https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

Hydrology and Earth System Sciences
Discussions

Stochastic Generation of Multisite Streamflow for Future Water Resources Vulnerability Assessments: Application over South Korea

Sukwang Ji[1] and Kuk-Hyun Ahn[2]

[1] Graduate Research Assistant, Department of Civil and Environmental Engineering, Kongju National University, Cheon-an, South Korea; e-mail: wltnrhkd123@gm.kongju.ac.kr

[2] Assistant Professor, Department of Civil and Environmental Engineering, Kongju National University, Cheon-an, South Korea; *Corresponding author;* e-mail: ahnkukhyun@kongju.ac.kr

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

ABSTRACT

Stochastically generated streamflow time series are increasingly used for various water management and hazard assessment applications. The sequences provide realizations, preserving the temporal and spatial characteristics observed in the historic data. However, the simulations are further desirable to represent nonstationarity to account for past and future interannual oscillations. This study proposes an approach for stochastically generating future multisite daily streamflow to evaluate future water security conditioned on a national-wide relationship between annual daily maximum temperature and annual streamflow. The approach is attractive since it can avoid limitations and uncertainties introduced during realization and bias correction processes for climate model-based rainfall information. Alternatively, this approach relies on high projection skills of temperature variability. While the approach is developed by coupling annual and daily simulations, it includes (1) a wavelet decomposition-based autoregressive simulation to impose the signal of regional climate covariate; (2) clustering-based spatial pattern recognition and simulation; and (3) block bootstrapping and vine copula-based simulation for multisite streamflow simulation. The approach is applied as an example to multiple basins in South Korea. Results show that the generated sequences properly preserve many of the historical characteristics across basins. For future streamflow simulations, significant decreases in streamflow are projected, likely resulting in nontrivial impacts on regional water security. Finally, we conclude with a discussion of possible improvements to further refine the approach.

Keywords: Streamflow simulations, Nonstationary-based simulation, Future streamflow generation, and South Korea.

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

1. Introduction

Water security is facing uncertainty in the near future but is increasingly perceived as a major threat to society and the economy (Grey et al., 2013; Wheater, 2015). Together with socio-economic and land-use changes, water-related losses in a warmer climate are projected to increase around the world (Klein et al., 2014). To recognize the possible threat, water planners need future hydrological scenarios that are utilized to evaluate the robustness of water resource systems and infrastructure. However, there is no universal procedure to generate the scenarios, which is still in need of further investigation.

The most common approach for projecting how future climate conditions affect local water security is using climate simulations (e.g., precipitation and evapotranspiration) of general climate models (GCMs) under future greenhouse gas emission scenarios. Projections from these numerical models over multidecadal timescales offer climate scenarios that can be utilized to investigate the impacts of anthropogenic climate change on hydrologic responses (Karlsson et al., 2016; Van Huijgevoort et al., 2014). While these impacts are often associated with the changes in altered atmospheric circulations at both global and regional scales (Gao et al., 2020; Seidel et al., 2008) and water cycle systems (Kundzewicz et al., 2008), they are primarily affected by the changes in water vapor content from a warm climate (Asadieh and Krakauer, 2015; Bao et al., 2017; Prein et al., 2017).

Although GCMs are valuable tools for projecting global changes in atmospheric dynamics, their projected scenarios have often been criticized due to low reliability (Merz et al., 2014; Salvi et al., 2017; Stephens et al., 2010). Also, the fidelity of regional climate in GCMs substantially varies across seasons (Gu et al., 2015; Tabari and Willems, 2018), regions (Bock et al., 2018; Jiang et al., 2016), and climate variables (Eghdamirad et al., 2017; Strobach and Bel, 2017) due to the differences in the physical processes and numerical limitations imposed by spatial resolutions. The outputs in GCMs occasionally provide improper scenarios for the purpose of local water resources management (Blöschl and Montanari, 2010; Kiem and Verdon-Kidd, 2011). This is particularly true for the direct use of climate-modeled precipitation simulations when they are employed for hydrological impact studies (Knighton et al., 2019). The simulation is further problematic if low frequency climate oscillations (e.g., multiyear droughts) are of interest or where multidecadal rainfall variability is significantly high (Kiem et al., 2016). Even though higher-resolution models can improve some aspects of modeled climate (Kendon et al., 2017), they are offset by being computationally intensive, which is inefficient for water supply agencies.

While efforts on the improvement of climate models are continually encouraged, representation of climate change signals in climate models is not straightforward due to the inherent chaotic nature of atmospheric and oceanic processes and their interactions (Aalbers et al., 2018; Hawkins et al., 2016). Moreover, when hydrological models are utilized to transfer climate simulations into hydrologic response (e.g., flow), underestimations of extreme events are frequently observed, hindering proper interpretation of climate change signals (Ahn and Kim, 2019). Alternatively, a few studies have proposed a projection approach indirectly incorporating climate change signals for hydrological impact studies by utilizing the signals from a regional climate covariate (Kiem et al., 2021; Stagge and Moglen, 2013; Wasko and Sharma, 2017). These approaches are developed based on stochastic modeling approaches. For example, Stagge and Moglen (2013) have developed an approach for stochastically generating future streamflow using GCM-based climate indicators and found that summer flows are

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

projected to decrease, caused by a shift to shorter, more sporadic rain events. Wasko and Sharma (2017) have utilized the parameters of a Neyman-Scott rectangular pulse model conditioned on monthly average temperatures and revealed a significant reduction in the medium-sized floods that contribute a great amount to local reservoir inflows.

Stochastic modeling-based projection not only offers more samples to represent hydrological variability but also requires less computational burden for evaluating regional water system performance (Borgomeo et al., 2015; Hirsch, 1979). Hence, it is commonly employed to pursue water resources decision-making including reservoir planning (Guimarães and Santos, 2011; Vogel and Stedinger, 1988), hydroelectric system operation (Lanini et al., 2014), environmental flow strategy (Aguilar et al., 2014). The typically used stochastic modeling approaches are classified into two classes known as parametric and nonparametric models. Parametric models include autoregressive moving average (ARMA) models, fractional Gaussian noise models and wavelet-based simulation (Brunner and Gilleland, 2020; Kirsch et al., 2013; Papalexiou, 2018). Nonparametric models include kernel density estimation and bootstrapping approaches (Herman et al., 2016; Salas and Lee, 2010; Sharma et al., 1997). More recently, semi-parametric approaches that use both parametric and non-parametric modules have been proven to be useful since the advantage of each class can be combined in a relatively simple model structure.

In this study, following the recent work in Kiem et al. (2021), we employ interannual temperature variability as a regional climate covariate. The projection approach by indirectly incorporating a signal from the temperature variable is attractive for hydrological impact studies since temperature simulations have relatively high projection skills and their projections are widely available throughout worldwide regions (Johnson and Sharma, 2009; Klein et al., 2014). Also, the covariate is strongly associated with regional streamflow variability. Numerous studies have reported that changes in temperature are often linked to the intensification of storms (Barbero et al., 2018; Utsumi et al., 2011) and rainfall duration (Herath et al., 2018; Panthou et al., 2014; Wasko et al., 2015). In this line, the changes can be significantly meaningful for interannual streamflow oscillations. To be specific, years with higher (lower) temperatures have drier (wetter) moisture due to more (less) evaporation, leading to decreased (increased) streamflow (Kiem et al., 2016; Sheffield et al., 2012; Van Loon, 2015).

Summing up, this study presents a new approach for stochastically generating future daily streamflow simulations at multiple sites for water supply security assessments over South Korea. The proposed approach is semi-parametric and includes [1] a wavelet decomposition-based autoregressive simulation to impose the signal of climate change; [2] clustering-based spatial pattern recognition and simulation; and [3] a block bootstrapping and vine copula-based simulation. Based on our new finding about the strong regional relationship between temperature and streamflow over the study area, we develop future streamflow simulations by using alternative climate model outputs rather than using precipitation variable, which is in line with recent works (Farnham et al., 2018; Kiem et al., 2021; Yu et al., 2018). Previous studies have used climate covariates to simulate hydroclimate responses under climate change (Kiem et al., 2021; Steinschneider et al., 2019; Wasko and Sharma, 2017; Yu et al., 2018; Zaerpour et al., 2021). The proposed approach has three novelties when compared to previous works: [1] we explicitly characterize the daily spatial pattern in the regional streamflow network and utilize it for simulating realistic regional streamflow occurrences; [2] we model inter-annual

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

variability in simulating regional streamflow based on the signal from the regional covariates, which could eventually lead to a suitable representation of hydrologic extreme events over long-term simulation periods (Sparks et al., 2018); [3] we identify a strong association between annual daily maximum temperature and regional streamflow over the study area at annual time scales and utilize it for nonstationary-based streamflow simulations.

The remainder of this paper is organized as follows. Section 2 presents the methodology of the nonstationary stochastic streamflow simulation model that is conditioned on a regional climate covariate. Section 3 provides the application information to the study area, focusing on the twelve primary basins in South Korea. The conditional stochastic model is evaluated in Section 4. Also, changes in streamflow from climate projection scenarios are addressed in the section. Finally, this paper concludes in Section 5 with a discussion of the limitations of our approach and future research needs.

2. A Multisite Stochastic Streamflow Simulation Conditioned on Climate Covariates

The stochastic model proposed for synthetic streamflow simulations couples annual and daily simulation modules. Regional annual streamflows are generated using a wavelet autoregressive model (Kwon et al., 2007) to allow for conditioning on climate covariates. Daily multisite streamflows are generated using the semi-parametric model akin to multisite weather generator models in Apipattanavis et al. (2007) and Steinschneider et al. (2019). For post-processing, daily simulations are reconstructed based on the realizations of annual streamflow. Figure 1 gives an overview of the input, modules and the simulation step, which are demonstrated in more detail in the following sub-sections.

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

2.1 Module I: Regional annual streamflow generation

Consider $\tilde{q}_t$ with time, $t = 1, \ldots, T$ represents a time series of annual regional-averaged streamflow. This time series is decomposed into K orthogonal component series $z_{k,t}$ that inform different frequency signals and a residual component ε_t.

$$\tilde{q}_t = \sum_{k=1}^{K} z_{k,t} + \varepsilon_t \quad \text{Eq. (1)}$$

A simulation of $\tilde{q}_t$ is generated with time series models of each frequency component and residual noise. To simulate the signals, we consider autoregressive (AR) models with adding the vector of climate covariate (ϕ_t) while only AR models are considered for the residuals:

$$\tilde{q}_t = \sum_{k=1}^{K}\left(\sum_{i=1}^{p_k} \alpha_{k,i} \times z_{k,t-i} + \beta_k \times \phi_t + \epsilon_{k,t}\right) + \sum_{i=1}^{p_r} \gamma_i \times \varepsilon_{t-i} + \zeta_t \quad \text{Eq. (2)}$$

where p_k is the order of the AR model for the kth frequency signals, p_r is the model order for the residual noises, $\alpha_{k,i}$, β_k, and γ_i are the AR model coefficients. $\epsilon_{k,t}$ and ζ_t are independent, and identically-distributed, white noise processes. Wavelet decomposition is used to generate the frequency component and noise term in equation (2). Also, in this study, ϕ_t is transformed to be approximately normally distributed using the Box-Cox transformation (Box and Cox, 1964) before being employed in equation (2). The decomposed time series are then summed together to synthesize a time series of regional annual streamflow $\tilde{q}_t$. Also, a variance correction factor is applied in $\tilde{q}_t$, following Nowak et al. (2011). Similar to Ahn (2020), this study simply utilizes a first order AR model for the orders of p_k and p_r although the AR orders can be determined using the penalized likelihood function (e.g., Bayesian information criterion (BIC) (Schwarz, 1978)). A more thorough exposition of the theoretical background of the

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

wavelet transformation approach can be found in Kwon et al. (2007) and Torrence and Compo (1998).

2.2 Module II: Multisite daily streamflow generation

2.2.1 Identification and generation of spatial pattern in streamflow

This study first determines spatial occurrence patterns (*s*) of daily streamflow over the study area using the self-organizing map (SOM; Kohonen, 1990). SOMs are neural network algorithms that utilize unsupervised classification to perform nonlinear mapping of high-dimensional datasets onto regularly arranged two-dimensional arrays referred to as SOMs (Kohonen, 1991). Here, each of the elements in the SOM array is denoted as a node. From the SOM analysis, each day is partitioned into one of the nodes (i.e., spatial output patterns). While the number of nodes is dependent on the level of detail desired in the analysis, a moderate-sized number of noises is preferred. To consider major spatial patterns, this study adopts 2×3 nodes. We also tested different grid sizes (2×2, 3×3 and 4×4) and found that 2×3 SOM most effectively captures the important heterogeneity (not shown).

Afterward, a synthetic daily time series of spatial patterns is modeled using the first-order Markov chain with a time-varying transition probability matrix ($TM_{\eta\varsigma}^{j}$) constructed in each Julian day j. To be specific, $TM_{\eta\varsigma}^{j}$ on simulation day *j* is estimated using the SOM patterns over 21 days centered on day j (i.e., s_{j-10} through s_{j+10}). Each $TM_{\eta\varsigma}$ has a size 6×6 with each coordinate showing the probability of a state occurring at day *j* and transitioning to another state at day $j + 1$. These conditional probabilities (*CP*) are computed using the following equation:

$$CP_{\eta\varsigma} = P[s_{j+1} = \eta | s_j = \varsigma] \quad \text{Eq. (3)}$$

where η and ς are the spatial patterns for the present and the next day, respectively. We generate the sequences of six spatial patterns for all the applicant period using the time-varying transitional matrices.

2.2.2 Generation of multisite streamflows conditioned on identified spatial patterns

Multisite streamflows are simulated based on the block bootstrapping technique and generated sequences of spatial patterns. Let the simulated spatial patterns from time t to n days contain the ωth ($\omega = 1, \ldots, 6$) pattern. While n substantially varies according to season, it can maintain longer than three months (presented in the Results section). To resample historical streamflows, this study confines the longest historical block day length (n^{**}) to 10 days. We thereby resample a n^*-day block of historical streamflow data that are classified into the ωth pattern, where n^* is the longest historical block length available such that $n^* \leq n^{**}$. A block is resampled from all H historical blocks of length n^* of which the central day is within a ϑ-day window of the day for simulation day j ($\vartheta = \pm 10$ day). Here, to resample a block, the H historical blocks are weighted using importance sampling based on the similarity between the streamflows on the first day of the historical blocks and the simulated streamflow in the simulation day j-1 to represent a more realistic fluctuation in the streamflow sequence. If the day length n^* of the resampled block is less than n^{**}, the remaining length $n^{**} - n^*$ is employed as a basis to resample another block for the ωth pattern, and this process is repeated until the data for the entire block of n days are resampled.

2.2.3 Multiple-dependence structure-based jittering to streamflow simulations

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

Hydrology and Earth System Sciences
Discussions

Based on the block bootstrapping described above, the multisite streamflows are generated but they are unable to simulate values outside the range of existing records. To alleviate this limitation, a vine-copula-based jittering approach is added to the daily generation model. Vine copulas are hierarchical models that describe multivariate copulas using a rich variety of bivariate copula (Aas et al., 2009). Let $u_{t,\varsigma}$ be the non-exceedance probability for simulated streamflow value (i.e., $u_{t,\varsigma} = F(q_{t,\varsigma}|\vartheta)$) from the block bootstrapping at time *t* and site ς. In this study, the non-exceedance probability is modeled on a monthly basis by using a Gamma distribution although a heavy-tailed distribution (e.g., an extended generalized Pareto distribution (Papastathopoulos and Tawn, 2013)) is more desirable. A new vector of u_t^* is generated based on the values of u_t that are centered but are not equal. The perturbations are simulated by using the conditional distribution functions ($F(q_{t,\varsigma}|\upsilon)$, also known as *h* functions (Ahn, 2021), with the following recursive relationship (Aas et al., 2009):

$$h(q_{t,\varsigma}|\upsilon) := F\left(q_j|\upsilon\right) = \frac{\partial C_{ji|\upsilon}\left(F\left(q_j|\upsilon\right), F(q_i|\upsilon)\right)}{\partial F(q_i|\upsilon)} \qquad \text{Eq. (4)}$$

where υ is the streamflow vector excluding $q_{t,\varsigma}$.

This study considers the basin-wide average of streamflow as a pivot variable. To do so, vine copulas are constructed by the non-exceedance probability of the simulated vector of $q_{t,1:\varsigma,avg} = [q_{t,1}, \ldots, q_{t,\varsigma}, q_{t,avg}]$ that contains all ς sites as well as the basin-wide average of streamflow. Conditional streamflow values for all ς sites are then estimated with the pivot variable $u_{t,avg}$ by using the inverse form of the conditional distribution function (i.e., Eq. 4).

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

This conditional simulation is substantially attractive since it enables the modeling of a wide range of complex dependencies from the pivot variable (Joe, 2014).

The final non-exceedance probability between the values of u_t and u_t^* is determined using the following conditional probabilities (Steinschneider et al., 2019):

$$\pi = \begin{cases} \Pr\left(Q > q_{t,\varsigma}^* | Q > q_{t,\varsigma}\right) = \frac{\Pr(Q > q_{t,\varsigma}^*)}{\Pr(Q > q_{t,\varsigma})} = \frac{1-u_{t,\varsigma}^*}{1-u_{t,\varsigma}}, & q_{t,\varsigma}^* > q_{t,\varsigma} \\ \Pr\left(Q \le q_{t,\varsigma}^* | Q \le q_{t,\varsigma}\right) = \frac{Pr(Q \le q_{t,\varsigma}^*)}{Pr(Q \le q_{t,\varsigma})} = \frac{u_{t,\varsigma}^*}{u_{t,\varsigma}}, & q_{t,\varsigma}^* \le q_{t,\varsigma} \end{cases} \quad \text{Eq. (5)}$$

$$u_t^{Final} = \begin{cases} u_{t,\varsigma}^* & \pi \le r \\ u_{t,\varsigma} & \pi \le r \end{cases} \quad \text{Eq. (6)}$$

where Q is the daily streamflow variable and r is a random sampling from a uniform distribution between 0 and 1. Finally, the simulated daily streamflow value is back-transformed to $F^{-1}(u_t^{Final}|\vartheta)$.

2.3 Module III: Coupling annual and daily simulations

To rearrange the daily streamflow simulations conditioned on the annual scale, annual regional-averaged streamflow simulation is employed to generate new daily simulation data for each simulation year that is comprised of the resampling of the generated daily simulation. Daily simulation is iteratively fit to each annual simulation and rearranged for a given simulation year. This procedure follows four steps: [1] Simulate a time series of annual, regional-averaged streamflow for the desirable length of year N_{annual} using Module I; [2] Simulate a time series of daily multisite streamflow for the desirable length of year N_{daily} using Module II. Here,

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

N_{daily} is set to be sufficiently greater than N_{annual} to ensure diversity in the final simulation set; [3] For a simulation year, determine the Euclidean distances $D_{t_{annual}} = \sqrt{(\tilde{q}_{t_{annual}} - \underline{q_{t_{annual}}})^2}$ between the annual simulation $\tilde{q}$ and the vector of annual regional-averaged $\underline{q}$ acquired by the multisite daily simulation; [4] Determine the 1-year daily sequences corresponding to the smallest distance for the target simulation year and exclude the 1-year daily simulation sequences from the daily streamflow data set (simulated in step [2]); [5] Repeat steps [3]-[5] for all simulation years of length N_{annual}.

3. Application to multiple watersheds over South Korea

3.1 Study area and data

The proposed nonstationary stochastic model for streamflow simulations is applied to the twelve basins in South Korea (Figure 2). The basins have four distinct seasons with a climate that is affected by the northeastern Asian and the western Pacific Ocean (Alcantara and Ahn, 2021). The basins receive two-thirds of their annual precipitation during summer (June-September) and occasionally experience water deficits during the remaining seasons (Cha et al., 2011). Flooding events in summer are often generated by extraordinarily high rainfall induced by typhoons passing close to or penetrating the regions in South Korea (Alcantara and Ahn, 2020). Although the basins have sufficient annual rainfall similar to other regions over South Korea, variability in intra- and interannual rainfall periodically causes floods and droughts. Thus, the long-term-based streamflow records to analyze hydrologic extremes are required for reliable hydrological decision making (e.g., reservoir operation policy), leading to the pursuit of the objective of this study.

Approximately 65% of the South Korean territory consists of mountainous regions, mainly located in the eastern and northern parts while the southern and western parts of the country have well-developed plains. The heterogeneous and asymmetrical topographic features provide nontrivial impacts on the study basins. For example, the river reaches of the five basins (BS1, BS2, BS10, BS11, and BS12) are short and have steep slopes. In addition, the precipitation varies substantially from north to south. For instance, B1 receives more than 1,100 mm annually, while B7, located in the southwestern part of the country receives a mere 1,550 mm of precipitation per year.

Historic maximum temperatures were gathered from the 0.15° × 0.15° K-Hidra version 2021 product (Noh and Ahn, 2021). K-Hidra properly describes the spatial- and temporal variabilities, particularly in areas consisting of complex mountainous topography through the utilization of a massive bias correction procedure (Noh and Ahn, 2021). We used watershed boundaries (see Figure 2) to identify the K-Hidra grid cells that overlap with each watershed. If multiple grid cells were found in a watershed, their average was employed as the climate data for that watershed. For our study, the climate products from 1 January 1998 to 31 December 2020 were used, while the original K-Hidra data was 48 years long (1973~2020). Daily streamflow data of the twelve basins were obtained from the Water Resources Management Information System webpage (http://www.wamis.go.kr/) from 1 January 1998 to 31 December 2020. Note that the streamflow data for the basins (BS5 and BS9) are only available starting from 1998.

3.2 Annual daily maximum temperature as a climate covariate

Previous work has found that annual daily maximum temperature represents well the annual streamflow variability in Australia (Kiem et al., 2021). Similar to their findings, this study identified that annual daily maximum temperature has a strong relationship with regional-averaged streamflow ($\tilde{q}$) over South Korea. Figure 3a shows scatter plots of regional-averaged streamflow and transformed annual daily maximum temperature while figure 3b presents correlations between transformed annual daily maximum temperature and at-site annual streamflow. For this analysis, the Box-Cox transformation is applied to the annual daily maximum temperature so that the distribution of the transformed temperature approximately follows Gaussian. The figures show a strong negative correlation (rho = -0.53 for regional-averaged streamflow with p-value < 0.001) although one basin (i.e., BS 8) exhibits insignificant positive correlation (rho = 0.18). This different relationship for BS8 may be due to the small size of its basin area, hence it is substantially affected by local factors. However, the relationship is also manageable in our model since the model is developed based on a multiple-dependence structure (see Section 2.2). Overall, the analysis supports that the annual relationship between streamflow and maximum temperature is significant over South Korea, and confirms the usefulness of temperature as a climate covariate in the proposed model.

3.3 Performance evaluation

To evaluate the performance of the stochastic simulations, we compare the observed and simulated distributional statistics, as well as the temporal and spatial characteristics. For the distributional statistics, we consider the average, standard deviation, skewness, and maximum based on daily, seasonal (winter: December-February, spring: March-May, summer: June-August, fall: September-November), regional-averaged values. For the temporal and spatial characteristics, the autocorrelation, and cross-correlation functions, and Hurst coefficient are employed to measure short- and long-term dependence for different time scales. Moreover, spatial dependencies in extremes are investigated using the F-madogram (Cooley et al., 2006). The F-madogram (F) compares the ordering of extreme events between two-time series and is expressed as follows:

$$F(d) = \frac{1}{2}E\left|F\big(Z(\varsigma + d)\big) - F(Z(\varsigma))\right| \qquad \text{Eq. (7)}$$

where $Z(\varsigma)$ are transformed to have $Fr\acute{e}chet$ margins so that $F(\varsigma) = exp(-\frac{1}{\varsigma})$, and d is the distance between a pair of basins (Ribatet, 2008).

4. Results

The stochastic model is used to simulate 200 simulations of a 23-year time series to match the length of the historical time series over the study area. We examine the model particularly focusing on (1) the recognized spatial patterns, including their transition probabilities and interannual variability; (2) the reproduced statistical characteristics for individual sites as well as regional consistencies; (3) confirming the usefulness of coupling annual and daily simulations; and (4) Exploring future climate change-informed streamflow simulations. Here, the third analysis is based on the two separate models, one using the full model (i.e., module I, II, and III) and one conditioned on the partial model with daily streamflow generation (i.e., module II) to better isolate the strengths of coupling simulations with a regional covariate.

4.1 Identified spatial patterns of streamflow

Figure 4 presents composites of daily streamflow for all days classified into each node. While Figure S1 shows the assigned node for all days, Table 1 shows the transition probability matrix

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

between nodes. Node (1,1) represents relatively dry conditions (i.e., base flow conditions) across the study area that is the most common and persistent node (see Table 1). Node (1,3) illustrates a moderate amount of streamflow along the northern part (e.g., BS1, BS2, and BS10) of South Korea, with intensive streamflow from the southwestern (e.g., BS5 and BS6). Node (2,2) shows a similar pattern, but with intensive streamflow from the southern part (e.g., BS7 and BS8). Node (2,1) is associated with streamflow that occurred moderately across the entire study area. The transition probability from node (1,3) or node (2,2) to node (2,1) is much higher than the persistence in their nodes, indicating that substantial cases in node (1,3) or node (2,2) represents flood occurrence at the forward end of a flood event under node (2,1). On the other hand, node (2,3) is associated with a flood event that is oriented farther along the coast, particularly in the northern region. It may be related to local mesoscale convective systems developed toward the southern region. Thus, it has high probabilities to transfer into node (2,1), otherwise, it persists by itself. Overall, the nodes represent two possible mechanisms (i.e., southern- and northern-oriented storm events) for flood occurrence over South Korea.

The interannual variability of node frequency is shown in Figure 5. A trend line is also presented if the p value of the slope is significant (< 0.10) based on student's t-test. South Korea has experienced a multiyear drought in 2014-2016 (Bae et al., 2019; Myoung et al., 2020). The dry period is properly represented in variations in the frequencies of the node (1,1), but is also manifested in other frequencies. Moreover, while other nodes exhibit no clear linear trend in their frequencies, node (1,1) shows a significant upward trend over the application period that is mirrored by a downward trend in nodes (1,2) and (2,1). These results indicate that streamflows in South Korea have significantly changed over time and the dry condition is expected to be more prevalent in the near future, implying that stochastic simulations over the study area are only adequate for exploring current and future hydrologic risks if nonstationary in the simulation is accounted for.

4.2 Assessing the performance of the stochastic simulations

To evaluate the performance of the stochastic simulations, the historical temperature time series is firstly employed based on the full model (i.e., module I, II, and III). Figure 6 presents observed and simulated streamflow statistics for all twelve basins, as well as for the regional-averaged performance. The 45° line indicates perfect model performance for figures in the first two rows. Overall, the results describe that the stochastic simulations properly represent the historic daily and seasonal characteristics for individual sites, including daily average and standard deviation although there are some underestimations (e.g., daily maximum and skewness) the majority of them are still within the acceptable range (i.e., 95% confidence level). Similarly, the regional-averaged daily statistics are also compared. In general, the results suggest good performance for the regional statistics but the simulations slightly underestimate the skewness in October. The model exhibits bias with regards to maximum streamflow for August, which can be seen in the at-site statistics.

Statistical comparisons for the annual streamflows are presented in Figure 7. The average and skewness fields are well preserved on the annual scale. The standard deviation is slightly underestimated by our simulations, although we note that there is significant uncertainty in the observed values due to the small number of available annual observations. The standard deviation is underestimated for those basins with larger values. This particular discrepancy may be due to the fact that regionally-averaged streamflows are being used to drive the model over the entire country and somewhat heterogeneous study area. Also, the spread of lag-1 autocorrelation and Hurst coefficients are compared for individual sites as well as the regional-averaged streamflow. There is a negative bias in some cases (e.g., BS9 and BS12) but most biases are within the acceptable range.

We further examine the reproduction of spatial dependence across all twelve basins. Figure S2 shows the shape and magnitude of observed and simulated cross-correlation functions for daily streamflow across the six basins. We note that the results of cross-correlation functions for all twelve basins are almost identical but are not shown due to the space limitation. Also, spatial dependencies in extremes are explored using the F-madogram (Figure 8). Both results suggest that spatial dependencies are properly preserved. To be specific, Figure S2 confirms that our simulations properly capture the shape and magnitude of spatial correlations at the daily scale. Figure 8 suggests that the simulated spatial dependences in extremes suitably capture observed dependences, even though a slight overestimation is observed in the stochastic simulations. Overall, the results show that our simulations are suitable for reproducing observed temporal and spatial characteristics.

4.3 Evaluating the usefulness of coupling annual and daily simulations

To assess the usefulness of utilizing historic temperature by coupling annual and daily simulations, stochastic simulations are additionally developed conditioned on the partial model with daily streamflow generation (i.e., module II). In other words, the simulations from the partial model only contain the historical fluctuations, assuming there is no temporal changing signal. Figure 9 presents the absolute differences between the observed and simulated median results obtained by annual time series from the two separate models (the full model and partial model) for the lag-1 autocorrelation and Hurst coefficients. Smaller difference in each

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

coefficient indicates more a realistic simulation of historical variability. Although a large difference from the full model is also observed (i.e., BS6), the differences for the full model are much smaller in most cases than those for the partial model. In particular, it is true for regional-averaged streamflows (see the rightmost boxes in Figure 9a), indicating that coupling annual and daily simulations effectively replicates short- and long-term persistence in the observed time series.

The usefulness of utilizing temperature as a covariate is also investigated using the simulations in the sub-periods. Figure 10 shows annual average streamflow simulations for the first 8 years and last 8 years, respectively, using the two separate models (the full model and partial model). The first (last) period is comparable to the period of 1998 ~ 2005 (2013 ~ 2020) in the observed time series. While the full model produces results similar to the observations for each sub-period, the partial model does not properly represent the recent decreases in streamflows. To be specific, the observed annual streamflow for the last period is expected to decrease by 16.3% from the first period. While the median decrease is represented from the full model by 14.1%, no decrease is found from the partial model. Rather, an insignificant increase is found from the partial model due to sampling uncertainty. This analysis may be critical for the utility of streamflow simulations in regional water management. In particular, the recent multi-year (2014-2016) drought drew water managers' attention in South Korea because it was an exceptional event considering the regularly recurring flood season (June to August) every year. Our results inform that even though the partial model (i.e., stationary-based model) also employs the recent water deficit years when the model is calibrated, the drought shortage condition is not significantly dealt with since the event was a substantially exceptional case, supporting the usefulness of coupling annual and daily simulations.

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

4.4 Exploring climate change-informed streamflow scenarios

The streamflow simulations proposed in this study are designed directly conditioned on annual daily maximum temperature scenarios. This section illustrates future streamflow projections based on a 2.0 °C increase in projected annual daily maximum temperature for 30 years in the future. For this analysis, the historical annual daily maximum temperature from 1991 to 2020 is employed as the baseline temperature. The streamflow projection is compared against at-site observed streamflow with differences in observation expressed as a percent change in the empirical quantiles of the data from both scenarios (see Figure 11). Also, Figure S3 shows the streamflow projections on the annual and two seasonal scales across the 200-ensemble simulations. Here, the wet season includes four months (June-September) whereas the dry season covers the other months.

In the warming scenario, annual precipitations are decreased as expected in Figure 3a. However, when we closely analyze the results on a daily scale, the upper quantiles are notably increased whereas the remaining lower quantiles are decreased. It indicates the streamflow distributions to stretch, implying that extreme events, particularly for drought, will occur more frequently. The inference is consistent with previous studies demonstrating how the rainfall events will be altered in a warming condition (Fischer and Knutti, 2016; Lenderink and Attema, 2015). In addition, the projected changes considerably vary by basins. For example, the streamflows in BS3 and BS8 are less sensitive than others, so that the streamflows for the basins may not be significantly changed in a warming condition. On the other hand, BS1, BS2, BS10, and BS11, located in the northern parts of the country, are significantly affected by temperature changes.

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

The non-trivial changes in streamflows can affect the water security of the regional water supply system. The water supply system in South Korea consists of a network of reservoirs and weirs. Our study basins also have major reservoirs with an operating capacity of 2800 MCM (1 MCM is equal to 10^6 m^3) in total, while their streamflows are utilized as inflows for the reservoirs. To evaluate the reservoirs' performance under the projected streamflows, four reservoir systems (Daecheong, Boryeong, Buan, and Hapcheon reservoirs) in BS3, BS4, BS5, and BS9 are modeled (see Text S1 in the supporting information).

For this analysis, the common drought security metric, reliability, is used to assess the reservoir system performance using the 200 ensembles of 30 year-length future simulations (Figure 12). The reliability metric is simply defined as the success probability of the system by counting the days that the reservoir is in a "safety zone" compared to the total period. The Daecheong reservoir may expect a minimal loss in the projected performance when compared to the historical performance. This is because the reservoir is located in BS3, of which streamflow is projected to be less sensitive to warming conditions (see Figure 11). However, other reservoirs are substantially affected by warming conditions. For example, the reliability for the Hapcheon reservoir decrease from 98.51% to 95.01. To sum up, this analysis informs that future water security for many regions over our study area is considerably vulnerable in a warming condition and may need some remedies to augment water supply sources or reduce demand.

5. Conclusions

This study presents a new approach for generating multisite synthetic streamflows for water resources vulnerability assessment in the current and near-future climate change-informed conditions. The nonstationary simulation is achieved with an identified regional covariate by

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

Hydrology and Earth System Sciences Discussions

coupling annual and daily simulation models. The model has a hierarchical structure, including clustering-based spatial pattern simulation, block bootstrapping, and vine copula-based jittering to simulate multisite streamflows. In order to confirm the usefulness of our nonstationary-based model, we first used the historical temperature time series, and the results illustrated that our simulations were proper to reproduce statistical characteristics for individual sites as well as for the regional performance. The analyses were further extended for future streamflow projections by employing future temperature scenarios and we confirmed substantial reductions in streamflow, which could be critical in regional water security.

While the approach builds on some previously proposed methods in the stochastic hydrology field, it also has several significant contributions to the existing approaches. First, our approach enables the modeling of a wide range of complex multisite dependencies by adopting the spatial pattern recognition-based modeling and vine-copula-based jittering approach. To our best knowledge, this is the first streamflow simulation approach emphasizing multiple spatial dependence using those techniques. Second, compared to climate model-based projections, our simulated streamflows properly reproduce the primary characteristics observed in historical records. The validity is essential, particularly in evaluating the risk of water supply under water deficit conditions (Ahn, 2020; Johnson and Sharma, 2009). Third, our simulation can account for nonstationary alterations (e.g., periodic oscillation and monotonic trend) in the historical record. It is strongly beneficial since streamflow records are often limited to the recent period in many regions over the world including our study area. This limited period can be extended to the historic period in which climate records were measured. In many regions, the climate-recorded period is much longer than the period for streamflow records. Accordingly, the simulation is useful to reconstruct historical streamflow scenarios for the climate-recorded

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

period. Lastly, our simulation can take advantage of the high reliability of temperature change for future projection when compared to projected precipitation in climate models (Perez et al., 2014).

While the proposed nonstationary-based model has many benefits to simulate future streamflows, several relevant limitations need to be further addressed. Most importantly, embedded in the underlying model structure is the assumption that the historical association between the identified climate covariate (i.e., annual daily maximum temperature in this study) and regional streamflows will remain unchanged in the future. Further analysis would be required for the proper relationship under climate change. If nonlinear change were identified for the relationship, the nonlinear-based regressive model would be useful when future streamflows are generated. Our model is then limited in regions with insufficient historical records because the streamflow generations are ultimately rooted in historical sequences from the bootstrapping technique. Insufficient historical records make it difficult to fully represent the distribution of streamflows (Alcantara and Ahn, 2021). In addition, we modeled the non-exceedance probability for the jitters by using a Gamma distribution, often leading to an underestimation of extreme precipitation events (Papalexiou and Koutsoyiannis, 2013). Alternatively, the "heavy-tailed" distribution such as extended generalized Pareto distribution could be employed in future studies.

There are several opportunities to improve the simulation model. For example, this study only uses a daily scale, which is a time scale widely used but could neglect detail processes occurring over shorter time scales particularly relevant to the evolution of flooding events. Flooding in many basins over South Korea is sensitive to rainfall characteristics occurring on sub-daily scales (Lee et al., 2021). The recognized spatial pattern in this work may have less effect on sub-daily extremes, which are often driven by localized rainfall events. It is worth expanding the scope of our model to consider 6-hourly to hourly data to determine if additional streamflow patterns can be found at those time scales. Also, our analysis demonstrates two possible pathways for flood occurrence over South Korea, but the mechanisms may suffer from insufficient historical records of past events. Further, increasing the number of SOM nodes could more accurately represent localized spatial patterns of streamflow and extremes but prior to the analysis, streamflow data with long-term records from the diverse basins than the twelve basins are preferentially required, which is an obstacle for the current analysis over the study area. Lastly, this study explores future streamflow simulations based on changes in a regional climate covariate. However, exploration of other changes is also feasible from our model. For example, a certain type of flooding-induced mechanism could change under a warming climate. Given the expected alteration, decision-makers may take a certain action that could emphasize the effects of the mechanism to make plans that are more robust. After assigning appropriate changes for the transition probability matrix to reflect the effects of the mechanism, the newly generated sequences of spatial patterns can be generated and then evaluated similar to the methodology in Alcantara and Ahn (2021).

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2019R1C1C1002438).

Reference

Aalbers, E.E., Lenderink, G., van Meijgaard, E., van den Hurk, B.J., 2018. Local-scale changes

in mean and heavy precipitation in Western Europe, climate change or internal variability? Clim. Dyn. 50, 4745–4766.
Aas, K., Czado, C., Frigessi, A., Bakken, H., 2009. Pair-copula constructions of multiple dependence. Insur. Math. Econ. 44, 182–198.
Aguilar, C., Egüen, M., Perales, J.M., Losada, M.A., Polo, M.J., 2014. Stochastic Assessment of environmental flows in semiarid environments.
Ahn, K.-H., 2021. Streamflow estimation at partially gaged sites using multiple-dependence conditions via vine copulas. Hydrol. Earth Syst. Sci. 25, 4319–4333.
Ahn, K.-H., 2020. Coupled annual and daily multivariate and multisite stochastic weather generator to preserve low-and high-frequency variability to assess climate vulnerability. J. Hydrol. 581, 124443.
Ahn, K.-H., Kim, Y.-O., 2019. Incorporating Climate Model Similarities and Hydrologic Error Models to Quantify Climate Change Impacts on Future Riverine Flood Risk. J. Hydrol.
Alcantara, A.L., Ahn, K.-H., 2021. Future flood riverine risk analysis considering the heterogeneous impacts from tropical cyclone and non-tropical cyclone rainfalls: Application to daily flows in the Nam River Basin, South Korea. Adv. Water Resour. 103983.
Alcantara, A.L., Ahn, K.-H., 2020. Probability Distribution and Characterization of Daily Precipitation Related to Tropical Cyclones over the Korean Peninsula. Water 12, 1214.
Apipattanavis, S., Podestá, G., Rajagopalan, B., Katz, R.W., 2007. A semiparametric multivariate and multisite weather generator. Water Resour. Res. 43.
Asadieh, B., Krakauer, N.Y., 2015. Global trends in extreme precipitation: climate models versus observations. Hydrol. Earth Syst. Sci. 19, 877–891.
Bae, H., Ji, H., Lim, Y.-J., Ryu, Y., Kim, M.-H., Kim, B.-J., 2019. Characteristics of drought propagation in South Korea: relationship between meteorological, agricultural, and hydrological droughts. Nat. Hazards 99, 1–16.
Bao, J., Sherwood, S.C., Alexander, L.V., Evans, J.P., 2017. Future increases in extreme precipitation exceed observed scaling rates. Nat. Clim. Change 7, 128–132.
Barbero, R., Westra, S., Lenderink, G., Fowler, H., 2018. Temperature-extreme precipitation scaling: A two-way causality? Int. J. Climatol. 38, e1274–e1279.
Blöschl, G., Montanari, A., 2010. Climate change impacts—throwing the dice? Hydrol. Process. Int. J. 24, 374–381.
Bock, A.R., Farmer, W.H., Hay, L.E., 2018. Quantifying uncertainty in simulated streamflow and runoff from a continental-scale monthly water balance model. Adv. Water Resour. 122, 166–175.
Borgomeo, E., Farmer, C.L., Hall, J.W., 2015. Numerical rivers: A synthetic streamflow generator for water resources vulnerability assessments. Water Resour. Res. 51, 5382–5405.
Box, G.E., Cox, D.R., 1964. An analysis of transformations. J. R. Stat. Soc. Ser. B Methodol. 211–252.
Brunner, M.I., Gilleland, E., 2020. Stochastic simulation of streamflow and spatial extremes: a continuous, wavelet-based approach. Hydrol. Earth Syst. Sci. 24, 3967–3982.
Cha, Y.J., Shim, M.-P., Kim, S.K., River, N., 2011. The four major rivers restoration project, in: UN-Water International Conference, Zaragoza. pp. 3–5.
Cooley, D., Naveau, P., Poncet, P., 2006. Variograms for spatial max-stable random fields, in: Dependence in Probability and Statistics. Springer, pp. 373–390.
Eghdamirad, S., Johnson, F., Sharma, A., 2017. How reliable are GCM simulations for different atmospheric variables? Clim. Change 145, 237–248.

Farnham, D.J., Doss-Gollin, J., Lall, U., 2018. Regional extreme precipitation events: robust inference from credibly simulated GCM variables. Water Resour. Res. 54, 3809–3824.
Gao, T., Zhang, Q., Luo, M., 2020. Intensifying effects of El Niño events on winter precipitation extremes in southeastern China. Clim. Dyn. 54, 631–648.
Grey, D., Garrick, D., Blackmore, D., Kelman, J., Muller, M., Sadoff, C., 2013. Water security in one blue planet: twenty-first century policy challenges for science. Philos. Trans. R. Soc. Math. Phys. Eng. Sci. 371, 20120406.
Gu, H., Yu, Z., Wang, J., Wang, G., Yang, T., Ju, Q., Yang, C., Xu, F., Fan, C., 2015. Assessing CMIP5 general circulation model simulations of precipitation and temperature over China. Int. J. Climatol. 35, 2431–2440.
Guimarães, R.C., Santos, E.G., 2011. Principles of stochastic generation of hydrologic time series for reservoir planning and design: Case study. J. Hydrol. Eng. 16, 891–898.
Hawkins, E., Smith, R.S., Gregory, J.M., Stainforth, D.A., 2016. Irreducible uncertainty in near-term climate projections. Clim. Dyn. 46, 3807–3819.
Herath, S.M., Sarukkalige, R., others, 2018. Evaluation of empirical relationships between extreme rainfall and daily maximum temperature in Australia. J. Hydrol. 556, 1171–1181.
Herman, J.D., Zeff, H.B., Lamontagne, J.R., Reed, P.M., Characklis, G.W., 2016. Synthetic drought scenario generation to support bottom-up water supply vulnerability assessments. J. Water Resour. Plan. Manag. 142, 04016050.
Hirsch, R.M., 1979. Synthetic hydrology and water supply reliability. Water Resour. Res. 15, 1603–1615.
Jiang, D., Tian, Z., Lang, X., 2016. Reliability of climate models for China through the IPCC Third to Fifth Assessment Reports. Int. J. Climatol. 36, 1114–1133.
Joe, H., 2014. Dependence modeling with copulas. CRC press.
Johnson, F., Sharma, A., 2009. Measurement of GCM skill in predicting variables relevant for hydroclimatological assessments. J. Clim. 22, 4373–4382.
Karlsson, I.B., Sonnenborg, T.O., Refsgaard, J.C., Trolle, D., Børgesen, C.D., Olesen, J.E., Jeppesen, E., Jensen, K.H., 2016. Combined effects of climate models, hydrological model structures and land use scenarios on hydrological impacts of climate change. J. Hydrol. 535, 301–317.
Kendon, E.J., Ban, N., Roberts, N.M., Fowler, H.J., Roberts, M.J., Chan, S.C., Evans, J.P., Fosser, G., Wilkinson, J.M., 2017. Do convection-permitting regional climate models improve projections of future precipitation change? Bull. Am. Meteorol. Soc. 98, 79–93.
Kiem, A.S., Johnson, F., Westra, S., van Dijk, A., Evans, J.P., O'Donnell, A., Rouillard, A., Barr, C., Tyler, J., Thyer, M., others, 2016. Natural hazards in Australia: droughts. Clim. Change 139, 37–54.
Kiem, A.S., Kuczera, G., Kozarovski, P., Zhang, L., Willgoose, G., 2021. Stochastic generation of future hydroclimate using temperature as a climate change covariate. Water Resour. Res. 57, 2020WR027331.
Kiem, A.S., Verdon-Kidd, D.C., 2011. Steps toward "useful" hydroclimatic scenarios for water resource management in the Murray-Darling Basin. Water Resour. Res. 47.
Kirsch, B.R., Characklis, G.W., Zeff, H.B., 2013. Evaluating the impact of alternative hydro-climate scenarios on transfer agreements: Practical improvement for generating synthetic streamflows. J. Water Resour. Plan. Manag. 139, 396–406.
Klein, R.J., Midglev, G., Preston, B., Alam, M., Berkhout, F., Dow, K., Shaw, M., 2014. Climate change 2014: impacts, adaptation, and vulnerability. IPCC Fifth Assess. Rep.

Stockh. Swed.
Knighton, J., Pleiss, G., Carter, E., Lyon, S., Walter, M.T., Steinschneider, S., 2019. Potential Predictability of Regional Precipitation and Discharge Extremes Using Synoptic-Scale Climate Information via Machine Learning: An Evaluation for the Eastern Continental United States. J. Hydrometeorol. 20, 883–900.
Kohonen, T., 1991. Self-organizing maps: ophmization approaches, in: Artificial Neural Networks. Elsevier, pp. 981–990.
Kohonen, T., 1990. The self-organizing map. Proc. IEEE 78, 1464–1480.
Kundzewicz, Z.W., Mata, L., Arnell, N.W., Döll, P., Jimenez, B., Miller, K., Oki, T., Şen, Z., Shiklomanov, I., 2008. The implications of projected climate change for freshwater resources and their management. Hydrol. Sci. J. 53, 3–10.
Kwon, H.-H., Lall, U., Khalil, A.F., 2007. Stochastic simulation model for nonstationary time series using an autoregressive wavelet decomposition: Applications to rainfall and temperature. Water Resour. Res. 43.
Lanini, J.S., Dozier, A.Q., Furey, P.R., Kampf, S.K., 2014. Stochastic method for examining vulnerability of hydropower generation and reservoir operations to climate change: Case study of the Dworshak Reservoir in Idaho. J. Water Resour. Plan. Manag. 140, 05014004.
Lee, T., Kim, J.-S., Chen, J., 2021. Spatiotemporal characteristics and hydrological implications of downscaled hourly precipitation climate scenarios for South Korea. Int. J. Climatol.
Merz, B., Aerts, J., Arnbjerg-Nielsen, K., Baldi, M., Becker, A., Bichet, A., Blöschl, G., Bouwer, L., Brauer, A., Cioffi, F., others, 2014. Floods and climate: emerging perspectives for flood risk assessment and management.
Myoung, B., Rhee, J., Yoo, C., 2020. Long-Lead Predictions of Warm Season Droughts in South Korea Using North Atlantic SST. J. Clim. 33, 4659–4677.
Noh, G.-H., Ahn, K.-H., 2021. New gridded rainfall dataset over the Korean peninsula: Gap infilling, Reconstruction, and Validation. J. Int. Climatol.
Nowak, K.C., Rajagopalan, B., Zagona, E., 2011. Wavelet Auto-Regressive Method (WARM) for multi-site streamflow simulation of data with non-stationary spectra. J. Hydrol. 410, 1–12.
Panthou, G., Mailhot, A., Laurence, E., Talbot, G., 2014. Relationship between surface temperature and extreme rainfalls: A multi-time-scale and event-based analysis. J. Hydrometeorol. 15, 1999–2011.
Papalexiou, S.M., 2018. Unified theory for stochastic modelling of hydroclimatic processes: Preserving marginal distributions, correlation structures, and intermittency. Adv. Water Resour. 115, 234–252.
Papalexiou, S.M., Koutsoyiannis, D., 2013. Battle of extreme value distributions: A global survey on extreme daily rainfall. Water Resour. Res. 49, 187–201.
Papastathopoulos, I., Tawn, J.A., 2013. Extended generalised Pareto models for tail estimation. J. Stat. Plan. Inference 143, 131–143.
Perez, J., Menendez, M., Mendez, F.J., Losada, I.J., 2014. Evaluating the performance of CMIP3 and CMIP5 global climate models over the north-east Atlantic region. Clim. Dyn. 43, 2663–2680.
Prein, A.F., Rasmussen, R.M., Ikeda, K., Liu, C., Clark, M.P., Holland, G.J., 2017. The future intensification of hourly precipitation extremes. Nat. Clim. Change 7, 48–52.
Ribatet, M., 2008. Spatialextremes: Ar package for modelling spatial extremes. R Package Version 2–0.

Salas, J.D., Lee, T., 2010. Nonparametric simulation of single-site seasonal streamflows. J. Hydrol. Eng. 15, 284–296.
Salvi, K., Villarini, G., Vecchi, G.A., 2017. High resolution decadal precipitation predictions over the continental United States for impacts assessment. J. Hydrol. 553, 559–573.
Schwarz, G., 1978. Estimating the dimension of a model. Ann. Stat. 461–464.
Seidel, D.J., Fu, Q., Randel, W.J., Reichler, T.J., 2008. Widening of the tropical belt in a changing climate. Nat. Geosci. 1, 21–24.
Sharma, A., Tarboton, D.G., Lall, U., 1997. Streamflow simulation: A nonparametric approach. Water Resour. Res. 33, 291–308.
Sheffield, J., Wood, E.F., Roderick, M.L., 2012. Little change in global drought over the past 60 years. Nature 491, 435–438.
Sparks, N.J., Hardwick, S.R., Schmid, M., Toumi, R., 2018. IMAGE: a multivariate multi-site stochastic weather generator for European weather and climate. Stoch. Environ. Res. Risk Assess. 32, 771–784.
Stagge, J., Moglen, G., 2013. A nonparametric stochastic method for generating daily climate-adjusted streamflows. Water Resour. Res. 49, 6179–6193.
Steinschneider, S., Ray, P., Rahat, S.H., Kucharski, J., 2019. A weather-regime-based stochastic weather generator for climate vulnerability assessments of water systems in the western United States. Water Resour. Res. 55, 6923–6945.
Stephens, G.L., L'Ecuyer, T., Forbes, R., Gettelmen, A., Golaz, J.-C., Bodas-Salcedo, A., Suzuki, K., Gabriel, P., Haynes, J., 2010. Dreary state of precipitation in global models. J. Geophys. Res. Atmospheres 115.
Strobach, E., Bel, G., 2017. The contribution of internal and model variabilities to the uncertainty in CMIP5 decadal climate predictions. Clim. Dyn. 49, 3221–3235.
Tabari, H., Willems, P., 2018. Seasonally varying footprint of climate change on precipitation in the Middle East. Sci. Rep. 8, 1–10.
Torrence, C., Compo, G.P., 1998. A practical guide to wavelet analysis. Bull. Am. Meteorol. Soc. 79, 61–78.
Utsumi, N., Seto, S., Kanae, S., Maeda, E.E., Oki, T., 2011. Does higher surface temperature intensify extreme precipitation? Geophys. Res. Lett. 38.
Van Huijgevoort, M., Van Lanen, H., Teuling, A., Uijlenhoet, R., 2014. Identification of changes in hydrological drought characteristics from a multi-GCM driven ensemble constrained by observed discharge. J. Hydrol. 512, 421–434.
Van Loon, A.F., 2015. Hydrological drought explained. Wiley Interdiscip. Rev. Water 2, 359–392.
Vogel, R.M., Stedinger, J.R., 1988. The value of stochastic streamflow models in overyear reservoir design applications. Water Resour. Res. 24, 1483–1490.
Wasko, C., Sharma, A., 2017. Continuous rainfall generation for a warmer climate using observed temperature sensitivities. J. Hydrol. 544, 575–590.
Wasko, C., Sharma, A., Johnson, F., 2015. Does storm duration modulate the extreme precipitation-temperature scaling relationship? Geophys. Res. Lett. 42, 8783–8790.
Wheater, H.S., 2015. Water Security–science and management challenges. Proc. Int. Assoc. Hydrol. Sci. 366, 23–30.
Yu, Z., Miller, S., Montalto, F., Lall, U., 2018. The bridge between precipitation and temperature–Pressure Change Events: Modeling future non-stationary precipitation. J. Hydrol. 562, 346–357.
Zaerpour, M., Papalexiou, S.M., Nazemi, A., 2021. Informing Stochastic Streamflow Generation by Large-Scale Climate Indices at Single and Multiple Sites. Adv. Water

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

Resour. 104037.

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

List of Figures

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

Hydrology and Earth System Sciences
Discussions

Figure 12. Results of the reservoir system performance based on the historical (dotted line) and projected (solid line) scenarios. For the projected scenario, the median performance across the 200 members is presented for each reservoir.

List of Tables

Table 1 Transition probabilities for the six nodes in the SOM. The self-transitions are presented in bold.

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

Hydrology and Earth System Sciences
Discussions
Open Access
EGU

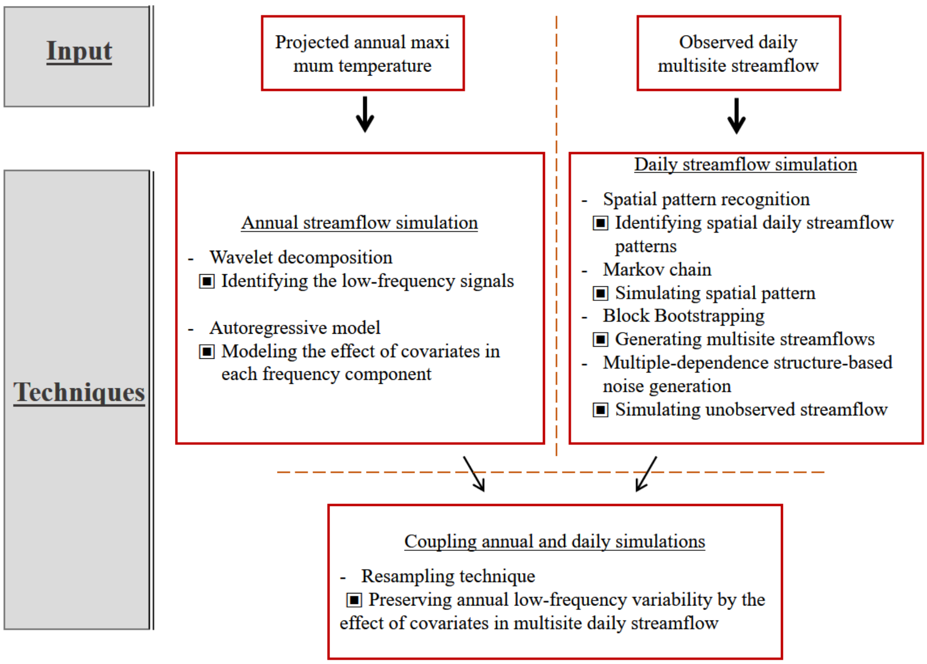

Figure 1. Schematic flowchart of the future multisite streamflow simulation.

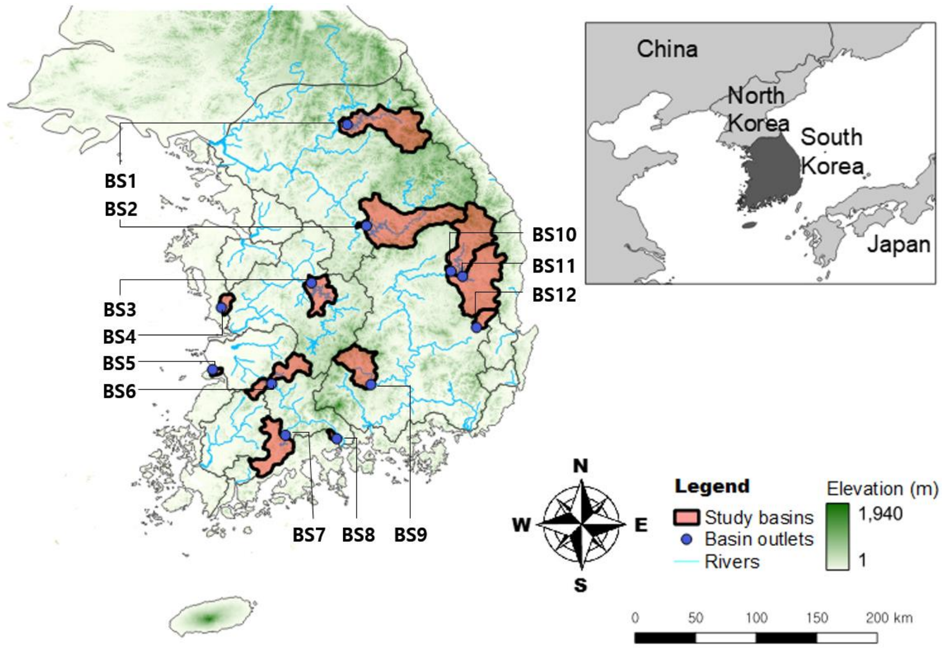

Figure 2. Map of the selected twelve basins in South Korea.

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

Hydrology and Earth System Sciences
Open Access
EGU
Discussions

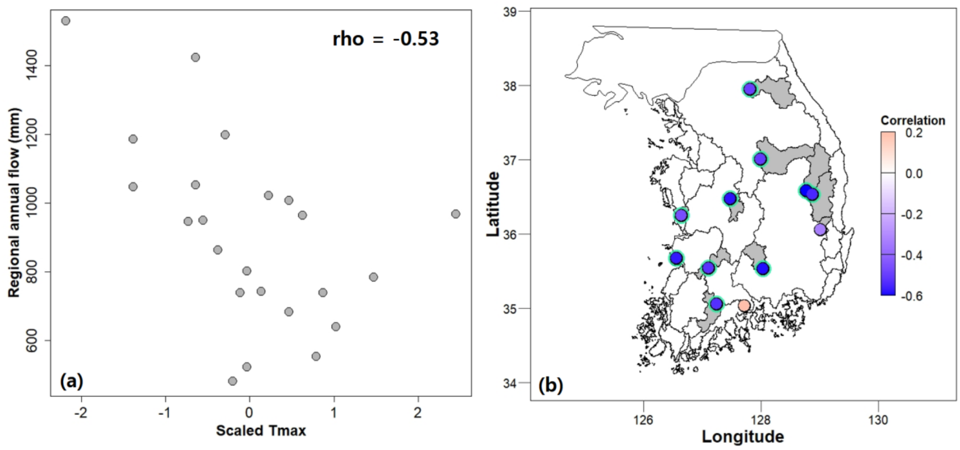

Figure 3. (a) Scatter plot between scaled annual daily maximum temperature and regional-averaged annual streamflow (mm), and (b) Pearson correlation coefficient between annual daily maximum temperature and at-site annual streamflow.

Hydrology and Earth System Sciences
Discussions

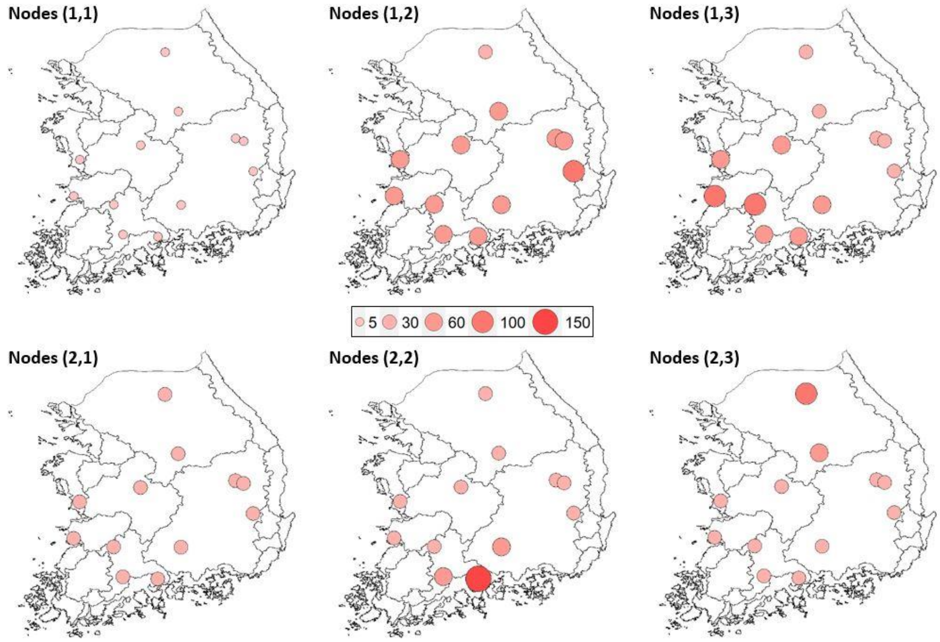

Figure 4. Map of the average daily streamflow (mm) for each of the six nodes.

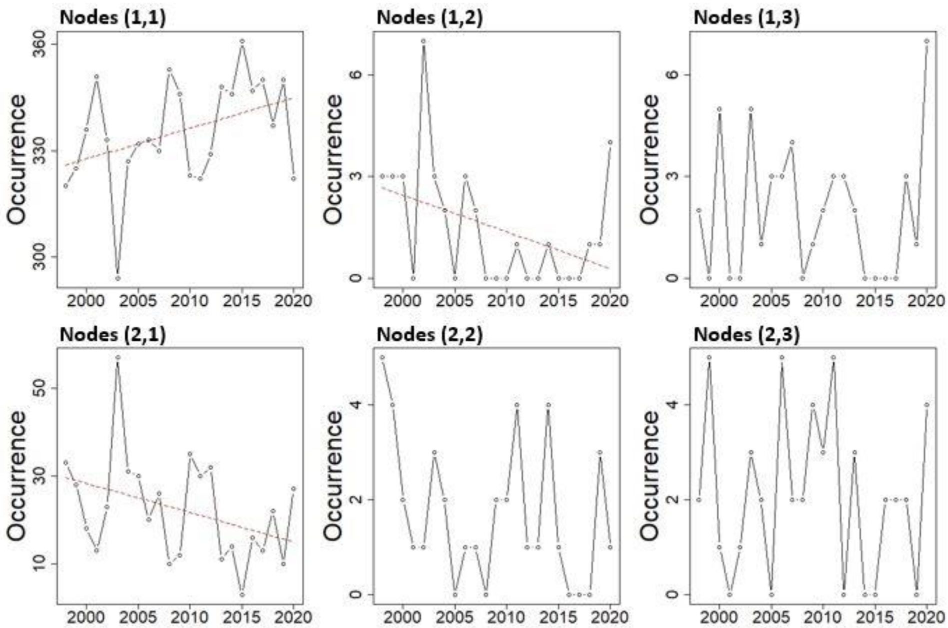

Figure 5. The number of node occurrences for each year. The trend line is shown if a Student's t-test has a *p* value < 0.10.

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

Hydrology and Earth System Sciences
Open Access
Discussions

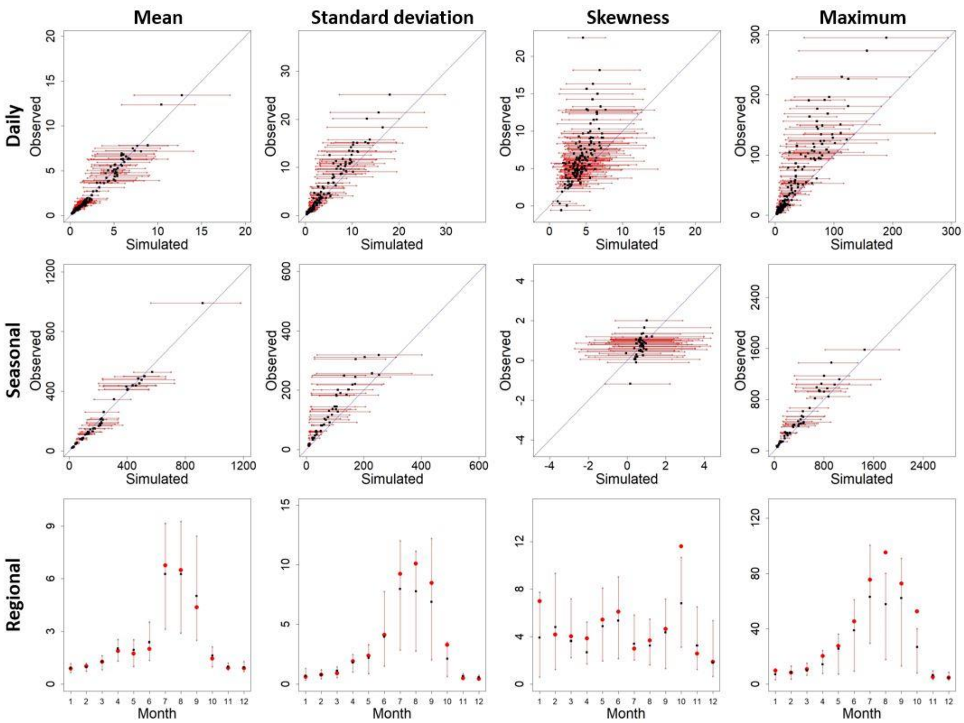

Figure 6. Observed and simulated statistics of (top) daily streamflow for all basins and months, (middle) seasonal streamflow for all basins and months, and (bottom) regional-averaged daily streamflow for each month. Here, the 95% range and the median value for simulated statistics across the 200 members are represented by whiskers and a black dot, respectively. For regional-averaged statistics, observed statistics are shown as red dots.

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

Hydrology and Earth System Sciences
Discussions

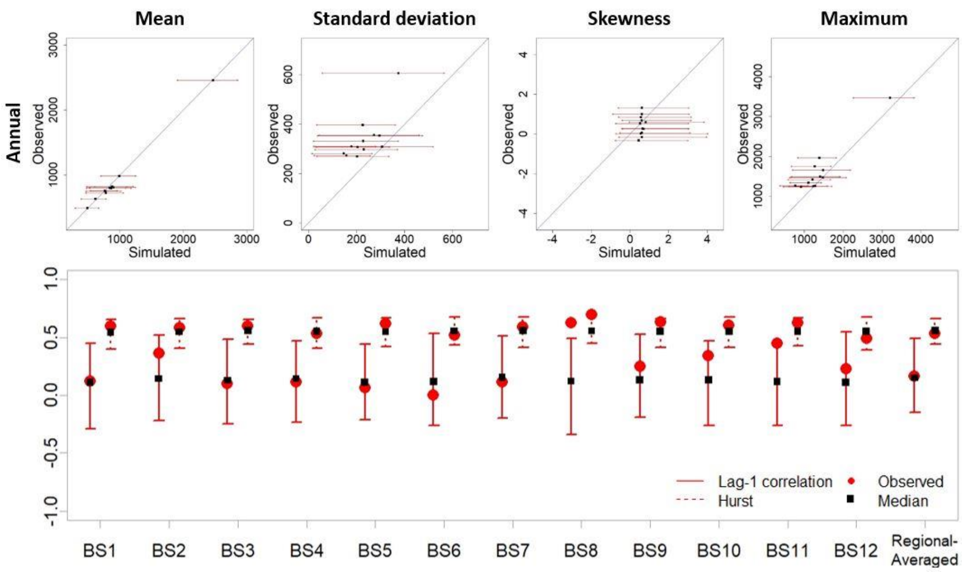

Figure 7. Observed and simulated annual statistics for all basins and regional-averaged streamflow, including the mean, standard deviation, and skewness. Also, in the bottom, lag-1 correlation (solid line) and Hurst coefficients (dotted line) are presented. The 95% range and median value for simulated statistics across the 200 members are represented by whiskers and a black dot, respectively.

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

Hydrology and Earth System Sciences
Discussions
Open Access

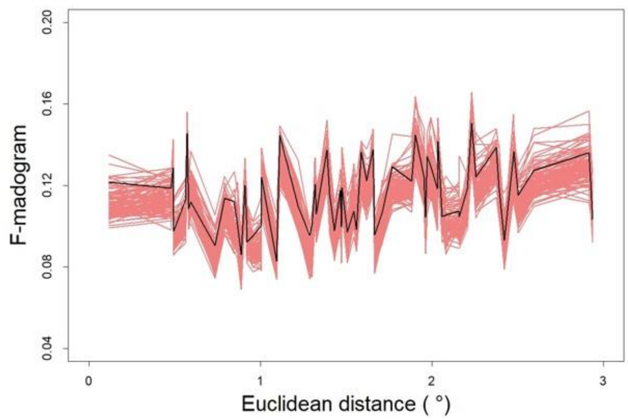

Figure 8. Observed (black) and simulated (red) F-madograms based on distances. The simulated results are obtained by using 200 different ensemble members. The lower value indicates higher dependence between a pair of basins.

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

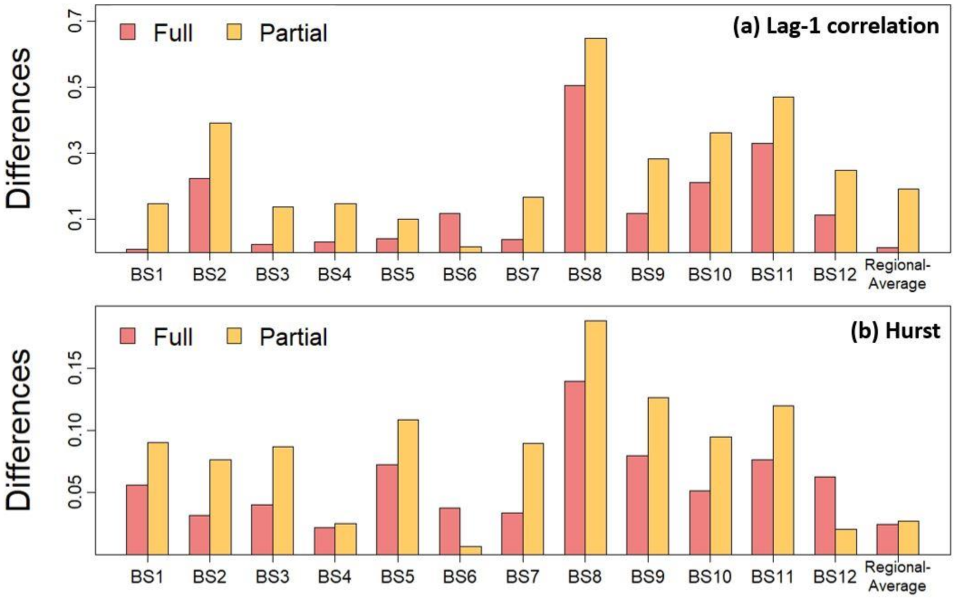

Figure 9. Absolute differences between the observed and simulated median results obtained by annually integrating time series from the two separate models (the full model and partial model) for (a) lag-1 autocorrelation and (b) Hurst coefficients.

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

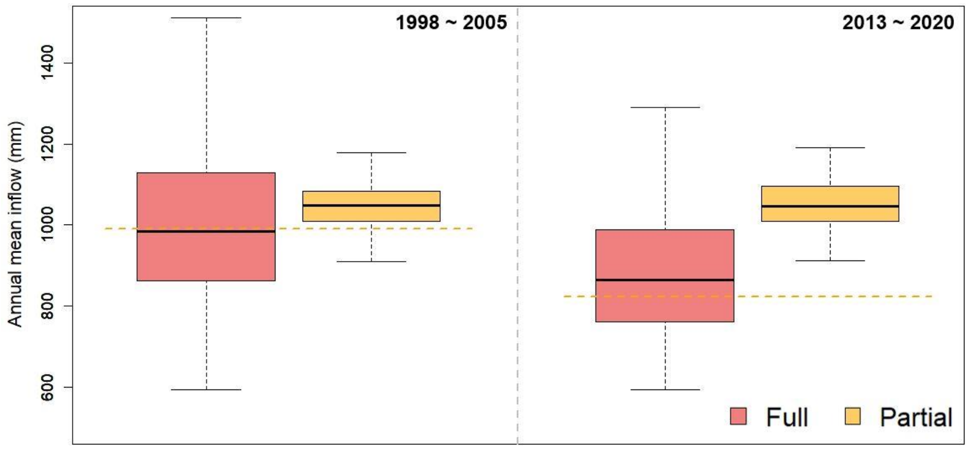

Figure 10. Distribution of regional-averaged annual streamflow for (left) the first years (1998 ~ 2005) (right) the last 8 years (2013 ~ 2020) to replicate historical simulations from the two models (the full model and partial model). The dotted line indicates the historical average for each period.

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

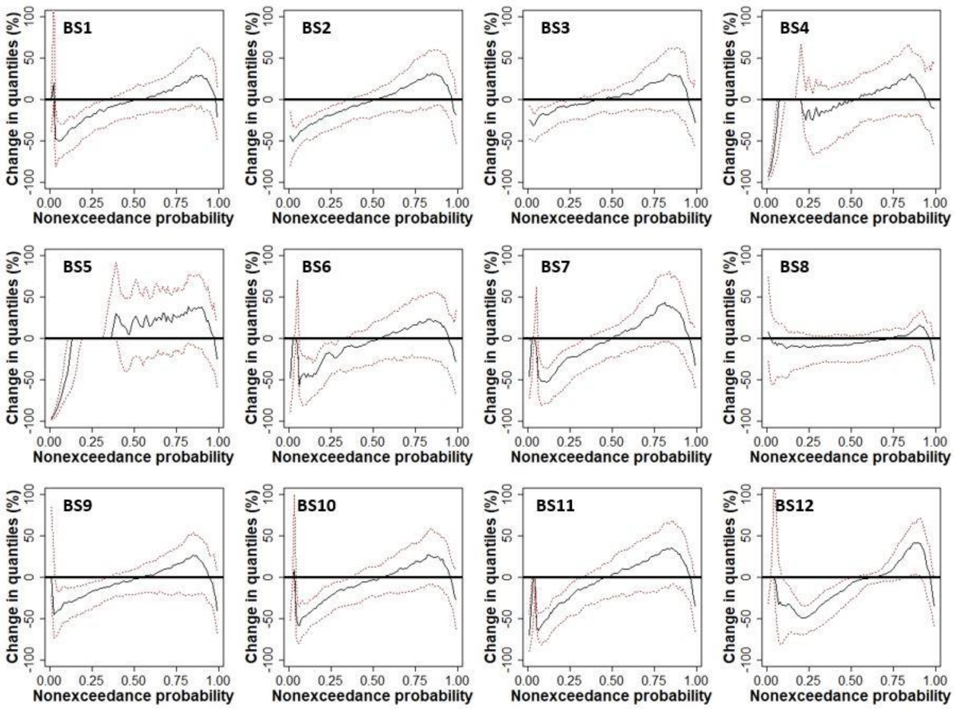

Figure 11. The change in daily streamflow distributions between projected scenarios for 2°C increase in maximum temperature and historic streamflow, presented as a percent difference from historic streamflow for each study basin. The median change (black) is illustrated along with 95% confidence intervals (red) using the 200-ensemble members.

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

Hydrology and Earth System Sciences
Discussions

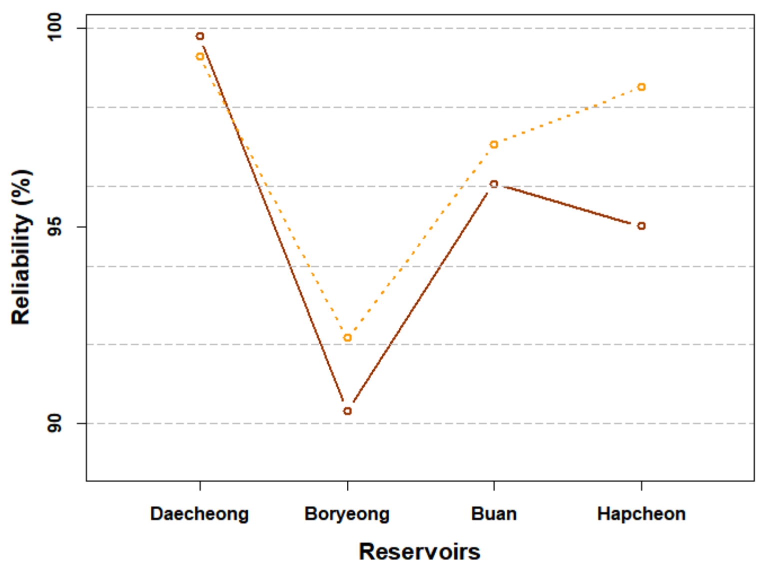

Figure 12. Results of the reservoir system performance based on the historical (dotted line) and projected (solid line) scenarios. For the projected scenario, the median performance across the 200 members is presented for each reservoir.

https://doi.org/10.5194/hess-2021-576
Preprint. Discussion started: 17 January 2022

Table 1. Transition probabilities for the six nodes in the SOM. The self-transitions are presented in bold.

	To node					
From node	(1,1)	(1,2)	(1,3)	(2,1)	(2,2)	(2,3)
(1,1)	**0.975**	0.001	0.001	0.02	0.002	0.001
(1,2)	0.029	**0.265**	0.089	0.559	0.029	0.029
(1,3)	0.000	0.111	**0.200**	0.556	0.089	0.044
(2,1)	0.369	0.012	0.039	**0.527**	0.016	0.037
(2,2)	0.026	0.154	0.077	0.564	**0.179**	0.000
(2,3)	0.063	0.063	0.063	0.416	0.020	**0.375**