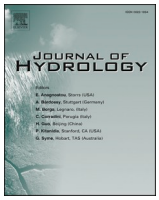
Journal of Hydrology 641 (2024) 131818

Contents lists available at ScienceDirect

Journal of Hydrology

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



Research papers

Understanding the variability of large-scale statistical downscaling methods under different climate regimes

Seon-Ho Kim, Jeongwoo Hwang *, A. Sankarasubramanian

Department of Civil, Construction, and Environmental Engineering, North Carolina State University, Raleigh, NC, USA

ARTICLE INFO

This manuscript was handled by Emmanouil Anagnostou, Editor-in-Chief, with the assistance of Yiwen Mei, Associate Editor

Keywords:
Large-scale downscaling
Comparison of downscaling
Climate regimes
Spatial uncertainty of downscaling

ABSTRACT

Large-scale downscaling plays an important role in assessing global impacts on hydrological sphere due to climate changes. In such downscaling efforts, it is essential to consider the various climate regimes. Although previous studies have indirectly suggested that the accuracy of downscaling might differ among climate regimes, research that systematically understands or quantifies the variability of this accuracy remains scarce. This study addresses this gap by systematically quantifying the performance of five different large-scale downscaling methods across various climate regimes in the context of downscaling hydroclimatic indicators. Our findings indicate that large-scale downscaling yields the highest accuracy on average when applied to temperature, precipitation, and runoff in tropical, arid, and temperate climate regimes, respectively, while showing poor accuracy in polar regimes for all variables. The maximum difference of normalized root mean squared errors for hydroclimate indicators is 69 % across climate zones, and the spatial distribution of downscaling accuracy aligns with spatial distribution of climate zones. The variation of downscaling accuracy is particularly significant in temperature, precipitation, and seasonal runoff indicators. Furthermore, linkages between accuracy of climate and hydrological indicators differ by climate zones. The underlying reasons for the different accuracy of downscaling are spatially different accuracy of global climate models (GCMs) and interaction of downscaling structure and climate regimes. This study articulated the source of spatially different accuracy/uncertainties for large-scale downscaling that have never been addressed before. The findings of this study provide valuable support in selecting appropriate downscaling methods, ultimately enhancing the spatial reliability and accuracy of large-scale downscaling methods.

1. Introduction

For decades, studies have been conducted to quantify the impacts of climate change, as it is expected to impact various climate zones (IPCC, 2007a; IPCC, 2014; IPCC, 2022). In particular, impact assessment of climate change on water resources has received significant attention due to its close relationship with various fields, such as agriculture, energy, and ecology (Kirchhoff et al., 2019; Kim et al., 2022a; Kumar et al., 2023). For such impact assessment studies, global climate models (GCMs) and hydrological models are often linked to quantify the change in runoff/streamflow (Hundecha et al., 2016). However, a common challenge encountered in linking GCMs and hydrological models is the disparity in scale between the two. This mismatch can be addressed by using a downscaling method (Fowler et al., 2007), which serves as a key component linking GCMs and hydrological models (Wilby and Wigley, 1997; Maraun, 2016).

Downscaling methods can be categorized into either dynamic or statistical approaches (Gutmann et al., 2014). Dynamic downscaling mainly relies on utilizing large-scale circulation information from GCMs as boundary conditions to downscale climate variables at a finer spatial resolution by using the conservation equations (i.e., mass, momentum and energy) within a focus area using regional climate models (RCMs). Nevertheless, the implementation of RCMs demands substantial computational resources (Eum et al., 2020), posing significant challenges when we consider multiple GCMs for analysis. On the other hand, statistical downscaling methods can be implemented quickly with limited/no computational cost and have been widely adopted in the application of GCM outputs over various sectors (Zhang et al., 2020).

Comparative studies of statistical downscaling approaches have been attempted (Chen et al., 2012). These studies are widely used to identify

* Corresponding author at: Department of Civil, Construction, and Environmental Engineering, North Carolina State University, 915 Partners Way, Raleigh, NC, USA.
E-mail address: jhwang24@ncsu.edu (J. Hwang).

https://doi.org/10.1016/j.jhydrol.2024.131818
Received 8 March 2024; Received in revised form 27 July 2024; Accepted 1 August 2024
Available online 11 August 2024

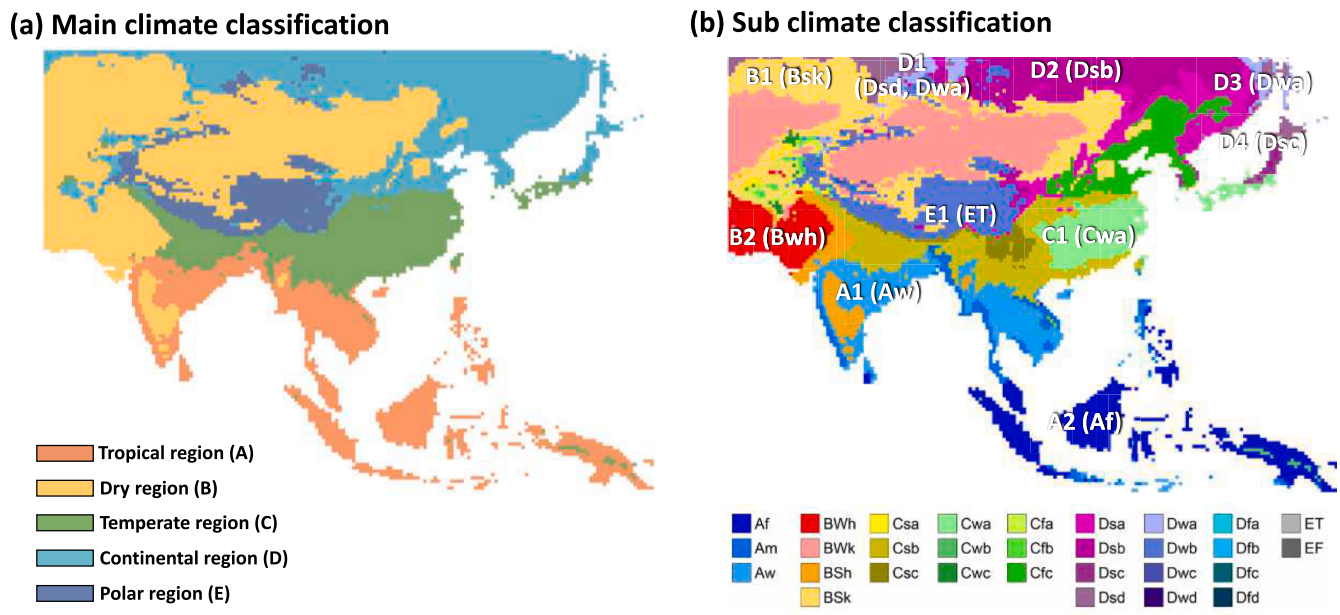


Fig. 1. The Köppen's Climate classification in the Asian Monsoon region. Fig. 1(a) ~ (b) provides Köppen's main and sub climate classification. A1 ~ D4 are spatial 3 clusters of the sub classification where significant spatial pattern of downscaling 4 accuracy arises along with climate regimes.

the appropriate downscaling methods, aiming to minimize the uncertainties from the downscaling process. Many studies have compared the reproducibility of climate indicators (Eum and Cannon, 2017; Gutiérrez et al., 2018; Gutmann et al., 2014) and hydrological indicators (Hundecha et al., 2016; Maurer et al., 2010; Teng et al., 2012) by various downscaling methods. However, the majority of these studies have focused only on a regional scale. Only a small number of studies have examined variations in downscaling performances over various regions across continents or globe (Fowler et al., 2007; Bürger et al., 2013).

When dealing with downscaling at a larger spatial scale, a diverse array of climate regimes can be encompassed within the study area. Several studies have indirectly implied that the accuracy of GCMs (Cai et al., 2009) or downscaled outputs (Fowler et al., 2007; Hundecha et al., 2016; Hou et al., 2019; Zhang et al., 2019; Kim et al., 2022b) varies across different climate regimes. However, most of these studies have primarily focused on comparing GCMs or downscaling methods, rather than understanding the varying performances of downscaled outputs across different climate regimes. Without such understanding, it would be difficult to select the appropriate downscaling method that works well for a particular hydroclimatic indicator in a given region. To the best of our knowledge, very few studies have been undertaken to quantify the variability or systematically identify the underlying cause of the varying performances of downscaling across different climate regimes.

This study aims to systematically compare the performance of large-scale downscaling methods under the different climates in Asia. The research questions addressed in this study are as follows: 1) Does the accuracy of the downscaled output vary by climate regimes? If so, which large-scale downscaling method is optimal for each climate regime? 2) Does the selection of a downscaling method for climate variables impact the accuracy of hydrological indicators to be estimated under each climate regime? 3) Which indicators/metrics of precipitation, temperature, and runoff have significant variations in accuracy under different climate regimes? 4) What are the underlying causes for differences in performance of downscaling methods under various climate regimes? The rest of the manuscript is organized as follows: The next section provides the study area and data employed in this study. Section 3 describes the methodology, including downscaling methods, hydrological modeling, hydroclimate indicators, evaluation measure, and analysis of variance method. The comparative results of different downscaling methods are presented in Section 4. Finally, the discussion and conclusion are provided in Section 5.

2. Study area and data collection

2.1. The Asian Monsoon region

The Asian Monsoon region, covering latitudes 9.75°S-54.75°N and longitudes 30.75°E-149.75°E, was selected as the study area. Given that the Asian Monsoon region is recognized as one of the climate change hotspots (Hirabayashi et al., 2021; Li and Li, 2022), it was selected as the focus of this study.

In this study, we utilize Köppen's climate classification for regional climate regime categorization. The Köppen climate classification is one of the most frequently used climate classification (Beck et al., 2018). Originally designed to empirically map vegetation (Köppen, 1936), but it has since been employed in studies focused on differences in climate regimes and assessing the impacts of climate changes (Beck et al., 2018; Kim and Bae, 2021; Son and Bae, 2015). The classification comprises five main classes (Fig. 1(a)) and 30 sub classes (Fig. 1(b)), based on thresholds and seasonality of precipitation and temperature. Asian Monsoon region includes tropical (16 %), dry (36 %), temperate (14 %), continental (26 %), and polar (8 %) climates. For clarity, we will refer to each of these climate zones as A~E, respectively, in accordance with Köppen's main classification.

2.2. Data collection

All GCMs include biases in representing predictors, and these biases are propagated to downscaling results. This propagation can significantly affect the evaluation of downscaling accuracy. Consequently, evaluating downscaling methods based on multiple GCMs with varying levels of bias complicates the derivation of generalizable information valid across all GCMs, as it is difficult to distinguish between biases originating from GCMs and those from downscaling skill (Abatzoglou and Brown, 2011). The standard practice to isolate and examine downscaling skill, regardless of bias in the predictors, is to use reanalysis data, as they are less influenced by these biases compared to other GCM outputs (Maraun et al., 2014). Therefore, evaluation of downscaling skill based on reanalysis data are likely to be more valid and applicable to

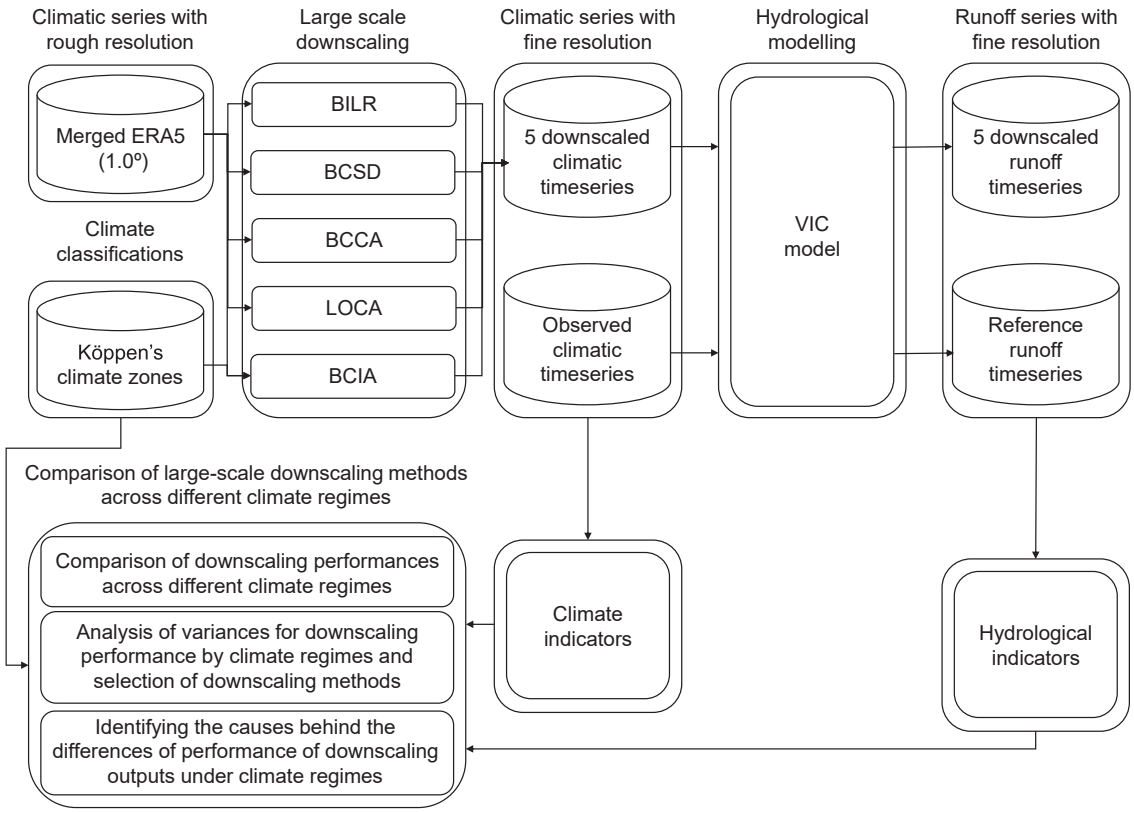


Fig. 2. Methodological framework for comparing large-scale downscaling methods 9 across different climate regimes.

other GCMs. For this reason, this study has chosen reanalysis data to represent the predictors.

The European Centre for Medium-Range Weather Forecasts reanalysis version 5 (ERA5) data is considered to represent the predictors in this study because it closely matches observed precipitation and temperature compared to GCMs, minimizing the influence of propagated biases in the large-scale predictors. ERA5 data is aggregated from the original 0.25° to 1° resolution through a simple average to align with the spatial resolution of the Coupled Model Intercomparison Project Phase 6 (CMIP6) GCMs. However, it is important to note that reanalysis data may still have biases in representing precipitation or temperatures because it involves models simulating circulations rather than directly focusing on precipitation or temperature. We consider regridded observed data of precipitation, maximum temperature and minimum temperature at 0.5 degree from Asian Precipitation–Highly Resolved Observational Data Integration Toward Evaluation (APHRODITE) and the Unified Gauged-Based Analysis of the Climate Prediction Center as the predictand for downscaling, so that the performance of different downscaling methods can be evaluated. We consider reanalysis and observational datasets on a daily timescale from 1979 to 2015, as Kim et al. (2022b) have shown that the five different large-scale downscaling methods considered in this study can be successfully applied to this period over the Asian Monsoon region.

3. Methodology

Fig. 2 illustrates the research framework designed to address the proposed research questions. Initially, daily precipitation, maximum temperature, and minimum temperature data from ERA5 were transformed from their original scale (0.25°) to a coarser scale (1.0°). This coarser-scale reanalysis data from ERA5 is considered to represent predictors. Subsequently, this coarse-scale reanalysis data was downscaled to a finer scale (0.5°) using five different large-scale downscaling methods, including bilinear interpolation (BILR), bias correction spatial disaggregation (BCSD), bias corrected constructed analogue (BCCA), localized constructed analogue (LOCA) and bias corrected climate informed analogue (BCIA). Utilizing climate variables downscaled by each of these five downscaling methods, as well as observed climate data at a 0.5° resolution, runoff simulations were conducted using the Variable Infiltration Capacity (VIC) model. The simulated runoff based on the observed climate data is hereafter referred to as the reference. Based on the five downscaled and one observed climatic timeseries, as well as the five downscaled and one reference runoff timeseries, various climate and hydrological indicators were calculated.

To assess the performance variability of five distinct downscaling methods across diverse climate regimes, we examined and compared the hydroclimate indicators derived through the application of these downscaling methods. We also evaluated how the use of different downscaling methods affects the accuracy of hydrological indicators. Moreover, we investigated the variability in downscaling performance with respect to both climate regimes and downscaling methods for each hydroclimate indicator. Finally, we identified the causes for difference in performance of downscaling methods under various climate regimes.

3.1. Large-scale downscaling methods

Statistical downscaling methods are typically grouped into two primary categories: model output statistics (MOS) and perfect prognosis (PP) (Maraun et al., 2010). MOS methods involve using identical variables for both predictors and predictands, allowing direct adjustments of coarse-scale GCM outputs to match fine-scale observations more closely. On the other hand, PP methods establish a dependable relationship between large-scale synoptical predictors and fine-scale predictands. While MOS exhibits significant downscaling capabilities when used with various GCMs, it may not fully account for the physical relationships between the predictors and predictands. On the other hand, PP can

Table 1
List of hydroclimate variables used for downscaling experiments.

Variable	Level	Predictor/Predictand(Sources)
Maximum temperature	Surface	Predictor (reanalysis) & predictand (observation)
Minimum temperature	Surface	Predictor (reanalysis) & predictand (observation)
Average temperature	Surface	Predictor (reanalysis)
Precipitation	Surface	Predictor (reanalysis) & predictand (observation)
Mean sea level pressure	Sea level	Predictor (reanalysis)
Geopotential height	250, 500, 700, 850, and 1000 mb pressure	Predictor (reanalysis)
Temperature	250, 500, 700, 850, and 1000 mb pressure	Predictor (reanalysis)
Wind speed-u direction	250, 500, 700, 850, and 1000 mb pressure	Predictor (reanalysis)
Wind speed-v direction	250, 500, 700, 850, and 1000 mb pressure	Predictor (reanalysis)
Specific humidity	250, 500, 700, 850, and 1000 mb pressure	Predictor (reanalysis)

consider these physical relationships, but its performance can be significantly influenced by the choice of GCMs. In this study, we employed the widely used MOS method, known for its simplicity, and the more sophisticated MOS-PP method, which combines the MOS and PP concepts (Kim et al., 2022b).

While BILR and BCSD are MOS type methods, BCCA, LOCA, and BCIA are considered MOS-PP type methods. BILR only considers spatial distance during the downscaling process, while BCSD incorporates both distance and bias correction through quantile mapping (QM). On the other hand, BCCA, LOCA, and BCIA use the Analogue method in combination with QM. Distance-based method and QM are included in the MOS concept, but Analogue is included in the PP concept. QM is a correction method grounded in cumulative probability distribution functions (CDFs) of simulation and observation. For this study, CDFs for precipitation and temperature were modeled using Gamma distribution and Gaussian distribution, respectively with maximum likelihood estimation. The Analogue method leverages historical observations to downscale climate outputs from GCMs. This involves identifying synoptical similarities between the target date and historical periods (analogue library), selecting similar dates (analogue dates) from the library, and using climate data from those analogue dates for downscaling at the target date. Degrees of synoptical similarity are estimated based on all predictors suggested in Table 1.

The BILR method stands out as the simplest approach, utilizing linear interpolation based on distance. This approach involves selecting the nearest four coarse grids to the target grid with finer resolution and applying inverse weighted interpolation based on distance. This method is not practical for large-scale downscaling due to the simple linear structure but for comparison with other methods.

BCSD employs an integrated approach combining BILR and QM (Abatzoglou and Brown, 2012). After applying BILR, QM is subsequently utilized to correct the biases in the BILR output based on the relationship between the BILR-derived output and observations. This method has a powerful ability in representing normal conditions, but it has limitation to represent predictands located close to the tails of their CDF. To address this limitation, corrections in the tails are performed by utilizing roughly interpolated or extrapolated CDFs due to the scarcity of samples available for estimating cumulative probabilities in those extreme tail sections. This approach can potentially introduce critical biases (Holthuijzen et al., 2022).

BCCA method combines Constructed Analogue (CA) and QM methods (Hidalgo et al., 2008). In this method, the coarse-scale predictand is corrected using QM and then downscaled using the CA method. The CA method, as it linearly combines climate data from multiple analogue dates, has the potential to overestimate drizzle days and smooth out extremes (Pierce et al., 2014). It also has limitations in identifying regionally appropriate analogue dates since it lacks localization. These limitations often lead to a relatively poor downscaling performance (Kim et al., 2022b).

LOCA method is a localized analogue method (Pierce et al., 2014). Study region is classified into local regions based on distance, and the analogue domain – i.e., the spatial domain for comparing regional synoptic similarities – is determined through spatial correlation of predictand. For each local region, analogue dates are identified on a coarse scale using the analogue domain. From these analogue dates, the most similar analogue date to the target date is assigned to each fine grid cell for downscaling. The degree of similarity can be estimated by calculating the distance between the predictand values of analogue dates and the bias-corrected values of the target dates at a fine scale through QM. This method can overcome the issues related to representing the number of drizzle days and extremes that arise with BCCA. However, it has limitations in classifying local regions and analogue domains; thus, the local regions have low consistency, and the analogue domain can be discontinuous.

BCIA method is another localized method suggested by Kim et al. (2022b). It has a similar structure to that of LOCA, but there is a difference in the localization and definition of the analogue domain. It adopts a climate classification for the localization and definition of analogue domains, overcomes the limitations of LOCA, and performs better in terms of climate downscaling (Kim et al., 2022b).

Except for the BILR method, all downscaling methods require a calibration period for fitting the probability distribution or library for extracting analogue dates. The data period was separated into calibration (1979 ~ 2000) and testing (2001 ~ 2015) periods following Kim et al. (2022b), which successfully applied these downscaling methods over Asian Monsoon region. All the results presented in this study are derived from the testing period.

3.2. Hydrological modeling

VIC is a global hydrological model that is widely used for hydrological analyses at continental and global scales. Bae et al. (2013) has developed a VIC model with a spatial resolution of 0.5° for the Asian Monsoon region. They calibrated model parameters for various gauged headwater basins and regionalized these parameters to other basins based on the Köppen climate classification across the Asian Monsoon region. Notably, they used the same observed meteorological forcings as this study to develop the VIC model. More detailed results about model performances can be found in Fig. S1, Table S2, and Table S3.

VIC model was selected to simulate runoff in this study. The forcing input variables for the VIC model are daily precipitation, maximum and minimum temperature, and wind speed. We conducted runoff simulations using the VIC model and climate forcings from the downscaled outputs and observations, except for the wind speed due to the lack of observed data. The reanalysis data with fine-scale was used for wind speed forcing.

3.3. Hydroclimate indicators

For the climate aspect, twenty-nine indicators from the Expert Teams on Climate Change Detection and Indices (ETCCDI), which are commonly used for detecting climate change and for evaluating downscaling methods (Eum et al., 2016; Werner and Cannon, 2016), were selected. For the hydrological aspect, we extracted thirty-three indicators from the Indicators of Hydrologic Alterations (IHAs), specifically designed for detecting flow alterations, such as changes in the monthly and annual extremes, high and low pulses, and water conditions (Mathews and Richter, 2007; Pandey et al., 2021). These sixty-two indicators were classified into three main group and sixteen subgroups, as shown in Table 2. Further detailed definitions can be found in

Table 2
List of hydroclimate indicators.

Main Group	Subgroup	Indicators
Temperature indicators	Duration of cold and hot days(Cold and hot)	- FD (frost days) - TR (tropical nights) - ID (days with ice) - GSL (growing season length) - SU (summer days)
	Duration based on quantile of temperature (T quantile)	- CSDI (cold spell duration), WSDI (warm spell duration) - TN10p, TX10p, TN50p, TX50p, TN90p, TX90p (durations when the daily maximum or minimum temperature falls below or exceeds the certain threshold based on quantiles)
	Intensity of temperature(T intensity)	- TNn, TNx, Txn, TXx (monthly minimum or maximum of daily maximum or minimum temperatures) - DTR (daily temperature range)
Precipitation indicators	Intensity of heavy precipitation (Heavy P intensity)	- Rx1day, Rx5day (monthly maximum 1-day or 5-days precipitation) - R95Ptot, R99pTOT (total precipitation when precipitation is larger than threshold quantile)
	Intensity of annual precipitation (Annual P)	- SDII (average precipitation of wet days) - PRCPTOT (annual total precipitation of wet days)
	Frequency of heavy precipitation (Heavy P frequency)	- R10 mm, R20 mm (annual number of days when daily precipitation is larger than the thresholds)
	Duration of wet and dry days (Wet and dry)	- CDD (maximum length of dry spell) - CWD (maximum length of wet spell) - WD (wet days)
Runoff indicators	Monthly runoff in MAM (MAM)	- MARF, APRF, MAYF (monthly mean runoff in March. April and May)
	Monthly runoff in JJA (JJA)	- JUNF, JULF, AUGF (monthly mean runoff in June, July and August)
	Monthly runoff in SON (SON)	- SEPF, OCTF, NOVF (monthly mean runoff in September, October, and November)
	Monthly runoff in DJF (DJF)	- DECF, JANF, FEBF (monthly mean runoff in December, January, and February)
	Magnitude and duration of low runoff (Low runoff)	- MI1F, MI3F, MI7F, MI30F, MI90F (annual minimum runoff of a day, 3 days, 7 days, 30 days and 90 days) - NOZR (number of zero-runoff days) - RABA (ratio of 7-days runoff to mean runoff)
	Magnitude of high runoff (High runoff)	- MA1F, MA3F, MA7F, MA30F, MA90F (annual maximum runoff a day, 3 days, 7 days, 30 days, and 90 days)
	Timing of annual extreme (Timing)	- TMAX, TMIN (Julian date of each annual daily maximum and minimum)
	Frequency and duration of high and low pulses (Pulse)	- NLP, DLP (number of low pulses and mean duration of low pulses within each water year), - NHP, DHP (number of high pulses and mean duration of high pulses within each water year),
	Rate and frequency of water condition changes (Changes)	- MPD, MND (rise and fall rate) - NREV (number of hydrological reversals)

Table S1 in the supplementary material.

3.4. Evaluation of downscaled outputs and analysis of variance of downscaled accuracy across climate zones

The normalized root mean squared error (NRMSE) was used to evaluate and compare the downscaling methods for each hydroclimate indicator. The NRMSE was calculated for each grid as a function of the root mean squared error and standard deviation, as shown in Eq. (1).

$$NR_i = \left(\frac{\sum_{k=1}^{n}\left(S_{i,k}-O_{i,k}\right)^2}{n}\right)^{\frac{1}{2}} \Big/ \left(\frac{\sum_{k=1}^{n}\left(O_{i,k}-\overline{O}_{i,\cdot}\right)^2}{n}\right)^{\frac{1}{2}} \quad (1)$$

where NR_i is the NRMSE of i-th grid for a certain hydroclimate indicator, k is the index of sample, n is the number of samples, $S_{i,k}$ is the indicator based on downscaled outputs, $O_{i,k}$ is the indicator based on observation, the bar over the variable and the dot(•) in subscript denotes average of the variable for the subscript term. For instance, $\overline{O}_{i,\bullet}$ represents the averaged value of observed indicator for the k term.

This study also adopted the two-way analysis of variance (ANOVA) to quantify the variance in downscaling performance attributable to climate regimes, downscaling methods, and their interactions, as reflected in the NRMSE of the downscaled outputs. ANOVA is a statistical method used to evaluate differences in the means of dependent variables classified into multiple nominal groups based on certain factors. It allows for comparing between-group variances and within-group variances to statistically test the influence of such factors (Lee et al., 2021). It is widely adopted in studies to understand the contribution of the factors to the variability of the variables of interest (Wang et al., 2020; Zhang et al., 2020) because it enables the analysis of interaction effects between factors and accommodates different sample sizes of the nominal groups in the analysis (Anaraki et al., 2021).

In this study, climate zones and downscaling methods are considered as the grouping factors, and we conducted a variance analysis of the NRMSE corresponding to each group. Based on the variances of between-groups and within-group of climate regimes and downscaling methods, the variance caused by climate regime groups (hereafter referred to as VCR) and downscaling method selection (hereafter referred to as VDM) on downscaling performance were quantified, respectively. Also, the variances caused by the interaction between climate regimes and downscaling method selection (hereafter referred to as VCD) were quantified. The VCD represents how the selection of downscaling methods can change the variations in downscaling performances across different climate regimes. For instance, if a downscaling method with limitations in representing heavy rainfall (referred to as Method A) is applied to tropical regions where heavy rainfall frequently occurs, the accuracy of Method A is expected to underperform compared to other downscaling methods in that region and compared to its performance in other climate regimes. This variability in accuracy cannot be explained solely by the climate regimes or the downscaling methods; therefore, it can be used for the 'interaction' term in ANOVA.

A detailed mathematical description of this process is provided by the equations presented in Eqs. (2) ~ (9). The total variance of

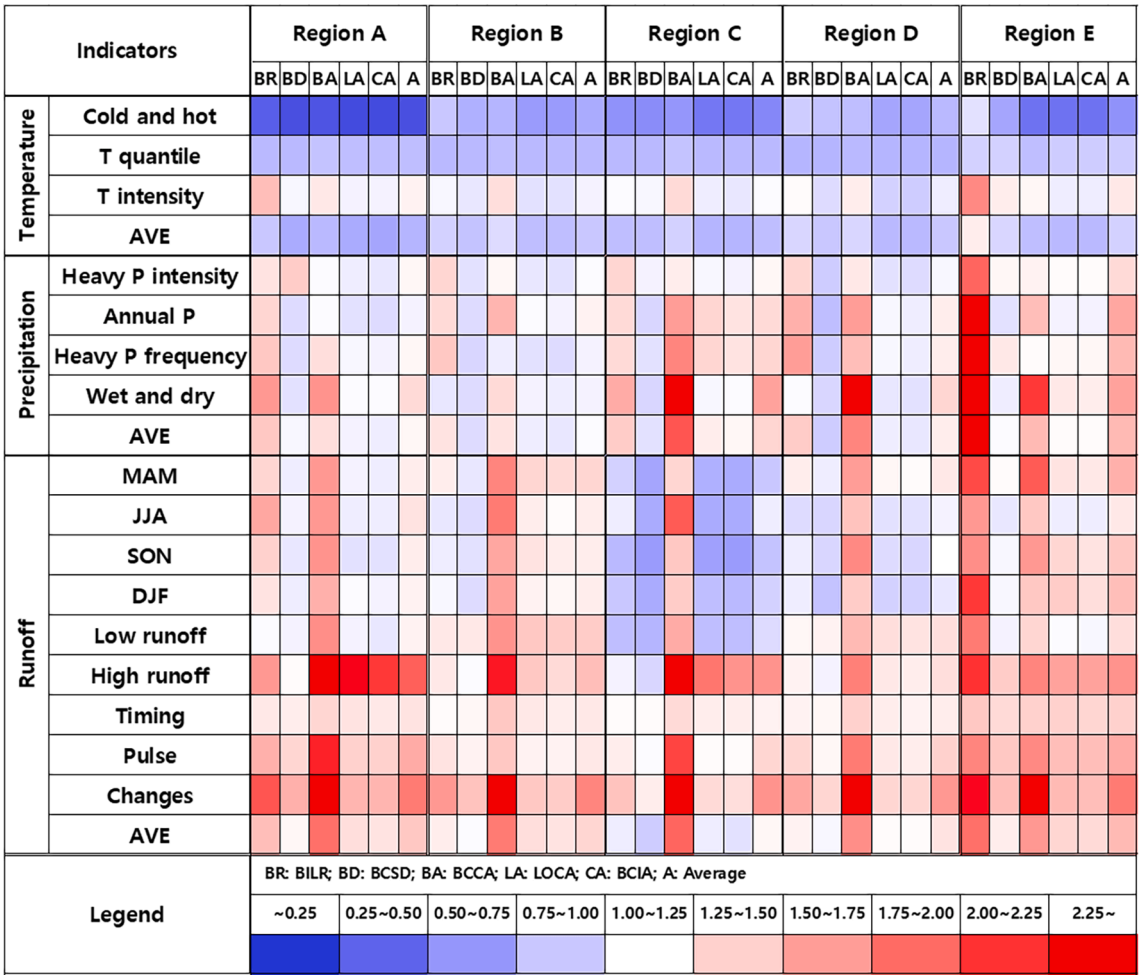


Fig. 3. Average NRMSEs of hydroclimate indicators. The color within each cell 13 represents the average NRMSEs of the corresponding indicators for the climate zone or 14 downscaling method. 'AVE' and 'A' denote the column-wise average value and the row-15 wise average value for each indicator, respectively.

dependent variables (SS_{total}) is equal to the sum of variances of between-groups and within-groups as:

$$SS_{total} = SS_{cr} + SS_{ds} + SS_{cd} + SS_{error} \tag{2}$$

where SS_{cr} is the variance of between-group among climate regimes (VCR), SS_{ds} the variance of between-group among downscaling methods (VDM), SS_{cd} the variance of interactions between climate regimes and downscaling methods (VCD), and SS_{error} the variance of within-group. This equation can be further specified as:

$$\sum_{c=1}^{nc}\sum_{d=1}^{nd}\sum_{g=1}^{ng_{c,d}}\left(y_{c,d,g}-\bar{y}_{\bullet,\bullet,\bullet}\right)^2 = \sum_{c=1}^{nc}\left(\bar{y}_{c,\bullet,\bullet}-\bar{y}_{\bullet,\bullet,\bullet}\right)^2 + \sum_{d=1}^{nd}\left(\bar{y}_{\bullet,d,\bullet}-\bar{y}_{\bullet,\bullet,\bullet}\right)^2 \tag{3}$$

$$+\sum_{c=1}^{nc}\sum_{d=1}^{nd}\left(\bar{y}_{c,d,\bullet}-\bar{y}_{c,\bullet,\bullet}-\bar{y}_{\bullet,d,\bullet}+\bar{y}_{\bullet,\bullet,\bullet}\right)^2 + \sum_{g=1}^{ng_{c,d}}\left(\bar{y}_{\bullet,\bullet,g}-\bar{y}_{\bullet,\bullet,\bullet}\right)^2$$

where $y_{c,d,g}$ is the NR_i based on dth downscaling method for a certain hydroclimate indicator. Here, the subscript i that denotes the grid in Eq. (1) is indicated by c and g, where c represents the climate zone and g represents the index of grids belonging to that particular climate zone. In Eq. (3), nc is the number of climate zones (which equals to five in this study), nd is the number of downscaling methods (which equals to five in this study), and $ng_{c,d}$ is the number of grids in climate zone c using the d-th downscaling method. The differences in the means by groups and interaction effect can be tested by using the F statistics:

$$F_{cr} = SS_{cr}/(nc-1) \tag{4}$$

$$F_{ds} = SS_{ds}/(nd-1) \tag{5}$$

$$F_{cd} = SS_{cd}/(nc-1)(nd-1) \tag{6}$$

The contributions η of each factor and their interactions to the total variance can be quantified as:

$$\eta_{cr} = SS_{cr}/SS_{total} \tag{7}$$

$$\eta_{ds} = SS_{ds}/SS_{total} \tag{8}$$

$$\eta_{cd} = SS_{cd}/SS_{total} \tag{9}$$

According to Cohen (2013), the contribution of factor can be classified as small ($\eta > 0.01$), medium ($\eta > 0.06$), and large ($\eta > 0.14$).

4. Results

4.1. Comparison of downscaling performances across different climate regimes

One of the simplest ways to assess variations in downscaling performance across different climate regimes involves comparing the average errors. As mentioned in section 2.2, we utilize the merged ERA5 to represent predictors; henceforth all mention of 'GCM' indicates the merged ERA5. The averaged NRMSEs, based on downscaling methods

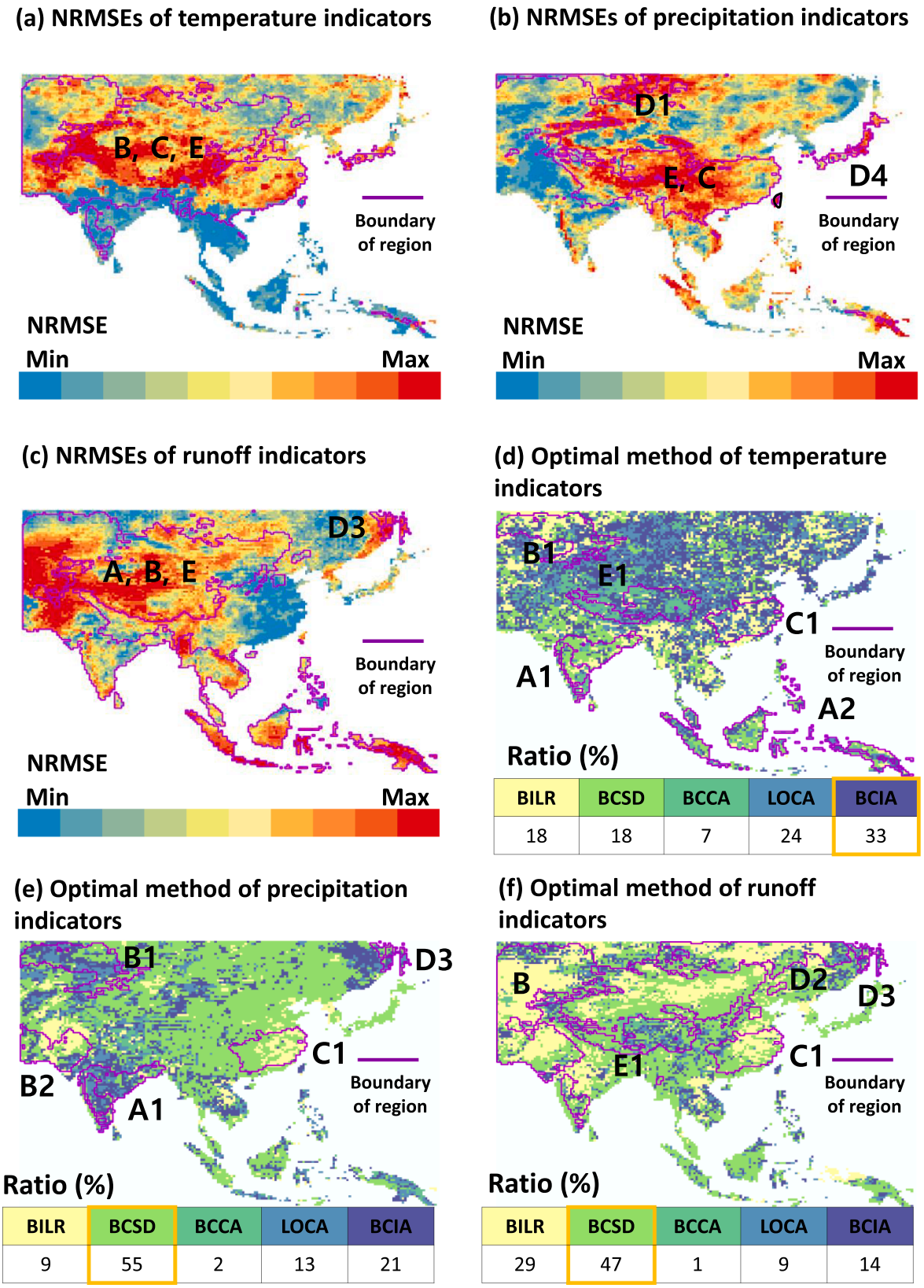


Fig. 4. Spatial distribution of NRMSEs and optimal downscaling method. In Fig. 4(a) ~ (f) Ratio (%) represents the percentage of grids where each respective downscaling 20 method demonstrates the best performance. Purple boundaries and labels are sub-21 classifications of Köppen, as shown in Fig. 1. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

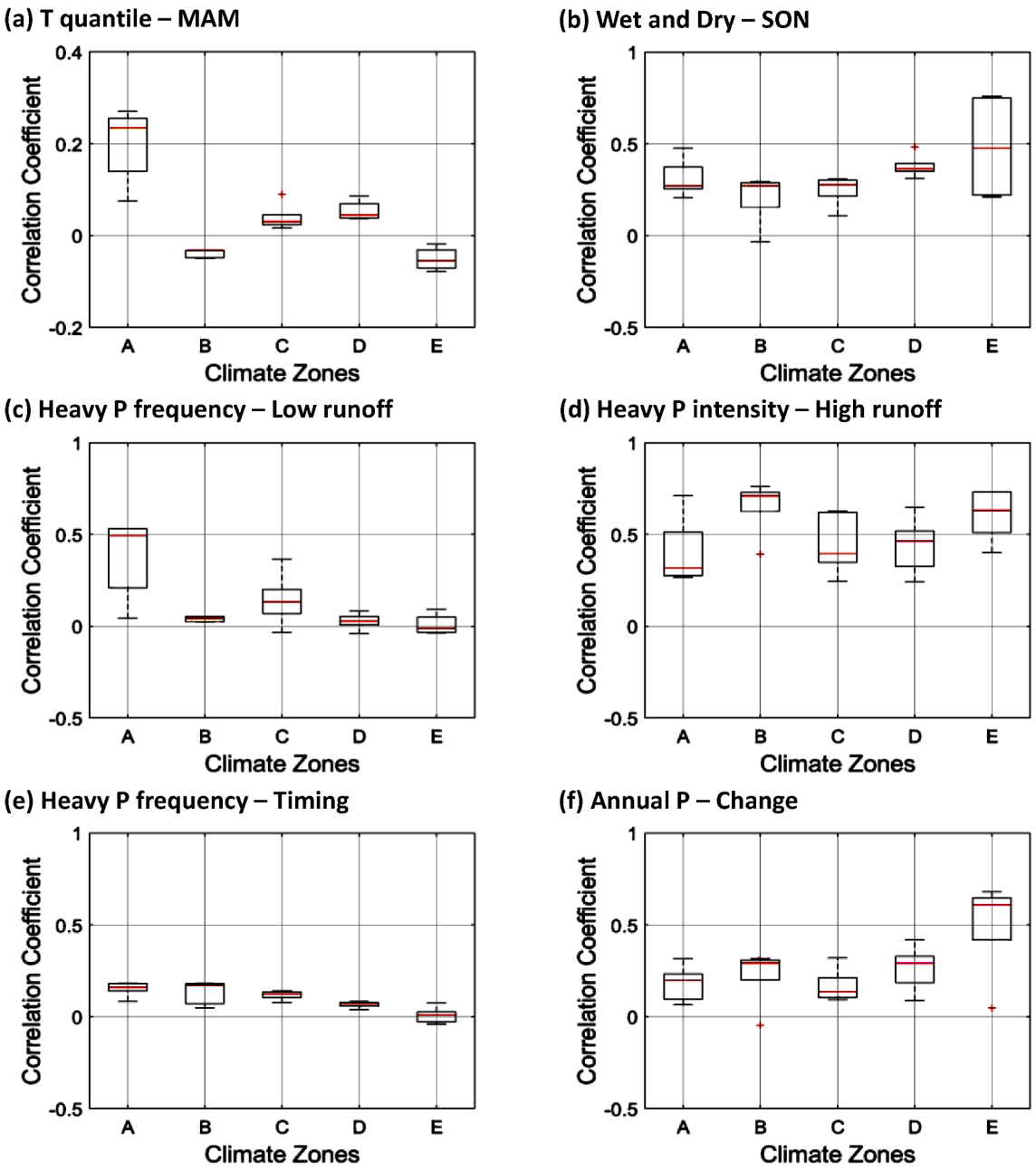


Fig. 5. Correlation Coefficients between NRMSEs of Climate and Hydrological 25 indicators. X-axis and Y-axis denote climate zones and correlation coefficients. The 26 range of the box plot means the correlation variability from different downscaling 27 methods.

and specific indicators, exhibit variation across climate regimes, with a range spanning from 7 % to 69 %. When averaging NRMSEs for the main indicator groups, identified in Table 2, and downscaling methods, the average variations for temperature, precipitation, and runoff indicators, based on climate zones, are 16 %, 28 %, and 21 %, respectively. The NRMSEs of region E show the highest values for all indicators, while regions A, B, and C, on average, demonstrate the lowest values for temperature, precipitation, and runoff indicators, respectively (Fig. 3).

Downscaling methods that exhibit the lowest NRMSEs are mostly not different across climate regimes. The BCIA shows the lowest averaged NRMSEs in all climate zones for the temperature indicators (Fig. 3). For the T quantile indicators, however, all downscaling methods show comparable NRMSEs. These similar levels of errors arise from the inability of all methods to correct the quantiles of GCM during the downscaling process. The BCSD shows the lowest averaged NRMSEs for precipitation and runoff indicators, except in region A for the precipitation indicators. However, the BCSD shows relatively high NRMSEs for the High P Intensity indicator in region A and for the High P Frequency indicator in region E.

BILR method shows relatively large errors the most due to its lack of consideration for bias correction within the scheme. Among the downscaling methods that incorporate bias correction, the BCCA shows the largest errors. It shows limitations in representing most of the precipitation indicators, as discussed in Section 3.1. The errors in precipitation indicators directly influence the High runoff and Changes in runoff indicators. The BCCA shows a relatively lower accuracy for these indicators compared to other indicators and methods. The LOCA method shows similar results with BCIA, as both methods employ the localized analogue concept.

Averaged NRMSEs across different climate zones can represent overall variation in NRMSE under different climate regimes. However, it is difficult to detect whether the average is affected by the outliers. Thus,

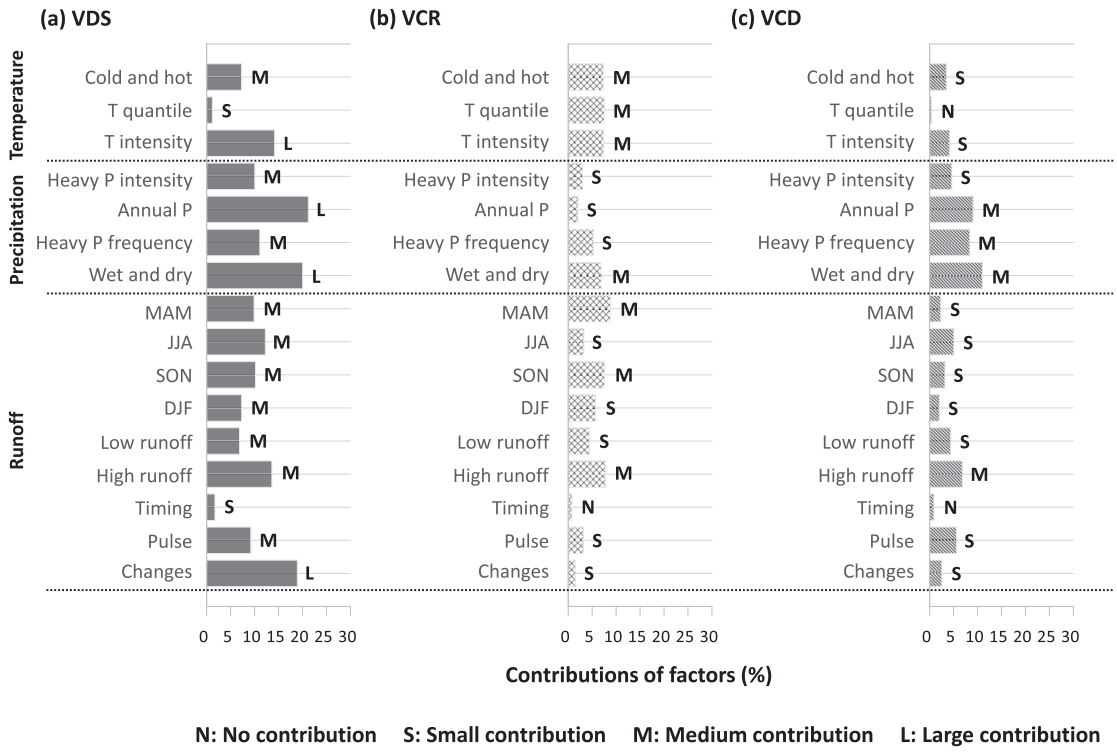


Fig. 6. Contributions of factors to variance of NRMSEs in downscaling output. (a) ~ (c) 31 indicate the contributions of downscaling selection, climate regimes, interaction between 32 climate regimes, and downscaling method, respectively. N, S, M, and L denote 33 contribution levels of Cohen (2013).

we visualize the spatial distribution of NRMSEs for all grids, as shown in Fig. 4. Additionally, we visualize the optimal downscaling methods, which have the lowest errors, to check whether there is a potential relationship between climate regimes and downscaling selections. It is noteworthy that spatial patterns of NRMSEs and optimal downscaling methods closely correspond to the spatial boundaries of Köppen's climate zone.

Spatial variations of NRMSEs are related to climate zones. Boundaries that indicate regions with high NRMSE (HNR) or low NRMSE (LNR) mostly align with climate zone boundaries. For temperature indicators, almost all HNRs are in regions B, C, and E in Fig. 4(a). For precipitation indicators, HNRs are in regions C, D1, D4, and E in Fig. 4 (b). For runoff indicators, HNRs are in regions A, B, E, and D3 within Fig. 4(c).

A single dominant downscaling method does not prevail across the entire study region. The optimal downscaling method shows spatial clusters that align with climate zone boundaries. Regarding temperature indicators, BCIA shows the lowest errors in 33 % of the total grids, while BILR shows the lowest errors in regions C1 and B1, according to Fig. 4 (d). Furthermore, BCSD shows the lowest errors in regions A1 and A2, and BCCA shows the lowest errors in region E1. For precipitation indicators, BCSD shows the lowest errors in 55 % of the total grids, while BILR shows the lowest errors in regions B2 and C1, as shown in Fig. 4(e). BCIA and LOCA show the lowest errors in regions A1, B1, and D3. As for the runoff indicators, BCSD shows the lowest errors in 47 % of the total grids, while BILR shows the lowest errors in regions B, C1, and D2 in Fig. 4(f). BCIA shows the lowest errors in regions D3 and E1.

4.2. Evaluating the impact of downscaling method selection on the accuracy of hydrological indicators

Here, we explore the correlations between the NRMSEs of climate and hydrological indicators to assess how the performance of climate downscaling affects the accuracy of the resultant hydrological variables. Higher correlation values may suggest that the errors in the downscaled climate variables are more likely to affect the downscaling accuracy of hydrological variables. Within each climate zone, these correlation values can vary depending on the downscaling methods employed. Such variation could indicate the sensitivity of the accuracy of hydrological variables to the performance of climate downscaling. The more significant the variation in correlation, the greater the indication that potential errors in hydrological indicators can be managed by selecting appropriate downscaling methods for each climate zone. While our analysis examines the relationship for each pair of climate and hydrological indicators (Fig. S2), we primarily focus on those that demonstrate significant patterns, as illustrated in Fig. 5.

The average correlation coefficient between the NRMSEs of temperature-related climate indicators and hydrological indicators across all climate zones is estimated to be 0.05. The estimated variability in these correlation coefficients, as represented by their maximum range, is 0.10. Meanwhile, the average correlation coefficient between the NRMSEs of precipitation-related climate indicators and hydrological indicators is estimated to be 0.10, with a higher estimated variability in these coefficients, represented by a maximum range of 0.24. These findings indicate that errors in precipitation-related climate indicators are more likely to impact the downscaling accuracy of hydrological indicators compared to errors in temperature-related climate indicators.

As shown in Fig. 5, regions A and E demonstrate notably high correlations between the NRMSEs of climate indicators and those of hydrological indicators, on average. These regions also display greater variability in these correlation coefficients compared to other regions. Meanwhile, in region B, the errors in hydrological indicators show a relatively low correlation with errors in both temperature-related and precipitation-related climate indicators, with overall less variability in these correlations. This suggests that the choice of downscaling methods for climate variables is likely to significantly impact the accuracy of hydrological variables in regions A and E – where the correlations and variability between the errors of climate and hydrological indicators are pronounced. Conversely, in region B, the impact may be minimal due to the lower correlations and variability between the errors in climate and

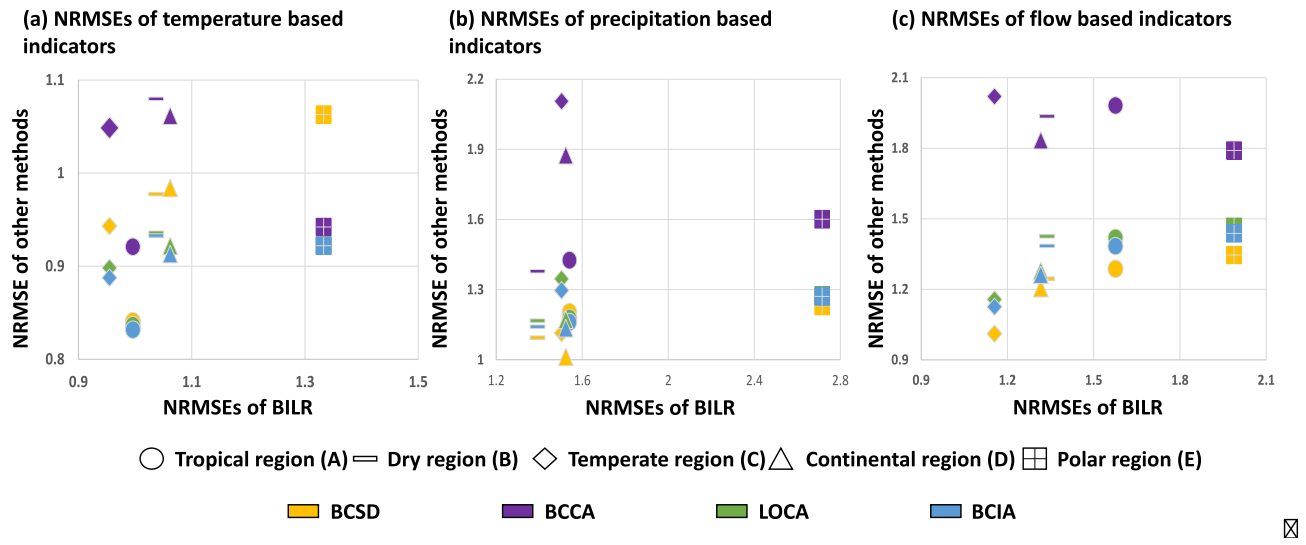


Fig. 7. Comparison of NRMSEs for hydroclimate indicators across different climate 37 zones and downscaling methods. Every symbol represents NRMSEs of BILR on x-axis 38 and the NRMSEs of other downscaling methods on y-axis.

hydrological indicators.

Furthermore, as illustrated in Fig. S1, the errors in High runoff indicator generally show high correlations with errors in climate indicators across most climate zones, accompanied by high variability in these correlations, suggesting the accuracy of High runoff indicator is closely tied to the accuracy of climate downscaling. Conversely, the errors in Timing indicator exhibit low correlations and low correlation variability in most zones, indicating it may be less affected by downscaling accuracy.

4.3. Analysis of variance for downscaling performance by climate regimes and selection of downscaling methods

ANOVA is used to quantify the variability in downscaling performance that can be explained by different climate regimes and downscaling methods. The results from ANOVA indicate that significant variations in NRMSEs of the downscaled hydroclimate indicators are attributable to climate regimes (VCR), downscaling methods (VDS), and their interactions (VCD), each demonstrating statistical significance at the 0.01 level for all indicators in F-test (Table S4). VDS, VCR, and VCD collectively explain 17.7 %, 28.2 %, and 18.4 % of the averaged variance in NRMSEs for temperature, precipitation, and runoff indicators, respectively. VDS accounts for 11.1 %, VCR for 5.5 %, and VCD for 4.9 % of the average explained variation in the NRMSEs of hydroclimate indicators. A more detailed breakdown of the contributions of VDS, VCR, and VCD to the variation in the downscaling performance for each indicator is shown in Fig. 6.

VDS effectively explains the variations in the downscaling errors for all hydroclimate indicators, with exception of the T quantile and Timing indicators. Meanwhile, VCR is highly related to the variations in the downscaling errors of all temperature-related indicators and certain hydrological indicators, including Wet and dry, MAM, SON, and High runoff indicators. VCD effectively explains the variations in the downscaling errors for High runoff and precipitation indicators except Heavy P intensity. Considering VCR and VCD, all temperature and precipitation indicators have a large variability of downscaling errors depending on climate regimes. Moreover, MAM, SON, and High runoff indicators also have large variations by climate regimes.

4.4. Potential causes behind the variance of downscaling performances under different climate regimes

We establish and validate two hypotheses to understand the sources of the variation in the downscaling performances across different climate zones. The first hypothesis posits that this variation is due to spatial variability in the performance of the GCM across the climate zones. The second hypothesis suggests that the interaction between the downscaling method and climate regimes leads to the differences in the downscaling performances across these climate zones.

To test the first hypothesis, the downscaling performance of BILR and others are compared (Fig. 7). BILR performance can represent the GCM performance on the same scale as the downscaled output since it does not involve any bias correction. This comparison allows us to discern whether the variation in the downscaling performances across different climate zones are due to the spatially variant GCM performance or not. A positive correlation between the performances of the downscaling method and GCM suggests that variations in downscaling performances are influenced by the GCM performance.

Fig. 7 illustrates a positive correlation between the downscaling errors of BILR and those achieved by other downscaling methods across different climate zones. Here, the downscaling error of each method in different climate zones is represented by the mean NRMSE specific to each respective climate zone. The positive relationship can be found in regions A, D, and E for temperature indicators, and regions A, B, and E for precipitation indicators, and in all regions for runoff indicators. In comparison to BCSD, LOCA, and BCIA, BCCA shows a weak positive relationship due to the limitations of BCCA described in subsection 3.1.

The Pearson correlation analysis results between NRMSEs of BILR and other downscaling methods show a strong positive correlation coefficient of 0.6 or higher, with significance levels ranging from 0.01 to 0.02 for all methods except BCCA. NRMSEs of BCCA also show a positive relationship with NRMSEs of BILR, but this relationship is relatively weak, as the correlation coefficient and significance level are 0.39 and 0.15, respectively. By confirming a similar performance of BILR and other downscaling methods in estimating hydroclimatic indicators, we validate our first hypothesis that the spatial variability in the GCM performance is the key factor that leads to the spatially varying performances of downscaling across different climate zones. This can account for most of VCR, as positive relationships are found in most downscaling methods.

The second hypothesis is that the downscaling performances vary by climate regimes due to interaction between climate regimes and downscaling methods (VCD). This hypothesis is substantiated when a particular downscaling method shows unusual accuracy in specific climatic zones compared to other climatic zones. We investigate the CDF of indicators with high VCD to identify the combinations of climate zones

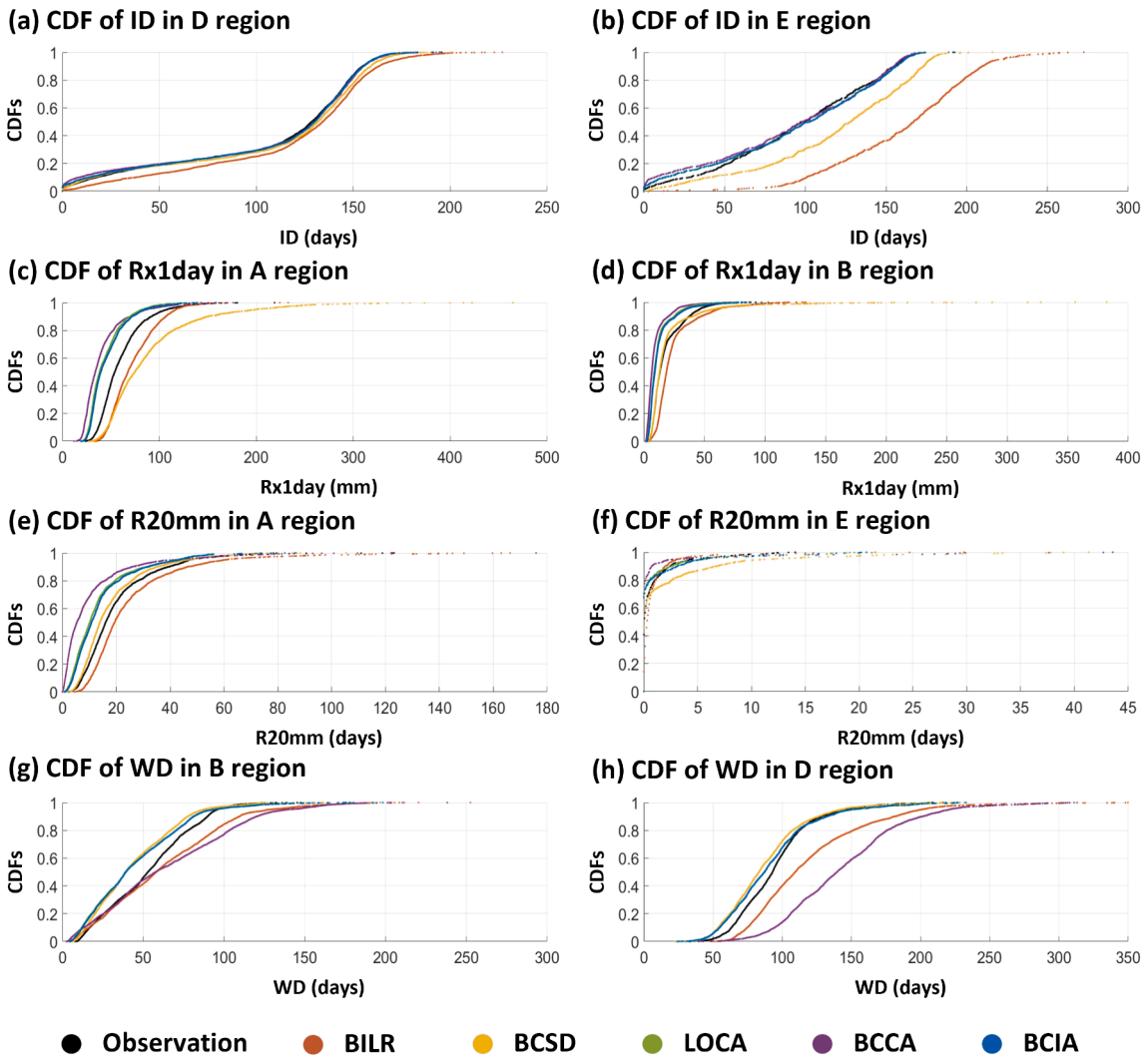


Fig. 8. Comparison of empirical cumulative distribution functions (ECDFs) for the 42 indicators where the interactions between climate regimes and downscaling methods 43 exhibit significant effects on explaining the variability of their NRMSEs. The CDFs are 44 derived from grids belonging to each climate zone.

and downscaling methods that exhibit unusual accuracy (hereafter interaction case). Fig. 8 includes CDFs of indicators that show interaction cases, such as ID, Rx1day, R20mm, and WD.

Most of the interaction cases are found in these indicators when they are derived from variables downscaled by the BCSD and BCCA methods. The BCSD method tends to result in greater errors in these indicators compared to other methods, especially when the BILR output significantly deviates from the observed values (Fig. 8a and b). The BCSD, which is based on only QM, is more sensitive to the performance of GCMs compared to other methods that involve combining QM and analogue.

BCSD exhibit a large uncertainty when downscaling extreme values as shown in Fig. 8(c) ~ Fig. 8(f). For example, the Rx1day, which represents 1-day maximum rainfall (Rx1), exhibits a larger error in region A compared to region B. This is attributable to the high variability and heavy-tailed distribution of Rx1 in region A, as opposed to the low variability and a lighter-tailed distribution in region B. Similarly, the R20mm, which indicates the number of days rainfall exceeding 20 mm, exhibits higher errors in region E. In this region, the 20 mm rainfall may be considered extreme. However, in region A, where the 20 mm threshold does not coincide with the tail of the precipitation CDF, the BCSD method demonstrates greater accuracy for this indicator.

Meanwhile, the BCCA method, which has limitations in reproducing rainfall days and tends to underestimate rainfall amounts, demonstrates higher accuracy for the R20mm indicator in region E compared to region A. Notably, the number of rainfall days exceeding 20 mm is relatively lower in region E than in region A. Similarly, in region B, which has fewer rainfall days, the BCCA method shows greater accuracy compared to region D, which has more rainfall days and greater variability.

The second hypothesis is validated for interaction cases as Fig. 8 shows evidence for the unusual performances of downscaling output interacting with climate regimes. Characteristics of climate regimes amplify or diminish the influence of structural limitations of downscaling methods, resulting in varying downscaling performance by climatic regimes.

5. Discussion

5.1. Heterogeneous climate characteristics leading to varying regional uncertainties in large-scale downscaling results

This study addresses the first research question by demonstrating significant variability in downscaling performance across different climate regimes, with the optimal downscaling methods being spatially aligned with these regimes. For large-scale GCM-based analysis, downscaling methods that outperform across the entire study area are generally selected. However, uncertainty levels of downscaling outputs vary by region due to the different climate regimes (Figs. 3 and 4). The

performance of the selected downscaling method can dramatically decrease within certain climate regions (Fig. 7). Hence, conducting inter-regional comparisons of downscaling outputs based on a single downscaling method is not a fair comparison under identical conditions. There is a need to consider how to account for varying levels of uncertainty in inter-regional comparisons of downscaling results, especially for sensitive indicators identified during this study.

The discussion can be further expanded regarding the third research question. This study analyzed VCR, VDM, and VCD. Indicators exhibiting high VCR and VCD are sensitive to climate regimes. When considering VCR and VCD collectively, indicators such as Wet and dry, High runoff, and Heavy P frequency are shown to be particularly sensitive to different climate regimes. These indicators are mostly associated with extreme precipitation events, implying that greater attention must be paid to the spatially different uncertainty levels across climate regimes when comparing downscaled precipitation extremes. A potential approach to minimize these spatial differences in uncertainty levels is to select downscaling methods tailored to each climate zone, thereby reducing VCR, as the spatial distribution of downscaling errors aligns with climate zones (Fig. 4). Additionally, employing a sufficient number of downscaling methods can help reduce VCD by smoothing the interactive impacts between the climate regime and downscaling structure.

5.2. Supporting the selection of evaluation measures for choosing a climate downscaling method for hydrological applications

This study demonstrates the varying impacts of different climate downscaling methods on the accuracy of hydrological indicators across various climate regimes, addressing our second research question. Hydrological indicators are simulated through the interaction between precipitation and temperature (Das Bhowmik et al., 2017; Seo et al., 2019). Hence, the accuracy of a downscaling method for climate variables, as evaluated using only the climate indicators, may not directly correspond to its accuracy in predicting hydrological indicators (Fig. 5). The linkages between errors of climate and hydrological indicators vary by climate regimes. Therefore, hydrological indicators are deemed optimal for evaluating and selecting downscaling methods for hydrological applications. However, assessing hydrological indicators through hydrological models requires significant time and resources. Due to these costs, climate indicators are still commonly used for evaluating downscaling methods. Our findings suggest that this approach may be unsuitable in regions B, C, and D, where the linkage of errors of hydrological and climate indicators is not robust.

5.3. Transferability of identified causes for varying performances of downscaling output depending on climate regimes

In this study, we identified the spatially varying performances of GCMs and the interaction between downscaling methods and climate regimes as key sources of the different downscaling performances across climate regimes as related to the fourth research question. The observed interactions are transferable findings since the relationship between climate regimes and the downscaling methods may not dynamically change. However, the spatially varying performance of GCMs across climate regimes may depend on the GCM employed.

Cai et al. (2009) assessed the regional variability of GCM simulations and suggested that the GCM zones, consisting of grids with similar performance levels, are closely correlated with the Köppen climate classifications. Although there have been no other comprehensive studies assessing the performance of GCMs across various climate zones, some studies have evaluated the performance of GCMs within specific regions. Liu et al., (2018) found that reanalysis shows better performance in lower elevated regions (<1000 m above sea level) than in higher elevated regions within the Tibetan Plateau. Moreover, IPCC (2007b) and Li et al. (2021) indicated that GCMs generally perform poorly in arid and high-altitude regions due to the complex topography and, associated mesoscale weather system of high altitudes. These findings collectively suggest that the spatially varying performance across different climate regimes is a common characteristic of GCMs.

5.4. Comparison of downscaling methods for large-scale analysis

MOS-based methods show better downscaling performance for precipitation and runoff indicators compared to other methods based on MOS-PP, which has a more complex structure (Table S5). This contradicts the expectation that methods with a complicated structure would outperform methods with a simpler structure. This discrepancy may be attributed to the difficulty of fine-tuning the parameters of the PP approach in large-scale applications. The performances of PP methods are sensitive to parameters, such as predictors and domains (Bettolli, 2021). In large-scale applications, it is difficult to optimize the parameters for each region due to catastrophic costs. As an alternative, large-scale studies often adopt regional parameters, potentially leading to a decline in the performance of MOS-PP methods. The MOS-PP method offers valuable advantages for large-scale applications despite its limitations. The accuracy of the MOS-PP method, excluding BCCA, remains stable across various climate regimes. It shows low sensitivity to the regionally varying performance of GCM and exhibits minimal interaction effects (Fig. 8). In addition, it can provide interpretability of physical relationships, enhancing our understanding of the highly complex large-scale climate.

5.5. Uncertainty of results from climate classification systems

Climate classifications are utilized for regional planning of water resources, ecology, agriculture, and for understanding the differences in climate systems or changes brought about by these classifications. Consequently, they are determined based on the user's needs, taking into account factors such as precipitation, temperature, biotemperature, humidity, etc. (Bergeron, 1928; Holdridge, 1967; Köppen, 1936; Thornthwaite, 1948). While most climate classification systems adopt precipitation and temperature as criteria, they differ in the additional variables considered and the thresholds used for classification. Therefore, the variability in downscaling performance across climate regimes, as suggested in this study, differs based on the climate classification used. Moreover, variability can be overestimated in regions where climate conditions are close to the classification thresholds because Köppen's climate classification employs discrete thresholds to delineate zones. Despite these limitations, the results of this study are valuable and useful since the Köppen classification is the most commonly used method for classifying climates, and our results are statistically significant.

6. Conclusions

This study articulates the sources of spatially different accuracy/uncertainty of large-scale downscaling that have never been addressed before. The variability of downscaling performance across different climate regimes was quantified and the underlying causes of such variability were identified. Understanding the variability can provide valuable information for assessing the reliability and applicability of large-scale downscaling methods in various climate regimes. By exploring the interactions between downscaling methods and climate regimes across a wide geographic area, the study examined the strengths and weaknesses of downscaling methods tailored to different climate zones. Moreover, the optimal downscaling method and regional error distribution associated with the selection of the methods for specific regions were presented. This information can be useful for regional experts and policymakers when conducting analyses based on GCM products, such as climate change impact assessments. The insights provided in this study are likely to contribute to enhancing the regional applicability of future large-scale downscaling methods. Although this

study already provides such advantages, further research could benefit from different climate classifications. This refinement could further enhance our understanding of how other climate conditions influence downscaling performances. Furthermore, it is advisable to explore a variety of downscaling methods, including stochastic weather generators, to better comprehend their applicability in various climate conditions.

CRediT authorship contribution statement

Seon-Ho Kim: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jeongwoo Hwang:** Writing – review & editing, Conceptualization. **A. Sankarasubramanian:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

ERA5: https://cds.climate.copernicus.eu/#!/search?text=ERA5&type=dataset.

APHRODITE: https://www.chikyu.ac.jp/precip/english/downloads.html.

Unified gauged analysis: https://psl.noaa.gov/data/gridded/data.cpc.globalprecip.html.

Expert Team on Climate Change Detection and Indices (Program): etccdi.pacificclimate.org/index.shtml.

Indicators of Hydrologic Alteration (Program): https://www.conservationgateway.org/ConservationPractices/Freshwater/EnvironmentalFlows/MethodsandTools/IndicatorsofhydrologicAlteration/Pages/indicators-hydrologic-alt.aspx.

Acknowledgments

This research is based upon work supported in part by the NSF within the framework of CAS-Climate: Understanding the Changing Climatology, Organizing Patterns and Source Attribution of Hazards of Floods over the Southcentral and Southeast US, award No. 2208562 (PI: Sankar Arumugam). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of NSF.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2024.131818.

References

Abatzoglou, J.T., Brown, T.J., 2012. A comparison of statistical downscaling methods suited for wildfire applications. Int. J. Climatol. 32 (5), 772–780. https://doi.org/10.1002/joc.2312.

Anaraki, M.V., Farzin, S., Mousavi, S.F., Karami, H., 2021. Uncertainty analysis of climate change impacts on flood frequency by using hybrid machine learning methods. Water Resour. Manag. 35, 199–223. https://doi.org/10.1007/s11269-020-02719-w.

Bae, D.H., Rahman, M., Koike, T., Ahmad, B., 2013. Climate change impact assessment on the Asia-Pacific water resources under AWCI/GEOSS. Final Report of the APN ARCP Project: ARCP2011-05CMY-Bae. https://doi.org/10.30852/p.4289.

Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2018. Present and future Köppen-Geiger climate classification maps at 1-km resolution. Sci. Data 5, 180214. https://doi.org/10.1038/sdata.2018.214.

Bergeron, T., 1928. Über die dreidimensional verknüpfende Wetteranalyse. Erster Teil: Prinzipielle Einführung in das Problem der Luftmassen- und Frontenbildung, 5 (6), 111.

Bettolli, M.L., 2021. Analog models for empirical-statistical downscaling. Clim. Sci. 23, 1–25. https://doi.org/10.1093/acrefore/9780190228620.013.738.

Bürger, G., Sobie, S.R., Cannon, A.J., Werner, A.T., Murdock, T.Q., 2013. Downscaling extremes: an intercomparison of multiple methods for future climate. J. Clim. 26, 3429–3449. https://doi.org/10.1175/JCLI-D-12-00249.1.

Cai, X., Wang, D., Zhu, T., Ringler, C., 2009. Assessing the regional variability of GCM simulations. Geophys. Res. Lett. 36, L02706. https://doi.org/10.1029/2008GL036443.

Chen, H., Xu, C.Y., Guo, S., 2012. Comparison and evaluation of multiple GCMs statistical downscaling and hydrological models in the study of climate change impacts on runoff. J. Hydrol. 434–435, 36–45.

Cohen, J., 2013. Statistical power analysis for the behavioural sciences. Routledge.

Das Bhowmik, R., Sankarasubramanian, A., Sinha, T., Patskoski, J., Mahendrakar, G., Kunkel, K.E., 2017. Multivariate downscaling approach preserving cross correlations across climate variables for projecting hydrologic fluxes. J. Hydrometeor. 18, 2187–2205. https://doi.org/10.1175/JHM-D-16-0160.1.

Eum, H.I., Cannon, A.J., 2017. Intercomparison of projected changes in climate extremes for South Korea: application of trend preserving statistical downscaling methods to the CMIP5 ensemble. Int. J. Climatol. 37 (8), 3381–3397. https://doi.org/10.1002/joc.4924.

Eum, H.I., Cannon, A.J., Murdock, T.Q., 2016. Intercomparison of multiple statistical downscaling methods: multi-criteria model selection for South Korea. Stoch. Environ. Res. Risk Assess. 31, 683–703. https://doi.org/10.1007/s00477-016-1312-9.

Eum, H.I., Gupta, A., Dibike, Y., 2020. Effects of univariate and multivariate statistical downscaling methods on climatic and hydrologic indicators for Alberta. Canada. J. Hydrol. 588, 125065 https://doi.org/10.1016/j.jhydrol.2020.125065.

Fowler, H.J., Blenkinsop, S., Tebaldi, C., 2007. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. Int. J. Climatol. 27, 1547–1578.

Gutiérrez, J.M., Maraun, D., Widmann, M., Huth, R., Hertig, E., Benestad, R., Roessler, O., Wibig, J., Wilcke, R., Kotlarski, S., San Martín, D., Herrera, S., Bedia, J., Casanueva, A., Manzanas, R., Iturbide, M., Vrac, M., Dubrovsky, M., Ribalaygua, J., Pórtoles, J., Räty, O., Räisänen, J., Hingray, B., Raynaud, D., Casado, M.J., Ramos, P., Zerenner, T., Turco, M., Bosshard, T., Štěpánek, P., Bartholy, J., Pongracz, R., Keller, D.E., Fischer, A.M., Cardoso, R.M., Soares, P.M.M., Czernecki, B., Pagé, C., 2018. An Intercomparison of a large ensemble of statistical downscaling methods over Europe: results from the VALUE perfect predictor cross-validation experiment. Int. J. Climatol. 39 (9), 3750–3785. https://doi.org/10.1002/joc.4562.

Gutmann, E., Pruitt, T., Clark, M.P., Brekke, L., Arnold, J.R., Raff, D.A., Rasmussen, R.M., 2014. An intercomparison of statistical downscaling methods used for water resources assessments in the United States. Water Resour. Res. 50 (9), 7167–7186. https://doi.org/10.1002/2014WR015559.

Hidalgo, H.G., Dettinger, M.D., Cayan, D.C., 2008. Downscaling with constructed analogues: daily precipitation and temperature fields over The United States: California Energy Commission Report.

Hirabayashi, Y., Tanoue, M., Sasaki, O., Zhou, X., Yamazaki, D., 2021. Global exposure to flooding from the new CMIP6 climate model projections. Sci. Rep. 11, 3740. https://doi.org/10.1038/s41598-021-83279-w.

Holdridge, L.R., 1967. Life zone ecology. Trop. Sci.

Holthuijzen, M., Beckage, B., Clemins, P.J., Higdon, D., Winter, J.M., 2022. Robust bias-correction of precipitation extremes using a novel hybrid empirical quantile-mapping method. Theor. Appl. Climatol. 149, 863–882. https://doi.org/10.1007/s00704-022-04035-2.

Hou, Y., He, Y., Chen, H., Xu, C.X., Chen, J., Kim, J.S., Guo, S.I., 2019. Comparison of multiple downscaling techniques for climate change projections given the different climatic zones in China. Theor. Appl. Climatol. 138, 27–45. https://doi.org/10.1007/s00704-019-02794-z.

Hundecha, Y., Sunyer, M.A., Lawrence, D., Madsen, H., Willems, P., Bürer, G., Kriaučiūnienė, J., Loukas, A., Martinkova, M., Osuch, M., Vasiliades, L., von Christierson, B., Vormmor, K., Yijcel, I., 2016. Inter-comparison of statistical downscaling methods for projection of extreme flow indices across Europe. J. Hydrol. 541, 1273–1286. https://doi.org/10.1016/j.jhydrol.2016.08.033.

Ipcc, 2007a. Climate change 2007: impacts, adaptation and vulnerability. In: Parry, M.L., Canziani, O.F., Palutikof, J.P., van der Linden, P.J., Hanson, C.E. (Eds.), Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY USA.

Ipcc, 2007b. Climate change 2007: the physical science basis. In: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., Miller, H.L. (Eds.), Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, p. 881.

Ipcc, 2014. Climate change 2014: Impacts, Adaptation, and Vulnerability. In: Field, C.B., Barros, V.R., Dokken, D.J., Mach, K.J., Mastrandrea, M.D., Bilir, T.E., Chatterjee, M., Ebi, K.L., Estrada, Y.O., Genova, R.C., Girma, B., Kissel, E.S., Levy, A.N., MacCracken, S., Mastrandrea, P.R., White, L.L. (Eds.), Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY USA.

Ipcc, 2022. Climate Change 2022: Impacts, Adaptation, and Vulnerability. In: Pörtner, H. O., Roberts, D.C., Tignor, M., Poloczanska, E.S., Mintenbeck, K., Alegría, A.,

Craig, M., Langsdorf, S., Löschke, S., Möller, V., Okem, A., Rama, B. (Eds.), Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK and New York, NY, USA.

Kim, J.B., Bae, D.H., 2021. The impacts of global warming on climate zone changes over asia based on CMIP6 projections. Earth Space Sci. 8, e2021EA001701 https://doi.org/10.1029/2021EA001701.

Kim, S.H., Kim, J.B., Park, D.R., Bae, D.H., 2022a. Detection of hydropower change points under future climate conditions based on technical hydropower potential changes in Asia. J. Hydrol.: Reg. Stud. 44, 101258 https://doi.org/10.1016/j.ejrh.2022.101258.

Kim, S.H., Kim, J.B., Bae, D.H., 2022b. Development of a climate-informed analog downscaling method for Asian region. Int. J. Climatol. 42 (12), 6148–6168. https://doi.org/10.1002/joc.7582.

Kirchhoff, C.J., Barsugli, J.J., Galford, G.L., Karmalkar, A.V., Lombardo, K., Stephenson, S.R., Barlow, M., Seth, A., Wang, G., Frank, A., 2019. Climate assessment for local action. Bull. Amer. Meteor. Soc. 100, 2147–2152. https://doi.org/10.1175/BAMS-D-18-0138.1.

Köppen, W. 1936. Das geographische System der Klimate, 1–44.

Kumar, H., Zhu, T., Sankarasubramanian, A., 2023. Understanding the food-energy-water nexus in mixed irrigation regimes using a regional hydroeconomic optimization modeling framework. Water Resour. Res. 59 (6), e2022WR033691.

Lee, M.H., Qui, L., Ha, S., Im, E.S., Bae, D.H., 2021. Future projection of low flows in the Chungju basin, Korea and their uncertainty decomposition. Int. J. Climatol. 42 (1), 157–174. https://doi.org/10.1002/joc.7237.

Li, X., Li, Z., 2022. Global water availability and its distribution under the coupled model intercomparison project phase six scenarios. Int. J. Climatol. 42 (11), 5748–5767. https://doi.org/10.1002/joc.7559.

Li, J., Miao, C., Wei, W., Zhang, G., Hua, L., Chen, Y., Wang, X., 2021. Evaluation of CMIP6 global climate models for simulating land surface energy and water fluxes during 1979–2014. J. Adv. Model. Earth Syst. 13, e2021MS002515 https://doi.org/10.1029/2021MS002515.

Liu, Z., Liu, Y., Wang, S., Yang, X., Wang, L., Baig, M.H.A., Chi, W., Wang, Z., 2018. Evaluation of spatial and temporal performances of ERA-interim precipitation and temperature in mainland China. J. Clim. 31 (11), 4347–4365. https://doi.org/10.1175/JCLI-D-17-0212.1.

Maraun, D., 2016. Bias correcting climate change simulations – a critical review. Curr. Clim. Change Rep. 2, 211–220. https://doi.org/10.1007/s40641-016-0050-x.

Maraun, D., Widmann, M., Gutiérrez, J.M., Kotlarski, S., Chandler, R.E., Hertig, E., Wibig, J., Huth, R., Wilcke, R.A.I., 2014. VALUE: A framework to validate downscaling approaches for climate change studies. Earth's Future 3 (1), 1–14. https://doi.org/10.1002/2014EF000259.

Mathews, R., Richter, B.D., 2007. Application of the indicators of hydrologic alteration software in environmental flow setting. J. Am. Water Resour. As. 43, 1400–1413. https://doi.org/10.1111/j.1752-1688.2007.00099.x.

Maurer, E.P., Hidalgo, H.G., Das, T., Dettinger, D., Cayan, D.R., 2010. The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California. Hydrol. Earth Syst. Sci. 14 (1125–1138), 2010. https://doi.org/10.5194/hess-14-1125-2010.

Pandey, V.P., Shretha, D., Adhikari, M., 2021. Characterizing natural drivers of water-induced disasters in a rain-fed watershed: Hydro-climatic extremes in the Extended East Rapti Watershed, Nepal. J. Hydrol. 598, 126383 https://doi.org/10.1016/j.jhydrol.2021.126383.

Pierce, D.W., Cayan, D.R., Thrasher, B.L., 2014. Statistical downscaling using localized constructed analogs (LOCA). J. Hydrometeorol. 15, 2558–2585. https://doi.org/10.1175/JHM-D-14-0082.1.

Seo, S.B., Bhowmik, R.D., Sankarasubramanian, A., Mahinthakumar, G., Kumar, M., 2019. The role of cross-correlation between precipitation and temperature in basin-scale simulations of hydrologic variables. J. Hydrol. 570, 304–314. https://doi.org/10.1016/j.jhydrol.2018.12.076.

Son, K.H., Bae, D.H., 2015. Drought analysis according to shifting of climate zones to arid climate zone over Asia monsoon region. J. Hydrol. 529, 1021–1029. https://doi.org/10.1016/j.jhydrol.2015.09.010.

Teng, J., Chiew, F.H.S., Timbal, B., Wang, Y., Vaze, J., Wang, B., 2012. Assessment of an analogue downscaling method for modelling climate change impacts on runoff. J. Hydrol. 472–473, 111–125. https://doi.org/10.1016/j.jhydrol.2012.09.024.

Thornthwaite, C.W., 1948. An approach toward a rational classification of climate. Geogr. Rev. 38 (1), 55–94.

Wang, F., Huang, G.H., Fan, Y., Li, Y.P., 2020. Robust subsampling ANOVA methods for sensitivity analysis of water resource and environmental models. Water Resour. Manag. 34, 3199–3217. https://doi.org/10.1007/s11269-020-02608-2.

Werner, A., Cannon, A.J., 2016. Hydrologic extremes – an intercomparison of multiple gridded statistical downscaling methods. Hydrol. Earth Syst. Sci. 20, 1483–1508. https://doi.org/10.5194/hess-20-1483-2016.

Wilby, R.L., Wigley, T.M.L., 1997. Downscaling general circulation model output: a review of methods and limitations. Prog. Phys. Geogr. 21 (4), 530–548. https://doi.org/10.1177/030913339702100403.

Zhang, J., Fan, H., He, D.M., Chen, J.W., 2019. Integrating precipitation zoning with random forest regression for the spatial downscaling of satellite-based precipitation: A case study of the Lancang-Mekong River basin. Int. J. Climatol. 39 (10), 3947–3961. https://doi.org/10.1002/joc.6050.

Zhang, X., Wang, H., Peng, A., Wang, W., Li, B., Huang, X., 2020b. Quantifying the uncertainties in data-driven models for reservoir inflow prediction. Water Resour. Manag. 34, 1479–1493. https://doi.org/10.1007/s11269-020-02514-7.

Zhang, L., Xu, Y., Meng, C.C., Li, X.H., Liu, H., Wang, C.G., 2020a. Comparison of statistical and dynamic downscaling techniques in generating high-resolution temperatures in China from CMIP5 GCMs. J. Appl. Meteorol. Climatol. 59 (2), 207–235. https://doi.org/10.1175/JAMC-D-19-0048.1.