Journal of Hydrology 530 (2015) 568–579

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Bias correction methods for regional climate model simulations considering the distributional parametric uncertainty underlying the observations

Kue Bum Kim [a], Hyun-Han Kwon [b,*], Dawei Han [a]

[a] Water and Environmental Management Research Centre, Department of Civil Engineering, University of Bristol, Bristol, UK
[b] Department of Civil Engineering, Chonbuk National University, Jeonju-si, Jeollabuk-do, South Korea

ARTICLE INFO

Article history:
Received 6 May 2015
Received in revised form 27 September 2015
Accepted 4 October 2015
Available online 13 October 2015
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Axel Bronstert, Associate Editor

Keywords:
Climate change
Internal climate variability
Uncertainty
Bayesian
Likelihood

SUMMARY

In this paper, we present a comparative study of bias correction methods for regional climate model simulations considering the distributional parametric uncertainty underlying the observations/models. In traditional bias correction schemes, the statistics of the simulated model outputs are adjusted to those of the observation data. However, the model output and the observation data are only one case (i.e., realization) out of many possibilities, rather than being sampled from the entire population of a certain distribution due to internal climate variability. This issue has not been considered in the bias correction schemes of the existing climate change studies. Here, three approaches are employed to explore this issue, with the intention of providing a practical tool for bias correction of daily rainfall for use in hydrologic models ((1) conventional method, (2) non-informative Bayesian method, and (3) informative Bayesian method using a Weather Generator (WG) data). The results show some plausible uncertainty ranges of precipitation after correcting for the bias of RCM precipitation. The informative Bayesian approach shows a narrower uncertainty range by approximately 25–45% than the non-informative Bayesian method after bias correction for the baseline period. This indicates that the prior distribution derived from WG may assist in reducing the uncertainty associated with parameters. The implications of our results are of great importance in hydrological impact assessments of climate change because they are related to actions for mitigation and adaptation to climate change. Since this is a proof of concept study that mainly illustrates the logic of the analysis for uncertainty-based bias correction, future research exploring the impacts of uncertainty on climate impact assessments and how to utilize uncertainty while planning mitigation and adaptation strategies is still needed.

1. Introduction

Quantifying uncertainty in estimates of future climate change for use in climate impact simulations is a necessary step for detection, attribution, and mitigation and adaptation strategies (Deser et al., 2012b). Hence, demand for more quantitative analyses of future climate change is increasing (Collins et al., 2012). These uncertainties are due to scenario uncertainty, model uncertainty, and internal climate variability (Hawkins and Sutton, 2009; Tebaldi and Knutti, 2007). Internal climate variability results from natural fluctuations. Some recent studies have drawn attention to the contributions of natural variability to climate change (Deser et al., 2012a; Hawkins and Sutton, 2009; Kendon et al., 2008; Knutti and Sedláček, 2013; Tebaldi et al., 2011; Zunz et al., 2013). The uncertainties of climate projections due to natural variability are considered to be irreducible (Fischer et al., 2013; Smith et al., 2007).

Although bias correction is controversial (Ehret et al., 2012; Muerth et al., 2013), bias correction methods have been successfully and widely applied in climate change studies (Dosio and Paruolo, 2011; Piani and Haerter, 2012; Rojas et al., 2011). In climate change studies, 30 years of observation data are generally used as a reference, as defined by the World Meteorological Organization (WMO). However, each observation is only one case out of many possibilities, rather than being sampled from the entire population of a certain distribution, due to distributional parametric uncertainties that have not been considered in the existing bias correction schemes used in the past climate change studies. This has particularly important implications for uncertainties associated with the parameters of the probability density function (PDF) used for correcting bias in climate model outputs, since in

* Corresponding author.
E-mail address: hkwon@jbnu.ac.kr (H.-H. Kwon).

http://dx.doi.org/10.1016/j.jhydrol.2015.10.015
0022-1694/

traditional bias correction schemes the statistics of the simulated model outputs are adjusted to those of the observation data, which is only one realization of many possibilities. The uncertainty in parameter estimates is directly related to the sample size and the quality of available information. In other words, distributional parametric uncertainty exists when limited amounts of hydrologic data are used to estimate the parameters of PDF. Both the natural randomness of hydrologic data and the distributional parametric uncertainty in bias correction may contribute to the uncertainty in future climate change projections. However, most studies of uncertainties in hydrologic studies have focused on measurement uncertainties mainly resulting from the spatial integration of measurements across different sites. Here we do not consider measurement error, but concentrate on distributional parametric uncertainty. In other words, measurement error and its impact on bias correction for future projections are not included in this study, but could be added in future analyses if needed.

In this paper, we explore the following questions:

(1) Can uncertainties in observation and regional climate model output with respect to distributional parametric uncertainty be modeled simultaneously and consistently?
(2) Climate change studies use Weather Generators (WG) informed by global climate model (GCM) or regional climate model (RCM) integration (forecast or climate change) for downscaling.
   A. Is it better to use precipitation sequences simulated from the WG as a prior distribution instead of a non-informative prior? Does the WG really add value?
   B. Does a combination of the WG and Bayesian model better inform uncertainty?
(3) Can a Bayesian-based bias correction model offer useful scenarios of daily precipitation for climate change studies?

In this study, three approaches are employed to explore these questions, with the intention of providing a practical tool for bias correction of daily rainfall for use in hydrologic models. The approaches are based on the quantile mapping method: (1) the conventional method and two new approaches, (2) the non-informative Bayesian (NIB) method and (3) the informative Bayesian (IB) method. In this study, we aim to quantify distributional parametric uncertainty and show some plausible range of precipitation after correcting for the bias of RCM precipitation. The proposed methodology is applied to three catchments located in the southwest of England. One emission scenario (A1B) and one-member (Q0) among 11 members of the HadRM3 model output driven by the GCM HadCM3 (Murphy et al., 2009) are used for the analysis because the purpose of this study is not to prove the two proposed bias correction methods (NIB and IB methods) or to determine which method is the best among the three approaches. The aim is mainly to introduce a new concept, the logic underlying uncertainty based bias correction and how this concept can be extended to conventional approaches. Hence, we believe three cases are sufficient.

2. Data and Weather Generator

Three catchments in the southwest of England are used in this study. The catchments have varying rainfall regimes (i.e., low, medium, and high rainfall) and are representative of the range of rainfall distributions in this region. The Avon River at Melksham (665.6 km^2) has low rainfall (797 mm/year), the Exe River at Thorverton (606 km^2) has medium rainfall (1260 mm/year), and the Tamar River at Gunnislake (916.9 km^2) has high rainfall (1751 mm/year). Daily time series of the observed precipitation data are obtained from the UK Met Office. For the model output, we have obtained the HadRM3 Perturbed Physics Experiment Dataset (HadRM3-PPE-UK, resolution 25×25 km), which provides time series data from 1950 to 2100. Among the data, only one ensemble member is used in the analysis. The UKCP09 Weather Generator (Jones et al., 2009) (WG) data is used for the prior distribution (Gamma distribution). The WG generates statistically plausible time series of nine climate variables (i.e., precipitation, temperature, vapor pressure, wind, sunshine, potential evapotranspiration, diffuse radiation and direct radiation) at a 5 km resolution. According to the official UK government guidelines on climate change, the UKCP09 WG is trained using the 5 km daily-observed baseline for 1961–1995. This means that the WG model baseline is fitted to the 1961–1995 historical observations. The WG data are the officially approved data for climate change studies in the UK. Precipitation is generated using the Neyman-Scott Rectangular Processes (NSRP) model (Cowpertwait et al., 1996; Jones et al., 2009), and other variables are then simulated given the simulated precipitation. The NSRP model is a clustered point process model comprised of clusters and rectangular impulse models for rainfall occurrence and amount (Onof et al., 2000). The NSRP model describes storm origins, durations and the intensity of each rain cell as a set of random variables. The more detailed structure of the NSRP model is described as follows: First, storm origins are represented by a Poisson distribution with parameter $\boldsymbol{\alpha}$ relating to the arrival times of the storms; Second, the storm origin randomly generates numbers $\boldsymbol{\nu}$ of rain cells departing from the storm origin at time intervals that are simulated by an exponential distribution with parameter $\boldsymbol{\beta}$; Third, the durations of the rain cells are generated by an exponential distribution with parameter $\boldsymbol{\gamma}$; Fourth, the intensities of the rain cells are again simulated by exponential distributions with parameter $\boldsymbol{\delta}$; Finally, rainfall intensity is calculated by summing the intensity of each rain cell. A schematic representation of the NSRP model is shown in Fig. 1.

The parameters of the NSRP model are estimated separately for each month to better characterize intra-annual rainfall variability. Expected values of rainfall statistics, such as the mean rainfall amount, the proportion of dry days, the variance and skewness of daily rainfall amounts, and the lag-1 autocorrelation (Cowpertwait et al., 2002), are analytically derived with respect to the five parameters of the NSRP model. These parameters are as follows: (1) the average waiting time between subsequent storm origins; (2) the average waiting time of the rain cells after the storm origin; (3) the average cell duration; (4) the average number of cells per storm; and (5) the average cell intensity. These expected parameter values are then used to optimize a set of parameters by minimizing an objective function using an optimization algorithm. The required rainfall statistics for UKCP09 are estimated based on a gridded rainfall dataset at 5×5 km resolution compiled by Perry and Hollis (2005a, b) that covers the UK for the period 1961–1990.

3. Methodology

3.1. Bias correction methods

Numerous studies have assessed the impacts of climate change on water resources using climate variables from global climate models (GCMs) and water resources models (Fung et al., 2011). However, because of the relatively low spatial resolution (100–250 km) of GCMs, regional climate models (RCMs) are widely used for regional impact studies at catchment scales (25–50 km) climate variables (Fowler et al., 2007; Qin et al., 2007). Although RCMs are able to simulate local climate at finer resolutions, outputs from RCMs cannot be used as direct input data for hydrological models