

This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use(<https://www.springernature.com/gp/open-research/policies/accepted-manuscript-terms>), but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: <http://dx.doi.org/10.1007/s00382-019-04702-7>

1 **Title Page:**

2 **Title:** Improving probabilistic hydroclimatic projections through high-resolution
3 convection-permitting climate modeling and Markov chain Monte Carlo simulations

4
5
6 **Author names and affiliations:** S. Wang^{1,*}, Y. Wang²

7
8 ¹Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic
9 University, Hong Kong, China

10 ²Department of Geosciences, Texas Tech University, Lubbock, Texas, USA

11
12
13 ^{*}Corresponding author. Phone: (852) 3400-3896; Email: shuo.s.wang@polyu.edu.hk

Abstract Understanding future changes in hydroclimatic variables plays a crucial role in improving resilience and adaptation to extreme weather events such as floods and droughts. In this study, we develop high-resolution climate projections over Texas by using the convection-permitting Weather Research and Forecasting (WRF) model with 4 km horizontal grid spacing, and then produce the Markov chain Monte Carlo (MCMC)-based hydrologic forecasts in the Guadalupe River basin which is the primary concern of the Texas Water Development Board and the Guadalupe-Blanco River Authority. The Parameter-elevation Regressions on Independent Slopes Model (PRISM) dataset is used to verify the WRF climate simulations. The Model Parameter Estimation Experiment (MOPEX) dataset is used to validate probabilistic hydrologic predictions. Projected changes in precipitation, potential evapotranspiration (PET) and streamflow at different temporal scales are examined by dynamically downscaling climate projections derived from 15 Coupled Model Intercomparison Project Phase 5 (CMIP5) general circulation models (GCMs). Our findings reveal that the Upper Coast Climate Division of Texas is projected to experience the most remarkable wetting caused by precipitation and PET changes, whereas the most significant drying is expected to occur for the North Central Texas Climate Division. The dry Guadalupe River basin is projected to become drier with a substantial increase in future drought risks, especially for the summer season. And the extreme precipitation events are projected to increase in frequency and intensity with a reduction in overall precipitation frequency, which may result in more frequent occurrences of flash floods and drought episodes in the Guadalupe River basin.

Keywords Convection permitting; High-resolution climate projection; Hydroclimatic changes; Markov chain Monte Carlo; Pseudo global warming

1 Introduction

Understanding future changes in hydroclimatic variables (e.g., precipitation, potential evapotranspiration and streamflow) is vital for improving resilience and adaptation to extreme weather events such as floods and droughts (Pfahl et al. 2017). Thus, hydroclimatic projections with a reasonable level of accuracy play a crucial role in water resources management, agricultural production, food security, and hazard mitigation (Yuan et al. 2013; Carvalho and Wang, 2019). The climate model-driven hydrologic prediction systems are recognized as a powerful means for assessing hydrologic regimes response to climate change, which have attracted increasing attention from the hydroclimatic science community in recent years (Leong and Donner 2015; Krysanova et al. 2017; Pechlivanidis et al. 2017).

General circulation models (GCMs) have been commonly used to study the climate response due to changes in anthropogenic greenhouse gas emissions (Ashfaq et al. 2016). The climate variables (e.g., precipitation and temperature) derived from GCMs can thus be used as forcing data to drive hydrologic models for predicting the climate-induced changes in hydrologic regimes. For instance, Hagemann et al. (2011) used three GCMs to produce daily time series of future temperature and precipitation under two emission scenarios, and then the bias-corrected outputs from GCM simulations were used to force two global hydrologic models for predicting river discharge in catchments located in different climate regions. Svensson et al. (2015) used a sophisticated conceptual hydrologic model to forecast winter river flows and groundwater levels in the UK catchments, and the hydrologic simulations were forced with winter rainfall forecasts derived from the Met Office Hadley Centre Global Environment Model version 3 (HadGEM3). Bennett et al.

(2016) used a coupled ocean-atmosphere GCM to produce rainfall forecasts, and then the calibrated rainfall forecasts were used to force a conceptual hydrologic model for predicting the response of runoff to rainfall in two Australian catchments. Although GCMs play an important role in assessing climate change impacts on hydrologic regimes, the spatial resolution of GCMs is too coarse to capture regional or local hydroclimatic characteristics. As a result, the dynamical downscaling simulations using regional climate models (RCMs) are necessary to represent the basin-scale features and dynamics.

Over the past decade, tremendous efforts have been made in the development and application of dynamical downscaling techniques for the high-resolution projections of hydroclimatic changes (Piao et al. 2010; Giorgi et al. 2014; Wang et al. 2014; Erler et al. 2015; Music et al. 2015; Rajbhandari et al. 2015; Amin et al. 2016; Li et al. 2016; Erler and Peltier 2017). The state-of-the-art RCMs have been extensively used to dynamically downscale the coarse-resolution global climate simulations for providing detailed climate information at a regional scale. However, RCMs heavily rely on the convection parameterization schemes due to their inability to explicitly resolve deep convection which is a key process in the climate system and the dominant source of precipitation (Kendon et al. 2012). The use of convection parameterization schemes is considered as a major source of model errors and uncertainties in future climate projections, which suffers from significant limitations including misrepresentation of the diurnal cycle of convective precipitation, overestimation of the frequency of low-precipitation events, and underestimation of hourly precipitation intensities (Prein et al. 2015). Thus, the convection-permitting climate models with horizontal grid spacing of ≤ 4 km have received great attention in recent years and have been proven to be powerful and promising tools for

86 producing more reliable and process-based climate information, especially for improving
87 the representation of extreme precipitation associated with mesoscale convective systems
88 (Westra et al. 2014; Fosser et al. 2015; Brisson et al. 2016; Kendon et al. 2017; Leutwyler
89 et al. 2017; Liu et al. 2017; Prein et al. 2017; Chan et al. 2018; Wang et al. 2018b).

90 The projected changes in precipitation patterns have a significant impact on future
91 runoff characteristics. Little effort has been devoted to examining the response of runoff to
92 rainfall at a basin scale based on the convection-permitting climate simulations. It is thus
93 desired to project future changes in hydroclimatic regimes at convection-permitting
94 resolutions (kilometer-scale grid spacing). Rainfall-runoff models can be used to predict
95 the hydrologic regimes in response to climate change. In recent years, a consensus view
96 from the hydrologic community is that probabilistic hydrologic predictions are superior to
97 deterministic predictions due to various sources of uncertainty (Ramos et al. 2013;
98 DeChant and Moradkhani 2015; Sadegh et al. 2015; Farmer and Vogel 2016; Roy et al.
99 2016; Wang et al. 2018a). It is thus necessary to address uncertainty in hydroclimatic
100 projections for improving the reliability and robustness of projected hydroclimatic
101 information.

102 In this work, we develop probabilistic high-resolution hydroclimatic projections over
103 Texas through convection-permitting climate simulations and Markov chain Monte Carlo
104 (MCMC)-based hydrologic predictions. Specifically, dynamical downscaling will be
105 performed using the convection-permitting Weather Research and Forecasting (WRF)
106 model with horizontal grid spacing of 4 km to produce the 21st century projections of
107 precipitation and potential evapotranspiration (PET) which are the most important
108 components of the hydrologic cycle. MCMC simulations will then be conducted by using

a well-known rainfall-runoff model (Hymod) to generate probabilistic predictions of daily and seasonal streamflow time series in response to climate change in the Guadalupe River basin which is one of the major river basins proposed by the Texas Water Development Board and is the Guadalupe-Blanco River Authority's primary concern. The Parameter-elevation Regressions on Independent Slopes Model (PRISM) dataset will be used to verify the convection-permitting WRF climate simulations. The Model Parameter Estimation Experiment (MOPEX) dataset and the data collected at the United States Geological Survey (USGS) river flow gauging stations will be used to calibrate and validate the hydrologic model.

This paper is organized as follows. Sect. 2 describes model configurations, algorithms, and datasets used to conduct climate and hydrologic simulations. Sect. 3 presents a thorough analysis and discussion on the evaluation of convection-permitting WRF climate experiments and the projection of future climate information as well as the uncertainty assessment of hydrologic predictions and probabilistic streamflow projections. Finally, conclusions drawn from this study are summarized in Sect. 4.

2 Models, algorithms and data sources

2.1 Convection-permitting climate modeling

The Weather Research and Forecasting (WRF) model v3.7.1 was used to conduct the convection-permitting climate simulations over Texas. The study domain is shown in Fig. 1a, which covers a region of $1,520 \text{ km} \times 1,400 \text{ km}$ (380×350 grid points) with 51 stretched vertical levels topped at 50 hPa. The horizontal grid spacing is 4 km, which is fine enough to allow for the explicit simulation of convection. Thus, no convection parameterization is

needed for performing the 4 km WRF climate simulations (Liu et al. 2011). The historical climate simulation spans the 15-year period of 1981–1995. The initial and lateral boundary conditions were provided by the National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR) that had a 6-hourly temporal resolution and a $0.5^\circ \times 0.5^\circ$ spatial resolution (Saha et al. 2010). Recent studies have shown that the convection-permitting WRF model can be properly configured to dynamically downscale reanalysis data to a high-resolution horizontal grid spacing of 4 km with no coarse grids in between (Liu et al., 2017; Wang et al., 2018b). The WRF model was configured with the Thompson cloud microphysics scheme (Thompson et al. 2008), the Yonsei University (YSU) planetary boundary layer scheme (Hong and Pan 1996), the revised Monin-Obukhov surface layer scheme (Jimenez et al. 2012), the Rapid Radiative Transfer Model (RRTMG) shortwave and longwave radiation scheme (Iacono et al. 2008), and the Noah-MP land surface scheme (Niu et al. 2011; Yang et al. 2011). The choice of the model configuration was made by closely following Liu et al. (2017) and Wang et al. (2018b). Results show that the model-simulated precipitation and near-surface temperature agree well with the PRISM data (see Sect. 3.1).

The high-resolution (4 km) climate simulations for a future period 2085–2099 were also forced with the CFSR data, but the initial and boundary conditions were continuously perturbed by using the Pseudo–Global Warming (PGW) technique (Liu et al. 2017). In the PGW method, initial and boundary conditions for regional model integrations are given by the sum of reanalysis data (observations) and a perturbation estimated from simulations with GCMs. A basic assumption of the PGW method is a linear coupling of observations and climate perturbation signals with neglect of nonlinear interactions between climate

change and the interannual variation of regional climate systems. The perturbed physical fields include temperature, geopotential, specific humidity, horizontal wind, sea surface temperature, sea level pressure, soil temperature, and sea ice. As shown in Eq. (1), the climate perturbation was estimated by a 30-year multi-model ensemble mean climate change signal. The historical (1976–2005) and future (2071–2100) climate projections were derived from the Coupled Model Intercomparison Project Phase 5 (CMIP5) under the Representative Concentration Pathway (RCP) 8.5 emission scenario.

$$\text{WRF}_{\text{input}} = \text{CFSR} + (\text{CMIP5}_{2071-2100} - \text{CMIP5}_{1976-2005}). \quad (1)$$

Since the climate change signal from a single GCM is not representative due to the large range of climate sensitivity and the large internal variability of current GCMs, a multi-model ensemble mean climate difference between past and future decades is used to quantify the climate change from the greenhouse gas forcing in dynamical downscaling. The use of a multi-model ensemble mean forcing improves upon the single-model-run method used in the PGW simulations (Rasmussen et al. 2014). A total of 15 CMIP5 GCMs were chosen in this study to produce the ensemble mean climate change signal based on their ability to simulate the 20th century climate over North America (Liu et al. 2017). The atmospheric grid spacing and the ensemble members of 15 GCMs are summarized in Table 1. The CFSR reanalysis data was perturbed every 6 hours by the derived climate change signal in order to provide the WRF model initial and boundary conditions for future climate projections.

Sea level pressure (SLP) and surface temperatures (ST) are the two most important parameters that are perturbed using the PGW technique. Fig. 2 shows the CMIP5 multi-model ensemble mean seasonal difference of SLP and ST between future (2071–2100) and

past (1976–2005) periods over the study domain. DJF, MAM, JJA, and SON represent the seasons of winter, spring, summer, and fall, respectively. In comparison, the projected SLP is lower during the winter (DJF) and spring (MAM) months (Figs. 2a and c). The domain-average reduction of SLP is projected to be 0.45 hPa in DJF and 0.56 hPa in MAM, respectively. The projected SLP increases for the summer months (JJA), and the domain-average increase of SLP is projected to be 0.23 hPa (Fig. 2e). The change of SLP will be relatively insignificant in fall (SON) with a domain-average decrease of 0.02 hPa. In addition, the summer ST is projected to become higher compared to the other seasons, and the domain-average increase of summer ST will be 4.6 °C (Fig. 2f). The domain-average ST increase is projected to be 3.8 °C, 4.1 °C, and 4.5 °C for DJF (Fig. 2b), MAM (Fig. 2d), and SON (Fig. 2h), respectively. The impact of climate change on the other important variables projected by the PGW simulations will be examined in Sect. 3.2.

2.2 Rainfall-runoff modelling

To examine the impact of climate change on hydrological regimes, we applied the conceptual hydrologic model (Hymod) for predicting daily streamflow in the Guadalupe River basin (Fig. 1b) based on the projected changes of precipitation and PET through the convection-permitting WRF climate simulations. PET was calculated by using the FAO-56 Penman–Monteith Equation (Allen et al. 1998). This equation has been widely used to calculate PET due to its ability to accurately estimate PET at different time scales (Boulard et al. 2016; Srivastava et al. 2016). Thus, PET was estimated based on the values of four WRF model outputs including air temperature, relative humidity, wind speed, and solar radiation.

The Hymod is a well-known rainfall-runoff model that has been extensively used to

predict streamflow in basins with different hydroclimatic regimes around the world (Bulygina and Gupta 2011; Herman et al. 2013; Sadegh and Vrugt 2013; Wang et al. 2015; Razavi and Gupta 2016; Abera et al. 2017; Roy et al. 2017). The Hymod consists of a soil moisture accounting module based on a probability-distributed soil storage capacity principle introduced by Moore (1985) and a routing module. The distribution function of storage capacity is defined as:

$$F(C) = 1 - \left(1 - \frac{C}{C_{\max}}\right)^{b_{\exp}} \quad 0 \leq C \leq C_{\max} . \quad (2)$$

where C_{\max} represents the maximum soil moisture storage capacity within the basin, and b_{\exp} describes the degree of spatial variability of the storage capacity. The routing module includes a sequence of three quick-flow tanks that represent surface flow, in parallel with a slow-flow tank that represents subsurface flow. The residence time parameters of quick- and slow-flow tanks are denoted by R_q and R_s , respectively. In the Hymod, the runoff generation mechanism is represented as a rainfall excess process that divides the excess rainfall into quick- and slow-flow tanks through a partitioning factor β , and the generated streamflow is the addition of the outputs from quick- and slow-flow tanks. Thus, the Hymod has five model parameters, including C_{\max} , b_{\exp} , β , R_s , and R_q . Since these parameters cannot be measured directly, they are usually specified with uncertainty ranges (as shown in Table 2) and then can be estimated by model calibration against observations.

2.3 Markov chain Monte Carlo simulation

The Markov chain Monte Carlo (MCMC) algorithm is recognized as a powerful tool for efficient calibration of hydrologic models within a Bayesian framework. The Bayes' theorem is used to infer the posterior distributions of model parameters by updating the

223 prior information on parameters when new observations become available:

$$224 \quad p(\theta | \tilde{Y}) = \frac{p(\theta) p(\tilde{Y} | \theta)}{p(\tilde{Y})}, \quad (3)$$

225 where θ denotes model parameters, \tilde{Y} denotes observations, $p(\theta)$ and $p(\theta | \tilde{Y})$
 226 represent prior and posterior distributions of model parameters, respectively.

227 $p(\tilde{Y} | \theta) \equiv L(\theta | \tilde{Y})$ is the likelihood function that quantifies the probability that the
 228 observed data can be generated by a particular parameter set, and $p(\tilde{Y})$ is the evidence
 229 that acts as a normalization constant. In practice, $p(\tilde{Y})$ can be removed from the Bayesian
 230 analysis for posterior parameter estimation, and thus the formulation of Eq. (3) can be
 231 simplified as:

$$232 \quad p(\theta | \tilde{Y}) \propto p(\theta) L(\theta | \tilde{Y}), \quad (4)$$

233 where $L(\theta | \tilde{Y})$ is the likelihood function that takes into account the difference between
 234 model simulations and the corresponding observations. If the error residuals are assumed
 235 to be uncorrelated, homoscedastic, and Gaussian-distributed with mean zero, the likelihood
 236 function can be formulated as (Sorooshian and Dracup 1980):

$$237 \quad L(\theta | \tilde{Y}) = \prod_{t=1}^n \frac{1}{\sqrt{2\pi\tilde{\sigma}^2}} \exp\left\{-\frac{1}{2}\tilde{\sigma}^{-2}[\tilde{y}_t - y_t(\theta)]^2\right\}, \quad (5)$$

238 where $\tilde{\sigma}$ is the estimated standard deviation of measurement error, \tilde{y}_t is the observation
 239 at time t , and $y_t(\theta)$ is the model simulation given parameter θ at time t . For algebraic
 240 simplicity and numerical stability, Eq. (5) can be logarithmically transformed to (Sadegh
 241 et al. 2017):

$$L(\theta | \tilde{Y}) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \tilde{\sigma}^2 - \frac{1}{2} \tilde{\sigma}^{-2} \sum_{i=1}^n [\tilde{y}_i - y_i(\theta)]^2. \quad (6)$$

When the prior parameter distributions are specified, the posterior distributions of model parameters can be estimated through the MCMC simulation (e.g., by repeated Monte Carlo sampling in the prior parameter space).

The theoretical basis of MCMC simulation is to construct a Markov chain that generates a random walk with an acceptance/rejection rule to converge to the stationary distribution (Andrieu and Thoms 2008). The fundamental block of many MCMC algorithms is the random walk Metropolis algorithm introduced by Metropolis et al. (1953), and the Metropolis algorithm operates as follows. First, a candidate point x_p is proposed at each step in the Markov chain. Next, the candidate can be either accepted or rejected by calculating the Metropolis acceptance probability:

$$p_{\text{accept}}(x_c \rightarrow x_p) = \min \left[1, \frac{p(x_p)}{p(x_c)} \right]. \quad (7)$$

where x_c and x_p denote the current point and the candidate point, respectively. $P(x_c)$ and $P(x_p)$ represent the probability density of the current point and the candidate point, respectively. If the candidate point is accepted, the chain moves to x_p ; otherwise, the chain remains at its current location (Vrugt 2016). The Markov process proceeds by randomly walking through the search space, and eventually the Markov chain converges to a stationary distribution with a posterior probability density function (PDF).

To estimate the posterior distributions of hydrologic model parameters, we adopted the Differential Evolution Adaptive Metropolis (DREAM) algorithm proposed by Vrugt et al. (2008) to implement the MCMC simulation. The DREAM algorithm is an advanced MCMC without unrealistic error assumptions, which can accurately estimate the posterior

distributions of model parameters and thus have superior predictive performance (Lu et al., 2017). The DREAM algorithm runs multiple Markov chains simultaneously in parallel to robustly explore the search space and uses the genetic algorithm for population evolution with a Metropolis selection rule to update the chains sequentially. The MCMC evolution proceeds until the convergence to the posterior distribution is achieved. The convergence of DREAM can be estimated by using the multi-chain \hat{R} diagnostic of Gelman and Rubin (1992). The \hat{R} diagnostic is a relatively robust estimator for convergence analysis as the DREAM algorithm runs multiple chains from different starting points. As declared by Vrugt (2016), the convergence of the posterior parameter distribution is achieved when the \hat{R} - statistic drops below a value of 1.2. More details about the DREAM algorithm can be found in Vrugt et al. (2008, 2009).

2.4 Data sources

The PRISM dataset was used to evaluate the historical WRF simulations over Texas. The PRISM dataset was developed by the Oregon State University's PRISM Climate Group (<http://www.prism.oregonstate.edu>) by using the "terrain-aware" interpolation techniques (Daly et al. 2008). PRISM is a gridded dataset with a 4×4 km grid resolution based on a number of gauge networks. It includes nearly 13,000 surface stations for precipitation and 10,000 for temperature. The two main sources of precipitation data used in PRISM include the SNOwpack TELemetry network and the National Weather Service Cooperative Observer Program (COOP) gauge network. The PRISM precipitation and temperature data have been used as a reference in a number of RCM studies (Silverman et al. 2013; Liu et al. 2017; Wang et al. 2018b). The PRISM precipitation and near-surface temperature data used in this study were regridded to the 4 km WRF domain by using the bilinear method.

The hydrologic model simulations were conducted using the MOPEX dataset described by Duan et al. (2006). A total of 16 years of data for the period from January 1980 to December 1995 were collected to predict daily streamflow in the Guadalupe River basin, south Texas. And the first year was used as a spin-up period to reduce sensitivity to the state-value initialization. The Hymod was forced with daily precipitation and potential evapotranspiration, and it was evaluated using daily streamflow data obtained from the USGS river flow gauging stations. Furthermore, the WRF-simulated forcing data including daily precipitation and PET were used to predict daily streamflow in order to perform a comparison of hydrologic simulations using different datasets.

The MCMC simulations using the DREAM algorithm were performed to calibrate the Hymod against streamflow observations over a period of 10 years from 1/1/1981 to 31/12/1990. And the hydrologic simulations with the derived posterior distributions of model parameters were then carried out to validate the Hymod by using streamflow observations over a period of 5 years from 1/1/1991 to 31/12/1995. When the Hymod was calibrated and validated by using historical data, the probabilistic daily streamflow projections of the Guadalupe River basin were produced based on the high-resolution projections of daily precipitation and PET by the end of the 21st century using the convection-permitting WRF climate simulations.

3 Results and discussion

3.1 Evaluation of convection-permitting WRF climate simulations

Fig. 3 shows a comparison of spatial patterns of the 2-m temperature generated from the PRISM observations, the WRF simulations, and model bias (WRF– PRISM). In general,

the WRF model captures the spatial distribution of the near-surface temperature reasonably well. The temperature pattern derived from PRISM can be reproduced by the WRF simulations for all seasons, and the model bias is smaller than 3 °C over the study domain. The blank space on PRISM and bias maps is generated due to the fact that the PRISM dataset only covers the continental United States.

As shown in Fig. 3, the WRF-simulated temperature generally shows a warm bias although the spatial pattern of model bias varies for different seasons. For the winter months (Fig. 3c), a cold temperature bias of approximately 2 °C appears in the northeastern domain, and a similar magnitude of cold bias is observed in the central and south Texas for the summer season (Fig. 3i). For spring (Fig. 3f) and fall seasons (Fig. 3l), a warm bias appears over the continental region except the northeastern corner of the domain where a cold bias of less than 1 °C is observed. As shown in Fig. 4, the spatial correlation coefficients are greater than 0.96 for all seasons, which indicates a good performance of the WRF model. In comparison, there is a relatively smaller domain-average temperature bias of 0.9 °C in winter (DJF) and summer (JJA). The derived smallest and largest root mean square error (RMSE) values are 1.1 °C for the summer season (JJA) and 1.7 °C for the spring season (MAM), respectively. In addition, the scatter plot displays that almost all blue dots that represent the grid cells within the Guadalupe River basin lie above the 1:1 line. This indicates that the temperature bias can lead to an overestimation of projected warming over the study basin. It should be noted that the warm temperature bias is less than 1.5 °C for all seasons, which is relatively small compared to the observational spread.

Fig. 5 presents a comparison of spatial distributions of the 15-year annual mean precipitation from PRISM and the WRF simulations. The spatial pattern of the PRISM-

derived precipitation can be reproduced by the WRF simulations reasonably well. In addition, the WRF-simulated precipitation tends to have a wet bias of less than 240 mm/year in the western domain, and the wet bias becomes a dry bias toward the southeast. The most significant wet bias is observed in the northeastern boundary although its relative bias is below 50%. Fig. 6 shows the spatial patterns of absolute and relative model biases of seasonal precipitation. The WRF model tends to overpredict the seasonal precipitation for the spring (MAM) and summer (JJA) months over the study domain. Specifically, the WRF-simulated spring precipitation has a wet bias over almost the entire domain, while the summer precipitation has a wet bias in the northern region and a dry bias in the southern region. In comparison, the model bias is smaller for the winter (DJF) and fall (SON) months. As shown in Fig. 7, the spatial correlation coefficients are greater than 0.80 for all seasons. This indicates that the WRF model is able to well reproduce the seasonal precipitation pattern over the study domain, especially for the winter season (DJF) with an RMSE value of 37.3 mm and the fall season (SON) with an RMSE value of 34.9 mm. In addition, the WRF model has a relatively small dry bias for the winter and fall seasons because most of the blue dots lie below the 1:1 line. Contrarily, a significant wet bias is observed for the warm seasons of spring (MAM) and summer (JJA). In general, the WRF model is skillful in simulating temperature and precipitation on the basis of spatial correlation over the study domain. Specifically, the WRF-simulated temperature has a relatively small domain-average bias for the winter and summer seasons, while the precipitation bias is relatively small for the winter and fall seasons.

3.2 Projection of high-resolution climate information

Since precipitation and PET are the most important components of the hydrologic cycle,

the high-resolution projection of future precipitation and PET plays a crucial role in examining the climate-induced changes in hydrologic regimes. Figs. 8a and b display the absolute and relative differences of the 15-year annual mean precipitation between past and future climates. The significant drying appears at the northeastern corner of the study domain, with a considerable reduction of 300 mm/year in the amount of precipitation. Most areas are expected to become wetter in south Texas, the eastern Mexico, and the Gulf of Mexico (Fig. 8a). As shown in Figs. 8c–j, the areas with the most significant drying and wetting caused by precipitation changes appear during the summer season. In general, the relative increase and reduction of annual precipitation will be less than 25% (Fig. 8b). The southwestern part of the domain is expected to become drier with a reduction of 150 mm/year in the amount of precipitation under global warming (Fig. 8a). The northern part of the domain is also projected to become drier except for a small region at the boundary between Texas and Oklahoma.

To examine seasonal precipitation changes in a warming climate, Figs. 8c–j present the differences of the 15-year seasonal mean precipitation between past and future climates. For the winter months, the study domain is dominated by a wetting climate except for the regions in Mexico or in the Gulf of Mexico (Fig. 8c). The most significant seasonal wetting with an increase of 60 mm in the amount of precipitation appears at the boundary between Oklahoma and Arkansas. Contrarily, the most significant seasonal drying with a precipitation decrease of 80 mm appears at the boundary between Louisiana and the Gulf of Mexico. The maximum relative reduction of seasonal precipitation is projected to be as large as 60% over the region of Mexico (Fig. 8d). In addition, the spatial patterns of the precipitation change for the spring (MAM) and fall (SON) seasons are more scattering in

comparison with the winter (DJF) and summer (JJA) seasons. For the spring season, the most significant wetting with a precipitation increase of approximately 60 mm appears at the boundary between Texas and Oklahoma. And a drying trend is projected in most regions of Texas. For the fall season, the amount of precipitation is projected to decrease as large as 120 mm at the northeastern corner of the domain (Fig. 8i). For the summer season, the most significant drying appears at the northeastern corner of the domain, whereas the most significant wetting with a magnitude larger than 120 mm is projected at the boundary between Texas and the Gulf of Mexico (Fig. 8g). In comparison, the second most significant wetting appears at the boundary between Texas and Mexico. The maximum relative increase of seasonal precipitation is projected to be as large as 50% over the wetting regions (Fig. 8h). In comparison, it can be seen that the most significant seasonal drying and wetting appear during the summer months.

According to Griffiths et al. (1990), the State of Texas can be divided into 10 climate divisions (Fig. 1a). In this study, the high-resolution WRF simulations allow a close look at the seasonal cycle of precipitation change over the 10 climate divisions of Texas. In general, the precipitation change is projected to be more significant for the summer season compared with the other seasons (Fig. 9), which is consistent with the results shown in Fig. 8. The seasonal cycle of precipitation change looks similar for climate divisions 1 (High Plains), 3 (North Central Texas), and 5 (Trans Pecos), where a significant drying appears for the summer months. The relative domain-averaged reduction of precipitation is projected to be as large as 30.7% for division 1 in June and 31.3% for division 5 in August. In comparison, the seasonal cycles of precipitation change for divisions 2 (Low Rolling Plains), 4 (East Texas), 6 (Edwards Plateau), and 10 (Lower Valley) are more scattering.

Divisions 7 (South Central Texas) and 8 (Upper Coast) appear to be similar for the most significant wetting in June, and the precipitation change is less significant for the other months. For division 9 (South Texas), most months exhibit a wetting signal except for January and March. In addition, an annual wetting is projected to occur for divisions 2, 7, 8, 9, and 10, whereas divisions 1, 3, 4, 5 and 6 are expected to experience an annual drying. The most significant wetting with a precipitation increase of 78.5 mm/year is projected to occur for division 8, whereas the most significant drying with a precipitation decrease of 46.8 mm/year is expected to occur for division 3.

In addition to the convection-permitting projection of precipitation changes, Fig. 10 depicts the comparison of the 15-year annual mean PET under past and future climates. In general, the values of PET range from 2 to 8 mm/day in the past (Fig. 10a). The future PET pattern (Fig. 10b) is similar to the past pattern although the magnitude of future PET is projected to increase over the study domain. The domain-average PET under past and future climate conditions are 4.6 mm/day and 5.3 mm/day, respectively. Thus, the absolute difference in PET between past and future climates is less than 1 mm/day (Fig. 10c), and it is relatively uniform over the domain (approximately 0.6 mm/day) except for the Gulf of Mexico and certain regions in Mexico (approximately 0.3 mm/day). The domain-average increase of PET over the whole domain ranges from 0.3 to 1.2 mm/day with a mean value of 0.6 mm/day (Table 3). In addition, the relative difference in PET is generally less than 10% in the southern part of the domain. And the smallest increase in PET is projected to be approximately 7% that appears over Mexico, whereas the relative difference in PET appears to be as much as 18% in the northern part of the domain (Fig. 10d). As shown in Table 3, the domain-average relative difference in PET is projected to be 13.0%.

To examine the seasonal cycles of PET, Fig. 11 depicts the spatial patterns of the 15-year seasonal mean PET for past and future climates as well as the absolute difference in PET between past and future climates. In general, PET is projected to increase for all seasons by the end of the 21st century. In comparison, PET is remarkably higher during the warm seasons of spring (MAM) and summer (JJA), especially for the summer months. And there will be a considerable increase of PET with a magnitude up to 120 mm for the warm seasons. The statistics of the domain-average PET values are provided in Table 3. Results verify that PET is more prevalent in summer, and the summer season experiences higher changes in PET (i.e. absolute difference in PET between past and future climates). Contrarily, the relative difference in PET is larger in winter because PET is smaller during the winter months. The seasonality of PET was further examined by comparing monthly time series of the domain-average PET under past and future climate conditions. Fig. 12 shows that PET increases from January to July and then decreases towards December under both past and future climate conditions. The future PET is projected to be greater than the past PET for every month. And the largest magnitude of increase in monthly PET appears in July, which is 0.82 mm/day. These results are consistent with those shown in Fig. 11 and Table 3. The comprehensive examination of precipitation and PET variables advances our understanding of the characteristics of hydroclimatic variability over the study domain. Moreover, the high-resolution projection of future precipitation and PET changes plays a crucial role in assessing the impacts of climate change on regional and local hydrologic regimes.

3.3 Uncertainty assessment of hydrologic model parameters and predictions

To examine the hydrologic response to climate change, the Hymod was used to predict

streamflow regimes in the Guadalupe River basin, Texas. The Hymod was calibrated by using 10 years of daily hydrometeorological data (1981–1990) provided by the MOPEX project. The MCMC simulations using the DREAM algorithm were performed to estimate the posterior distributions of the Hymod parameters. The trace plots of sampled parameter estimates in three randomly selected Markov chains are depicted in Fig. S1 of the supplementary material. A total of 5,000 samples are generated in each individual chain. These plots depict that all chains eventually converge to a stationary distribution for each individual parameter. And the convergence to the posterior parameter distribution can be checked by using the multi-chain \hat{R} diagnostic. As shown in Fig. 13, the convergence of the MCMC chains to the posterior distribution is achieved since the \hat{R} values drop below the critical threshold of 1.2 for all parameters during the evolution of DREAM with a total number of 15,000 evaluations in the model calibration experiment.

The marginal posterior distributions of the Hymod parameters are shown in Fig. S2 of the supplementary material. All posterior parameter distributions are well defined, which appear approximately normal in shape. Table 4 shows the estimated mean values and standard deviations of model parameters. Results indicate that the posterior variance becomes much smaller than the prior variance for all parameters, leading to a significant reduction of parameter uncertainty. To examine the predictive uncertainty resulting from the parameter uncertainty in hydrologic modeling, Fig. 14 presents the daily streamflow predictions with the 95% uncertainty range over a period of 15 years (1981–1995) for the Guadalupe River basin. The Hymod was calibrated during the 10-year period from 1981 to 1990, and then validated during the 5-year period from 1991 to 1995 by comparing against streamflow observations. Results indicate that 87.98% of the observed streamflow data can

be captured inside the 95% uncertainty range in the calibration period, and 87.95% of streamflow observations are captured in the validation period. As the probabilistic streamflow predictions match a substantial portion of the hydrograph, the Hymod is able to characterize the rainfall-runoff behavior in the Guadalupe River basin. Nevertheless, the predictive uncertainty only arises from the parameter uncertainty in this study, thereby resulting in narrow intervals with inaccurate predictions. Moreover, the inaccuracies in prediction intervals are relatively prevalent for high flows, which may result in an underestimation of extreme events. Thus, future studies would be undertaken to take into account uncertainties in model structure and forcing data for improving the accuracy of prediction intervals.

In addition to the MOPEX data used in hydrologic simulations, the WRF-simulated precipitation and PET were also used as the forcing data to drive the Hymod for enhancing the robustness of streamflow predictions. The marginal posterior distributions of the Hymod parameters derived by using the WRF-simulated forcing data are shown in Fig. S2 of the supplementary material. Results reveal that the posterior parameter distributions obtained by using the WRF-simulated dataset are similar in shape compared with those obtained by using the MOPEX dataset, and they all appear to be normal distributions approximately. However, the mean values and standard deviations of model parameters estimated by using the WRF-simulated dataset are different from those estimated by using the MOPEX dataset (Table 4). As a result, the Hymod driven by using the WRF-simulated forcing data was also calibrated and validated by comparing against streamflow observations over a period of 15 years (1981–1995) for the Guadalupe River basin. Fig. 15 reveals that 80.69% of the observed streamflow data can be captured inside the 95%

prediction interval in the calibration period, and 80.61% of streamflow observations are captured in the validation period. In comparison, the MOPEX dataset can be used to predict daily streamflow time series more accurately, and thus it can be adopted for probabilistic streamflow projections in the Guadalupe River basin.

3.4 Probabilistic streamflow projections

When the Hymod was calibrated using the DREAM algorithm and validated against historical observations, the dynamically downscaled daily outputs of precipitation and PET through the convection-permitting WRF simulations were used to force the Hymod for daily streamflow projections in the Guadalupe River basin. Since precipitation and PET are the most important hydrologic components affecting the runoff processes, the projected changes of precipitation and PET patterns lead to streamflow variations in the river basin. In addition, the probabilistic projections of future streamflow can be produced based on the estimated posterior distributions of model parameters, enhancing the credibility of hydrologic forecasts by taking into account parameter uncertainties.

Fig. 16 presents the forecasts of daily rainfall-runoff time series with the 95% uncertainty range over a 15-year period (2085–2099). The Guadalupe River basin is the driest of the MOPEX basins described in the study of Duan et al. (2006). The number of dry days with daily rainfall amounts less than 1 mm is 4,089 (i.e. 75% of the days are dry) during the historical 15-year period (1981–1995). In comparison, the number of dry days will increase by 9 percentage points (i.e. 84% of the days are dry) for the future 15-year period (2085–2099). Moreover, the average daily PET will increase by 14% for the future period 2085–2099 compared to the historical period 1981–1995. These results indicate that the dry river basin will become drier by the end of the 21st century. Nevertheless, the

number of very heavy rainfall events with daily rainfall amounts greater than 25 mm is projected to increase by 37% for the future period 2085–2099 compared against the historical period 1981–1995. This reveals that the overall precipitation frequency will decrease while extreme precipitation events will increase significantly in frequency and intensity by the end of this century, which may result in more frequent occurrences of flash floods and drought episodes in the Guadalupe River basin. As shown in Fig. 16, the results indicate an increase in the frequency and magnitude of high flows as well as in the duration of low flows in the future. Thus, the increased risk of extreme hydrologic events demands infrastructure improvements to divert heavy downpours for relieving devastating flooding and to conserve water for dry periods.

Seasonal hydroclimatic projections with a reasonable level of accuracy play a crucial role in the management of water resources systems. Fig. 17 presents the projected changes in seasonal precipitation, PET, and streamflow for the future period 2085–2099 compared to the historical period 1981–1995. As shown in Fig. 17a, the average precipitation is projected to decrease slightly during the winter months (DJF). In comparison, there will be a significant decrease in the projected precipitation of up to 21% and 18% for the summer (JJA) and fall (SON) months, respectively. Different from the projected changes of the winter, summer, and fall precipitation, the spring season (MAM) precipitation is projected to increase by 6% relative to the spring precipitation in the historical period. In addition, the average PET is projected to increase for all seasons. In particular, the projected spring and summer PET will increase by 28% and 15%, respectively. It should be noted that the pattern and magnitude of increase in PET over the Guadalupe River basin are similar to those over the entire study domain.

The significant increase in PET and decrease in the amount of precipitation reveal a substantial increase in future drought risks during the summer season. Thus, the Guadalupe River basin would suffer from more severe and frequent summer droughts by the end of the 21st century. Fig. 17b depicts the projected changes in seasonal streamflow compared to the historical period 1981–1995. It should be noted that extreme high flows are observed in the winter of 1987 and 1992, resulting in the relatively high mean of winter streamflow, as shown in Fig. S4 of the supplementary material. The historical streamflow time series fall within the 95% uncertainty range of the projected future streamflow. This reveals that there will not be a substantial change in the future seasonal streamflow volume. Nevertheless, the number of very heavy rainfall events with daily rainfall amounts greater than 25 mm is projected to increase by 37% for the future period 2085–2099 compared to the historical period 1981–1995. This indicates that the frequency of extreme precipitation events will increase significantly by the end of this century, which may result in more frequent occurrences of flash floods. In addition, the number of dry days with daily rainfall amounts less than 1 mm will increase by 9 percentage points (i.e. 84% of the days are dry) for the future period 2085–2099 compared to the historical period 1981–1995. The increase in the frequency of extreme events including flash floods and droughts is of particular concern to regional and local government authorities who play a pivotal role in developing climate change adaptation strategies. It should be noted that the future changes in hydroclimatic regimes are projected under RCP8.5 which is the high emissions pathway (emissions are assumed to continue increasing throughout the century). The projected frequency and severity of extreme events vary under different emissions pathways.

4. Summary and conclusions

We developed high-resolution climate projections over Texas through the convection-permitting WRF simulations, and then produced probabilistic hydrologic forecasts in the Guadalupe River basin through the MCMC simulations. A total of 15 CMIP5 GCMs were chosen to generate the multi-model ensemble mean climate change signal based on their ability to simulate the 20th century climate over North America. The PGW dynamical downscaling was performed using the WRF model with horizontal grid spacing of 4 km to produce the 21st century projections of precipitation and PET under the RCP8.5 emission scenario. The PRISM precipitation and near-surface temperature data were regridded to the 4 km resolution in order to evaluate the historical WRF climate simulations over Texas. Based on the projected changes of future precipitation and potential evapotranspiration, the Hymod rainfall-runoff model was forced to predict daily and seasonal streamflow time series in the Guadalupe River basin which is the Texas Water Development Board's primary concern. The MOPEX dataset and the data collected at the USGS gauging stations were used to calibrate and validate the hydrologic model. The adaptive MCMC simulations using the DREAM algorithm were performed to explicitly characterize uncertainties in streamflow predictions. The probabilistic hydrologic forecasts are useful for decision makers to carry out risk assessments of water resources systems and to develop effective water management plans.

Dynamical downscaling results indicate that the spatial patterns of temperature and precipitation from the PRISM observations can be reproduced reasonably well by the convection-permitting WRF simulations for all seasons over the study domain. Specifically, the WRF-simulated temperature generally shows a warm bias which is less than 1.5 °C for

all seasons, and it has a relatively small bias of 0.9 °C in winter and summer. The WRF-simulated precipitation has a relatively small dry bias for the winter and fall seasons but a significant wet bias for the warm seasons of spring and summer. Climate projections reveal that the most remarkable seasonal drying and wetting caused by precipitation changes will occur during the summer months, and the most significant wetting trend with a magnitude of precipitation changes larger than 120 mm is projected to occur at the boundary between Texas and the Gulf of Mexico. In addition, the Upper Coast Climate Division is projected to experience the most significant wetting with a precipitation increase of 78.5 mm/year, whereas the most significant drying with a precipitation decrease of 46.8 mm/year is expected to occur for the North Central Texas Climate Division. The spatial pattern of projected annual mean PET is similar to the past PET pattern although the magnitude of future PET is expected to increase for all regions by the end of the 21st century. The projected increase of PET ranges from 0.3 to 1.2 mm/day with a mean value of 0.6 mm/day, which appears remarkably large during warm seasons, especially for the summer months.

The high-resolution projections of future precipitation and PET changes play a crucial role in advancing our understanding of regional characteristics of hydroclimatic changes and consequently predicting the climate-induced changes in hydrologic regimes. This study is a first attempt to examine the response of runoff to rainfall at a basin scale based on the convection-permitting climate simulations. In addition, the MCMC-based streamflow predictions and high-resolution climate projections with 4 km horizontal grid spacing are incorporated into a probabilistic framework. The proposed computational framework can be applied to examine future changes in hydroclimatic regimes for any other regions around the world, which makes a considerable contribution to the advances in hydroclimatic

projections at the convection-permitting scales (horizontal grid spacing of ≤ 4 km). The MCMC-based hydrologic simulations reveal that 87.98% of the observed streamflow data can be captured inside the 95% prediction interval in the calibration period and 87.95% of streamflow observations are captured in the validation period. Although the probabilistic hydrologic model is able to generally capture the rainfall-runoff behavior in the Guadalupe River basin, it should be noted that the inaccuracy in prediction intervals may result in an underestimation of extreme events. Furthermore, misrepresenting predictive uncertainty may amplify the potential risk of failure. For instance, if a 95% interval is used for a 100-year flood event, a mitigation strategy might be designed with only a 5% chance of failure. If the interval captures 80% of flood events, such a misrepresentation of the predictive uncertainty might quadruple the risk of failure (from 5% to 20% failure probabilities). It is thus necessary to further enhance the accuracy of prediction intervals in future studies.

Our findings reveal that the number of future dry days with daily rainfall amounts less than 1 mm will increase by 9 percentage points (i.e. 84% of the days are dry) and the average daily PET will increase by 14% in the Guadalupe River basin. The considerable decrease in the amount of precipitation and the increase in PET imply that the dry river basin will become drier with a substantial increase in future drought risks, especially for the summer season. Furthermore, the number of very heavy rainfall events with daily rainfall amounts greater than 25 mm is projected to increase by 37% while there will not be a substantial change in the future streamflow volume. These findings reveal that the extreme precipitation events will increase significantly in frequency and intensity although the overall precipitation frequency will decrease, which may result in more frequent occurrences of flash floods and drought episodes in the Guadalupe River basin. The high-

resolution projections of future hydroclimatic changes are useful to facilitate a better understanding of the potential risks in climate-induced extreme events and to help Texas policymakers and stakeholders develop sound climate change adaptation and mitigation plans.

It should be noted that uncertainties are inevitable in hydroclimatic projections, especially in the high-resolution projections of future hydroclimatic changes. These uncertainties stem from various sources, mainly including the emission scenarios of future greenhouse gas emissions, the selection of GCMs, downscaling techniques, hydrologic model structures, and the estimated values of hydrologic parameters. Efforts have been made in this study to perform uncertainty assessment of hydrologic model parameters and predictions. To further improve the robustness and reliability of hydroclimatic projections, it is desired to characterize multiple sources of uncertainty by using an ensemble approach that takes into account combinations of multiple emission scenarios, downscaling methods and hydrologic models when more powerful computing resources and computationally efficient algorithms become available. Moreover, the conceptual hydrologic model cannot well represent the physical processes and the spatial heterogeneity in the river basin, which may result in inaccurate prediction intervals, especially for peak flow predictions. The inaccuracy of peak flow predictions may cause overconfidence in risk assessment of extreme events, which poses a potential challenge to the planning of climate change adaptation measures. It is thus desired to make use of more parametrized and complex physically-based hydrologic models in future studies for improving effectiveness and reliability of hydrologic predictions. In addition, bias correction was not applied in this paper to the dynamically downscaled climate variables (e.g., precipitation and temperature)

655 due to the stationarity assumption of bias-correction approaches (i.e. bias refers to errors
656 in downscaling model outputs). Bias correction is often performed based on a strong
657 assumption that model bias is time invariant (e.g., future precipitation error is the same to
658 the past precipitation error). Thus, the use of bias correction may bring additional sources
659 of uncertainty and assumption in streamflow prediction. Nevertheless, it is necessary to
660 examine how the bias correction of dynamically downscaled climate variables affects the
661 predicted streamflow in future studies, which can provide meaningful insights into error
662 correlations and assumptions.

663
664 **Acknowledgments** This research was supported by the National Natural Science
665 Foundation of China (Grant No. 51809223) and the Hong Kong Polytechnic University
666 Start-up Grant (Grant No. 1-ZE8S). The author Y. Wang was funded by the Texas Tech
667 Research Assistant Professorship Initiative. The authors would like to express their sincere
668 gratitude to the editor and three anonymous reviewers for their constructive comments and
669 suggestions.

671 **Supplementary material**

672 Supplementary material 1

References

- Abera W, Formetta G, Brocca L, Rigon R (2017) Modeling the water budget of the Upper Blue Nile basin using the JGrass-NewAge model system and satellite data. *Hydrol Earth Syst Sci* 21:3145–3165. doi:10.5194/hess-21-3145-2017
- Allen RG, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration: Guidelines for computing crop water requirements, Irrigation and Drainage Paper No. 56, Food and Agriculture Organization of the United Nations, Rome, Italy.
- Amin MZM, Shaaban AJ, Ohara N, Kavvas ML, Chen ZQ, Kure S, Jang S (2016) Climate change assessment of water resources in Sabah and Sarawak, Malaysia, based on dynamically-downscaled GCM projections using a regional hydroclimate model. *J Hydrol Eng* 21:05015015. doi:10.1061/(ASCE)HE.1943-5584.0001242
- Andrieu C, Thoms J (2008) A tutorial on adaptive MCMC. *Stat Comput* 18:343–373. doi:10.1007/s11222-008-9110-y
- Ashfaq M, Rastogi D, Mei R, Kao SC, Gangrade S, Naz BS, Touma D (2016) High-resolution ensemble projections of near-term regional climate over the continental United States. *J Geophys Res Atmos* 121:9943–9963. doi:10.1002/2016JD025285
- Bennett JC, Wang QJ, Li M, Robertson DE, Schepen A (2016) Reliable long-range ensemble streamflow forecasts: Combining calibrated climate forecasts with a conceptual runoff model and a staged error model. *Water Resour Res* 52:8238–8259. doi:10.1002/2016WR019193
- Boulard D, Castel T, Camberlin P, Sergent AS, Bréda N, Badeau V, Rossi A, Pohl B (2016) Capability of a regional climate model to simulate climate variables requested for water balance computation: A case study over northeastern France. *Clim Dyn*

701 46:2689–2716. doi:10.1007/s00382-015-2724-9

702 Brisson E, Weverberg KV, Demuzere M, Devis A, Saeed S, Stengel M, van Lipzig NPM
703 (2016) How well can a convection-permitting climate model reproduce decadal
704 statistics of precipitation, temperature and cloud characteristics? *Clim Dyn* 47:3043–
705 3061. doi:10.1007/s00382-016-3012-z

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

709 Carvalho KS, Wang S (2019) Characterizing the Indian Ocean sea level changes and
710 potential coastal flooding impacts under global warming. *J Hydrol* 569: 373–386.
711 doi.org/10.1016/j.jhydrol.2018.11.072

712 Chan SC, Kendon EJ, Roberts N, Blenkinsop S, Fowler HJ (2018) Large-scale predictors
713 for extreme hourly precipitation events in convection-permitting climate simulations.
714 *J Climate* 31:2115–2131. doi:10.1175/JCLI-D-17-0404.1

715 Daly C, Halbleib M, Smith JI, Gibson WP, Doggett MK, Taylor GH, Curtis J, Pasteris PP
716 (2008) Physiographically sensitive mapping of climatological temperature and
717 precipitation across the conterminous United States. *Int J Climat* 28:2031–2064.
718 doi:10.1002/joc.1688

719 DeChant CM, Moradkhani H (2015) On the assessment of reliability in probabilistic
720 hydrometeorological event forecasting. *Water Resour Res* 51:3867–3883.
721 doi:10.1002/2014WR016617

722 Duan Q et al (2006) Model Parameter Estimation Experiment (MOPEX): An overview of
723 science strategy and major results from the second and third workshops. *J Hydrol*

724 320:3–17. doi:10.1016/j.jhydrol.2005.07.031

725 Erler AR, Peltier WR (2017) Projected hydroclimatic changes in two major river basins at
 726 the Canadian west coast based on high-resolution regional climate simulations. *J*
 727 *Climate* 30:8081–8105. doi:10.1175/JCLI-D-16-0870.1

728 Erler AR, Peltier WR, D’Orgeville M (2015) Dynamically downscaled high-resolution
 729 hydroclimate projections for western Canada. *J Climate* 28:423–450. doi:
 730 10.1175/JCLI-D-14-00174.1

731 Farmer WH, Vogel RM (2016) On the deterministic and stochastic use of hydrologic
 732 models. *Water Resour Res* 52:5619–5633. doi:10.1002/2016WR019129

733 Fosser G, Khodayar S, Berg P (2015) Benefit of convection permitting climate model
 734 simulations in the representation of convective precipitation. *Clim Dyn* 44:45–60.
 735 doi:10.1007/s00382-014-2242-1

736 Gelman A, Rubin DB (1992) Inference from iterative simulation using multiple sequences.
 737 *Stat Sci* 7:457–472. doi:10.1214/ss/1177011136

738 Giorgi F et al (2014) Changes in extremes and hydroclimatic regimes in the CREMA
 739 ensemble projections. *Clim Change* 125:39–51. doi:10.1007/s10584-014-1117-0

740 Griffiths J, Zeitler J, Sedlar L, Bjornson D, Bjornson B (1990) A Decade of Texas Weather
 741 (1980–1989) Rep, 88 pp, Off of the State Climatol, Dep of Meteorol, Coll of Geosci,
 742 Texas A&M Univ, College Station, Tex

743 Hagemann S, Chen C, Haerter JO (2011) Impact of a statistical bias correction on the
 744 projected hydrological changes obtained from three gcms and two hydrology models.
 745 *J Hydrometeorol* 12:556–578. doi:10.1175/2011JHM1336.1

746 Herman JD, Reed PM, Wagener T. (2013) Time-varying sensitivity analysis clarifies the

747 effects of watershed model formulation on model behavior. *Water Resour Res*
 748 49:1400–1414. doi:10.1002/wrcr.20124

749 Hong SY, Pan HL (1996) Nonlocal boundary layer vertical diffusion in a medium-range
 750 forecast model. *Mon Wea Rev* 124:2322–2339. doi:10.1175/1520-
 751 0493(1996)124<2322:NBLVDI>2.0.CO;2

752 Iacono MJ, Delamere JS, Mlawer EJ, Shephard MW, Clough SA, Collins WD (2008)
 753 Radiative forcing by long-lived greenhouse gases: Calculations with the AER
 754 radiative transfer models. *J Geophys Res Atmos* 113:D13103.
 755 doi:10.1029/2008JD009944

756 Im ES, Choi YW, Ahn JB (2017) Robust intensification of hydroclimatic intensity over
 757 East Asia from multi-model ensemble regional projections. *Theor Appl Climatol*
 758 129:1241–1254. doi:10.1007/s00704-016-1846-2

759 Jimenez PA, Dudhia J, Gonzalez-Rouco JF, Navarro J, Montavez JP, Garcia-Bustamante E
 760 (2012) A revised scheme for the WRF surface layer formulation. *Mon Wea Rev*
 761 140:898–918. doi:10.1175/MWR-D-11-00056.1

762 Kendon EJ et al (2017) Do convection-permitting regional climate models improve
 763 projections of future precipitation change? *B Am Meteorol Soc* 79–93.
 764 doi:10.1175/BAMS-D-15-0004.1

765 Kendon EJ, Roberts NM, Senior CA, Roberts MJ (2012) Realism of rainfall in a very high-
 766 resolution regional climate model. *J Climate* 25:5791–5806. doi:10.1175/JCLI-D-11-
 767 00562.1

768 Krysanova V et al (2017) Intercomparison of regional-scale hydrological models and
 769 climate change impacts projected for 12 large river basins worldwide — A synthesis.

770 Environ Res Lett 12:105002. doi:10.1088/1748-9326/aa8359
 771 Leong DNS, Donner SD (2015) Climate change impacts on streamflow availability for the
 772 Athabasca Oil Sands. Clim Change 133:651–663. doi:10.1007/s10584-015-1479-y
 773 Leutwyler D, Lüthi D, Ban N, Fuhrer O, Schär C (2017) Evaluation of the convection-
 774 resolving climate modeling approach on continental scales. J Geophys Res Atmos
 775 122:5237–5258. doi:10.1002/2016JD026013
 776 Li Z, Huang GH, Wang X, Han J, Fan YR (2016) Impacts of future climate change on river
 777 discharge based on hydrological inference: A case study of the Grand River Watershed
 778 in Ontario, Canada. Sci Total Environ 548–549:198–210.
 779 doi:10.1016/j.scitotenv.2016.01.002
 780 Liu C, Ikeda K, Thompson G, Rasmussen R, Dudhia J (2011) High-resolution simulations
 781 of wintertime precipitation in the Colorado Headwaters region: Sensitivity to physics
 782 parameterizations. Mon Wea Rev 139:3533–3553. doi:10.1175/MWR-D-11-00009.1
 783 Liu C et al (2017) Continental-scale convection-permitting modeling of the current and
 784 future climate of North America. Clim Dyn 49:71–95. doi:10.1007/s00382-016-3327-
 785 9
 786 Lu D, Ricciuto D, Walker A, Safta C, Munger W (2017) Bayesian calibration of terrestrial
 787 ecosystem models: a study of advanced Markov chain Monte Carlo methods.
 788 Biogeosciences 14: 4295–4314. doi: 10.5194/bg-14-4295-2017
 789 Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equation of
 790 state calculations by fast computing machines. J Chem Phys 21:1087–1092.
 791 doi:10.1063/1.1699114
 792 Moore RJ (1985) The probability-distributed principle and runoff production at point and

793 basin scales. *Hydrol Sci J* 30:273–297. doi:10.1080/02626668509490989

794 Music B, Frigon A, Lofgren B, Turcotte R, Cyr JF (2015) Present and future Laurentian
 795 Great Lakes hydroclimatic conditions as simulated by regional climate models with
 796 an emphasis on Lake Michigan-Huron. *Clim Change* 130:603–618.
 797 doi:10.1007/s10584-015-1348-8

798 Niu GY et al (2011) The community Noah land surface model with multiparameterization
 799 options (Noah-MP): 1. Model description and evaluation with local-scale
 800 measurements. *J Geophys Res Atmos* 116:D12109. doi:10.1029/2010JD015139

801 Pechlivanidis IG et al (2017) Analysis of hydrological extremes at different hydro-climatic
 802 regimes under present and future conditions. *Clim Change* 141:467–481.
 803 doi:10.1007/s10584-016-1723-0

804 Pfahl S, O’Gorman PA, Fischer EM (2017) Understanding the regional pattern of projected
 805 future changes in extreme precipitation. *Nat Clim Change* 7:423–427. doi:
 806 10.1038/nclimate3287

807 Piao S et al (2010) The impacts of climate change on water resources and agriculture in
 808 China. *Nature* 467:43–51. doi:10.1038/nature09364

809 Prein AF, Rasmussen RM, Ikeda K, Liu C, Clark MP, Holland GJ (2017) The future
 810 intensification of hourly precipitation extremes. *Nat Clim Change* 7:48–52.
 811 doi:10.1038/nclimate3168

812 Prein AF et al (2015) A review on regional convection-permitting climate modeling:
 813 Demonstrations, prospects, and challenges. *Rev Geophys* 53:323–361.
 814 doi:10.1002/2014RG000475

815 Rajbhandari R, Shrestha AB, Kulkarni A, Patwardhan SK, Bajracharya SR (2015)

816 Projected changes in climate over the Indus river basin using a high resolution
817 regional climate model (PRECIS). *Clim Dynam* 44:339–357. doi:10.1007/s00382-
818 014-2183-8

819 Ramos MH, van Andel SJ, Pappenberger F (2013) Do probabilistic forecasts lead to better
820 decisions? *Hydrol Earth Syst Sci* 17:2219–2232. doi:10.5194/hess-17-2219-2013

821 Rasmussen R, Ikeda K, Liu C, Gochis D, Clark D (2014) Climate change impacts on the
822 water balance of the Colorado Headwaters: High-resolution regional climate model
823 simulations. *J Hydrometeorol* 15:1091–1116. doi:10.1175/JHM-D-13-0118.1

824 Razavi S, Gupta HV (2016) A new framework for comprehensive, robust, and efficient
825 global sensitivity analysis: 2. Application. *Water Resour Res* 52:440–455.
826 doi:10.1002/2015WR017559

827 Roy T, Serrat-Capdevila A, Gupta H, Valdes J (2016) A platform for probabilistic
828 multimodel and multiproduct streamflow forecasting. *Water Resour Res* 53:376–399.
829 doi:10.1002/2016WR019752

830 Roy T, Gupta HV, Serrat-Capdevila A, Valdes JB (2017) Using satellite-based
831 evapotranspiration estimates to improve the structure of a simple conceptual rainfall–
832 runoff model. *Hydrol Earth Syst Sci* 21:879–896. doi:10.5194/hess-21-879-2017

833 Sadegh M, Ragno E, AghaKouchak A (2017) Multivariate Copula Analysis Toolbox
834 (MvCAT): Describing dependence and underlying uncertainty using a Bayesian
835 framework. *Water Resour Res* 53:5166–5183. doi:10.1002/2016WR020242

836 Sadegh M, Vrugt JA, Xu C, Volpi E (2015) The stationarity paradigm revisited: Hypothesis
837 testing using diagnostics, summary metrics, and DREAM_(ABC). *Water Resour Res*
838 51:9207–9231. doi:10.1002/2014WR016805

839 Sadegh M, Vrugt JA (2013) Bridging the gap between GLUE and formal statistical
840 approaches: Approximate Bayesian computation. *Hydrol Earth Syst Sci* 17:4831–
841 4850. doi:10.5194/hess-17-4831-2013

842 Saha S et al (2010) The NCEP climate forecast system reanalysis. *Bull Amer Meteorol Soc*
843 91:1015–1057. doi:10.1175/2010BAMS3001.1

844 Silverman NL, Maneta MP, Chen SH, Harper JT (2013) Dynamically downscaled winter
845 precipitation over complex terrain of the Central Rockies of western Montana, USA.
846 *Water Resour Res* 49:458–470. doi:10.1029/2012WR012874

847 Sorooshian S, Dracup JA (1980) Stochastic parameter estimation procedures for hydrologic
848 rainfall-runoff models: Correlated and heteroscedastic error cases. *Water Resour Res*
849 16:430–442. doi:10.1029/WR016i002p00430

850 Srivastava PK, Han D, Islam T, Petropoulos GP, Gupta M, Dai Q (2016) Seasonal
851 evaluation of evapotranspiration fluxes from MODIS satellite and mesoscale model
852 downscaled global reanalysis datasets. *Theor Appl Climatol* 124:461–473.
853 doi:10.1007/s00704-015-1430-1

854 Svensson C et al (2015) Long-range forecasts of UK winter hydrology. *Environ Res Lett*
855 10:064006. doi:10.1088/1748-9326/10/6/064006

856 Thompson G, Field PR, Rasmussen RM, Hall WD (2008) Explicit forecasts of winter
857 precipitation using an improved bulk microphysics scheme. Part II: Implementation
858 of a new snow parameterization. *Mon Wea Rev* 136:5095–5115.
859 doi:10.1175/2008MWR2387.1

860 Vrugt JA (2016) Markov chain Monte Carlo simulation using the DREAM software
861 package: Theory, concepts, and MATLAB implementation. *Environ Modell Softw*

862 75:273–316. doi:10.1016/j.envsoft.2015.08.013

863 Vrugt JA, ter Braak CJF, Diks CGH, Robinson BA, Hyman JM, Higdon D (2009)

864 Accelerating Markov chain Monte Carlo simulation by differential evolution with

865 self-adaptive randomized subspace sampling. *Int J Nonlinear Sci Numer Simul*

866 10:273–290. doi:10.1515/IJNSNS.2009.10.3.273

867 Vrugt JA, ter Braak CJF, Clark MP, Hyman JM, Robinson BA (2008) Treatment of input

868 uncertainty in hydrologic modeling: doing hydrology backward with Markov chain

869 Monte Carlo simulation. *Water Resour Res* 44:W00B09.

870 doi:10.1029/2007WR006720

871 Wang S, Ancell BC, Huang GH, Baetz BW (2018a) Improving robustness of hydrologic

872 ensemble predictions through probabilistic pre- and post-processing in sequential data

873 assimilation. *Water Resour Res* 54:2129–2151. doi:10.1002/2018WR022546

874 Wang S, Huang GH, Baetz BW, Huang W (2015) A polynomial chaos ensemble hydrologic

875 prediction system for efficient parameter inference and robust uncertainty assessment.

876 *J Hydrol* 530:716–733. doi:10.1016/j.jhydrol.2015.10.021

877 Wang S, Huang GH, Lin QG, Li Z, Zhang H, Fan YR (2014) Comparison of interpolation

878 methods for estimating spatial distribution of precipitation in Ontario, Canada. *Int J*

879 *Climatol* 34:3745–3751. doi:10.1002/joc.3941

880 Wang Y, Geerts B, Liu C (2018b) A 30-year convection-permitting regional climate

881 simulation over the interior western United States. Part I: Validation. *Int J Climatol* 1–

882 21. doi:10.1002/joc.5527

883 Westra S et al (2014) Future changes to the intensity and frequency of short-duration

884 extreme rainfall. *Rev Geophys* 52:522–555. doi:10.1002/2014RG000464

885 Yang ZL et al (2011) The community Noah land surface model with multiparameterization
886 options (Noah–MP): 2. Evaluation over global river basins. J Geophys Res Atmos
887 116:D1211. doi:10.1029/2010JD015140

888 Yuan X, Wood EF, Roundy JK, Pan M (2013) CFSv2-based seasonal hydroclimatic
889 forecasts over the conterminous United States. J Climate 26:4828–4847.
890 doi:10.1175/JCLI-D-12-00683.