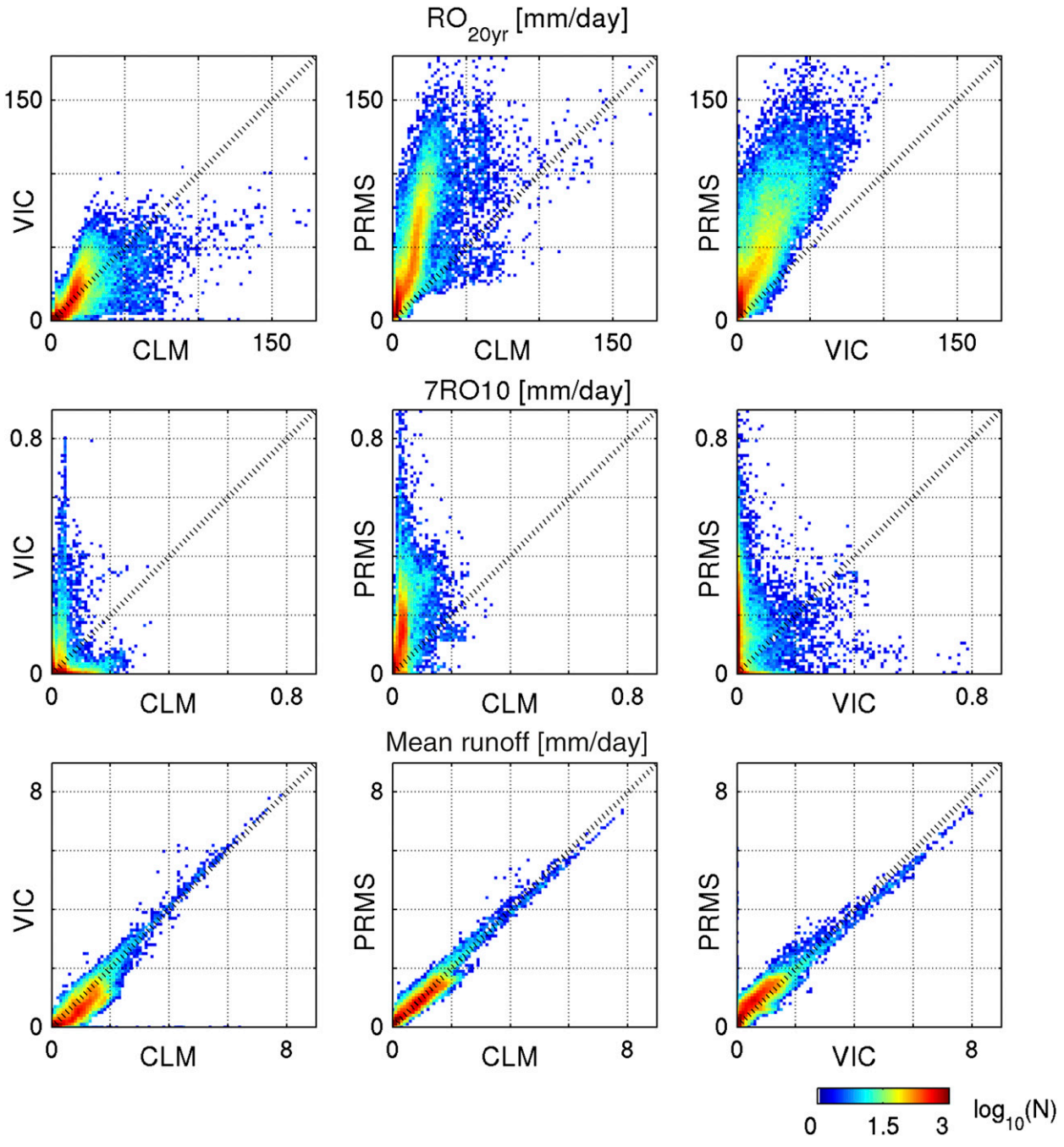

FIG. 10. Intermodel comparison of (top) RO_{20yr}, (middle) 7RO10, and (bottom) mean RO between (left) CLM and VIC, (center) CLM and PRMS, and (right) VIC and PRMS. The forcing data are from M02. Color bar indicates number of occurrences in a 2D bin (bin width is 2 mm for RO_{20yr} and 0.01 mm for 7RO10).

timing of peak SWE. Variable SWE_{CT} is computed using (Stewart et al. 2005)

$$SWE_{CT} = \frac{\sum day[SWE(day)]}{\sum SWE(day)}. \quad (2)$$

Figure 11 compares biases of peak SWE (Fig. 11, top) and SWE_{CT} (Fig. 11, bottom) for three models forced by each SD dataset. To construct a cumulative distribution function of biases over the CONUS shown in Fig. 11, the grid boxes where 19-yr climatological peak SWE is over 50 mm are selected. For peak SWE, all the

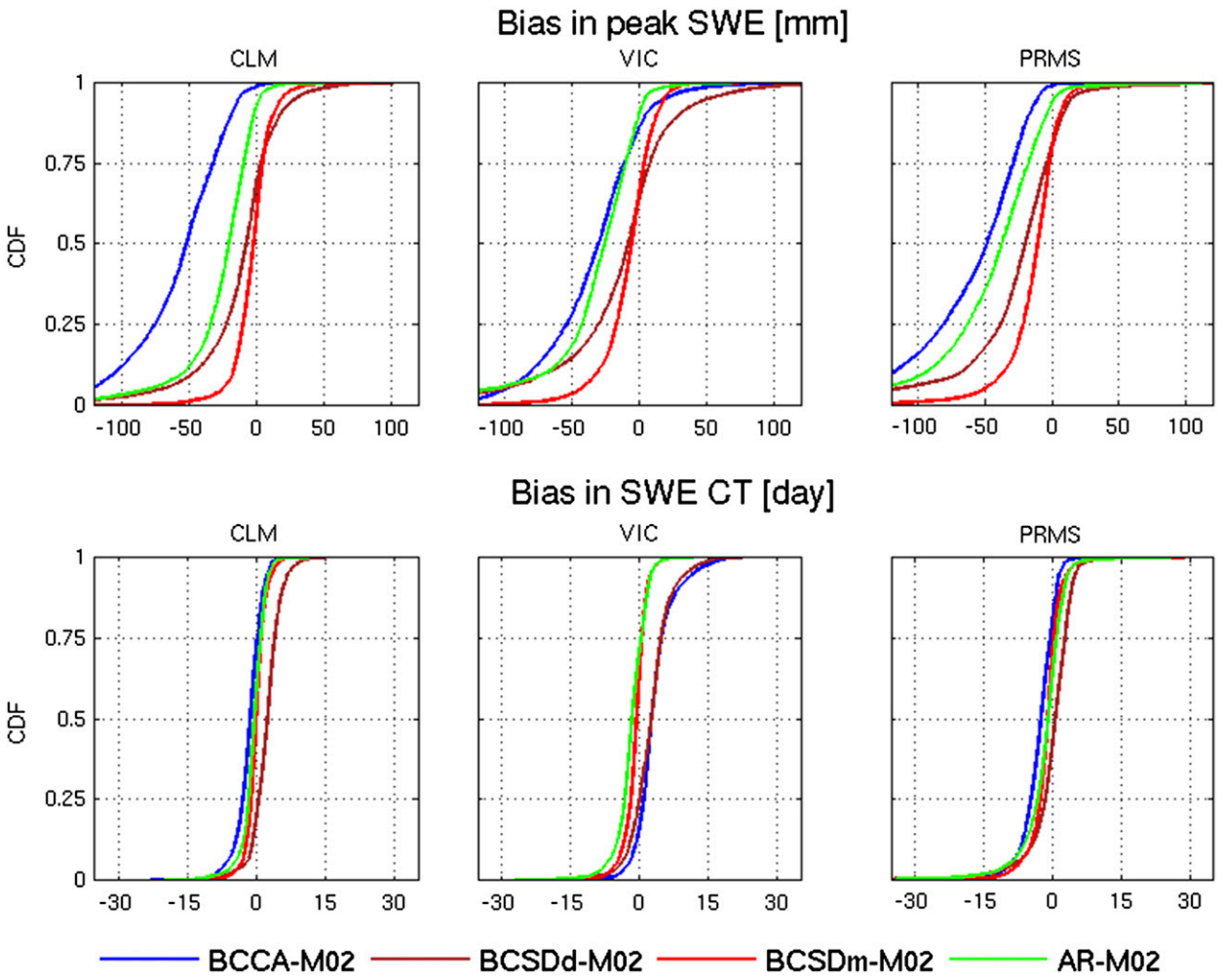

FIG. 11. Cumulative distribution of error in 19-yr mean annual (top) peak SWE and (bottom) SWE centroid over the CONUS for (left) CLM, (middle) VIC, and (right) PRMS. Pixels below 50 mm of 28-yr mean peak SWE are removed for the CDF construction.

SD forcing datasets generally produced an underestimation in all the hydrologic models. The order of the bias magnitude is generally preserved among the models. The causes of peak SWE underestimations depend on SD forcing datasets. For instance, the BCCA simulations produced the least snow accumulation because of severe precipitation underestimation while the AR simulations produced negative SWE bias, possibly because of the large bias of incoming shortwave radiation. The biases from BCSDm simulations are the smallest possibly because of the least underestimation of precipitation, particularly over the western United States and negative bias in shortwave radiation. VIC produces relatively consistent peak SWE bias pattern across all the SD forcing data compared to the other two models. For bias of SWE_{CT}, all the models forced by SD forcings produced within 15 days of bias over the CONUS. However, VIC-BCCA and VIC-BCSDd produced positive bias of SWE_{CT} for 75% of the total examined grid boxes, indicating later peak SWE timing compared to M02.

Figure 12 shows intermodel differences of peak SWE and SWE_{CT} for all the forcing data, indicating much larger differences than interforcing differences or bias (cf. Fig. 11). Since all models use the same method to partition precipitation into rain and snow, the intermodel differences in peak SWE occur because each model produces different snow ablation processes (sublimation and snowmelt). Such dissimilarity can stem from discrepancies in model parameterizations of canopy interception and sublimation, radiative transfer through canopy, within- and below-canopy turbulence, albedo decay computation and storage, and transmission of liquid water through the snowpack.

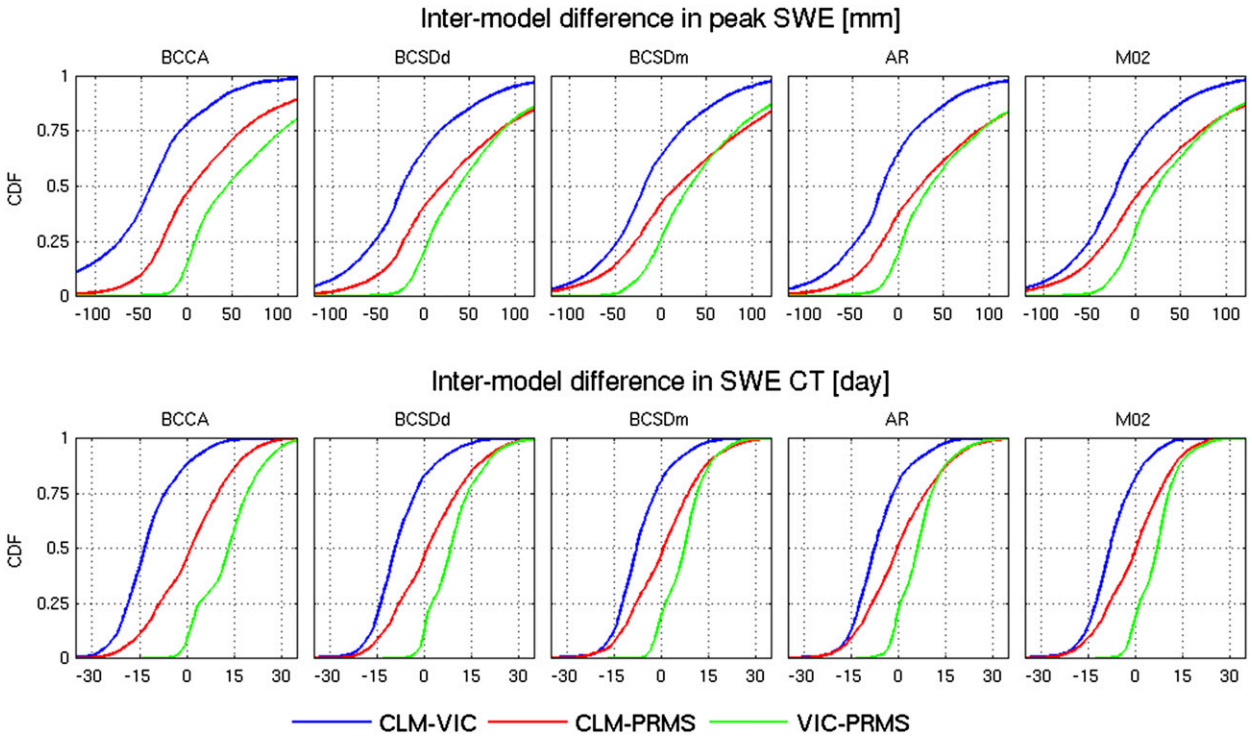

FIG. 12. Cumulative distribution of intermodel differences in 28-yr mean annual (top) peak SWE and (bottom) SWE centroid over the CONUS for (from left to right) BCCA, BCSDd, BCSDm, AR, and M02. Pixels below 50 mm of 28-yr mean peak SWE are removed for the CDF construction.

(ii) Soil moisture

Soil moisture dynamics are characterized by an autocorrelation of daily soil moisture series. This analysis used the *e*-folding time of the daily soil moisture autocorrelation as a measure of soil moisture persistence (Delworth and Manabe 1988). The *e*-folding time is the lag for which the autocorrelation falls below $1/e$. The larger the *e*-folding time is, the larger the soil moisture memory is. In hydrologic models, autocorrelation of soil moisture depends on soil storage, which differs substantially among the models. The analysis in this paper uses total column soil moisture in all the soil storages from which surface flow and base flow are generated. Figure 13 shows the *e*-folding time from the M02 simulation and its bias for each model. Figure 14 shows intermodel differences in soil moisture memory from M02 simulations. CLM and VIC have similar soil representations (i.e., layer structure), but CLM has much more memory than VIC, possibly because base flow is generated from an unconfined aquifer that can hold large (4800 mm at maximum) soil moisture (see Fig. 1). PRMS, which uses reservoirs as soil moisture storage, has the least soil moisture memory. The other factors that affect soil moisture memory are soil hydraulic parameters in the model that affect percolation rate and baseflow rate. These characteristics of soil moisture memory are consistent with runoff characteristics. For example, CLM, which has the largest stored soil moisture and soil moisture memory, produces the least runoff, particularly low runoff, while the least soil moisture memory in PRMS is related to the largest runoff. Since soil moisture depth is much more influenced by model soil representation, there are large intermodel differences as shown in Fig. 14.

4) INTERMODEL DIFFERENCE VERSUS INTERFORCING DIFFERENCE

Finally, to examine which of the methodological choices—the choice of models or SD—contributes more to the spread between the hydrologic simulations, we use the multimodel analysis approach used in past studies (Dirmeyer et al. 2006; Xia et al. 2012). The analysis involves the use of similarity measures to quantify the spread of ensemble simulations generated from multiple hydrologic models or multiple downscaled forcing datasets.

For this analysis, the spread of multiple simulations is defined as $R_{\mathrm{std}}^{\mathrm{factor}}$, where R_{std} denotes a ratio of two standard deviations and factor identifies the source of an ensemble. Here, the ensemble is formed by collection of multiple models ($N = 3$) or multiple SD forcing data

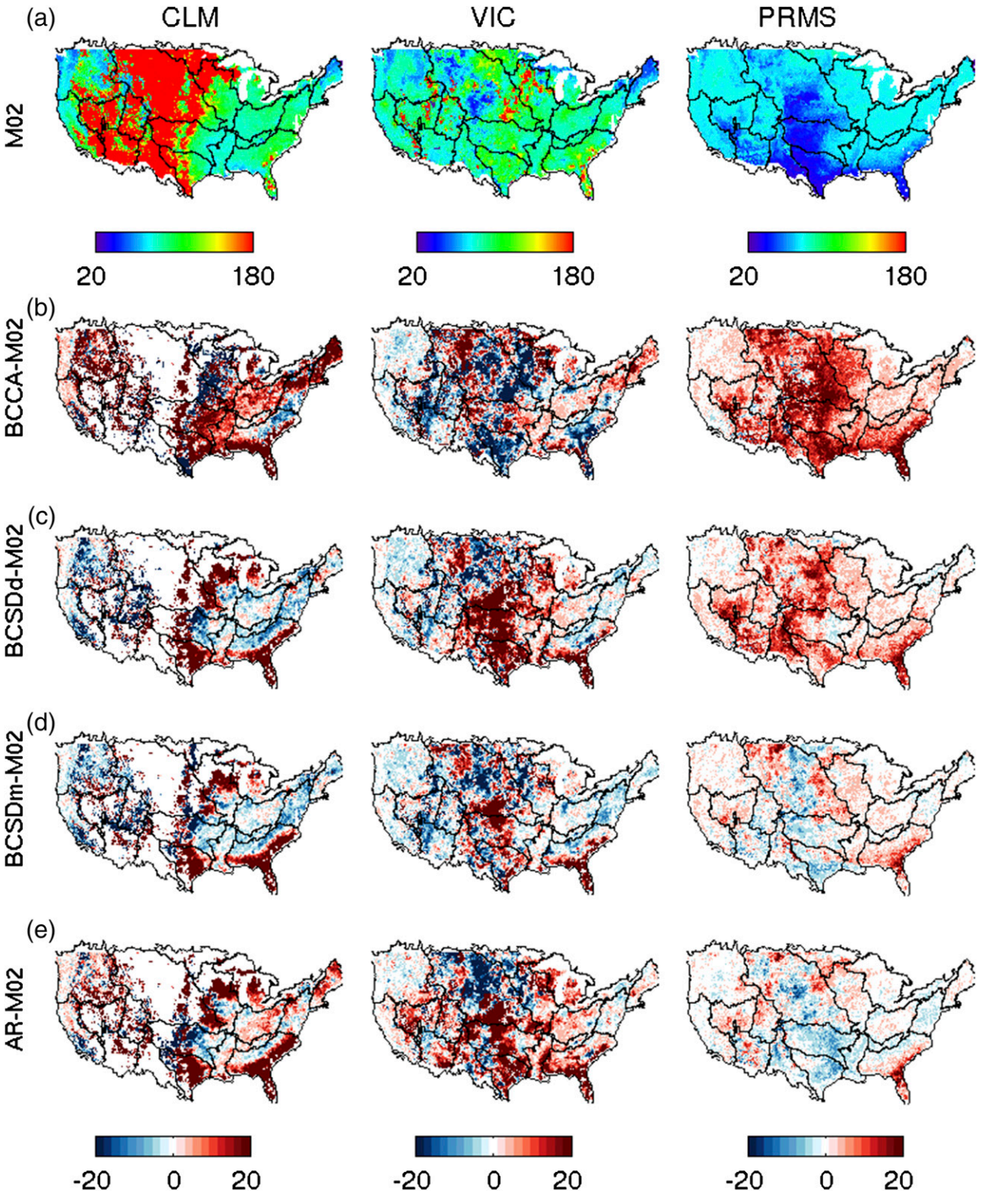

FIG. 13. (a) Value of *e*-folding decay days of total column soil moisture autocorrelation for three models forced by M02 and (b)–(e) bias for three models forced by the four SD methods relative to M02.

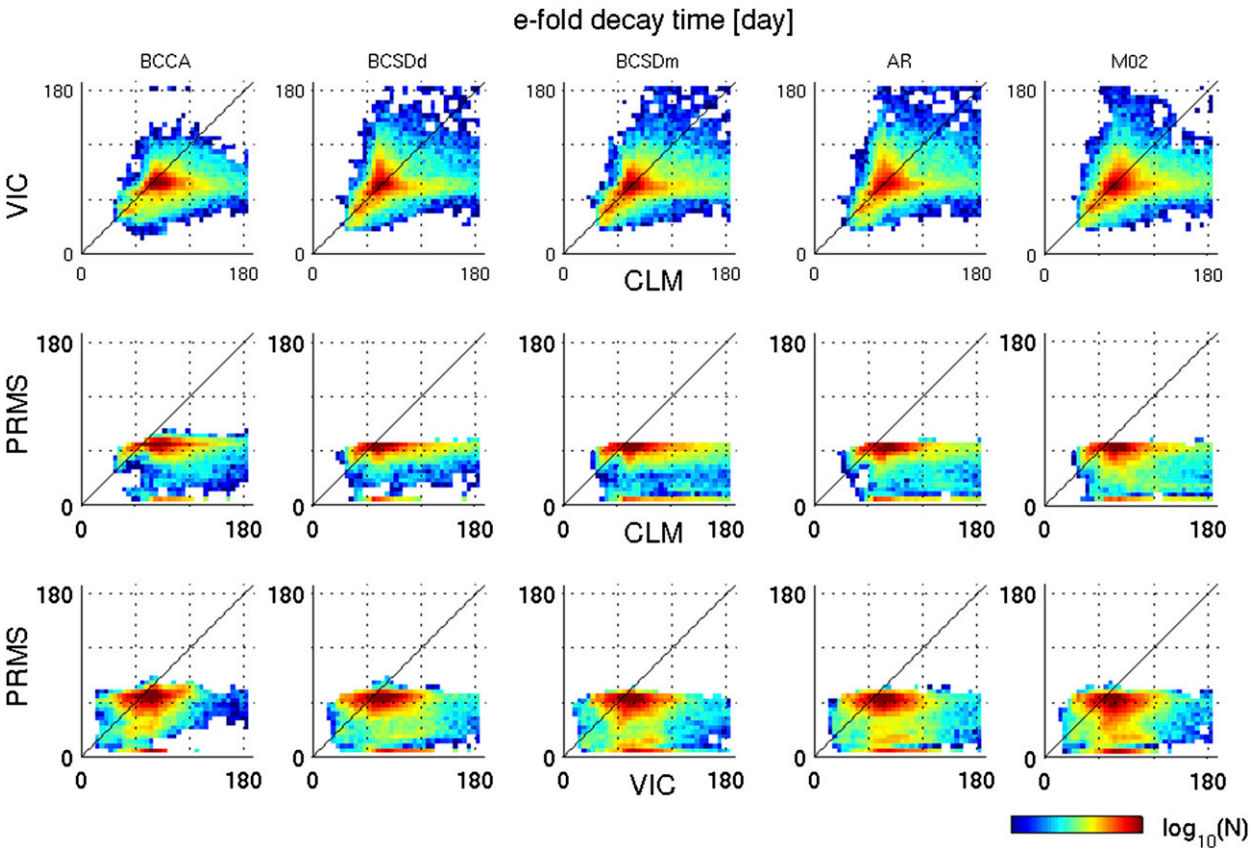

FIG. 14. Intermodel comparison of *e*-folding decay days of soil moisture autocorrelation for the different forcing data. Color code indicates $\log_{10}$ of the number of pixels that fall in *x*- and *y*-axis bins.

($N = 4$). First, the mean and standard deviation of the ensemble of a particular variable X is computed at each time step t (i.e., monthly step) as $\overline{X}_t$ and σ_t, respectively, and then temporal mean $\overline{\sigma}$ of σ_t is computed with

$$\overline{\sigma} = \frac{1}{M}\sum_{t=1}^{M}\sigma_t, \tag{3}$$

where M is total number of months from October 1980 through September 1999 (i.e., 228 months). Next, the temporal standard deviation of the ensemble mean σ_{total} is computed as

$$\sigma_{\text{total}} = \sqrt{\frac{1}{M}\sum_{t=1}^{M}(\overline{X}_t - \overline{X})^2} \tag{4}$$

and

$$\overline{X} = \frac{1}{M}\sum_{t=1}^{M}\sum_{n=1}^{N}X_{n,t}, \tag{5}$$

where $X_{n,t}$ is a variable of nth member at time t. Finally, the spread index $R_{\text{std}}^{\text{factor}}$ is given by

$$R_{\text{std}}^{\text{factor}} = \frac{\overline{\sigma}}{\sigma_{\text{total}}}. \tag{6}$$

The larger the value of $R_{\text{std}}^{\text{factor}}$ is, the larger the ensemble spread is. Taking the hydrologic model and SD forcing data as a factor of ensemble variability of the hydrologic simulations, $R_{\text{std}}^{\text{model}}$ and $R_{\text{std}}^{\text{forcing}}$ are computed with Eq. (6) for the monthly runoff anomalies for each individual grid box to examine spatial patterns of intermodel (or forcing) spread. Monthly anomaly time series were used to remove the seasonality of intermodel (or interforcing data) variability.

Figure 15a shows the $R_{\text{std}}^{\text{forcing}}$ histogram for each model. Since BCCA simulations appear to be an outlier (i.e., severe underestimations of all the water fluxes), $R_{\text{std}}^{\text{forcing}}$ values with and without BCCA were computed. In Fig. 15a, the solid (dashed) line indicates $R_{\text{std}}^{\text{forcing}}$ computed without (with) BCCA. Figure 15b displays a histogram containing occurrences of $R_{\text{std}}^{\text{model}}$ of the monthly runoff anomaly for each SD dataset as well as M02 over the CONUS. Overall, the BCCA-forced simulations produce the largest magnitudes of $R_{\text{std}}^{\text{model}}$. Compared to $R_{\text{std}}^{\text{model}}$, $R_{\text{std}}^{\text{forcing}}$ or the interforcing spread

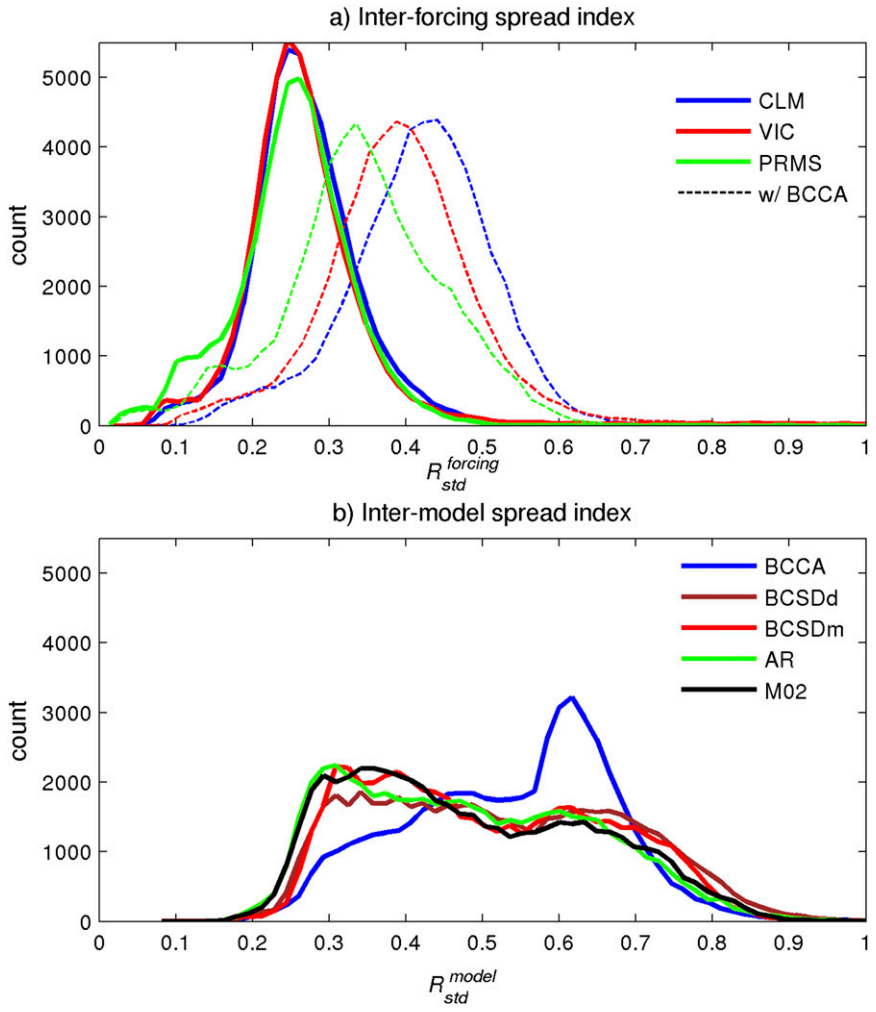

FIG. 15. Histograms of (a) interforcing spread index and (b) intermodel spread index of monthly runoff over the CONUS.

over the CONUS is quite consistent across the three models if BCCA is excluded. However, inclusion of BCCA in forcing ensembles increases the spread of runoff simulation for all the models, and the magnitude of spread becomes more inconsistent among the models.

The spatial variability of $R_{\text{std}}^{\text{forcing}}$ is much smaller than for $R_{\text{std}}^{\text{model}}$ as indicated by tight histograms of $R_{\text{std}}^{\text{forcing}}$. The distributions of $R_{\text{std}}^{\text{model}}$ are characterized as bimodal, which are caused by a large $R_{\text{std}}^{\text{model}}$ region ($R_{\text{std}}^{\text{model}} > 0.6$) over the Intermountain West and low $R_{\text{std}}^{\text{model}}$ region ($R_{\text{std}}^{\text{model}} \approx 0.3$) over the Great Plains through the Southeast (spatial distribution not shown here).

This analysis suggests that model choice is likely to produce more variability than the choice of SD methods for runoff, especially when BCCA is not included. However, this disparity has to be interpreted with caution because the analysis is based on a small sample of hydrologic models and SD methods. Therefore, the conservative conclusion drawn from this analysis is specific to models and forcing data used in this study. In addition, and more importantly, it is likely that the magnitude of intermodel differences is expected to be affected by hydrologic parameter estimates. As previously mentioned, Mendoza et al. (2015a) demonstrated that the spread of the runoff simulations among calibrated hydrologic models is substantially less than that among uncalibrated models based on three Colorado headwater basins.

5. Conclusions

Offline hydrologic simulations forced by SD global climate model output have become the practical method of choice for continental-scale hydrologic assessments. However, uncertainty in hydrologic simulations (e.g., runoff estimation) that comes solely from the choice of downscaling methods and hydrologic model has been

relatively unexplored relative to other uncertainty sources, such as the selection of emission scenarios and GCMs. The main goal of this study was to examine how differences in precipitation and temperature, derived by SD methods commonly used for practical application, propagate into hydrologic simulations and to assess the relative importance of both model choice and downscaling method choice. The following conclusions are summarized.

- Severe underestimation of precipitation from BCCA produces considerably different hydrologic portrayals than the other SD data examined in this paper.
- It is important to reproduce hydrologically relevant attributes (e.g., wet-day fraction, and diurnal temperature range) as well as precipitation amount or air temperature via SD. This paper showed that BCSDd, which has a much higher wet-day fraction, produced lower precipitation intensity and shortwave radiation than BCSDm and AR. Bias in those attributes affects evaporation processes differently among the models. Biases in the hydrologically relevant attributes caused by downscaling process are also propagated into estimations of other types of meteorological forcings (e.g., shortwave radiation) by empirical algorithm (e.g., MTCLIM).
- If BCCA is not considered, sensitivity of runoff estimates to choice of SD methods is by a factor of 1.5–3 smaller than the impacts of the hydrologic model choice. The discrepancy of particular hydrologic assessments (e.g., extreme runoff estimations, especially low runoff values) due to model choice becomes even greater than choice of SD methods.

The above conclusions on methodological choice impacts may be limited to historical hydrologic fidelity. Our current effort is to evaluate how the methodological choices affect projected hydrologic changes over the CONUS. Characteristics of differences in projected climate changes (i.e., temperature increase and precipitation changes) among the SD methods can be different than those of portrayal of historical climate conditions. Projected hydrologic changes are affected by how differently each SD method projects climate changes.

Finally, although this study does not intend to characterize the whole uncertainty range associated with choices of SD methods and hydrologic model choice, the results presented in this paper highlight that hydrologic assessments are more sensitive to choices of hydrologic models than SD method choices, especially if model parameters are not calibrated. To more fully characterize and understand uncertainty of hydrologic assessments from offline hydrologic modeling, the following research efforts are needed: 1) implementation of improved downscaling methods (Wilby et al. 2004; Maraun et al. 2010), 2) development of spatially consistent parameter estimations for continental-scale domains (Samaniego et al. 2010), and 3) analysis of more complete intermodel structural effect on hydrologic simulations (Clark et al. 2015a,b). Those efforts are expected to improve our understanding of uncertainty associated with hydrologic modeling and ultimately facilitate reduction of hydrologic uncertainty.

Acknowledgments. This work was financially supported by the U.S Bureau of Reclamation and the U.S Army Corps of Engineers. The authors thank Steven Markstrom, Steve Regan, and Roland Viger of the USGS for providing PRMS parameters and Youlong Xia and Justin Sheffield for providing VIC parameters. Finally, the authors are grateful to Stacey Archfield and three anonymous reviewers for their constructive comments.

REFERENCES

Anderson, E. A., 1976: A point energy and mass balance model of a snow cover. NOAA Tech. Rep. NWS 19, 150 pp. [Available online at http://amazon.nws.noaa.gov/articles/HRL_Pubs_PDF_May12_2009/HRL_PUBS_51-100/81_A_POINT_ENERGY_AND_MASS.pdf.]

Andreadis, K. M., P. Storck, and D. P. Lettenmaier, 2009: Modeling snow accumulation and ablation processes in forested environments. *Water Resour. Res.*, **45**, W05429, doi:10.1029/2008WR007042.

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.

Battaglin, W. A., L. E. Hay, and S. L. Markstrom, 2011: Simulating the potential effects of climate change in two Colorado basins and at two Colorado ski areas. *Earth Interact.*, **15**, 1–23, doi:10.1175/2011EI373.1.

Bennett, K. E., A. T. Werner, and M. Schnorbus, 2012: Uncertainties in hydrologic and climate change impact analyses in headwater basins of British Columbia. *J. Climate*, **25**, 5711–5730, doi:10.1175/JCLI-D-11-00417.1.

Bjerklie, D. M., T. J. Trombley, and R. J. Viger, 2011: Simulations of historical and future trends in snowfall and groundwater recharge for basins draining to Long Island Sound. *Earth Interact.*, **15**, doi:10.1175/2011EI374.1.

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. Meteor.*, **176**, 38–49, doi:10.1016/j.agrformet.2013.03.003.

Brekke, L., A. Wood, and T. Pruitt, 2014: Downscaled CMIP3 and CMIP5 hydrology projections: Release of hydrology projections, comparison with preceding information, and summary of user needs. USBR Tech Memo., 110 pp. [Available online at http://gdo-dcp.ucllnl.org/downscaled_cmip_projections/techmemo/BCSD5HydrologyMemo.pdf.]

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. USA*, **107**, 21 271–21 276, doi:10.1073/pnas.0912391107.

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.

——, A. W. Wood, N. Voisin, D. P. Lettenmaier, and R. N. Palmer, 2004: The effects of climate change on the hydrology and water resources of the Colorado River basin. *Climatic Change*, **62**, 337–363, doi:10.1023/B:CLIM.0000013684.13621.1f.

Christiansen, D. E., S. L. Markstrom, and L. E. Hay, 2011: Impacts of climate change on the growing season in the United States. *Earth Interact.*, **15**, doi:10.1175/2011EI376.1.

Clark, M. P., and Coauthors, 2015a: A unified approach for process-based hydrologic modeling: 1. Modeling concept. *Water Resour. Res.*, **51**, 2498–2514, doi:10.1002/2015WR017198.

——, and Coauthors, 2015b: A unified approach for process-based hydrologic modeling: 2. Model implementation and case studies. *Water Resour. Res.*, **51**, 2515–2542, doi:10.1002/2015WR017200.

Cosgrove, B. A., and Coauthors, 2003: Land surface model spin-up behavior in the North American Land Data Assimilation System (NLDAS). *J. Geophys. Res.*, **108**, 8845, doi:10.1029/2002JD003316.

Crane, R. G., B. Yarnal, E. J. Barron, and B. Hewitson, 2002: Scale interactions and regional climate: Examples from the Susquehanna River basin. *Hum. Ecol. Risk Assess.*, **8**, 147–158, doi:10.1080/20028091056782.

Dai, A., 2008: Temperature and pressure dependence of the rain–snow phase transition over land and ocean. *Geophys. Res. Lett.*, **35**, L12802, doi:10.1029/2008GL033295.

Delworth, T. L., and S. Manabe, 1988: The influence of potential evaporation on the variabilities of simulated soil wetness and climate. *J. Climate*, **1**, 523–547, doi:10.1175/1520-0442(1988)001<0523:TIOPEO>2.0.CO;2.

Déry, S. J., K. Stahl, R. D. Moore, P. H. Whitfield, B. Menounos, and J. E. Burford, 2009: Detection of runoff timing changes in pluvial, nival, and glacial rivers of western Canada. *Water Resour. Res.*, **45**, W04426, doi:10.1029/2008WR006975; Corrigendum, **45**, W06701, doi:10.1029/2009WR008244.

Dettinger, M., D. Cayan, M. Meyer, and A. Jeton, 2004: Simulated hydrologic responses to climate variations and change in the Merced, Carson, and American River basins, Sierra Nevada, California, 1900–2099. *Climatic Change*, **62**, 283–317, doi:10.1023/B:CLIM.0000013683.13346.4f.

Dirmeyer, P. A., X. Gao, M. Zhao, Z. Guo, T. Oki, and N. Hanasaki, 2006: GSWP-2: Multimodel analysis and implications for our perception of the land surface. *Bull. Amer. Meteor. Soc.*, **87**, 1381–1397, doi:10.1175/BAMS-87-10-1381.

Elsner, M., and Coauthors, 2010: Implications of 21st century climate change for the hydrology of Washington State. *Climatic Change*, **102**, 225–260, doi:10.1007/s10584-010-9855-0.

——, 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. Hydrometeor.*, **15**, 1384–1403, doi:10.1175/JHM-D-13-083.1.

Exbrayat, J.-F., W. Buytaert, E. Timbe, D. Windhorst, and L. Breuer, 2014: Addressing sources of uncertainty in runoff projections for a data scarce catchment in the Ecuadorian Andes. *Climatic Change*, **125**, 221–235, doi:10.1007/s10584-014-1160-x.

Feld, S. I., N. C. Cristea, and J. D. Lundquist, 2013: Representing atmospheric moisture content along mountain slopes: Examination using distributed sensors in the Sierra Nevada, California. *Water Resour. Res.*, **49**, 4424–4441, doi:10.1002/wrcr.20318.

Fritze, H., I. T. Stewart, and E. Pebesma, 2011: Shifts in western North American snowmelt runoff regimes for the recent warm decades. *J. Hydrometeor.*, **12**, 989–1006, doi:10.1175/2011JHM1360.1.

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**, 7167–7186, doi:10.1002/2014WR015559.

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.

Hanson, R. T., L. E. Flint, A. L. Flint, M. D. Dettinger, C. C. Faunt, D. Cayan, and W. Schmid, 2012: A method for physically based model analysis of conjunctive use in response to potential climate changes. *Water Resour. Res.*, **48**, W00L08, doi:10.1029/2011WR010774.

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.

——, S. L. Markstrom, and C. Ward-Garrison, 2011: Watershed-scale response to climate change through the twenty-first century for selected basins across the United States. *Earth Interact.*, **15**, doi:10.1175/2010EI370.1.

——, J. LaFontaine, and S. L. Markstrom, 2014: Evaluation of statistically downscaled GCM output as input for hydrological and stream temperature simulation in the Apalachicola–Chattahoochee–Flint River basin (1961–99). *Earth Interact.*, **18**, doi:10.1175/2013EI000554.1.

Hayhoe, K., and Coauthors, 2004: Emissions pathways, climate change, and impacts on California. *Proc. Natl. Acad. Sci. USA*, **101**, 12 422–12 427, doi:10.1073/pnas.0404500101.

——, and Coauthors, 2007: Past and future changes in climate and hydrological indicators in the US Northeast. *Climate Dyn.*, **28**, 381–407, doi:10.1007/s00382-006-0187-8.

Hungerford, R. D., R. Nemani, S. W. Running, and J. C. Coughlan, 1989: MTCLIM: A mountain microclimate simulation model. U.S. Forest Service Research Paper INT-414, 52 pp. [Available online at http://www.fs.fed.us/rm/pubs_int/int_rp414.pdf.]

Hurrell, J. W., and Coauthors, 2013: The Community Earth System Model: A framework for collaborative research. *Bull. Amer. Meteor. Soc.*, **94**, 1339–1360, doi:10.1175/BAMS-D-12-00121.1.

Idso, S. B., 1981: A set of equations for full spectrum and 8- to 14-μm and 10.5- to 12.5-μm thermal radiation from cloudless skies. *Water Resour. Res.*, **17**, 295–304, doi:10.1029/WR017i002p00295.

Jordan, R., 1991: A one-dimensional temperature model for a snow cover: Technical documentation for SNTHERM.89. Special Rep. 91-16, Cold Region Research and Engineers Laboratory, U.S. Army Corps of Engineers, Hanover, NH, 61 pp.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kimball, J. S., S. W. Running, and R. Nemani, 1997: An improved method for estimating surface humidity from daily minimum temperature. *Agric. For. Meteor.*, **85**, 87–98, doi:10.1016/S0168-1923(96)02366-0.

Koczot, K. M., S. L. Markstrom, and L. E. Hay, 2011: Effects of baseline conditions on the simulated hydrologic response to projected climate change. *Earth Interact.*, **15**, doi:10.1175/2011EI378.1.

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.

Lawrence, P. J., and T. N. Chase, 2007: Representing a new MODIS consistent land surface in the Community Land Model (CLM 3.0). *J. Geophys. Res.*, **112**, G01023, doi:10.1029/2006JG000168.

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, 281–310.

——, R. W. Lichty, B. M. Troutman, and L. G. Saindon, 1983: Precipitation–Runoff Modeling System: User's manual. Water-Resources Investigations Rep. 83-4238, 206 pp. [Available online at http://pubs.usgs.gov/wri/1983/4238/report.pdf.]

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, doi:10.1029/94JD00483.

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

Livneh, B., and D. P. Lettenmaier, 2012: Multi-criteria parameter estimation for the Unified Land Model. *Hydrol. Earth Syst. Sci.*, **16**, 3029–3048, doi:10.5194/hess-16-3029-2012.

Luce, C. H., and Z. A. Holden, 2009: Declining annual streamflow distributions in the Pacific Northwest United States, 1948–2006. *Geophys. Res. Lett.*, **36**, L16401, doi:10.1029/2009GL039407.

Maraun, D., and Coauthors, 2010: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Rev. Geophys.*, **48**, RG3003, doi:10.1029/2009RG000314.

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**, doi:10.1175/2010EI368.1.

Maurer, E. P., 2007: Uncertainty in hydrologic impacts of climate change in the Sierra Nevada, California, under two emissions scenarios. *Climatic Change*, **82**, 309–325, doi:10.1007/s10584-006-9180-9.

——, A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen, 2002: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States. *J. Climate*, **15**, 3237–3251, doi:10.1175/1520-0442(2002)015<3237:ALTHBD>2.0.CO;2.

——, L. Brekke, T. Pruitt, and P. B. Duffy, 2007: Fine-resolution climate projections enhance regional climate change impact studies. *Eos, Trans. Amer. Geophys. Union*, **88**, 504, doi:10.1029/2007EO470006.

——, H. G. Hidalgo, T. Das, M. D. Dettinger, and D. R. Cayan, 2010: The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California. *Hydrol. Earth Syst. Sci.*, **14**, 1125–1138, doi:10.5194/hess-14-1125-2010.

——, and Coauthors, 2014: An enhanced archive facilitating climate impacts and adaptation analysis. *Bull. Amer. Meteor. Soc.*, **95**, 1011–1019, doi:10.1175/BAMS-D-13-00126.1.

Meehl, G. A., C. Covey, B. McAvaney, M. Latif, and R. J. Stouffer, 2005: Overview of the Coupled Model Intercomparison Project. *Bull. Amer. Meteor. Soc.*, **86**, 89–93, doi:10.1175/BAMS-86-1-89.

Mendoza, P. A., and Coauthors, 2015a: Effects of hydrologic model choice and calibration on the portrayal of climate change impacts. *J. Hydrometeor.*, **16**, 762–780, doi:10.1175/JHM-D-14-0104.1.

——, M. P. Clark, N. Mizukami, E. D. Gutmann, J. R. Arnold, L. D. Brekke, and B. Rajagopalan, 2015b: How do hydrologic modeling decisions affect the portrayal of climate change impacts?. *Hydrol. Process.*, doi:10.1002/hyp.10684, in press.

Miller, W. P., R. 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. Plann. Manage.*, **138**, 412–420, doi:10.1061/(ASCE)WR.1943-5452.0000237.

——, G. M. DeRosa, S. Gangopadhyay, and J. B. Valdés, 2013: Predicting regime shifts in flow of the Gunnison River under changing climate conditions. *Water Resour. Res.*, **49**, 2966–2974, doi:10.1002/wrcr.20215.

Mizukami, N., V. Koren, M. Smith, D. Kingsmill, Z. Zhang, B. Cosgrove, and Z. Cui, 2013: The impact of precipitation type discrimination on hydrologic simulation: Rain–snow partitioning derived from HMT-West radar-detected bright-band height versus surface temperature data. *J. Hydrometeor.*, **14**, 1139–1158, doi:10.1175/JHM-D-12-035.1.

——, 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. Hydrometeor.*, **15**, 474–488, doi:10.1175/JHM-D-13-036.1.

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

Nasonova, O. N., Y. M. Gusev, and Y. E. Kovalev, 2011: Impact of uncertainties in meteorological forcing data and land surface parameters on global estimates of terrestrial water balance components. *Hydrol. Processes*, **25**, 1074–1090, doi:10.1002/hyp.7651.

Nicholas, R. E., and D. S. Battisti, 2012: Empirical downscaling of high-resolution regional precipitation from large-scale reanalysis fields. *J. Appl. Meteor. Climatol.*, **51**, 100–114, doi:10.1175/JAMC-D-11-04.1.

Nijssen, B., and Coauthors, 2003: Simulation of high latitude hydrological processes in the Torne–Kalix basin: PILPS Phase 2(e): 2: Comparison of model results with observations. *Global Planet. Change*, **38**, 31–53, doi:10.1016/S0921-8181(03)00004-3.

Niu, G.-Y., Z.-L. Yang, 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.

Oleson, K. W., and Coauthors, 2010: Technical description of version 4.0 of the Community Land Model (CLM). NCAR Tech. Note NCAR/TN-478+STR, 257 pp., doi:10.5065/D6FB50WZ.

Pierce, D. W., A. L. Westerling, and J. Oyler, 2013: Future humidity trends over the western United States in the CMIP5 global climate models and Variable Infiltration Capacity hydrological modeling system. *Hydrol. Earth Syst. Sci.*, **17**, 1833–1850, doi:10.5194/hess-17-1833-2013.

——, D. R. Cayan, and B. L. Thrasher, 2014: Statistical downscaling using localized constructed analogs (LOCA). *J. Hydrometeor.*, **15**, 2558–2585, doi:10.1175/JHM-D-14-0082.1.

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.

Regonda, S. K., B. Rajagopalan, M. Clark, and J. Pitlick, 2005: Seasonal cycle shifts in hydroclimatology over the western United States. *J. Climate*, **18**, 372–384, doi:10.1175/JCLI-3272.1.

Risley, J., H. Moradkhani, L. Hay, and S. Markstrom, 2011: Statistical comparisons of watershed-scale response to climate change in selected basins across the United States. *Earth Interact.*, **15**, doi:10.1175/2010EI364.1.

Sagarika, S., A. Kalra, and S. Ahmad, 2014: Evaluating the effect of persistence on long-term trends and analyzing step changes in streamflows of the continental United States. *J. Hydrol.*, **517**, 36–53, doi:10.1016/j.jhydrol.2014.05.002.

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.

Sharma, D., and M. Babel, 2013: Application of downscaled precipitation for hydrological climate-change impact assessment in the upper Ping River basin of Thailand. *Climate Dyn.*, **41**, 2589–2602, doi:10.1007/s00382-013-1788-7.

Sheffield, J., G. Goteti, F. Wen, and E. F. Wood, 2004: A simulated soil moisture based drought analysis for the United States. *J. Geophys. Res.*, **109**, D24108, doi:10.1029/2004JD005182.

——, B. Livneh, and E. F. Wood, 2012: Representation of terrestrial hydrology and large-scale drought of the continental United States from the North American Regional Reanalysis. *J. Hydrometeor.*, **13**, 856–876, doi:10.1175/JHM-D-11-065.1.

Stewart, I. T., D. R. Cayan, and M. D. Dettinger, 2005: Changes toward earlier streamflow timing across western North America. *J. Climate*, **18**, 1136–1155, doi:10.1175/JCLI3321.1.

Stoner, A. M. K., K. Hayhoe, X. Yang, and D. J. Wuebbles, 2013: An asynchronous regional regression model for statistical downscaling of daily climate variables. *Int. J. Climatol.*, **33**, 2473–2494, doi:10.1002/joc.3603.

Surfleet, C. G., 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.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, doi:10.1175/BAMS-D-11-00094.1.

Teutschbein, C., F. Wetterhall, and J. Seibert, 2011: Evaluation of different downscaling techniques for hydrological climate-change impact studies at the catchment scale. *Climate Dyn.*, **37**, 2087–2105, doi:10.1007/s00382-010-0979-8.

Thornton, P. E., and S. W. Running, 1999: An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation. *Agric. For. Meteor.*, **93**, 211–228, doi:10.1016/S0168-1923(98)00126-9.

——, H. Hasenauer, and M. A. White, 2000: Simultaneous estimation of daily solar radiation and humidity from observed temperature and precipitation: An application over complex terrain in Austria. *Agric. For. Meteor.*, **104**, 255–271, doi:10.1016/S0168-1923(00)00170-2.

Thrasher, B., E. P. Maurer, C. McKellar, and P. B. Duffy, 2012: Technical note: Bias correcting climate model simulated daily temperature extremes with quantile mapping. *Hydrol. Earth Syst. Sci.*, **16**, 3309–3314, doi:10.5194/hess-16-3309-2012.

U.S. Bureau of Reclamation, 2011: West-wide climate risk assessments: Bias-corrected and spatially downscaled surface water projections. Tech. Memo. 86-68210-2011-01, 138 pp. [Available online at http://www.usbr.gov/watersmart/docs/west-wide-climate-risk-assessments.pdf.]

Vano, J. A., T. Das, and D. P. Lettenmaier, 2012: Hydrologic sensitivities of Colorado River runoff to changes in precipitation and temperature. *J. Hydrometeor.*, **13**, 932–949, doi:10.1175/JHM-D-11-069.1.

——, and Coauthors, 2014: Understanding uncertainties in future Colorado River streamflow. *Bull. Amer. Meteor. Soc.*, **95**, 59–78, doi:10.1175/BAMS-D-12-00228.1.

Viger, R. J., 2014: Preliminary spatial parameters for PRMS based on the Geospatial Fabric, NLCD2001 and SSURGO. U.S. Geological Survey, accessed 9 November 2015, doi:10.5066/F7WM1BF7.

——, L. E. Hay, S. L. Markstrom, J. W. Jones, and G. R. Buell, 2011: Hydrologic effects of urbanization and climate change on the Flint River basin, Georgia. *Earth Interact.*, **15**, doi:10.1175/2010EI369.1.

Walker, J. F., L. E. Hay, S. L. Markstrom, and M. D. Dettinger, 2011: Characterizing climate-change impacts on the 1.5-yr flood flow in selected basins across the United States: A probabilistic approach. *Earth Interact.*, **15**, doi:10.1175/2010EI379.1.

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. Hydrometeor.*, **14**, 677–699, doi:10.1175/JHM-D-12-0102.1.

Wilby, R. L., H. Hassan, and K. Hanaki, 1998: Statistical downscaling of hydrometeorological variables using general circulation model output. *J. Hydrol.*, **205**, 1–19, doi:10.1016/S0022-1694(97)00130-3.

——, L. E. Hay, W. J. Gutowski Jr., R. W. Arritt, E. S. Takle, Z. Pan, G. H. Leavesley, and M. P. Clark, 2000: Hydrological responses to dynamically and statistically downscaled climate model output. *Geophys. Res. Lett.*, **27**, 1199–1202, doi:10.1029/1999GL006078.

——, S. P. Charles, E. Zorita, B. Timbal, P. Whetton, and L. O. Mearns, 2004: Guidelines for use of climate scenarios developed from statistical downscaling methods. IPCC Doc., 27 pp. [Available online at http://www.ipcc-data.org/guidelines/dgm_no2_v1_09_2004.pdf.]

Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier, 2004: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change*, **62**, 189–216, doi:10.1023/B:CLIM.0000013685.99609.9e.

Xia, Y., and Coauthors, 2012: 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.