A novel spatial downscaling approach for climate change assessment in regions with sparse ground data networks

Yong-Tak Kim[1], Hyun-Han Kwon[1]*, Carlos Lima[2] and Ashish Sharma[3]

10/28/2021

*Corresponding author: hkwon@sejong.ac.kr (https://orcid.org/0000-0003-4465-2708)

To be submitted to geophysical research letters

[1] Dept. of Civil and Env. Eng., Sejong University, Seoul, South Korea
[2] Dept. of Civil and Env. Eng., University of Brasília, Brasília, Brazil
[3] Dept. of Civil and Env. Eng., UNSW, Sydney, Australia

Key Points

- All parameters for the spatial downscaling and bias-correction can be simultaneously estimated and interpolated at the desired points
- The parameter interpolation is more effective than that of the precipitation in the context of spatial downscaling and bias-correction
- The BK-SDQDM effectively can reproduce spatial dependency in the interpolated parameters associated with bias-correction

Abstract

This study proposes a novel approach that expands the existing QDM (quantile delta mapping) to address spatial bias, using Kriging within a Bayesian framework to assess the impact of using a point reference field. Our focus here is to spatially downscale daily rainfall sequences simulated by regional climate models (RCMs), coupled to the proposed QDM-spatial bias-correction, in which the distribution parameters are first interpolated onto a fine grid (rather than the observed daily rainfall). The proposed model is validated through a cross-validatory (CV) evaluation using rainfall data from a set of weather stations in South Korea and climate change scenarios simulated by three alternate RCMs. The results demonstrate the efficacy of the proposed model to simulate the bias-corrected daily rainfall sequences over large regions at fine resolutions. A discussion of the potential use of the proposed approach in the field of hydrometeorology is also offered.

Keywords: quantile delta mapping, bias-correction, climate models, climate change scenario, statistical downscaling

Plain Language Summary

Climate models can simulate biased representations of atmospheric processes, necessitating procedures for correction before use in hydrological applications. Such spatial bias can be caused for many reasons, one of which is the use of point data in establishing a spatial reference field to compare model simulations against. The most straightforward way to address this bias is to interpolate the locally observed data at the weather station onto a fine grid and use as reference. Alternatively, one can define a bias-correction model that accounts for the systematic impact induced by the use of point data, of special importance when the point data field is sparse and unevenly distributed. Here, we develop a novel approach to better address spatial bias using the Bayesian Kriging model. The results demonstrate the efficacy of the proposed model to simulate the bias-corrected daily rainfall sequences over large regions at fine resolutions.

1. Introduction

Due to the increased variability in the climate caused by global warming, the number of natural disasters has risen over the past four decades (Brown et al., 2008). Extreme events beyond the historical record and the bounds of natural variability have led to human casualties, property damages, and socioeconomic problems, creating international disagreements (AghaKouchak et al., 2014; Meehl et al., 2000; Oki & Kanae, 2006). It has become increasingly important to consider the changes in extreme events to design for safety from natural disasters as the climate gets warmer.

Climate models (e.g., global climate models (GCMs) and regional climate models (RCMs)) represent the main tools used to assess the future climate and the associated changes in the hydrological circulation over a long-term planning horizon (Borgomeo et al., 2014; Haro-Monteagudo et al., 2020; Steinschneider et al., 2015). These climate models attempt to simulate accurately the current climate as well as the response of the climate system to projected greenhouse gas concentrations into the future (Kattsov et al., 2007; Kripalani et al., 2007). However, model simulations are known to exhibit systematic bias, which has limited the direct use of especially precipitation from climate models (S. Kim et al., 2020; Woldemeskel et al., 2016). GCMs often have a low spatial resolution (100-300 km), with which regional climate may not be well reproduced (Diaconescu et al., 2018). In this context, RCMs with higher resolutions of 50km or less can provide a better representation of localized extreme rainfall events at finer spatial scales (Hadjinicolaou et al., 2011; Hanel & Buishand, 2011; Hewitson & Crane, 1996; Kyselý & Beranová, 2009; Kyselý et al., 2011; Wilks & Wilby, 1999).

The dynamic downscaling models (e.g., GCMs and RCMs) could capture key features of the

large climate system, and their outputs are mainly used for statistical downscaling approaches as inputs. Nonetheless, a reliable assessment of the implications for regional climates is still difficult to evaluate due to large uncertainties associated with initial and boundary conditions, model inadequacy, and resolution. A bias-correction approach, therefore, is often required to remove systematic bias for climate change impact assessment (Ayar et al., 2016; Tan et al., 2020; Wilby et al., 1998). More importantly, the quantile mapping (QM) based bias-correction approach with a downscaling model is commonly employed to correct hydro-meteorological variables (e.g., temperature and precipitation) at the station levels (Khalil et al., 2010; Y. T. Kim et al., 2020; Kwon et al., 2011; Lima et al., 2016; Lima et al., 2018; So et al., 2017) and finer resolutions (Guo & Wang, 2016; Mamalakis et al., 2017). Operationally, the statistical downscaling with the bias-correction is preferred due to its ability and simplicity in downscaling the GCM outputs, especially precipitation. Statistical downscaling methods include the delta approach (Kuok et al., 2016; Simonovic et al., 2016), the disaggregation approach (Koutsoyiannis et al., 1998; Pui et al., 2012), the nonstationary frequency analysis for downscaling precipitation extremes (Cheng & AghaKouchak, 2014; Lehmann et al., 2016; Lima et al., 2016), and approaches that model long-term dependence in both space and time (Mehrotra & Sharma, 2006; Mehrotra et al., 2013). Recently, the spatial downscaling of daily rainfall with the bias-correction has been widely applied for downscaling the simulated precipitation at desired grids and points (Immerzeel et al., 2009; Kim et al., 2019; Kim et al., 2015; Kwon et al., 2012; Nahar et al., 2017, 2018).

The bias-correction procedure often relies on observed variables at the weather station and grid points, which limits the full use of precipitation information obtained from the GCMs (or RCMs). Therefore, in order to fully utilize spatially varying climate model simulations, it is desirable to have the grid-based observation data at finer spatial resolution (or at least the

same resolution) than those from climate models to construct transfer function for adjusting bias at any desired location in space and time. It is worth noting that high-resolution daily gridded rainfall observation data is not readily available in many countries, including South Korea. Under these circumstances, there are two ways of modeling bias-correction at the desired grid resolution. First, the observed data at the weather station can be interpolated onto a fine grid to allow proper comparison to proceed. Alternatively, a set of parameters associated with the bias-correction can be gridded to form a spatial bias-correction model. To the best of our knowledge, there is no reason to prefer one or the other approach, as there is no systematic comparison in the literature between them in the context of bias-correction and spatial downscaling. This work expects to shed some light on this problem by exploring the following issues within the bias-correction and spatial downscaling:

(1) *Can all parameters associated with the spatial downscaling and bias-correction be simultaneously estimated and interpolated at the desired points?*

(2) *Can the interpolation of the bias-correction parameters be more effective than the interpolation of the daily precipitation in the context of spatial downscaling and bias-correction?*

(3) *Can a Bayesian Kriging based bias-correction approach effectively reproduce spatial dependency over a network of weather stations in the interpolated parameters associated with bias-correction?*

We intend to contribute to the existing literature with a novel approach that incorporates the spatial downscaling and the QDM (quantile delta mapping) with a Bayesian Kriging method, which, as compared with the large variety of statistical downscaling and bias correction

methods, finds motivation by the need to better address spatial dependencies of distribution parameters over gauging stations, including parameter uncertainty representation, and consequently provide unbiased, interpolated rainfall simulations with sound GCM/RCM subgrid variability. The Bayesian parameter estimation has been widely employed in the field of hydrology (Haddad et al., 2012; Kwon et al., 2011; Kwon et al., 2008; Liang et al., 2011; Lima et al., 2018; Lima & Lall, 2010; Viglione et al., 2013). Our focus here is to spatially downscale daily rainfall sequences simulated by regional climate models (RCMs) at any desired higher resolution, fully coupled with the QDM-based bias-correction. Our approach assumes that persistence or low-frequency variability attributes in the precipitation simulations are unbiased, and the main bias resides in the probability distribution of the rainfall amounts, largely a result of the use of point observed data instead of compatible gridded fields. The proposed modeling framework is demonstrated through a Leave-One-Out CV (LOOCV) evaluation in South Korea and climate change scenarios simulated by three different RCMs informed by the HadGEM2-AO GCM (Kim et al., 2020; Magnusson et al., 2020; Park et al., 2016; Sivula et al., 2020).

The precipitation data, including climate change scenarios used in this study, are summarized in the following section. The proposed Bayesian Kriging Spatial Disaggregation Quantile Delta Mapping (SD-QDM) approach for bias-correction and spatial downscaling is described in Section 3. The modeling results and their efficacy are demonstrated and discussed in Section 4. Finally, concluding remarks and a summary of this work are provided.

2. Observation Data and Climate Change Scenario

The daily precipitation data was compiled from over 92 Automatic Synoptic Observation

System (ASOS) across South Korea, operated by the Korea Meteorological Administration (KMA). Here, daily precipitation data at 60 weather stations with more than 45 years, ranging from 1973 to 2018, were finally selected for the subsequent analysis. The data used in this study can be obtained from the KMA data library, and the locations of weather stations used in this work are reported in Table S1. It should be important to highlight that approximately 60% of the annual rainfall is attributed to the summer season, from mid-June to early September, and extreme rainfalls often occur in this season. In the present work, we explore a seasonal-varying model under the circumstances of high rainfall variability, without focusing on an annual basis model, in a more general context.

This study aims to develop a spatial downscaling model that is fully coupled with the QDM-based bias-correction. The proposed model is applied to climate change scenarios simulated by three different RCMs employed in the Coordinated Regional Climate Downscaling Experiment-East Asia (CORDEX-EA), covering the entire East Asian areas, including South Korea, as shown in Figure S1. CORDEX is an internationally coordinated framework with the use of multiple RCMs for providing high-resolution climate change projections (So et al., 2017). The three different RCMs in the CORDEX-EA Phase 2 considered in the study are Seoul National University Regional Climate Model (SNURCM, with a spatial resolution of 12.5km×12.5km) (Lee et al., 2004), Weather Research and Forecasting (WRF, with a spatial resolution of 25km × 25km) model (Skamarock & Klemp, 2008), version 3.7, and the Consortium for Small-scale Modeling (COSMO)-CLM (or CCLM, with a spatial resolution of 25km×25km) 5.0 (Huang et al., 2015; Huang et al., 2017; Wang et al., 2013) for downscaling from the Hadley Centre Global Environment Model version 2 (HadGEM2-AO) atmosphere-ocean coupled general circulation model (Baek et al., 2013; Ngai et al., 2017). For more details, please refer to the link for details of the models (http://cordex-ea.climate.go.kr). In this study,

[7]

the future precipitation simulation for 2006-2100 under the representative concentration pathways (RCP) 4.5 and 8.5 was used with the historical precipitation simulation for 1979-2005. The Bayesian Kriging based SD-QDM approach was applied to provide downscaled precipitation at finer scales of about 6.25km, 12.5km and 12.5km resolution for SNURCM, WRF, and CCLM, respectively, which is typically more relevant as input for hydrological model applications.

3. Quantile Delta Mapping with Bayesian Kriging Approach

3.1 Bayesian Kriging based SD-QDM (Quantile Delta Mapping)

This study assumes that the daily precipitation amounts follow a Gamma distribution, and the CDF for Gamma distribution with the shape (k_{oh}) and scale (θ_{oh}) parameters can be defined as follow:

$$F(x|k_{oh}, \theta_{oh}) = \frac{1}{\theta^k \Gamma(k)} \int_0^x t^{k-1} e^{-t/\theta} dt ; \quad x \geq 0; \quad k, \theta > 0 \qquad (1)$$

$$k_{oh} \sim Inv - G(shape: 0.01, \ scale: 0.01) \qquad (2)$$

$$\theta_{oh} \sim Inv - G(shape: 0.01, \ scale: 0.01) \qquad (3)$$

For the Gamma distribution parameters (shape (k_{oh}) and scale (θ_{oh})), this study adopts a weakly informative Gamma distribution with the shape (0.01) and scale (0.01) parameters, conjugate prior distribution (Gelman, 2006; Lima et al., 2018; Lunn et al., 2012).

More generally, let us say that S is our variable of interest, to which we want to perform the spatial interpolation. S can be, for instance, the shape and scale parameters (i.e., the k_{oh} and θ_{oh}) of the Gamma distribution representing daily rainfall. S has a dimension n, the

number of rainfall gauges. For simplicity, we assume that S follows a multivariate normal distribution (MVN)

$$S \sim MVN(\mu, \tau^2 \Sigma) \tag{4}$$

where μ is the mean vector, τ^2 is the overall variance and Σ is a squared positive-definite matrix of size n representing the spatial dependence present.

We want to model the spatial variability of S through a parametric formulation for Σ:

$$\Sigma_{ij} = f(d_{ij}|\theta) \tag{5}$$

where d_{ij} is the distance between sites i and j and θ represents the correlation parameters.

We adopt here the powered exponential family for the function f:

$$f(d_{ij}|\phi, \kappa) = exp\left[-(\phi \cdot d_{ij})^{\delta}\right] \tag{6}$$

in which $\phi > 0$ governs the decay of the spatial correlation with distance and $0 < \delta < 2$ is the smoothing degree for S.

The estimation of the set of parameters $k_{oh}, \theta_{oh}, \mu, \tau^2, \phi$ and δ will be performed here using Bayesian inference, which provides a better groundwork to deal with parameter uncertainties, particularly for a small data set, as compared with traditional methods in ordinary Kriging. In addition to the distance, the formulation of the spatial model could be expanded by including auxiliary variables, such as elevation. The first two Gamma distribution parameters (i.e., k_{oh}, θ_{oh}) are spatially varying, but the remainings (τ^2, ϕ and δ) are not.

We start by setting the prior distributions for the parameters. As for μ, since no relevant information is known about it, we define independent and weakly informative priors:

$$\mu_i \propto 1 \tag{7}$$

where $i = 1, \ldots, n$ denotes the rainfall gauge. In our case, we have $\mu_i > 0$; therefore, our prior should be restricted to the positive domain. For the purpose of coding the MCMC simulation, we adopted the gamma distribution with the values 100 and 0.01 for scale and shape parameters, respectively.

For the overall variance parameter τ^2, we adopt a weakly informative, conjugate prior distribution as done in other studies (Gelman, 2006; Lima et al., 2018; Lunn et al., 2012):

$$\tau^2 \sim Inv - G(0.01, \ 0.01) \tag{8}$$

The prior for ϕ should ideally consider the minimum and maximum correlations at the minimum and maximum distances d_{ij}, respectively. It results, therefore, in the following uniform prior:

$$\phi \sim U(0, \ 1.0) \tag{9}$$

The joint posterior distribution of $k_{oh}, \theta_{oh}, \mu, \tau^2, \phi$ and δ is obtained following the Bayesian rule, which combines the Likelihood function $L(S|k_{oh}, \theta_{oh}, \mu, \tau^2, \phi, \delta)$ with the prior distributions described above. The Markov Chain Monte Carlo (MCMC) approach is adopted to sample from the joint posterior distribution, as an analytical integration is not feasible. The module GeoBUGS from the free software OpenBUGS was used for the MCMC algorithm. We checked the convergence using the R coefficient as proposed in Gelman et al. (2013) and visually based on the mixture of 5 chains with 2,000 simulations each. The powered exponential function in (6) is already implemented in GeoBUGS and this was the main

motivation for its choice. Moreover, it has proved to facilitate the convergence of the MCMC chains. Other spatial models (e.g., spherical) could be tested as well, but their implementation under the Bayesian framework used here may not be straightforward and the convergence of the MCMC chains may be hard to achieve. The flowchart of the proposed modeling framework is illustrated in Figure S2.

3.2 Quantile Delta Mapping (QDM) Approach

The quantile delta mapping (QDM) has been widely adopted for bias-correction of the climate change scenarios due to its ability to preserve relative changes in quantiles over historical and future simulations. The QDM approach is based on the cumulative distribution functions (CDFs) obtained from precipitation sequences over two time periods that represent modeled and observed precipitation during the historical period (1979-2005) and projected precipitation during the future period (2006-2100). A graphical representation of the QDM is provided in Figure S3.

The cumulative probability of the projected precipitation x_{sf} can be obtained from the CDFs (F_{sf}) as follows:

$$Q_{sf}(t) = F_{sf}^{(t)}[x_{sf}(t)], \ Q_{sf}(t) \in \{0,1\} \tag{10}$$

Here, $\mathbf{Q_{sf}}$ represents non-exceedance probability at time t. The modeled quantiles during the historical and future periods corresponding to the Q_{sf} can be obtained from F_{sh}^{-1} and F_{sf}^{-1} (inverse cumulative distribution functions, ICDFs), respectively. The simulated precipitation x_{sf} is denoted by the subscript s during the future period, denoted by the subscript f, and similarly for the CDF F_{sf}. The relative quantile changes in precipitation over two time periods at time t is given as follows:

$$\Delta w_s(t) = \frac{F_{sf}^{(t)-1}[Q_{sf}(t)]}{F_{sh}^{(t)-1}[Q_{sf}(t)]} = \frac{x_{sf}(t)}{F_{sh}^{(t)-1}[Q_{sf}(t)]} \tag{11}$$

The bias-corrected quantile $\hat{x}_{sh}$ for the modeled Q_{sf} at time t can be obtained by adopting the inverse CDF (F_{oh}^{-1}) estimated from the observed precipitation during the historical period as written in equation (3). The CDF F_{sh} of the simulated precipitation during the historical period is denoted by the subscripts s, and h and similarly for the CDF of the observed precipitation during the historical period F_{oh}, marked by the subscripts o, and h.

$$\hat{x}_{sh}(t) = F_{oh}^{-1}(Q_{sf}(t)) \tag{12}$$

Lastly, the bias-corrected future scenario values $\hat{x}_{sf}$ can be computed by multiplying the relative quantile changes Δw_s and the bias-corrected quantile $\hat{x}_{sh}(t)$ during the historical period. Hence the relative changes can be preserved in the course of bias-correction.

$$\hat{x}_{sf}(t) = \hat{x}_{sh}(t) \cdot \Delta w_s(t) \tag{13}$$

4. Result and Discussion

4.1 Parameter Estimation and Cross Validation

The gamma distribution is often used to model daily rainfall (Aksoy, 2000; Kim et al., 2019; Lee et al., 2019; Yoo et al., 2005) in hydrological applications, particularly for the bias-correction of modeled rainfall sequences obtained from climate models (Bárdossy & Pegram, 2011; Heo et al., 2019; Johnson & Sharma, 2012; Kim et al., 2016; Lee et al., 2019; Piani et al., 2010; Volosciuk et al., 2017). The two-parameter gamma distribution was selected based on Bayesian Information Criterion (BIC) value as the best distribution for the observed and

modeled daily rainfall during the historical period (1979-2005) and projected daily rainfall during the future period (2006-2100).

First, we explored whether all parameters can be simultaneously estimated and interpolated at the desired points within a Bayesian Kriging modeling framework. In this perspective, the gamma distribution was fitted to daily rainfall series from 60 weather stations, and the distribution parameters were retained. More importantly, the distribution parameters were simultaneously interpolated by the Bayesian Kriging approach over the gauge locations in the parameter estimation process. The model performance was explored by testing the predictions of the parameters over the entire station set within a LOOCV framework. Here, the LOOCV scheme drops one gauge, and the remaining gauges are used to estimate all the parameters associated with both the gamma distribution and the Kriging model. The spatial location of the gauge that is not considered in the calibration process is then used as input to obtain its predictive gamma distribution parameters. The cross-validated results and the associated credible intervals for August are illustrated in Figure 1(a) for a graphical representation of the efficacy of the proposed model. The results showed that the predicted parameters (i.e., shape and scale parameters) fall within the 95% credible interval derived from the predictive posterior distribution. As illustrated in Figure 1(a), the predicted parameters are strongly correlated with those estimated directly from the observations (Pearson correlation equals to 0.95 for the shape and 0.98 for the scale parameter). The results for the remaining months are also similar to that of August and are displayed in Figure S4. As inferred from the Pearson correlation coefficient (not shown here), the proposed model yields slightly better predictions for the scale parameter as compared to those for the shape parameter. Moreover, a seasonal-varying model under the LOOCV scheme is investigated, and the model performance with four different performance metrics (Table S2)

(correlation coefficient (CC), Nash–Sutcliffe Efficiency (NSE), Index-of-Agreement (IoA), and Root Mean Square Error (RMSE)) is summarized in Table S3. For the shape parameter, the model showed slightly lower performance in the spring season from February to April in terms of the NSE. Still, the model predictability can be regarded as "Very good: NSE ≥ 0.7" according to the given criteria suggested by Kalin et al. (2010), and other performance measures are largely comparable to that of different seasons. For the scale parameter, the predicted parameters appear to be almost identical to that of the observed, confirming the efficacy of the model.

[Insert Figure 1]

Following the previous step, we investigated whether the grid generated by direct interpolating the gamma parameters can be more reliable than the grid obtained by first interpolating the observed daily precipitation onto the grid and thereafter estimating the gamma parameters over these points. The directly interpolated gamma distribution parameters through the proposed Bayesian Kriging approach were then compared to the parameters obtained from the interpolated daily precipitation that will serve as a baseline model. Note that daily precipitation was also interpolated by the Bayesian Kriging approach and all the results presented here were achieved under the LOOCV scheme. Figure 1(b) shows the gamma parameters interpolated onto the locations of the rainfall gauge set using both approaches (displayed along the y-axis) and the local estimates of the gamma parameters (displayed along the x-axis), obtained from the rainfall gauge data. It is clear that the proposed model outperformed the baseline model, which is based on the interpolated daily rainfall series. This finding is likely a direct result of the difficulty to effectively interpolate daily precipitation sequences due to the nature of rainfall intermittency in both

space and time, which is closely related to the binary process representing rainfall occurrence (i.e., rain and no rain) (Hasenauer et al., 2003; Kleiber et al., 2012; Militino et al., 2015). Furthermore, due to the large spatial and temporal variability in precipitation, it is believed that there are limitations in accurately reproducing the underlying distribution of precipitation in the interpolation stage (Y.-T. Kim et al., 2020; Ly et al., 2013; Mandapaka et al., 2009; Obled et al., 1994). The bias in estimating the parameters in the interpolation process eventually leads to incorrect estimation of the probability density function, as illustrated in Figure 1. Here, for the graphical representation, gamma distributions over two approaches are compared with the representative stations (i.e., ST. No 100 and 105). Based on these results and under these contexts, we suggest avoiding the estimate of the parameters onto the fine grid from the interpolated daily precipitation for the bias-correction and their use to the spatial downscaling. The results for the remaining months are all comparable to that of August and are displayed in the supplementary information section (Figure S5).

4.2 Interpolation of Parameters Using Bayesian Kriging Approach

Furthermore, the gridded estimates using the Bayesian Kriging approach (right panel) are illustrated in Figure 2 for August, with the point estimates (left panel) for the weather stations considered in this study. The relatively large shape parameters are identified in the southeastern regions, while relatively large scale parameter values are concentrated in the southern coastal area. On the other hand, the smaller shape parameters are mainly distributed in the mid-western region, while the lower scale parameter values are largely seen in the southern part of South Korea. Overall, the proposed Bayesian Kriging approach is capable of reproducing the main spatial patterns seen in the direct point estimates of both shape and scale parameters.

[Insert Figure 2]

This study further tested the efficacy of the model in effectively reproducing spatial dependency over a network of weather stations in the interpolated gamma parameters. The semivariogram of gamma parameters estimated from the posterior distributions obtained from the Bayesian Kriging approach was then compared to that obtained from local estimates based on gauged rainfall data. Figure 3 (a) shows, for August, the efficacy of the proposed model to reproduce the bias-correction parameters while preserving the spatial variability observed in the historical data-based estimates, where both semivariograms are almost identical. The semivariogram directly obtained from the interpolated precipitation is significantly biased from the *observed* one. Similarly, the spatial pattern of the parameters for the remaining months is well captured by the Bayesian Kriging based SD-QDM approach, as displayed in the supplementary information, Figure S6. Moreover, the Bayesian Kriging based SD-QDM model was compared with the ordinary Kriging approach, which is widely adopted in spatial interpolation. The results confirmed that the proposed approach showed better performance to estimate the Gamma distribution parameters in the context of cross-validation, as illustrated in Figure S7.

[Insert Figure 3]

4.3 Spatial Downscaling of Climate Change Scenarios

Finally, the interpolated parameters shown in Figure 2 can then be used to construct the transfer functions at the fine grid for the bias-correction and spatial downscaling of simulated daily precipitation. More specifically, in order to illustrate the use of the proposed Bayesian Kriging based SD-QDM approach, this work downscaled the historical and the future daily precipitation simulated by RCMs in the CORDEX-EA Phase 2 for 1979-2005 and 2006-

2100, respectively, under the RCP 4.5 and RCP 8.5 scenarios shown in Figure 3(b-c) for WRF RCM as a representative example. Further, Figure S8 displays the mean annual precipitation compiled from three RCMs (i.e., SNURCM, WRF and CCLM) without bias-correction (Figure S8 (a)) and with bias-correction based on the SD-QDM approach (Figure S8(b)). Here, spatial downscaling was done at the fine grid by interpolating the pointwise estimation of QDM parameters onto the same grid points (or same spatial resolutions) of three RCMs, with resolutions of 12.5km, 25km and 25km for SNURCM, WRF, and CCLM, respectively. Further, the proposed Bayesian Kriging based SD-QDM approach provides downscaled precipitation at finer scales of about 6.25km, 12.5km and 12.5km resolution for SNURCM, WRF, and CCLM, respectively, which could be more relevant for hydrological models as input (Figure S8(c)). As illustrated in Figure S8, the spatial patterns of mean annual precipitation with Bayesian Kriging SD-QDM are largely similar and comparable to that of the three RCMs without bias-correction, confirming that the proposed model can preserve the spatial variability after bias-correction. But more importantly, there is a substantial increase in the amount of precipitation with bias-correction due to the significant underestimation of precipitation simulated from RCMs without bias-correction.

5. Concluding Remarks

The bias-correction of precipitation simulated by GCMs (or RCMs) is often dependent on observed precipitation at the weather station and grid points, limiting the full use of climate information obtained from climate models. In particular, daily gridded rainfall observation data at high resolution is not readily available in many countries. In this perspective, this study proposed a Kriging Bayesian SD-QDM approach to obtain distribution and bias-correction parameters over a fine grid, representing the appropriate spatial dependencies observed over gauging stations. We illustrated the efficiency and applicability of the

proposed model through a cross-validatory experiment (LOOCV scheme) using observed

rainfall data from several stations covering the entire South Korea and historical and future

rainfall scenarios generated by 3 RCMs. The key findings from this work are provided as

follows:

(1) We investigated whether all parameters associated with the SD-QDM approach can

be simultaneously estimated and gridded at the desired points within a Bayesian

Kriging modeling framework. In particular, a gamma distribution was fitted to daily

rainfall series over 60 weather stations, and the associated parameters were

simultaneously interpolated by a Bayesian Kriging approach onto a fine grid. The

cross-validated results under the LOOCV scheme showed that the predicted

(interpolated) parameters at the locations of the gauge stations are almost identical to

that of local estimates obtained directly from the fitting of the gamma distribution to

the rainfall gauge data parameters, confirming the efficacy of the model.

(2) Under the LOOCV scheme, we also found that the directly interpolated gamma

parameters through the proposed Bayesian Kriging approach outperformed the

baseline model based on interpolated daily rainfall, which produced a substantial bias

that leads to an incorrect representation of the probability density function. Under

these circumstances, the direct estimation of the distribution parameters from the

interpolated daily precipitation for bias-correction and spatial downscaling should be

cautious. This study further investigated whether spatial dependency over the

interpolated gamma parameters can be effectively preserved. In this regard, the

semivariogram of gamma parameters obtained from the proposed model was

evaluated. The results confirmed that the proposed model could effectively reproduce

the spatial variability of parameters estimated from gauging stations, given that the semivariogram of the interpolated parameters estimated from the Bayesian Kriging based SD-QDM approach was almost identical to that of the local parameters estimated from gauged rainfall data. In contrast, the semivariogram directly obtained from the interpolated precipitation highly deviated from the observed.

The proposed Bayesian Kriging based SD-QDM approach could apply to various applications with different temporal scales in hydrometeorology to establish a spatial reference field to compare model simulations against. More specifically, the bias-correction and spatial downscaling for other climate variables, including temperature, soil moisture, solar radiation and wind field, rely on observed variables at the weather station (or grid points), limiting the full use of climate information obtained from the climate models. Although the proposed modeling framework provides an important basis for the spatial downscaling of climate model outputs, the variability of precipitation areal reduction factors are not fully incorporated and explored in this study. These aspects will be further investigated in future work.

Acknowledgments

References

AghaKouchak, A., Cheng, L., Mazdiyasni, O., & Farahmand, A. (2014). Global warming and changes in risk of concurrent climate extremes: Insights from the 2014 California drought. *Geophysical Research Letters*, *41*(24), 8847-8852.

Aksoy, H. (2000). Use of gamma distribution in hydrological analysis. *Turkish Journal of Engineering and Environmental Sciences*, *24*(6), 419-428.

Ayar, P. V., Vrac, M., Bastin, S., Carreau, J., Déqué, M., & Gallardo, C. (2016). Intercomparison of statistical and dynamical downscaling models under the EURO-and MED-CORDEX initiative framework: present climate evaluations. *Climate dynamics*, *46*(3-4), 1301-1329.

Baek, H.-J., Lee, J., Lee, H.-S., Hyun, Y.-K., Cho, C., Kwon, W.-T., Marzin, C., Gan, S.-Y., Kim, M.-J., & Choi, D.-H. (2013). Climate change in the 21st century simulated by HadGEM2-AO under representative concentration pathways. *Asia-Pacific Journal of Atmospheric Sciences*, *49*(5), 603-618.

Bárdossy, A., & Pegram, G. (2011). Downscaling precipitation using regional climate models and circulation patterns toward hydrology. *Water Resources Research*, *47*(4).

Borgomeo, E., Hall, J. W., Fung, F., Watts, G., Colquhoun, K., & Lambert, C. (2014). Risk-based water resources planning: Incorporating probabilistic nonstationary climate uncertainties. *Water Resources Research*, *50*(8), 6850-6873.

Brown, S., Caesar, J., & Ferro, C. A. (2008). Global changes in extreme daily temperature since 1950. *Journal of Geophysical Research: Atmospheres*, *113*(D5).

Cheng, L., & AghaKouchak, A. (2014). Nonstationary precipitation intensity-duration-frequency curves for infrastructure design in a changing climate. *Scientific reports*, *4*, 7093.

Diaconescu, E. P., Mailhot, A., Brown, R., & Chaumont, D. (2018). Evaluation of CORDEX-Arctic daily precipitation and temperature-based climate indices over Canadian Arctic land areas. *Climate Dynamics*, *50*(5-6), 2061-2085.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian analysis*, *1*(3), 515-534.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. CRC press.

Guo, D., & Wang, H. (2016). Comparison of a very-fine-resolution GCM with RCM dynamical downscaling in simulating climate in China. *Advances in Atmospheric Sciences*, *33*(5), 559-570.

Haddad, K., Rahman, A., & Stedinger, J. R. (2012). Regional flood frequency analysis using Bayesian generalized least squares: a comparison between quantile and parameter regression techniques. *Hydrological Processes*, *26*(7), 1008-1021.

Hadjinicolaou, P., Giannakopoulos, C., Zerefos, C., Lange, M. A., Pashiardis, S., & Lelieveld, J. (2011). Mid-21st century climate and weather extremes in Cyprus as projected by six regional climate models. *Regional Environmental Change*, *11*(3), 441-457.

Hanel, M., & Buishand, T. A. (2011). Analysis of precipitation extremes in an ensemble of transient regional climate model simulations for the Rhine basin. *Climate Dynamics*, *36*(5-6), 1135-1153.

Haro-Monteagudo, D., Palazón, L., & Beguería, S. (2020). Long-term sustainability of large water resource systems under climate change: A cascade modeling approach. *Journal of Hydrology*, *582*, 124546.

Hasenauer, H., Merganicova, K., Petritsch, R., Pietsch, S. A., & Thornton, P. E. (2003). Validating daily climate interpolations over complex terrain in Austria. *Agricultural and Forest Meteorology*, *119*(1-2), 87-107.

Heo, J.-H., Ahn, H., Shin, J.-Y., Kjeldsen, T. R., & Jeong, C. (2019). Probability distributions for a quantile mapping technique for a bias correction of precipitation data: A case study to precipitation data under climate change. *Water*, *11*(7), 1475.

Hewitson, B. C., & Crane, R. G. (1996). Climate downscaling: techniques and application. *Climate Research*, *7*(2), 85-95.

Huang, B., Polanski, S., & Cubasch, U. (2015). Assessment of precipitation climatology in an ensemble of CORDEX-East Asia regional climate simulations. *Climate Research*, *64*(2), 141-158.

Huang, J., Wang, Y., Fischer, T., Su, B., Li, X., & Jiang, T. (2017). Simulation and projection of climatic changes in the Indus River Basin, using the regional climate model COSMO-CLM. *International Journal of Climatology*, *37*(5), 2545-2562.

Immerzeel, W., Rutten, M., & Droogers, P. (2009). Spatial downscaling of TRMM precipitation using vegetative response on the Iberian Peninsula. *Remote Sensing of Environment*, *113*(2), 362-370.

Johnson, F., & Sharma, A. (2012). A nesting model for bias correction of variability at multiple time scales in general circulation model precipitation simulations. *Water Resources Research*, *48*(1).

Kalin, L., Isik, S., Schoonover, J. E., & Lockaby, B. G. (2010). Predicting water quality in unmonitored watersheds using artificial neural networks. *Journal of environmental quality*, *39*(4), 1429-1440.

Kattsov, V. M., Walsh, J. E., Chapman, W. L., Govorkova, V. A., Pavlova, T. V., & Zhang, X. (2007). Simulation and projection of Arctic freshwater budget components by the IPCC AR4 global climate models. *Journal of Hydrometeorology*, *8*(3), 571-589.

Khalil, A. F., Kwon, H.-H., Lall, U., & Kaheil, Y. H. (2010). Predictive downscaling based on non-homogeneous hidden Markov models. *Hydrological Sciences Journal–Journal des Sciences Hydrologiques*, *55*(3), 333-350.

Kim, D.-I., Kwon, H.-H., & Han, D. (2019). Bias correction of daily precipitation over South Korea from the long-term reanalysis using a composite Gamma-Pareto distribution approach. *Hydrology Research*, *50*(4), 1138-1161.

Kim, K. B., Kwon, H.-H., & Han, D. (2015). Bias correction methods for regional climate model simulations considering the distributional parametric uncertainty underlying the observations. *Journal of Hydrology*, *530*, 568-579.

Kim, K. B., Kwon, H.-H., & Han, D. (2016). Precipitation ensembles conforming to natural variations derived from a regional climate model using a new bias correction scheme. *Hydrology & Earth System Sciences*, *20*(5).

Kim, S., Eghdamirad, S., Sharma, A., & Kim, J. H. (2020). Quantification of Uncertainty in Projections of Extreme Daily Precipitation. *Earth and Space Science*, *7*(8). https://doi.org/ARTN e2019EA001052
10.1029/2019EA001052

Kim, Y.-T., Park, M., & Kwon, H.-H. (2020). Optimization of PRISM parameters using the SCEM-UA algorithm for gridded daily time series precipitation. *Journal of Korea Water Resources Association*, *53*(10), 903-915.

Kim, Y. T., So, B. J., Kwon, H. H., & Lall, U. (2020). A multiscale precipitation forecasting framework: Linking teleconnections and climate dipoles to seasonal and 24-hr extreme rainfall prediction. *Geophysical Research Letters*, *47*(3), e2019GL085418.

Kleiber, W., Katz, R. W., & Rajagopalan, B. (2012). Daily spatiotemporal precipitation simulation using latent and transformed Gaussian processes. *Water Resources Research*, *48*(1).

Koutsoyiannis, D., Kozonis, D., & Manetas, A. (1998). A mathematical framework for studying rainfall intensity-duration-frequency relationships. *Journal of Hydrology*, *206*(1-2), 118-135.

Kripalani, R., Oh, J., & Chaudhari, H. (2007). Response of the East Asian summer monsoon to doubled atmospheric CO 2: Coupled climate model simulations and projections under IPCC AR4. *Theoretical and applied climatology*, *87*(1-4), 1-28.

Kuok, K. K., Mah, Y. S., Imteaz, M. A., & Kueh, S. M. (2016). Comparison of future intensity duration frequency curve by considering the impact of climate change: case study for Kuching city. *International Journal of River Basin Management*, *14*(1), 47-55.

Kwon, H.-H., Sivakumar, B., Moon, Y.-I., & Kim, B.-S. (2011). Assessment of change in design flood frequency under climate change using a multivariate downscaling model and a precipitation-runoff model. *Stochastic Environmental Research and Risk Assessment*, *25*(4), 567-581.

Kwon, H. H., Brown, C., & Lall, U. (2008). Climate informed flood frequency analysis and prediction in Montana using hierarchical Bayesian modeling. *Geophysical Research Letters*, *35*(5).

Kwon, H. H., de Assis de Souza Filho, F., Block, P., Sun, L., Lall, U., & Reis Jr, D. S. (2012). Uncertainty assessment of hydrologic and climate forecast models in Northeastern Brazil. *Hydrological Processes*, *26*(25), 3875-3885.

Kyselý, J., & Beranová, R. (2009). Climate-change effects on extreme precipitation in central Europe: uncertainties of scenarios based on regional climate models. *Theoretical and applied climatology*, *95*(3-4), 361-374.

Kyselý, J., Gaál, L., Beranová, R., & Plavcová, E. (2011). Climate change scenarios of precipitation extremes in Central Europe from ENSEMBLES regional climate models. *Theoretical and applied climatology*, *104*(3-4), 529-542.

Lee, D.-K., Cha, D.-H., & Kang, H.-S. (2004). Regional climate simulation of the 1998 summer flood over East Asia. *Journal of the Meteorological Society of Japan. Ser. II*, *82*(6), 1735-1753.

Lee, M. H., Im, E. S., & Bae, D. H. (2019). Impact of the spatial variability of daily precipitation on hydrological projections: A comparison of GCM-and RCM-driven cases in the Han River basin, Korea. *Hydrological Processes*, *33*(16), 2240-2257.

Lehmann, E. A., Phatak, A., Stephenson, A., & Lau, R. (2016). Spatial modelling framework for the characterisation of rainfall extremes at different durations and under climate change. *Environmetrics*, *27*(4), 239-251.

Liang, Z., Li, B., Yu, Z., & Chang, W. (2011). Application of Bayesian approach to hydrological frequency analysis. *Science China Technological Sciences*, *54*(5), 1183-1192.

Lima, C. H., Kwon, H.-H., & Kim, J.-Y. (2016). A Bayesian beta distribution model for estimating rainfall IDF curves in a changing climate. *Journal of Hydrology*, *540*, 744-756.

Lima, C. H., Kwon, H.-H., & Kim, Y.-T. (2018). A local-regional scaling-invariant Bayesian GEV model for estimating rainfall IDF curves in a future climate. *Journal of Hydrology*, *566*, 73-88.

Lima, C. H., & Lall, U. (2010). Spatial scaling in a changing climate: A hierarchical bayesian model for non-stationary multi-site annual maximum and monthly streamflow. *Journal of Hydrology*, *383*(3-4), 307-318.

Lunn, D., Jackson, C., Best, N., Thomas, A., & Spiegelhalter, D. (2012). *The BUGS book: A practical introduction to Bayesian analysis*. CRC press.

Ly, S., Charles, C., & Degré, A. (2013). Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale: a review. *Biotechnologie, Agronomie, Société et Environnement*, *17*(2), 392-406.

Magnusson, M., Vehtari, A., Jonasson, J., & Andersen, M. (2020). Leave-one-out cross-validation for Bayesian model comparison in large data. International Conference on Artificial Intelligence and Statistics,

Mamalakis, A., Langousis, A., Deidda, R., & Marrocu, M. (2017). A parametric approach for simultaneous bias correction and high-resolution downscaling of climate model rainfall. *Water Resources Research*, *53*(3), 2149-2170.

Mandapaka, P. V., Krajewski, W. F., Mantilla, R., & Gupta, V. K. (2009). Dissecting the effect of rainfall variability on the statistical structure of peak flows. *Advances in Water Resources*, *32*(10), 1508-1525.

Meehl, G. A., Zwiers, F., Evans, J., Knutson, T., Mearns, L., & Whetton, P. (2000). Trends in extreme weather and climate events: issues related to modeling extremes in projections of future climate change. *Bulletin of the American Meteorological Society*, *81*(3), 427-436.

Mehrotra, R., & Sharma, A. (2006). A nonparametric stochastic downscaling framework for daily rainfall at multiple locations. *Journal of Geophysical Research-Atmospheres*, *111*(D15101), doi:10.1029/2005JD006637.

Mehrotra, R., Sharma, A., Kumar, D. N., & Reshmidevi, T. V. (2013). Assessing future rainfall projections using multiple GCMS and a multi-site stochastic downscaling model. *Journal of Hydrology*, *488*, 84-100. http://www.scopus.com/inward/record.url?eid=2-s2.0-84886300952&partnerID=40&md5=5398b57a0e59b4eae48134face11c17d

Militino, A., Ugarte, M., Goicoa, T., & Genton, M. (2015). Interpolation of daily rainfall using spatiotemporal models and clustering. *International Journal of Climatology*, *35*(7), 1453-1464.

Nahar, J., Johnson, F., & Sharma, A. (2017). A rank-based approach for correcting systematic biases in spatial disaggregation of coarse-scale climate simulations [Article]. *Journal of Hydrology*, *550*, 716-725. https://doi.org/10.1016/j.jhydrol.2017.05.045

Nahar, J., Johnson, F., & Sharma, A. (2018). Addressing Spatial Dependence Bias in Climate Model Simulations—An Independent Component Analysis Approach [Article]. *Water Resources Research*, *54*(2), 827-841. https://doi.org/10.1002/2017WR021293

Ngai, S. T., Tangang, F., & Juneng, L. (2017). Bias correction of global and regional simulated daily precipitation and surface mean temperature over Southeast Asia using quantile mapping method. *Global and Planetary Change*, *149*, 79-90.

Obled, C., Wendling, J., & Beven, K. (1994). The sensitivity of hydrological models to spatial rainfall patterns: an evaluation using observed data. *Journal of Hydrology*, *159*(1-4), 305-333.

Oki, T., & Kanae, S. (2006). Global hydrological cycles and world water resources. *science*, *313*(5790), 1068-1072.

Park, C., Min, S.-K., Lee, D., Cha, D.-H., Suh, M.-S., Kang, H.-S., Hong, S.-Y., Lee, D.-K., Baek, H.-J., & Boo, K.-O. (2016). Evaluation of multiple regional climate models for summer climate extremes over East Asia. *Climate dynamics*, *46*(7-8), 2469-2486.

Piani, C., Haerter, J., & Coppola, E. (2010). Statistical bias correction for daily precipitation in regional climate models over Europe. *Theoretical and applied climatology*, *99*(1-2), 187-192.

Pui, A., Sharma, A., Mehrotra, R., Sivakumar, B., & Jeremiah, E. (2012). A comparison of alternatives for daily to sub-daily rainfall disaggregation. *Journal of Hydrology*, *470*, 138-157.

Simonovic, S. P., Schardong, A., Sandink, D., & Srivastav, R. (2016). A web-based tool for the development of intensity duration frequency curves under changing climate. *Environmental modelling & software*, *81*, 136-153.

Sivula, T., Magnusson, M., & Vehtari, A. (2020). Unbiased estimator for the variance of the leave-one-out cross-validation estimator for a Bayesian normal model with fixed variance. *arXiv preprint arXiv:2008.10859*.

Skamarock, W. C., & Klemp, J. B. (2008). A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. *Journal of computational physics*, *227*(7), 3465-3485.

So, B.-J., Kim, J.-Y., Kwon, H.-H., & Lima, C. H. (2017). Stochastic extreme downscaling model for an assessment of changes in rainfall intensity-duration-frequency curves over South Korea using multiple regional climate models. *Journal of Hydrology*, *553*, 321-337.

Steinschneider, S., McCrary, R., Wi, S., Mulligan, K., Mearns, L. O., & Brown, C. (2015). Expanded decision-scaling framework to select robust long-term water-system plans under hydroclimatic uncertainties. *Journal of Water Resources Planning and Management*, *141*(11), 04015023.

Tan, Y., Guzman, S. M., Dong, Z., & Tan, L. (2020). Selection of Effective GCM Bias Correction Methods and Evaluation of Hydrological Response under Future Climate Scenarios. *Climate*, *8*(10), 108.

Viglione, A., Merz, R., Salinas, J. L., & Blöschl, G. (2013). Flood frequency hydrology: 3. A Bayesian analysis. *Water Resources Research*, *49*(2), 675-692.

Volosciuk, C. D., Maraun, D., Vrac, M., & Widmann, M. (2017). A combined statistical bias correction and stochastic downscaling method for precipitation. *Hydrology and Earth System Sciences*, *21*, 1693-1719.

Wang, D., Menz, C., Simon, T., Simmer, C., & Ohlwein, C. (2013). Regional dynamical downscaling with CCLM over East Asia. *Meteorology and Atmospheric Physics*, *121*(1-2), 39-53.

Wilby, R. L., Wigley, T., Conway, D., Jones, P., Hewitson, B., Main, J., & Wilks, D. (1998). Statistical downscaling of general circulation model output: A comparison of methods. *Water resources research*, *34*(11), 2995-3008.

Wilks, D. S., & Wilby, R. L. (1999). The weather generation game: a review of stochastic weather models. *Progress in physical geography*, *23*(3), 329-357.

Woldemeskel, F. M., Sharma, A., Sivakumar, B., & Mehrotra, R. (2016). Quantification of precipitation and temperature uncertainties simulated by CMIP3 and CMIP5 models. *Journal of Geophysical Research-Atmospheres*, *121*(1), 3-17. https://doi.org/10.1002/2015JD023719

Yoo, C., Jung, K. S., & Kim, T. W. (2005). Rainfall frequency analysis using a mixed Gamma distribution: evaluation of the global warming effect on daily rainfall. *Hydrological Processes: An International Journal*, *19*(19), 3851-3861.
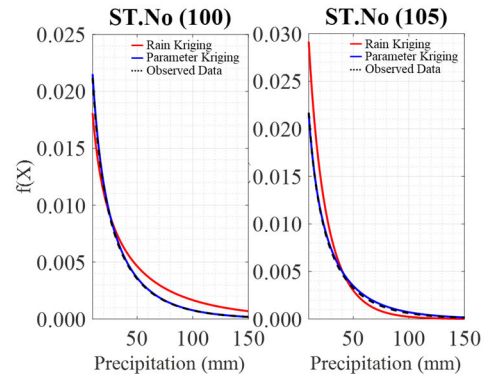
Figures

(a)
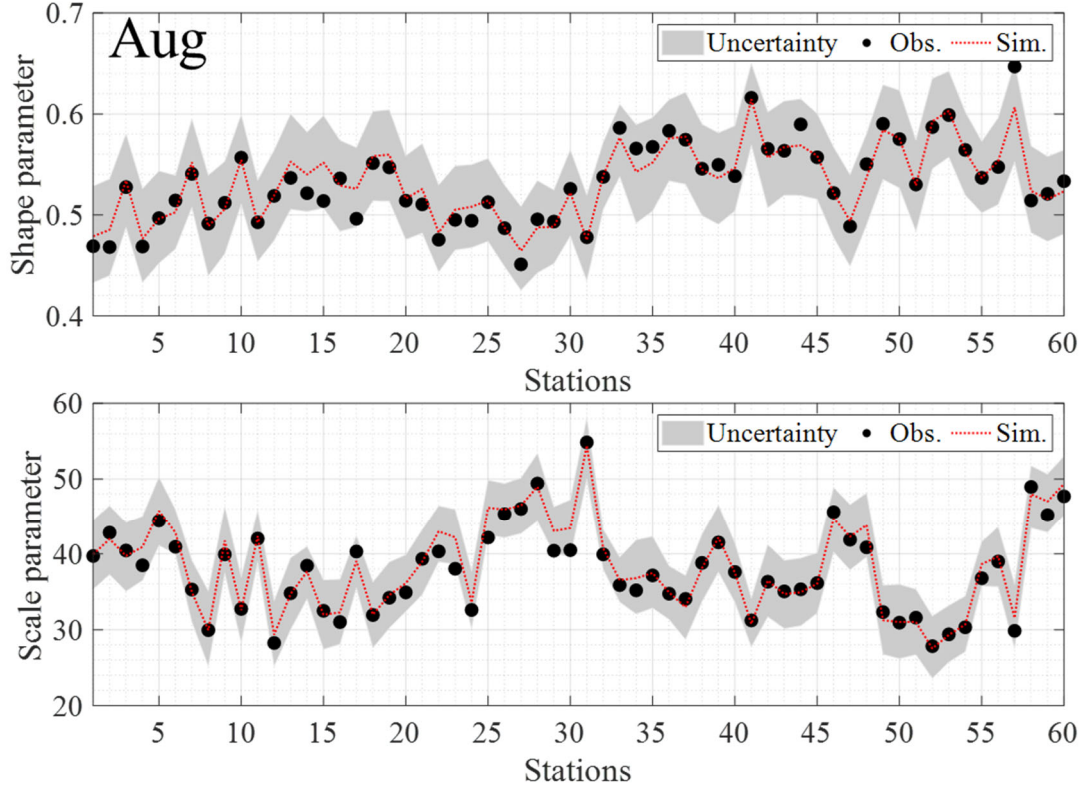
(b)
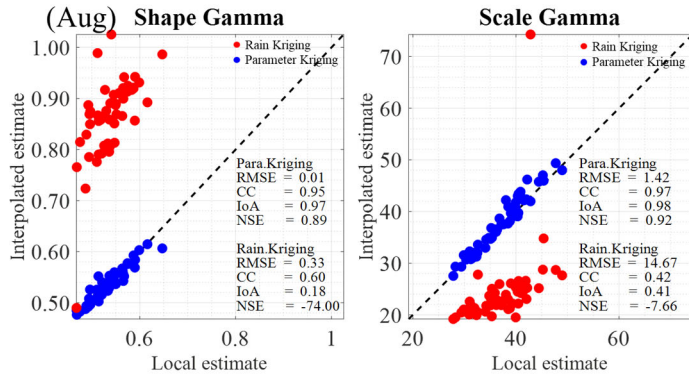
(c)

Figure 1. (a) The cross-validated results and their 95% credible intervals (shaded region) of the Gamma parameters within a LOOCV assessment in August. The black-filled circles are the parameters estimated from observed precipitation, while the red-dotted lines represent the median values estimated from the predictive posterior distribution. The number on the X-axis represents the gauge code, summarized in the supplementary information. (b) Comparison of gamma parameters directly interpolated from the proposed model (blue-filled circles) and obtained from the interpolated precipitation field (red-filled circles). Both estimates are shown along the y-axis, while the x-axis refers to the local estimates, obtained from the gauged data. (c) Comparison of gamma distributions directly interpolated from the proposed model (blue line) and obtained from the interpolated precipitation field (red line), along with the observed gamma distributions obtained from the representative stations (i.e., ST. No 100 and 105)
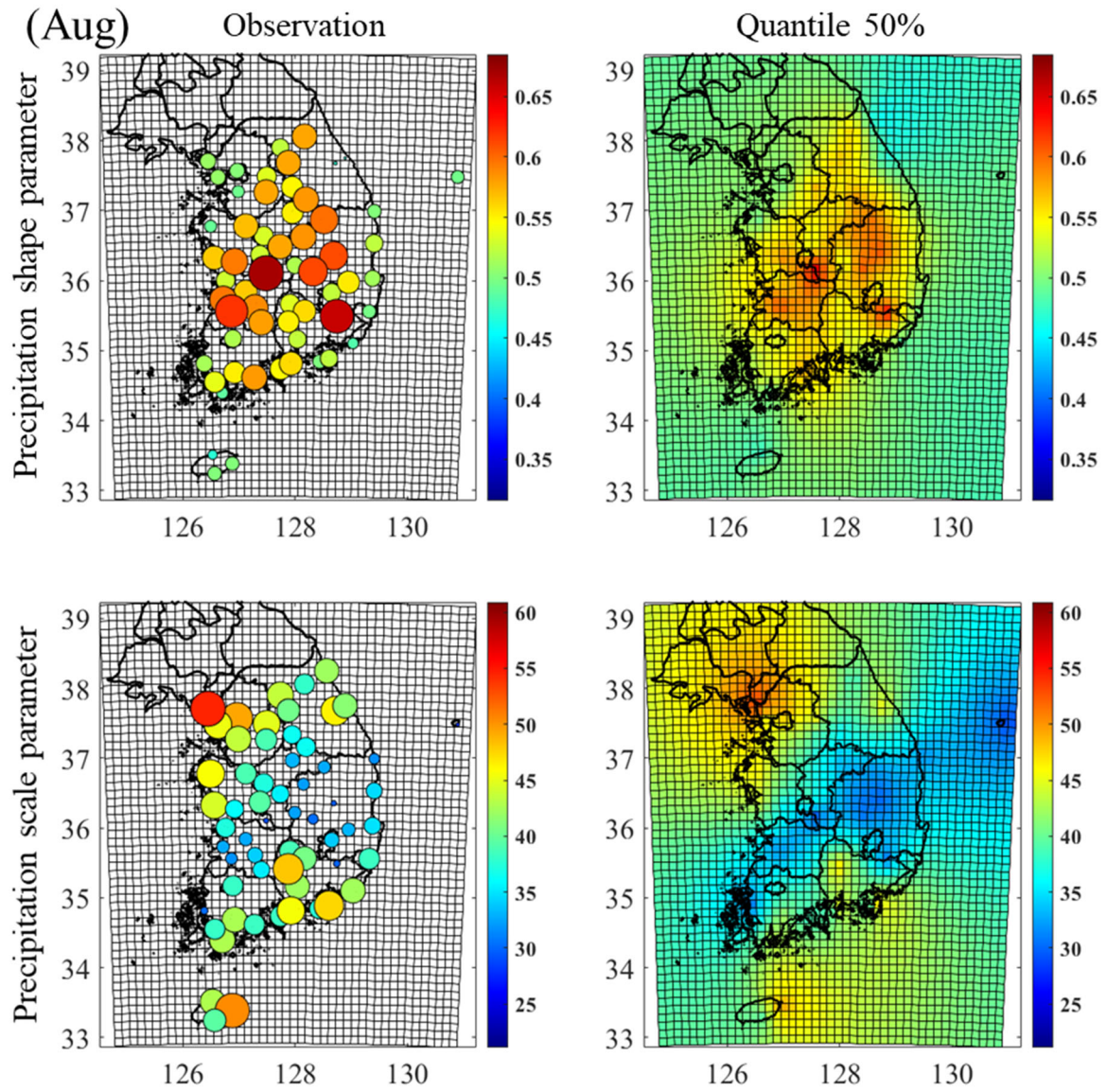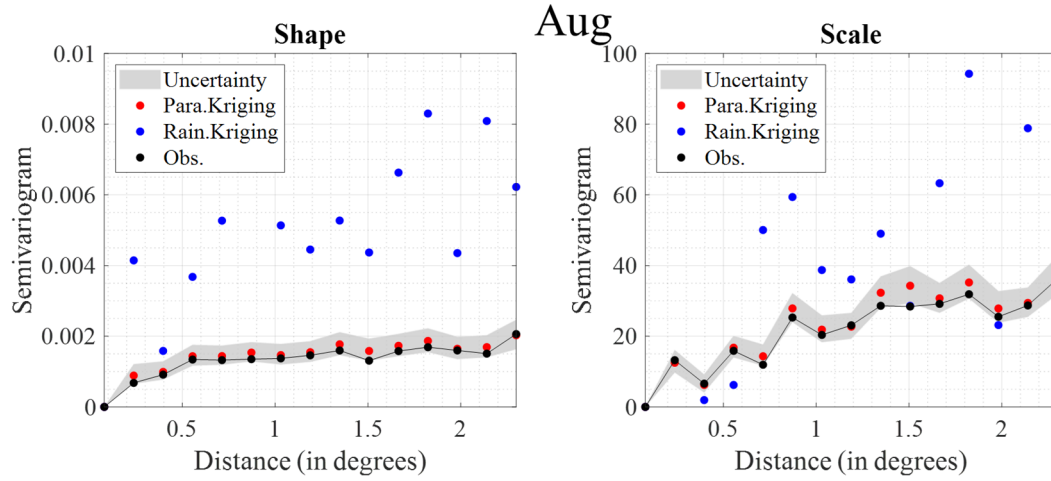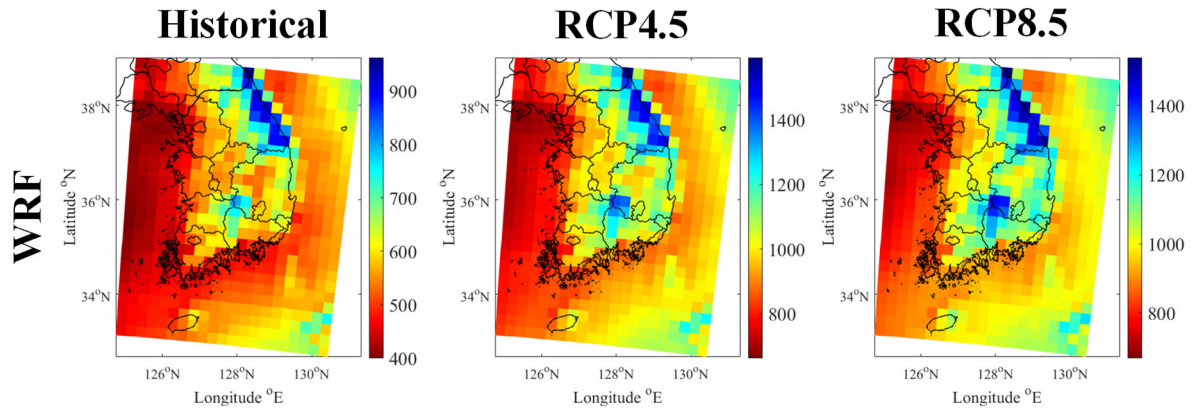
[26]

Figure 2. Gamma parameters estimated from observed precipitation at weather stations (left panel) and gridded parameters (12.5km × 12.5km) obtained from the Bayesian Kriging approach (median estimates from the predictive posterior distribution). The estimates refer to the period 1979 – 2005
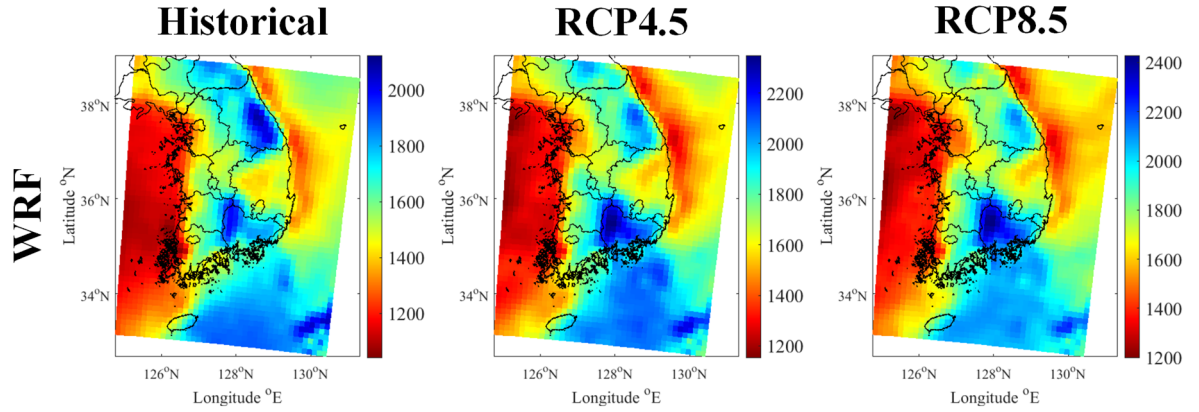
Figure 3. (a) Comparison of semivariogram estimated from the directly interpolated gamma parameters via the Bayesian Kriging approach (red filled circle) and the baseline model based on interpolated daily rainfall (blue filled circle). The semivariogram obtained from the local estimates of the gamma parameters is shown as black-filled circles. The reference month is August. (b) Mean annual precipitation compiled from WRF (25km×25km) without bias-correction. (c) Spatially downscaled mean annual precipitation (12.5km×12.5km) from WRF (25km×25km) through the proposed Bayesian SD-QDM approach. Here, the historical and the future daily precipitation are simulated by WRF in the CORDEX-EA Phase 2 for 1979-2005 ("Historical") and 2006-2100 ("Future") under the RCP 4.5 and 8.5