Advances in Water Resources 94 (2016) 278–290

Contents lists available at ScienceDirect

Advances in Water Resources

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

Hybrid approach in statistical bias correction of projected precipitation for the frequency analysis of extreme events

Myoung-Jin Um[a], Hanbeen Kim[a], Jun-Haeng Heo[a,*]

[a] *School of Civil and Environmental Engineering, Yonsei University, Seoul 120-749, Korea*

ARTICLE INFO

Article history:
Received 24 September 2015
Revised 30 May 2016
Accepted 30 May 2016
Available online 1 June 2016

Keywords:
Bias correction
Extreme value analysis
Hybrid process
General circulation model

ABSTRACT

A general circulation model (GCM) can be applied to project future climate factors, such as precipitation and atmospheric temperature, to study hydrological and environmental climate change. Although many improvements in GCMs have been proposed recently, projected climate data are still required to be corrected for the biases in generating data before applying the model to practical applications. In this study, a new hybrid process was proposed, and its ability to perform bias correction for the prediction of annual precipitation and annual daily maxima, was tested. The hybrid process in this study was based on quantile mapping with the gamma and generalized extreme value (GEV) distributions and a spline technique to correct the bias of projected daily precipitation. The observed and projected daily precipitation values from the selected stations were analyzed using three bias correction methods, namely, linear scaling, quantile mapping, and hybrid methods. The performances of these methods were analyzed to find the optimal method for prediction of annual precipitation and annual daily maxima. The linear scaling method yielded the best results for estimating the annual average precipitation, while the hybrid method was optimal for predicting the variation in annual precipitation. The hybrid method described the statistical characteristics of the annual maximum series (AMS) similarly to the observed data. In addition, this method demonstrated the lowest root mean squared error (RMSE) and the highest coefficient of determination (R^2) for predicting the quantiles of the AMS for the extreme value analysis of precipitation.

1. Introduction

Understanding future changes in extreme precipitation is necessary for preparing citizens for storms and floods. Evidence for climate change has been observed in hydrological data and climate models such as general circulation models (GCMs) that have been developed to predict climate variables in the future (IPCC, 2013). The usage of GCMs to predict future climatic factors, such as precipitation and atmosphere temperature, is rapidly increasing in the study of hydrological and environmental climate change. Although many recent improvements in GCMs have been proposed, climate data are still required to correct for biases before practical application of the models (Bergstrom et al., 2001; Graham et al., 2007a; Graham et al., 2007b; Themebl et al., 2010; Teutschbein and Seibert, 2012).

Dynamic and statistical downscaling methods have been developed to address issues of scale (Chen et al., 2011). Regional climate models (RCMs) based on dynamic downscaling methods are used to generate more detailed information than are obtainable with GCMs (Rummukainen, 2010), but despite having a relatively high resolution, the resolution of RCMs is still coarse and retains uncertainty that causes biases (Marinucci and Giorgi, 1992; Chen et al., 2011). Moreover, RCM precipitation simulations involve bias problems similar to GCM simulations, making it difficult to use them directly. Some methods for correcting biases exist, which compare the simulation data of RCMs to observed data.

Bias correction for the data generated by RCM precipitation simulations was used for various regional studies (Hay et al., 2002; Fowler and Ekstrom, 2009; Quintana Segui et al., 2010; Teng et al., 2015). However, the higher computational cost, depending on the resolution of the climate model, is a major issue for the application of RCMs (Solman and Nunez, 1999; IPCC, 2013). Alternatively, statistical downscaling is an effective technique for reducing the computational cost associated with bias correction of RCM/GCM simulation data (Ahmed et al., 2013). A number of statistical downscaling methods have been researched to determine the appropriate bias correction procedure (Fowler et al., 2007; Burger et al., 2012, 2013; Hanel et al., 2013; Maraun et al., 2013; Teutschbein and Seibert, 2013).

For climate models, the statistical downscaling method still requires improvement to properly analyze the projected data

* Corresponding author.
E-mail address: jhheo@yonsei.ac.kr (J.-H. Heo).

http://dx.doi.org/10.1016/j.advwatres.2016.05.021
0309-1708/

(Sunyer et al., 2015). Leander and Buishand (2007) proposed a bias correction of the mean that adjusts the mean of the RCM simulation outputs. Furthermore, bias correction of the mean and the variance that employs exponential transformation was applied to correct the precipitation output of the RCM (Leander et al., 2008; Hanel et al., 2013). The quantile mapping method uses the distribution function to correct RCM outputs. It can preserve the frequency of precipitation (dry/wet days) of simulated data while individually considering the mean and extreme precipitation (Sunyer et al., 2015). Additionally, the proper distribution that represents the data characteristics can be applied to estimate the climate data. For example, the gamma distribution and the Gaussian distributions are appropriate for describing precipitation data (Katz, 1999; Piani et al., 2010) and temperature data (Schoenau and Kehrig, 1990), respectively. Maurer and Pierce (2014) and Piani et al. (2010) used the gamma distribution for quantile mapping of daily precipitation data.

The aim of this study was to propose a bias correction process for estimating the appropriate extreme daily precipitation in a global climate model. A hybrid process based on quantile mapping with gamma and generalized extreme value (GEV) distributions was proposed to correct the bias of the distribution. The observed and projected daily precipitation data from eight stations were analyzed using three bias correction methods: linear scaling, quantile mapping, and hybrid methods. Subsequently, the performance of the methods was analyzed to determine the optimal method for estimating annual precipitation and the annual daily maximum series.

2. Methods and materials

2.1. Bias correction

A number of statistical bias correction methods are usually applied to correct the systematic biases in climate change models because these methods do not require computing power as expensive as that of dynamical methods such as using regional climate models with high-resolution grids. The former methods require a transformation process to modify the climate model's results. Statistical bias corrections have several limitations such as the assumption of similar climate conditions for historical and projected future precipitation. Consequently, many studies have been conducted to correct these problems.

In this study, two popular previously tested methods, linear scaling and quantile mapping with gamma distribution, were considered. Additionally, a new hybrid method focusing on extreme events for hydrological impacts in the future was examined to correct for biases in the daily projected precipitation. These methods are referenced and summarized hereafter (Ines and Hansen, 2006; Li et al., 2010; Teutschbein and Seibert, 2012; Watanabe et al., 2012).

2.1.1. Linear scaling method

Lenderink et al. (2007) used the linear scaling method for predicting monthly precipitation and temperature using the differences between the observed time series and the historically projected time series. This method uses the ratio of the means of observed and projected data for precipitation and the difference between the averages of observed and projected data for temperature.

In this study, linear scaling models were applied to estimate the corrected projected daily precipitation as shown in Eqs. (1) and (2),

$$P_{hist}^{bc} = P_{hist}\,\mu(P_{obs})/\mu(P_{hist}) \tag{1}$$

$$P_{scen}^{bc} = P_{scen}\,\mu(P_{obs})/\mu(P_{scen}) \tag{2}$$

where P is the precipitation, μ is the mean, bc is the bias-corrected data, obs is the observed data, $hist$ is the projected data during the historical period, and $scen$ is the projected data during the scenario period (future period).

2.1.2. Quantile mapping

The quantile mapping method uses the statistical distribution of climate data to correct the bias of the projected climate values. This method is also known as probability mapping, distribution mapping, and the cumulative distribution function (CDF) matching method (Panofsky and Brier, 1968; Boe et al., 2007; Block et al., 2009; Sennikovs and Bethers, 2009; Piani et al., 2010; Johnson and Sharma, 2011; Teutschbein and Seibert, 2012).

For predicting precipitation, the quantile mapping method is expressed in Eq. (3),

$$P_{hist}^{bc} = F_{obs}^{-1}(F_{hist}(P_{hist})) \tag{3}$$

where F and F^{-1} are the cumulative distribution function (CDF) and inverse CDF, respectively, for the historical period (HIST). Eq. (3) can be applied to the scenario (SCEN) data with P_{scen} instead of P_{hist}. For a precipitation time series, the gamma distribution is frequently used (Eq. (4)) (Thom, 1958),

$$F_{gamma}(x) = \frac{1}{(\beta^{\gamma}\Gamma(\gamma))}x^{(\gamma-1)}exp\left(-\frac{x}{\beta}\right) \tag{4}$$

where F_{gamma} is the CDF of the gamma distribution, $\Gamma(\gamma)$ is the gamma function, and β and γ are the parameters of the distribution. The parameters of the gamma distribution in this study were estimated with L-moments (Hosking and Wallis, 1997).

2.1.3. Hybrid quantile mapping

In environmental science, hydrology, and climatology, the block maxima series are normally applied to estimate the quantile of precipitation and flooding (e.g., 100-year flood), and the block maxima approach frequently uses the generalized extreme value (GEV) distribution for annual maximum series (AMS) of climate data. The GEV distribution for the AMS of climate data is given by Eq. (5):

$$F_{GEV}(x) = exp\left\{-[1+\gamma(x-\alpha)/\beta]^{-1/\gamma}\right\},\ \ 1+\gamma(x-\alpha)/\beta > 0 \tag{5}$$

where F_{GEV} is the CDF of the GEV distribution and α, β, and γ are the parameters of the distribution. The GEV distribution was introduced in this study to correct for the bias of projected precipitation. This hybrid method is based on the classical quantile mapping method in Section 2.1.2. However, this method also applies an additional probability distribution, which is the GEV distribution, to the AMS data to improve the model performance for the extreme events. The gamma distribution was used to estimate the bias of the data without the AMS data in case 1 in Fig. 1(a); subsequently, the GEV distribution was applied to analyze the bias of the AMS data for each year (e.g., N−1, N, N+1 year) in case 2 in Fig. 1(a), as given by Eq. (6). This method can be applied to the scenario (SCEN) data with P_{scen} instead of P_{hist} in Eq. (6).

$$P_{hist}^{bc} = \begin{matrix} F_{obs,\ GEV}^{-1}\left(F_{hist,\ GEV}(P_{hist})\right)\ for\ AMS \\ F_{obs,\ gamma}^{-1}\left(F_{hist,\ gamma}(P_{hist})\right)\ for\ the\ data\ without\ AMS \end{matrix} \tag{6}$$

However, the gamma distribution in this hybrid approach might have been estimated higher values than those by the GEV distribution with AMS in an individual year, i.e., Fig. 1(b–2), because we used two different distributions for AMS data and the non-AMS data. In this case, the spline interpolation method (Wolberg and Alfy, 1999) was applied to maintain the order of the data for the corresponding year in Fig. 1(b). The parameters of the GEV distribution in this study were estimated with L-moments (Hosking and Wallis, 1997).