https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

A two-stage Bias-Correction and Super-Resolution Framework for Post-Processing Climate Model Outputs

Abba Ibrahim[1,5], Aimrun Wayayok[1,2,3,b], Helmi Zulhaidi Mohd Shafri[4,] Noorellimia Mat Toridi[1], Wada Idris Muhammad[6].

[1]Department of Biological and Agricultural Engineering, Faculty of Engineering, Universiti Putra Malaysia. 43400 UPM Serdang, Selangor DE, Malaysia.

[2]SMART Farming Technology Research Center (SFTRC), Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor DE, Malaysia.

[3]International Institute of Aquaculture and Aquatic Sciences (I-AQUAS), Universiti Putra Malaysia, Mile 7, Kemang Rd. 6. Kemang Bay, Si Rusa, Port Dickson, Negeri Sembilan 71050, Malaysia.

[4]Department of Civil Engineering, Faculty of Engineering, Universiti Putra Malaysia. 43400 UPM Serdang, Selangor DE, Malaysia.

[5]Department of Agricultural and Environmental Engineering, Faculty of Engineering, Bayero University, Kano, Nigeria.

[6]College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, China.

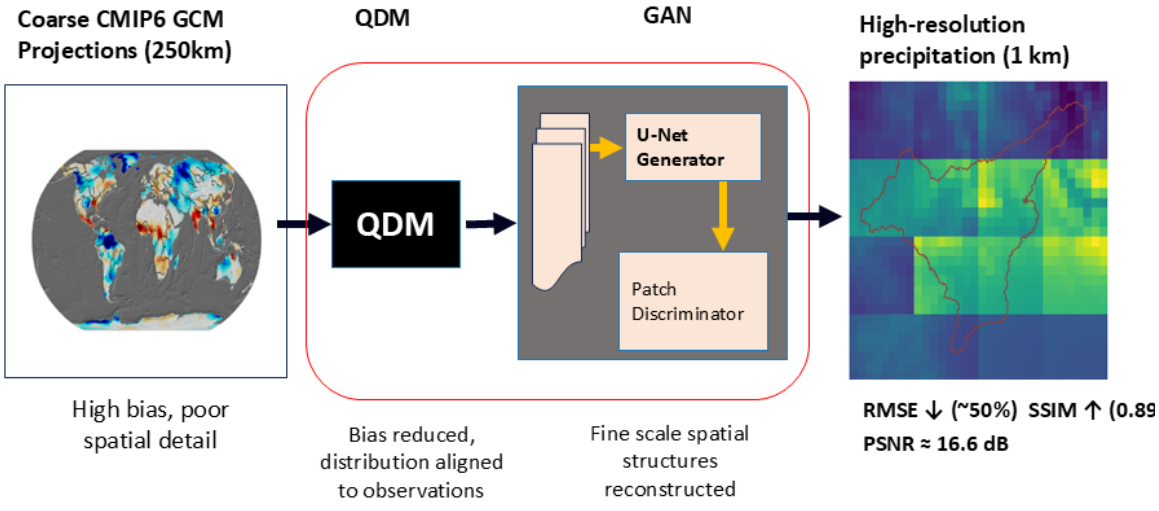
Correspondence to: A. Wayayok (aimrun@upm.edu.my)

Abstract

General circulation models (GCMs) underpin climate change assessments, yet their coarse spatial resolution and systematic biases constrain their direct use in regional applications. Post-processing approaches such as bias correction and statistical downscaling are therefore widely applied, yet these steps are often implemented independently, leading to inconsistencies between corrected statistics and spatial structure. This study presents a reproducible two-stage framework that integrates Quantile Delta Mapping (QDM) for bias correction with a deep learning-based super-resolution method to improve the statistical fidelity and spatial detail of climate model outputs. The framework is evaluated using precipitation, runoff, and evapotranspiration from three Coupled Model Intercomparison Project Phase 6 (CMIP6) GCMs (Canadian Earth System Model (CanESM5), Hadley Centre Global Environment Model (HadGEM3-GC31-LL), and Max Planck Institute Earth System Model (MPI-ESM1-2-HR)), over the Hadejia-Jama'are River Basin in northern Nigeria. We first demonstrate that raw model outputs exhibit substantial biases, with domain-averaged root mean square errors (RMSE) of 48.6-57.3 mm month^{-1} for precipitation, 0.93-7.51 mm month^{-1} for runoff, and 33.8-58.7 mm month^{-1} for evapotranspiration. QDM substantially reduces systematic errors, lowering precipitation RMSE to 23.8-

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

27.8 mm month^{-1}, runoff RMSE to 0.24-1.85 mm month^{-1}, and evapotranspiration RMSE to 3.7-4.3 mm month^{-1}, while preserving projected distributional changes, as confirmed by Kolmogorov-Smirnov ($D \leq 0.072$) and Wasserstein ($\leq$ 11.86) metrics. In the second stage, a conditional Generative Adversarial Network (GAN) super-resolves the bias-corrected precipitation fields from ~250 km to 14 km (regridded to 1 km), with performance assessed using complementary pixel-wise, structural, and distributional metrics, and outperforming bilinear and bicubic interpolation in terms of spatial coherence and structural similarity. The proposed QDM-GAN framework is fully documented and reproducible, with openly available code and data sources, and is intended as a modular post-processing tool that can support downstream modeling applications requiring bias-corrected, high-resolution climate inputs.



Key words:
Generative Adversarial Networks (GANs), Precipitation, CMIP6, GCMs, RCMs, sub-km spatial resolution.

1. Introduction

General circulation models (GCMs) and regional climate models (RCMs) form the backbone of contemporary climate simulations, providing physically consistent representations of large-scale atmospheric and land-surface processes (Flato et al., 2013). However, the spatial resolution of most GCMs, typically on the order of 100-250 km, remains insufficient for many regional and local-scale applications (Yu et al., 2018, Thao et al., 2022; Wang 2023). At these resolutions, sub-grid heterogeneity associated with topography, land cover, and convective processes is poorly resolved, and systematic biases relative to observations persist (Maraun et al., 2017; Thao et al., 2022). When GCM outputs are used directly as forcings for impact or Earth-system component models, these limitations can propagate into downstream

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

analyses, motivating the widespread use of post-processing techniques (Labeurthre et al., 2024). At the same time, evaluating whether post-processed outputs preserve both statistical fidelity and realistic spatial structure remains an important challenge in climate model assessment.

Downscaling approaches aimed at bridging the scale gap between global simulations and local climate information can be broadly categorized into dynamical and statistical methods (Calel and Stainforth, 2017; Hudson, 2021). Dynamical downscaling, implemented through RCMs, explicitly resolves regional processes and has achieved grid spacings as fine as a few kilometers in recent studies (Afrizal and Surussavadee, 2018; Ajjur, 2024; Merks et al., 2020). Despite these advances, RCM simulations remain computationally expensive and inherit uncertainties from both the driving GCM and the regional model configuration, including sensitivity to boundary conditions and parameterizations (Ghimire et al., 2018; Goodwin and Cael, 2021; Wang and Tian, 2022). These uncertainties can be particularly consequential for hydrological variables such as runoff and evapotranspiration (Teng et al., 2015; Wang and Tian, 2022).

Statistical approaches to downscaling offer computational efficiency and flexibility and therefore remain central to many climate modeling workflows, particularly where ensemble processing or long time series are required (Hudson, 2021). Bias correction is a key component of statistical post-processing, aiming to reduce systematic discrepancies between modeled and observed climatology (Cannon et al., 2015; Maraun et al. 2017). Quantile-based bias-correction methods are among the most widely used and have been shown to effectively adjust mean state and variability (Maraun, 2016). However, classical quantile mapping approaches typically assume stationary biases and may distort projected climate change signals, particularly for extremes (Maraun et al., 2017; Rajulapati and Papalexiou, 2023).

Quantile Delta Mapping (QDM) was introduced to address this limitation by preserving relative or absolute changes in quantiles between historical and future simulations while correcting baseline biases (Cannon et al., 2015; Pierce et al., 2015) QDM has been applied successfully across multiple hydroclimatic variables and regions and is widely used in climate impact studies (Tong et al., 2021; Ugolotti et al., 2023). Nevertheless, QDM operates independently at each grid point and does not address the loss of spatial detail inherent in coarse-resolution climate model outputs (Gergel et al., 2024).

In parallel, recent developments in machine learning have enabled new approaches for statistical downscaling through image super-resolution techniques. Convolutional neural networks and generative adversarial networks (GANs) have demonstrated skill in reconstructing fine-scale spatial structure from coarse inputs by learning mappings from paired low- and high-resolution data (Stengel et al., 2020; Wang and Tian, 2022). GAN-based approaches, in particular, have been shown to recover sharp gradients, terrain-induced variability, and localized extremes that are smoothed by conventional interpolation methods (Oyama et al., 2023). However, many machine-learning-based downscaling studies focus primarily on spatial enhancement and apply limited or no bias correction, which can result in high-resolution fields that retain systematic statistical errors (Besombes et al., 2021; Rampal et al., 2024).

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

Despite progress in both bias correction and machine-learning-based downscaling, relatively few studies have integrated these components within a unified and reproducible evaluation workflow. As a result, bias correction and super-resolution are often treated as separate steps, with limited assessment of how corrections applied in one stage influence statistical fidelity and spatial realism in the other (Gergel et al., 2024; Rajulapati and Papalexiou, 2023; Ugolotti et al., 2023). This fragmentation complicates reproducibility and comparison across studies, particularly in the context of Earth system model post-processing.

This paper addresses this gap by developing and evaluating a two-stage post-processing framework that combines Quantile Delta Mapping with a conditional GAN-based super-resolution model. The framework is applied to multiple CMIP6 GCMs and assessed using complementary statistical, distributional, and structural metrics following established practices in bias-correction and super-resolution literature (Cannon et al., 2015; Stengel et al., 2020). Every processing step is documented and supported by open-access code and data to facilitate transparency and reproducibility. Consequently, this study provides a transferable workflow for generating and evaluating bias-corrected, high-resolution climate model outputs intended for downstream hydrological and environmental applications.

2. Data and Methods

2.1 Study Area Description

The two-stage framework is demonstrated over Hadejia-Jama'are River Basin (HJRB) in northern Nigeria, geographically bounded by 10°-13° N and 7°-11.5° E (Figure 1). The basin serves as a critical water resource hub for agricultural production, domestic water supply, and ecological stability. Despite its importance, the basin suffers from notable data scarcity, which complicates both hydrological modeling and long-term resource planning. The region's climate is characterized by high evapotranspiration, marked wet and dry seasons, and pronounced rainfall variability, rendering local communities and ecosystems highly susceptible to climate the climate change impacts.

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

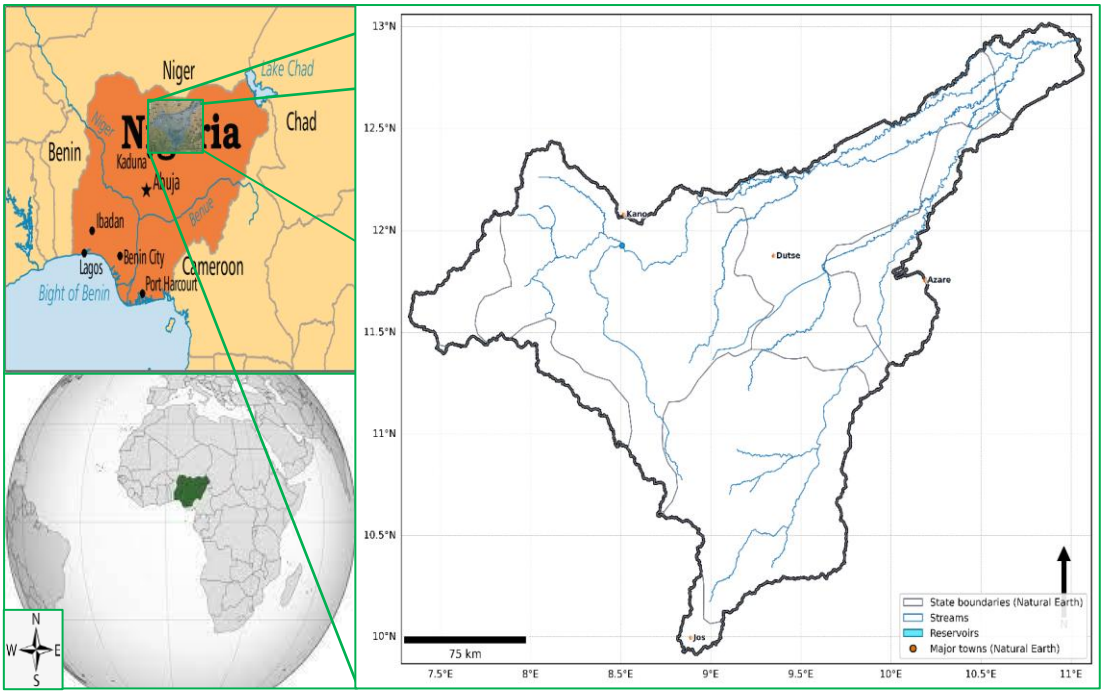


Figure 1: Study area location

2.2 Data Sources

Observed satellite data

Precipitation:

Precipitation estimates for the study area were obtained from the Climate Hazards Group InfraRed Precipitation with Stations (CHIRPS) product (Funk et al., 2015) which merges satellite infrared data with in-situ station observations to capture high-resolution rainfall variability. It provides global coverage at a spatial resolution of 0.05° (~5 km) from 1981 to near-real-time, updated monthly or more frequently (Horn et al., 2018). CHIRPS integrates infrared Cold Cloud Duration (CCD) measurements from geostationary satellites with quality-controlled rain-gauge observations, thereby reducing biases often found in purely satellite-based products (Bandara et al., 2021). Its relatively fine spatial resolution and long-term temporal record make CHIRPS well-suited for hydrological applications in data-scarce regions like the Hadejia-Jama'are River Basin. CHIRPS has been validated and found to be effective in monitoring extreme precipitation events, contributing to a better understanding of rainfall variability in various regions (Duan et al., 2019). Studies have shown that CHIRPS performs comparably to other satellite-based precipitation products, making it a reliable choice for hydrological modeling (Bandara et al., 2021; Bichet and Diedhiou, 2018). Furthermore,

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

CHIRPS has been utilized in various hydrological evaluations, demonstrating its utility in regions with limited ground-based rainfall data (Belay et al., 2019; Pandey et al., 2021). The integration of CHIRPS data into hydrological models enhances the accuracy of simulations, particularly in areas where traditional rain gauge networks are sparse which is typical of HJRB.

Evapotranspiration data:

The Evapotranspiration (ET) datasets were obtained from MODIS/061/MOD16A2GF product, accessed through the NASA Earthdata portal (Running et al., 2021). This dataset, referred to as MODIS ET, employs a Penman-Monteith-type algorithm that integrates various meteorological inputs, such as air temperature and solar radiation, along with MODIS land cover and leaf area index (LAI) products). The dataset is as an 8-day composite with a nominal spatial resolution of 500 meters, upscaled to a 1-kilometer resolution and aggregated to monthly time steps in this work. The physically based approach of the MODIS ET product captures dynamic vegetation conditions, accounting for both energy-driven and water-driven controls on ET, which is crucial for accurate modeling of water balance (Kim et al., 2012). However, it is important to note that discrepancies have been observed between MODIS ET estimates and ground measurements, often attributed to biases in the input data from both MODIS and global meteorological datasets (Aguilar et al., 2018; Kim et al., 2012). The MODIS ET dataset has been widely adopted in global and regional water balance studies, providing a consistent and regularly updated resource since its inception in 2000. Its application extends to various fields, including hydrology, ecology, and environmental science, making it a vital tool for researchers aiming to understand evapotranspiration dynamics on both regional and global scales (Liu et al., 2020).

Surface Runoff:

Surface runoff estimates for this study were sourced from the Global Land Data Assimilation System (GLDAS), a NASA-led initiative that integrates satellite-based and ground-based observations into sophisticated land surface models (Qing et al., 2021). The datasets were accessed through the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) (Beaudoing and Rodell, 2024) .GLDAS outputs are available at various spatial resolutions, typically 0.25° or 0.10°, and temporal resolutions ranging from 3-hourly to monthly. For this research, a monthly aggregated surface runoff (Qs) dataset was utilized, covering the period from 2000 to the present, which aligns with the available MODIS ET data. A physically consistent approach was employed by GLDAS to produce land-surface fluxes through the assimilation of satellite-derived precipitation, snow cover, and soil moisture products. This integration was designed to mitigate biases that may arise from purely modeled runoff estimates, thus enhancing the reliability of the data (Gruber et al., 2019).

Table 1 summarized the features of each of the satellite-observed datasets. Each of these offers distinct advantages for hydrological modeling, especially in regions with limited ground-based observations. CHIRPS supplies high-resolution

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

rainfall estimates critical for driving surface-water dynamics, while MODIS ET captures seasonally evolving water use by vegetation, and GLDAS runoff helps constrain near-surface hydrological fluxes under changing precipitation regimes. Together, they enable a more robust quantification of water balance components across the HJRB, supporting both validation of bias-corrected GCM outputs and subsequent downscaling efforts.

Table1: Summary of satellite-observed variables

Variable	Data Source	Temporal Coverage	Spatial Resolution	Temporal Resolution	References
Precipitation (mm)	CHIRPS (Climate Hazards Group)	1981-Present, near real-time	0.05° (~5 km)	Daily, Pentadal, Monthly	(Center, 2025) and (Funk et al., 2015)
Evapotranspiration (mm)	MODIS (MOD16A2GF)	2000-Present	500 m to 1 km	8-day composites (aggregable)	(Running et al., 2021)
Surface Runoff (mm)	GLDAS (e.g., Noah LSM)	2000-Present (varies by version)	0.25° or 0.10° (~28-11 km)	3-hourly, Daily, or Monthly	(Beaudoing and Rodell, 2024)

CMIP6 GCs Model Datasetssn

Three advanced Earth System Models from the Coupled Model Intercomparison Project Phase 6 (CMIP6) (MPI-ESM1-2-HR, HadGEM3-GC31-LL, and CanESM5) were selected to represent both historical (pre-2015) and future scenarios under distinct Shared Socioeconomic Pathways (SSPs). The data were downloaded from Earth System Grid Federation (ESGF) under Analytics and Informatics Management Systems (AIMS) at Lawrence Livermore National Laboratory data portal (ESGF, 2025) which archives and distributes CMIP6 outputs in compliance with standardized metadata and file formats, specifically NetCDF.

The selection of these CMIP6 models was justified by a recent study conducted by Wada et al. (Wada et al., 2023) which examined the performance of CMIP6 models (100 ensembles) using an advanced envelop-based selection approach in the same study area. The study emphasized that MPI-ESM1-2-HR, HadGEM3-GC31-LL, and CanESM5 exhibit lower biases and better seasonal cycle representation, making them suitable for climate impact assessments in the area. Furthermore, the selected models have been shown to effectively simulate extreme climate events, which is key for understanding potential future impacts. For instance, HadGEM3-GC31-LL has been recognized for its robust performance in simulating extreme precipitation and temperature events (Nishant et al., 2022).

Monthly mean fields for precipitation (pr), surface runoff (mrros), and evapotranspiration (evspsbl) were utilized, employing the r1i1p1f1 realization for each model to ensure consistent forcing protocols and initial conditions across all scenarios. Table 2 presented a summary of the model and detail of the variables used which are publicly accessible via the CMIP6 license and data usage policies.

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

Table 2: Summary of CMIP6 Datasets used

Models	Historical (years)	Future (2015-2100) SSPs	Variables	Native Resolution
MPI-ESM1-2-HR (from Planck Institute for Meteorology, Germany)	1850-2014	SSP126, SSP245, SSP585	pr, mrros, evspsbl	~1.0° atmosphere (approx)
HadGEM3-GC31-LL (from Met Office Hadley Centre, United Kindom)	1850-2014	SSP126, SSP245, SSP585	pr, mrros, evspsbl	~1.25° × 1.875°
CanESM5 (from Canadian Centre for Climate Modelling and Analysis, Canada)	1850-2014	SSP126, SSP245, SSP585	pr, mrros, evspsbl	~2.8° (atmosphere)

In this study, the historical data (2000-2014) were used as a baseline, while future projections (2015-2100) under *SSP126*, *SSP245*, and *SSP585* are used to explore a range of emission and socioeconomic pathways.

2.3 Methodology

2.3.1 Quantile Delta Mapping (QDM) for Bias Correction

Biases in Global Climate Model (GCM) outputs were corrected using Quantile Delta Mapping (QDM), which extends the standard quantile mapping (QM) approach by aligning the quantile structure of simulated data with observations while preserving the model-projected distributional changes (Cannon et al., 2015; Pierce et al., 2015). Unlike classical QM, which can suppress or overcorrect projected changes, QDM ensures that relative or absolute shifts in quantiles are retained.

Let $X_m(t)$ denote the model simulation at time t, and $X_o(t)$ represent the corresponding observed value. The percentile of a model value under the future distribution is computed as:

$$p(t) = F_{mod,fut}\left(X_{m,fut}(t)\right) \quad (1)$$

where $F_{\text{mod, fut}}$ is the future model cumulative distribution function (CDF). The corresponding baseline mapped value is obtained as:

$$\tilde{X}(t) = F^{-1}_{obs,ref}(p(t)), \quad (2)$$

with $F^{-1}_{\text{obs, ref}}$ denoting the inverse CDF of the observed reference distribution. QDM then applies a quantiles specific adjustment (Eq. 3, 4).

Additive form (e.g., temperature): $X^*(t) = \tilde{X}(t) + \left[F^{-1}_{mod,fut}(p) - F^{-1}_{mod,ref}(p)\right]$ (3)

Multiplicative form (e.g., precipitation, runoff): $X^*(t) = \tilde{X}(t) \times \frac{F^{-1}_{mod,fut}(p)}{F^{-1}_{mod,ref}(p)}$ (4)

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

The QDM procedure was applied independently to precipitation (pr), evapotranspiration (evspsbl), and runoff (mrros). Both observed and model datasets were interpolated to a common 128 × 128 grid. Following (Cannon et al., 2015) and (Tong et al., 2021), the implementation comprised three steps: (i) estimation of empirical CDFs for observed and modeled data, (ii) derivation of quantile-wise correction factors, and (iii) application of these factors to future CMIP6 simulations under SSP126, SSP245, and SSP585.

Validation of the QDM-corrected simulations against historical observations involved:

a) Quantitative assessment using Root Mean Squared Error (RMSE), Nash-Sutcliffe Efficiency (NSE), and coefficient of determination (R^2).
b) Qualitative assessment through time-series plots, scatter diagrams, and quantile-quantile (Q-Q) plots.

This ensured both statistical improvements and the preservation of distributional integrity in corrected simulations.

2.3.2 GAN-Based Downscaling

To further enhance spatial resolution**,** Generative Adversarial Networks (GANs) (Goodfellow et al., 2014) were employed. In this framework, a **Generator (G)** learns to transform coarse-resolution GCM fields into high-resolution climate fields, while a **Discriminator (D)** learns to distinguish between real high-resolution data and generated outputs. A conditional GAN (cGAN) structure was implemented (Isola et al., 2017; Mirza and Osindero, 2014), conditioning both G and D on low-resolution model input.

Loss Functions:

The generator was trained using a composite objective (Eq. 3, 4 and 5)

$$L_G = \alpha L_{MSE} + \beta L_{adv}, \tag{3}$$

where L_{MSE} penalizes pixel-wise errors, and L_{adv} enforces realism. The discriminator optimized a binary cross-entropy objective:

$$L_D = -E[\log D(y_{\text{true}})] - E\left[\log\left(1 - D\left(G(x_{\text{coarse}})\right)\right)\right], \tag{4}$$

where $D(y_{\text{true}})$ denotes the probability assigned to real high-resolution data. A perceptual loss (Ledig et al., 2017) was also integrated:

$$L_G = L_{GAN} + \lambda_{L1}\|\hat{y} - y\|_1 + \lambda_{perc}\|\Phi(\hat{y}) - \Phi(y)\|_1, \tag{5}$$

where $\Phi(\cdot)$ denotes VGG-16 feature maps, $\lambda_{L1} = 100$, and $\lambda_{\text{perc}} = 0.1$.

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

Data and Preprocessing:

Monthly precipitation from bias-corrected CMIP6 models ($\sim 250km$ resolution) and 14 km reanalysis data were paired for 180 months. Each pair was stored as a 128×128 array, normalized to $[-1,1]$ following Radford et al. (2016). Data were split into 80% training and 20% validation sets, with augmentation (rotations, flips) applied only to training samples.

Network Architecture and Training:

- Generator: A U-Net architecture with four encoder-decoder levels and skip connections.
- Discriminator: 64×64 PatchGAN with spectral normalization.
- Optimization: Adam optimizer ($\beta_1 = 0.5, \beta_2 = 0.999$); learning rates initialized at 2×10^{-4} (generator) and 4×10^{-4} (discriminator), annealed to 10^{-6}.

Training was performed on an NVIDIA GeForce GTX 1050 GPU with a batch size of 4, using early stopping based on minimum validation loss.

Assessment Framework and Evaluation Metrics:

Downscaling performance was assessed using three complementary and widely accepted metrics in super-resolution (SR): Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM). MSE was computed on the rescaled rainfall fields to assess amplitude errors in the physical variable after normalization; PSNR (a monotonic transform of MSE) provided a distortion measure in decibels that is standard in SR benchmarks; and SSIM captured luminance/contrast/structure consistency to better reflect perceived and structural quality than pixel-wise errors alone. The joint use of MSE/PSNR and SSIM follows common SR practice and addresses the known distortion-perception trade-off (Arabboev et al., 2024).

Results were additionally benchmarked against bicubic interpolation, the conventional “no-learning” baseline in SR (including remote-sensing SR) because it is deterministic, fast, and grounded in cubic convolution theory. Demonstrating improvements over bicubic therefore isolates the added value of learned (adversarial) models in reconstructing high-frequency detail beyond interpolation’s inherent smoothing (Keys, 1981; Wang et al., 2022).

Finally, inclusion of an adversarially trained model (GAN) enabled assessment of the extent to which perceptual optimization improves spatial realism relative to conventional interpolation and pixel-wise optimization approaches, showing that adversarial/perceptual losses can restore sharper, more realistic textures relative to pure MSE optimization, while acknowledging the perception-distortion trade-off in evaluation (Ledig et al., 2017).

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

Computational Environment and Reproducibility

All data preprocessing, bias correction, model training, and evaluation were conducted in a Windows 10 (64-bit) environment using Python (v3.9). The computational platform was a Dell G3 3579 workstation equipped with an Intel® Core™ i5-8300H CPU (2.30 GHz; 4 cores, 8 threads) and 24 GB RAM. Deep learning experiments were executed on an NVIDIA GeForce GTX 1050 GPU (4 GB VRAM) using NVIDIA driver version 560.94 and CUDA version 12.6.

Bias-correction and data-processing steps were implemented using NumPy (v1.23), pandas (v1.5), xarray (v2023.1), and SciPy (v1.9). Geospatial operations and visualization were performed using rasterio (v1.3), GeoPandas (v0.12), Matplotlib (v3.7), and Cartopy (v0.21). The deep learning architecture was implemented in PyTorch (v2.0) with CUDA acceleration enabled.

Training of the conditional GAN was performed for 200 epochs using a batch size of 4. On the GTX 1050 GPU (4 GB VRAM), training required approximately 8-10 hours of wall-clock time per model. Inference for a single monthly precipitation field required less than 3 seconds. These runtimes demonstrate that the framework can be executed on mid-range consumer hardware without requiring high-performance computing infrastructure.

To ensure reproducibility, fixed random seeds were applied across NumPy and PyTorch libraries. Deterministic training settings were enabled where supported by the backend framework. Training and validation splits were generated deterministically using predefined seeds to guarantee consistent data partitioning across runs. All hyperparameters, including learning rates, optimizer parameters ($\beta_1 = 0.5$, $\beta_2 = 0.999$), loss-function weights (λ_L1 and λ_perc), batch size, and epoch count, are explicitly defined in configuration files within the project repository.

The complete source code, including preprocessing scripts, QDM implementation, neural network architectures, configuration files, training routines, and evaluation utilities, is publicly available at https://doi.org/10.5281/zenodo.18746067 (Ibrahim et al., 2026). All input datasets (CMIP6 via ESGF, CHIRPS, MODIS, and GLDAS) are publicly accessible through the official portals cited in the Data Availability section.

3. Results and Discussion

The assessment framework adopted in this study evaluates post-processing performance across three complementary dimensions: (i) statistical fidelity, measured using bias, RMSE, Kolmogorov-Smirnov, and Wasserstein metrics; (ii) structural fidelity, evaluated using SSIM and visual reconstruction diagnostics; and (iii) spatial enhancement capability, assessed relative to interpolation baselines. This multi-metric strategy is necessary because no single metric adequately

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

characterizes the performance of climate-model post-processing methods, particularly for hydroclimatic fields containing sharp gradients and localized extremes.

3.1 QDM Results

Pre- and Post-Correction Model Performance

Evaluation of raw GCM outputs revealed significant discrepancies relative to observed hydroclimatic records, particularly for precipitation, surface runoff, and evapotranspiration (Table 3). Precipitation exhibited strong correlations ($R^2 = 0.84 - 0.92$) but large negative biases, with MPI-ESM1-2-HR showing the highest temporal correspondence ($R^2 = 0.917$) yet underestimating rainfall by $31.68mm/$ month, and CanESM5 yielding the smallest bias ($-4.57mm/$ month) but with slightly weaker correlation. Surface runoff was best captured by MPI-ESM1-2-HR ($R^2 = 0.742$, bias $= +0.45mm/month$), while HadGEM3 overestimated runoff substantially ($+4.22mm/$ month). Evapotranspiration was consistently overpredicted, with CanESM5 producing the highest bias ($+44.98mm/$ month). These findings confirm that raw GCMs, despite moderate-to-high correlations, contained substantial magnitude and distributional errors that limited their suitability for hydrological applications.

Table 3: Monthly domain-averaged evaluation metrics for the raw outputs of HadGEM3, MPI-ESM1-2-HR, and CanESM5.

Model	Variable	Correlation (R^2)	RMSE (mm)	Bias (mm/month)
HadGEM3	pr	0.855	57.253	-26.764
	mrros	0.467	7.505	+4.222
	evspsbl	0.577	42.749	+28.190
MPI-ESM1-2-HR	pr	0.917	52.442	-31.68
	mrros	0.742	0.930	+0.452
	evspsbl	0.844	34.279	+23.562
CanESM5	pr	0.840	48.569	-4.574
	mrros	0.334	2.265	+0.067
	evspsbl	0.632	58.102	+44.980

Following QDM correction, model performance improved markedly (Table 4). Precipitation biases narrowed to $+5.60 - 10.90mm/$ month and RMSE values dropped by nearly half ($23.79 - 27.84mm$). Surface runoff errors were reduced to near zero, with RMSE consistently below $2mm/month$. Evapotranspiration showed the most dramatic improvement, with mean biases converging to $\sim +1.7mm/$ month and RMSE reduced by over $85\%(3.71 - 4.30mm)$.

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

Table 4: Comparison of monthly domain-averaged Bias and RMSE values for raw (pre-correction) and QDM-adjusted (post-correction) outputs of HadGEM3, MPI-ESM1-2-HR, and CanESM5.

Model	Variable	Bias Pre (mm/month)	Bias Post (mm/month)	RMSE Pre (mm)	RMSE Post (mm)
HadGEM3	pr	−26.76	+5.60	57.25	24.91
	mrros	+4.22	−0.072	7.51	0.24
	evspsbl	+28.19	+1.69	42.75	4.30
MPI-ESM1-2-HR	pr	−31.68	+6.23	52.44	23.79
	mrros	+0.45	+0.068	0.93	0.27
	evspsbl	+23.56	+1.69	34.28	3.86
CanESM5	pr	−4.57	+10.90	48.57	27.84
	mrros	+0.07	+1.14	2.27	1.85
	evspsbl	+44.98	+1.69	58.10	3.71

QDM correction markedly reduced systematic errors across precipitation (**pr**), surface runoff (**mrros**), and evapotranspiration (**evspsbl**). Relative to the raw GCM fields, precipitation RMSE was nearly halved; runoff RMSE declined to < 2 mm month^{-1}; and evapotranspiration RMSE decreased by > 85%. Distributional agreement also improved: both the Kolmogorov-Smirnov statistic (D) and the 1-Wasserstein distance were reduced across variables (see Table 5).

Table 5: Domain-averaged performance metrics (Bias, RMSE, Kolmogorov-Smirnov D, and Wasserstein distance) after QDM correction for HadGEM3, MPI-ESM1-2-HR, and CanESM5.

Model	Variable	Mean_Obs	Mean_Corr	Domain_Bias	KS_D	KS_p	Wasserstein	RMSE
MPI-ESM1-2-HR	pr	71.867	78.092	6.226	0.0441	2.63E-05	6.263	23.785
	mrros	0.236	0.304	0.068	0.0601	1.74E-09	0.068	0.266
	evspsbl	12.842	12.842	1.693	0.0108	0.889	0.0178	3.862
HadGEM3-GC31-LL	pr	71.867	77.471	5.604	0.0579	0.0155	5.625	24.91
	mrros	0.236	0.164	-0.072	0.2319	1.56E-34	0.122	0.24
	evspsbl	12.842	12.841	1.692	0.0111	0.9999	0.0174	4.295
CanESM5	pr	71.867	82.761	10.895	0.0717	0.0468	11.862	27.842
	mrros	0.236	1.373	1.137	0.2696	1.35E-23	1.1379	1.845
	evspsbl	12.842	12.84	1.691	0.0111	1	0.0188	3.706

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

The corrected simulations reproduced seasonal cycles and interannual variability more faithfully, and Q-Q plots. Figures 2 showed closer adherence to the 1:1 line across quantiles. Collectively, these findings underscore the inadequacy of uncorrected GCM outputs for hydrological applications and demonstrate the efficacy of QDM in correcting both mean state and variability, thereby providing a robust basis for subsequent downscaling.

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

327

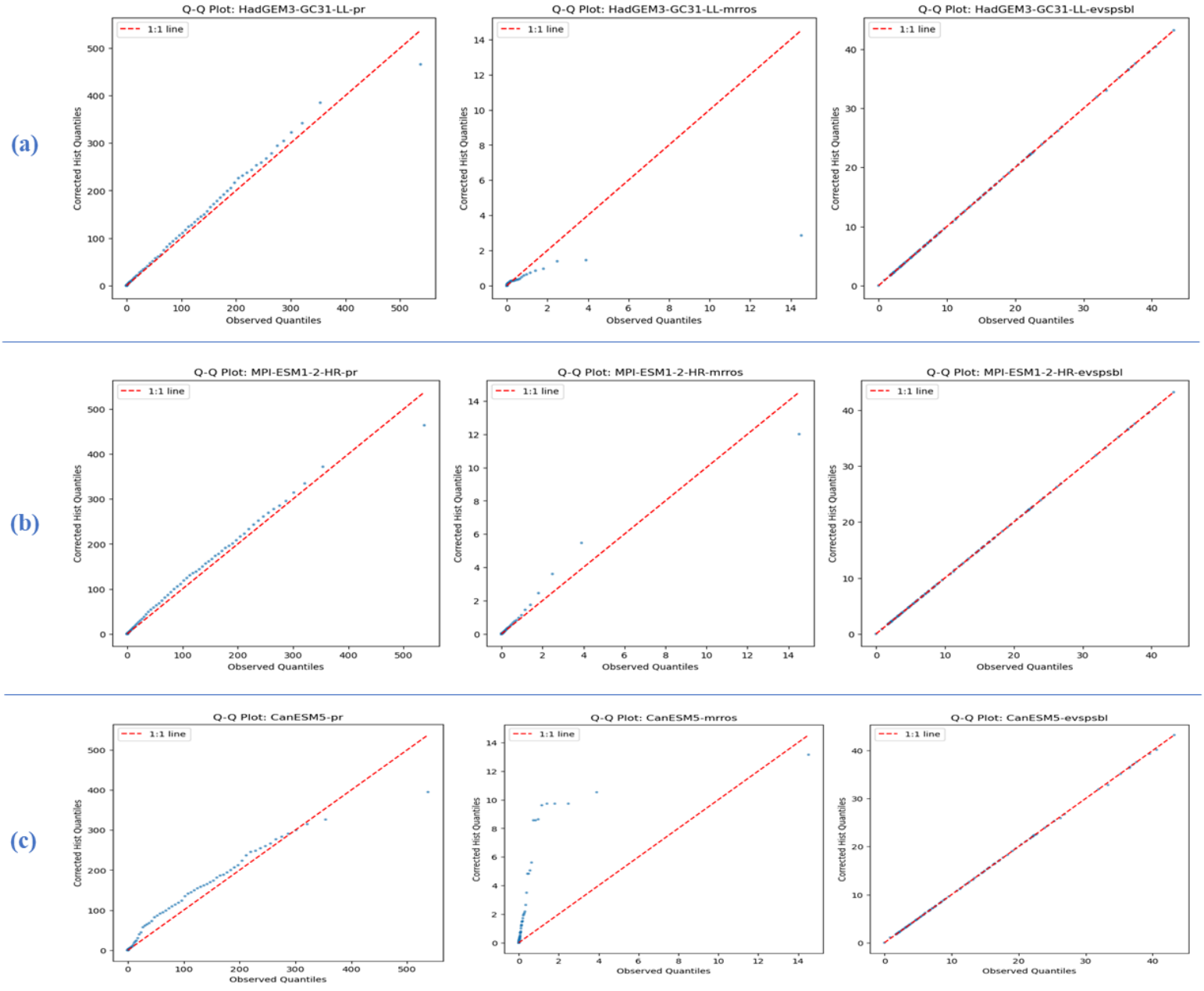


Figure 2: Quantile-quantile (Q-Q) plots comparing observed and QDM-corrected precipitation (mm month⁻¹), surface runoff (mm month⁻¹), and evapotranspiration (mm month⁻¹) for the (a) HadGEM3-GC31-LL, (b) MPI-ESM1-2-HR, and (c) CanESM5 models.

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

3.2 GAN Downscaling results

GAN Training and Evaluation Summary

Figure 3 show the training loss curves of the generator and discriminator. Over 200 epochs, the generator's loss steadily declined from ~88.7 to ~21.6, while discriminator loss remained consistently low (<0.08), indicating stable adversarial dynamics and the absence of mode collapse. Validation loss improved from 0.864 to 0.172, confirming good generalization. These trends suggest effective convergence and balanced adversarial training.

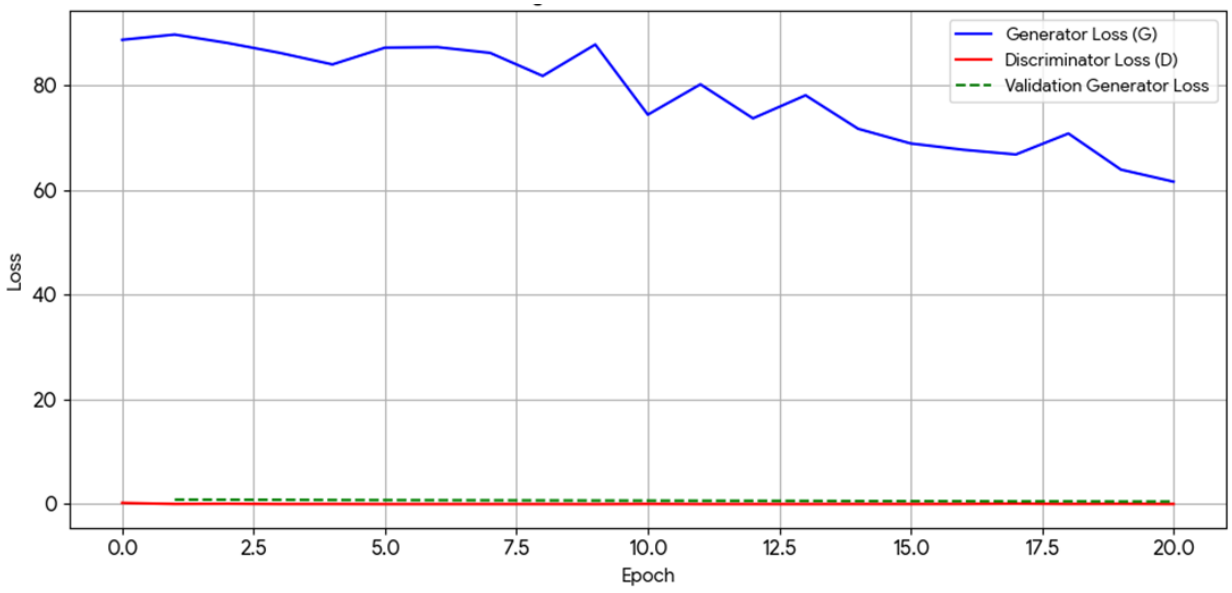


Figure 3 Training and Validation Loss Curve

Final validation on 180 LR-HR sample pairs yielded strong results: MSE = 0.028, PSNR = 16.6 dB, and SSIM = 0.888. The relatively high SSIM indicates strong structural preservation, while PSNR values, although moderate, are typical of perceptual super-resolution methods that prioritize realism over strict pixel fidelity.

As shown in Table 6, the GAN clearly outperformed interpolation baselines, reducing RMSE by ~22% relative to bicubic (9.1 to 7.0 mm/day) and improving SSIM from 0.65 to 0.89.

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

Table 6: ***Summary of validation metrics for bilinear, bicubic, and GAN-based downscaling.***

Method	RMSE (mm/day)	SSIM
Bilinear	9.4	0.62
Bicubic	9.1	0.65
GAN (proposed)	7.0	0.89

Although PSNR remained moderate (~16.6 dB), this outcome is expected when models optimize perceptual or adversarial objectives. Theory and experiments describe a perception versus distortion trade-off: methods that favor spatial realism often score lower on pixelwise distortion metrics such as MSE and PSNR, yet they produce outputs that and are more physically plausible spatial structure (Blau and Michaeli, 2018; Johnson et al., 2016).

The high SSIM (approximately 0.89) supports this interpretation because SSIM emphasizes structural similarity rather than exact pixel matching. In geophysical images, small phase shifts of narrow rain bands or orographic edges can inflate MSE and depress PSNR without degrading structural fidelity, which SSIM captures more faithfully (Wang et al., 2004). Preservation of structural organization is particularly important for hydroclimatic applications because runoff generation, infiltration, and hydrological connectivity are highly sensitive to the spatial coherence of precipitation fields rather than pixelwise agreement alone.

Adversarial and perceptual super resolution methods have also been shown to recover coherent small-scale features in Earth system data that interpolation misses, including topography aligned gradients and mesoscale organization. Studies report sharper, more realistic convective structures than baseline interpolation, consistent with our qualitative gains (Oyama et al., 2023; Reddy et al., 2023; Stengel et al., 2020). These results justify the claim that the GAN prioritized perceptual and structural realism and successfully reconstructed orographic gradients and convective patterns despite only moderate PSNR.

Interpolation baselines were included because they represent the operational standard for spatial resampling in many climate-impact workflows. Improvements beyond interpolation therefore indicate genuine recovery of unresolved spatial variability rather than simple grid refinement.

Visual and Structural Analysis

Figure 4 compares downscaled precipitation fields produced by bilinear, bicubic, and GAN methods against the 1 km reference. Interpolation approaches produced over-smoothed outputs, with blurred rainbands and diffuse convective patterns. These artifacts suppressed extremes and degraded spatial coherence, yielding low SSIM and PSNR (0.28 and 3.5 dB in selected cases).

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

In contrast, the GAN restored sharper gradients, convective cells, and textural variability. Minor textural artifacts, typical of GAN outputs, were observed but did not substantially degrade structural coherence. These results highlight the limitation of PSNR as a sole evaluation metric, as it penalizes perceptual detail. The GAN's higher SSIM confirms its superiority in reproducing spatial coherence and variability.

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

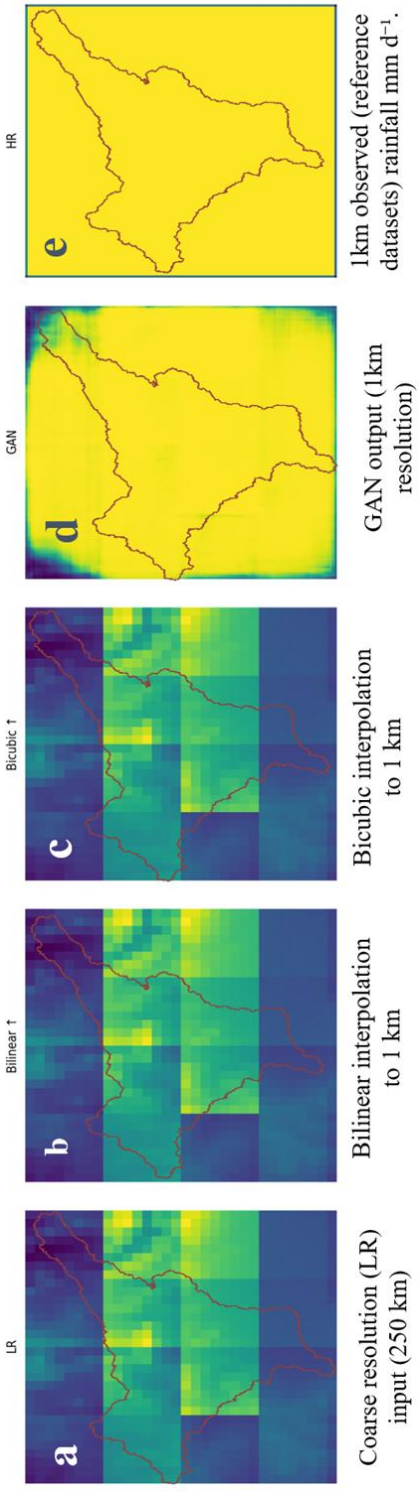


Figure 4. Visual comparison over selected validation patches: (a) input low-resolution GRACE field; (b) bilinear interpolation; (c) bicubic interpolation; (d) GAN-super-resolved (SR) output; and (e) high-resolution (HR) reference precipitation (mm month^{-1}).

373

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

Validation Against Observed Data

Further validation against reference datasets reinforced these results. As summarized in table 7, the GAN consistently outperformed bicubic interpolation across metrics, achieving a ~1.4 dB gain in PSNR and a substantial improvement in SSIM (0.62 to 0.888).

Table 7: ***Comparative validation metrics for bicubic interpolation and GAN-based downscaling.***

Method	MSE	PSNR (dB)	SSIM
Bicubic interpolation	0.037	15.2	0.62
GAN (proposed)	0.028	16.6	0.888

The slight reduction in validation PSNR relative to training indicates minimal overfitting, while the high SSIM demonstrates consistent structural preservation.

Limitations and Future Improvements

Two main limitations were observed. First, the moderate PSNR reflects a common tradeoff in adversarial super-resolution, where some pixel-level accuracy is exchanged for sharper and more realistic textures. Second, the model was trained and applied on single images, so temporal consistency was not enforced, even though the source GCM fields are spatiotemporal. As a result, consecutive maps can show small frame-to-frame jitter and less coherent evolution of storms. A third limitation concerns the absence of a formal component-wise ablation analysis. While the present study evaluates the integrated QDM-GAN framework against interpolation baselines, the isolated contributions of QDM-only and GAN-only configurations were not explicitly quantified. Although the complementary roles of bias correction and spatial super-resolution are well established in the literature, controlled experiments separating their marginal effects would provide a more rigorous assessment of additive benefits. This aspect is therefore identified as a priority for future investigation.

Future improvements should address these points directly. Spatiotemporal architectures should be adopted to learn dynamics and maintain consistency across time, and temporal discriminators or coherence losses should be introduced so that sequences evolve smoothly. Component-wise benchmarking and sensitivity analyses should be conducted to better quantify the relative contributions of statistical correction and learned spatial enhancement. Pixel losses should be combined with perceptual or feature-matching losses to reduce artifacts while keeping fine details, and the weights on L1 and adversarial terms should be adjusted during training to balance accuracy and realism. Physics-aware constraints and simple mass-balance checks should be included to preserve physical plausibility, and probabilistic outputs should be considered to quantify uncertainty.

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

Integrated Discussion

The integrated QDM and GAN framework highlights the complementary strengths of statistical bias correction and machine learning based spatial enhancement. QDM corrected climatological biases, and the GAN restored high frequency spatial features that interpolation methods could not reproduce. In quantitative terms, monthly evapotranspiration bias decreased from about +45 mm to approximately +1.7 mm after QDM, and the GAN yielded a 22% reduction in precipitation RMSE together with a clear increase in SSIM relative to bicubic interpolation.

These outcomes are consistent with studies that combine trend-preserving bias adjustment with deep learning for super-resolution downscaling, which report concurrent gains in mean-state accuracy, distributional fidelity, and spatial structure (Gergel et al., 2024; Oyama et al., 2023). Other work (Besombes et al., 2021; Rampal et al., 2024) also notes that GAN based downscaling can underrepresent extremes unless the training objective and conditioning signals are tuned to emphasize tails, which aligns with the mixed performance observed for extremes in this study.

Residual biases in precipitation and runoff, particularly for CanESM5, therefore remain a challenge for the reliable simulation of extremes, consistent with documented CMIP6 biases in precipitation and extremes as well as known CanESM5 regional biases (Abdelmoaty et al., 2021; Li et al., 2022; Swart et al., 2019; Teng et al., 2015).

Even with these limitations, the hybrid approach constitutes a substantial improvement over conventional interpolation and raw GCM outputs. The modular design facilitates extension to additional variables and regions and supports uncertainty aware hydrological impact assessments, which makes the method suitable for climate adaptation and water resources planning in data scarce environments.

The framework was evaluated across three structurally different CMIP6 GCMs to reduce dependence on any single model configuration and to assess methodological robustness across varying climatological biases. Although the framework improves deterministic reconstruction skill, uncertainty associated with GCM structural differences and adversarial downscaling remains non-negligible and should be propagated in downstream hydrological applications.

4. Conclusion

This study developed and evaluated a two-stage post-processing framework that integrates Quantile Delta Mapping (QDM) for non-stationary bias correction with a conditional Generative Adversarial Network (GAN) for statistical super-resolution of climate model outputs. The objective was to provide a reproducible and modular workflow capable of improving both distributional fidelity and spatial structure in coarse-resolution GCM simulations.

Application of QDM substantially reduced systematic biases in precipitation, runoff, and evapotranspiration, improving mean-state alignment and distributional consistency relative to observational datasets. These improvements were confirmed using conventional error metrics as well as Kolmogorov-Smirnov and Wasserstein distance diagnostics. Subsequent application of the GAN enhanced spatial realism by reconstructing finer-scale gradients and localized

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

variability that were not recoverable through interpolation alone. Compared with interpolation baselines, the integrated framework achieved lower RMSE and higher structural similarity for precipitation fields.

The results demonstrate that statistical bias correction and learned super-resolution address complementary dimensions of model error: QDM improves climatological consistency, while the GAN enhances spatial heterogeneity. When applied sequentially, these components provide improvements in both statistical and structural performance relative to raw GCM outputs and interpolation-based approaches.

The study emphasizes methodological transparency and reproducibility. All processing steps, hyperparameters, and evaluation procedures are documented, and the complete implementation is publicly available to support replication and extension. The framework is designed to be modular and adaptable, and additionally provides a transferable basis for assessing the statistical fidelity and spatial realism of machine learning-based climate downscaling approaches across different variables and regions, although further evaluation, including controlled ablation analyses, sensitivity testing, and incorporation of physical consistency constraints would strengthen future applications. The study therefore contributes not only a post-processing workflow, but also an assessment-oriented framework for evaluating statistical and structural consistency in climate-model refinement.

Overall, the proposed QDM-GAN workflow provides a reproducible post-processing tool for generating bias-corrected, high-resolution climate model fields suitable for downstream modeling tasks that require improved statistical consistency and spatial detail.

Code and data availability

The code used in this study is openly available via Zenodo at https://doi.org/10.5281/zenodo.18746067 (Ibrahim et al., 2026). The repository contains the full reproducible workflow implemented as documented Jupyter notebooks. Input datasets are publicly available and are not redistributed.

Author contributions

Abba Ibrahim: Conceptualization and Writing-original draft. Aimrun Wayayok: Conceptualization, review, and editing. Helmi Zulhaidi Mohd Shafri: Conceptualization, supervision. Noorellimia Mat Toridi: Conceptualization and supervision. Idris Wada Muhammd: Data collection and preprocessing.

Competing interests

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

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

Disclaimer

Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

Acknowledgements

The CMIP6 modeling groups are gratefully acknowledged for producing and making their simulations available. We also thank the providers of CHIRPS, GLDAS, and MODIS datasets for open access to their products. Finally, we acknowledge the open-source software community particularly contributors to Python's scientific ecosystem (e.g., xarray, NumPy, pandas, GeoPandas, rasterio, scikit-learn, and PyTorch/TensorFlow), whose tools were integral to this work.

Financial support

This research has been supported by the Petroleum Technology Development Fund (PTDF), Federal Republic of Nigeria, which funded the first author's PhD program.

References

Abdelmoaty, H. M., Papalexiou, S. M., Rajulapati, C. R., and AghaKouchak, A.: Biases Beyond the Mean in CMIP6 Extreme Precipitation: A Global Investigation, Earth's Futur., 9, e2021EF002196, https://doi.org/10.1029/2021EF002196, 2021.

Afrizal, T. and Surussavadee, C.: High-Resolution Climate Simulations in the Tropics With Complex Terrain Employing the CESM/WRF Model, Adv. Meteorol., 2018, 1–15, https://doi.org/10.1155/2018/5707819, 2018.

Aguilar, A., Flores, H. E. M., Crespo, G., Marin, M., Campos, I., and Calera, A.: Performance Assessment of MOD16 in Evapotranspiration Evaluation in Northwestern Mexico, Water, 10, 901, https://doi.org/10.3390/w10070901, 2018.

Ajjur, S. B.: Runoff Uncertainty Associated With Global Climate Model Chosen in Regional Climate Modeling, Proc. Int. Assoc. Hydrol. Sci., 385, 333–337, https://doi.org/10.5194/piahs-385-333-2024, 2024.

Arabboev, M., Begmatov, S., Rikhsivoev, M., Nosirov, K., and Saydiakbarov, S.: A comprehensive review of image super-resolution metrics: classical and AI-based approaches, Acta IMEKO, 13, 1–8, https://doi.org/10.21014/ACTAIMEKO.V13I1.1679, 2024.

Bandara, U., Agarwal, A., Srinivasan, G., Shanmugasundaram, J., and Jayawardena, I. M. S.: Intercomparison of Gridded Precipitation Datasets for Prospective Hydrological Applications in Sri Lanka, Int. J. Climatol., 42, 3378–3396,

https://doi.org/10.1002/joc.7421, 2021.
Beaudoing, H. and Rodell, M.: GLDAS Noah Land Surface Model L4 3 hourly 0.25 x 0.25 degree V2.1, https://doi.org/10.5067/E7TYRXPJKWOQ, 15 April 2024.
Belay, A. S., Fenta, A. A., Yenehun, A., Nigate, F., Tilahun, S. A., Moges, M. M., Dessie, M., Adgo, E., Nyssen, J., Chen, M., Griensven, A. v., and Walraevens, K.: Evaluation and Application of Multi-Source Satellite Rainfall Product CHIRPS to Assess Spatio-Temporal Rainfall Variability on Data-Sparse Western Margins of Ethiopian Highlands, Remote Sens., 11, 2688, https://doi.org/10.3390/rs11222688, 2019.
Besombes, C., Pannekoucke, O., Lapeyre, C., Sanderson, B., and Thual, O.: Producing realistic climate data with generative adversarial networks, Nonlinear Process. Geophys., 28, 347–370, https://doi.org/10.5194/npg-28-347-2021, 2021.
Bichet, A. and Diedhiou, A.: West African Sahel Has Become Wetter During the Last 30 Years, but Dry Spells Are Shorter and More Frequent, Clim. Res., 75, 155–162, https://doi.org/10.3354/cr01515, 2018.
Blau, Y. and Michaeli, T.: The Perception-Distortion Tradeoff, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 6228–6237, https://doi.org/10.1109/CVPR.2018.00652, 2018.
Calel, R. and Stainforth, D. A.: On the physics of three integrated assessment models, Bull. Am. Meteorol. Soc., 98, 1199–1216, https://doi.org/10.1175/BAMS-D-16-0034.1, 2017.
Cannon, A. J., Sobie, S. R., and Murdock, T. Q.: Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes?, J. Clim., 28, 6938–6959, https://doi.org/10.1175/jcli-d-14-00754.1, 2015.
Center, C. H.: Climate Hazards Center Infrared Precipitation with Stations version 3 (CHIRPS3) Data Repository, https://doi.org/10.15780/G2JQ0P, 1 April 2025.
Duan, Z., Tuo, Y., Liu, J., Gao, H., Song, X., Zhang, Z., Yang, L., and Mekonnen, D. F.: Hydrological Evaluation of Open-Access Precipitation and Air Temperature Datasets Using SWAT in a Poorly Gauged Basin in Ethiopia, J. Hydrol., 569, 612–626, https://doi.org/10.1016/j.jhydrol.2018.12.026, 2019.
ESGF, A. L.: Earth System Grid Federation AIMS2 Data Portal: CMIP6 Model Output Archive, https://aims2.llnl.gov/search, 1 April 2025.
Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., and Rummukainen, M.: Evaluation of climate models BT - Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Doschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, UK, 741–882, https://doi.org/10.1017/CBO9781107415324.020, 2013.
Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., and Michaelsen, J.: The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes, Sci. Data, 2, 150066, https://doi.org/10.1038/sdata.2015.66, 2015.
Gergel, D. R., Malevich, S. B., McCusker, K. E., Tenezakis, E., Delgado, M. T., Fish, M. A., and Kopp, R. E.: Global Downscaled Projections for Climate Impacts Research (GDPCIR): Preserving quantile trends for modeling future climate impacts, Geosci. Model Dev., 17, 191–227, https://doi.org/10.5194/gmd-17-191-2024, 2024.
Ghimire, U., Srinivasan, G., and Agarwal, A.: Assessment of Rainfall Bias Correction Techniques for Improved Hydrological Simulation, Int. J. Climatol., 39, 2386–2399, https://doi.org/10.1002/joc.5959, 2018.
Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y.: Generative adversarial nets, Adv. Neural Inf. Process. Syst., 27, 2014.
Goodwin, P. and Cael, B. B.: Bayesian Estimation of Earth's Climate Sensitivity and Transient Climate Response From Observational Warming and Heat Content Datasets, Earth Syst. Dyn., 12, 709–723, https://doi.org/10.5194/esd-12-709-2021, 2021.
Gruber, A., Scanlon, T., Schalie, R. v. d., Wagner, W., and Dorigo, W.: Evolution of the ESA CCI Soil Moisture Climate Data Records and Their Underlying Merging Methodology, Earth Syst. Sci. Data, 11, 717–739, https://doi.org/10.5194/essd-11-717-2019, 2019.
Horn, L., Hajat, A., Sheppard, L., Quinn, C., Colborn, J., Zermoglio, M. F., Gudo, E. S., Marrufo, T., and Ebi, K. L.:

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

Association Between Precipitation and Diarrheal Disease in Mozambique, Int. J. Environ. Res. Public Health, 15, 709, https://doi.org/10.3390/ijerph15040709, 2018.
Hudson, B.: Computationally-Efficient Climate Predictions Using Multi-Fidelity Surrogate Modelling, https://doi.org/10.48550/arxiv.2109.07468, 2021.
Ibrahim, A., Wayayok, A., Mohd Shafri, H. Z., Mat Toridi, N., and Wada, I.: A two-stage Bias-Correction and Super-Resolution Framework for Post-Processing Climate Model Outputs, https://doi.org/10.5281/zenodo.18746068, 2026.
Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A. A.: Image-to-Image Translation with Conditional Adversarial Networks, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 5967–5976, https://doi.org/10.1109/CVPR.2017.632, 2017.
Johnson, J., Alahi, A., and Fei-Fei, L.: Perceptual losses for real-time style transfer and super-resolution, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 694–711, https://doi.org/10.1007/978-3-319-46475-6_43, 2016.
Keys, R. G.: Cubic Convolution Interpolation for Digital Image Processing, IEEE Trans. Acoust., 29, 1153–1160, https://doi.org/10.1109/TASSP.1981.1163711, 1981.
Kim, H. W., Hwang, K., Mu, Q., Lee, S. O., and Choi, M.: Validation of MODIS 16 Global Terrestrial Evapotranspiration Products in Various Climates and Land Cover Types in Asia, Ksce J. Civ. Eng., 16, 229–238, https://doi.org/10.1007/s12205-012-0006-1, 2012.
Labeurthre, D., Reffet, A., and Schrapffer, A.: An Overview of Statistical Downscaling Methods : Techniques , Applications , and Advances State of the Art on Bias Correction Methods, 1–28, 2024.
Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., and Shi, W.: Photo-realistic single image super-resolution using a generative adversarial network, in: Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 105–114, https://doi.org/10.1109/CVPR.2017.19, 2017.
Li, Z., Liu, T., Huang, Y., Peng, J., and Ling, Y.: Evaluation of the CMIP6 Precipitation Simulations Over Global Land, Earth's Futur., 10, e2021EF002500, https://doi.org/https://doi.org/10.1029/2021EF002500, 2022.
Liu, Z., Huang, Y., Liu, T., Li, J., Xing, W., Akmalov, S., Peng, J., Pan, X., Guo, C., and Duan, Y.: Water balance analysis based on a quantitative evapotranspiration inversion in the Nukus irrigation area, Lower Amu River Basin, Remote Sens., 12, https://doi.org/10.3390/rs12142317, 2020.
Maraun, D.: Bias Correcting Climate Change Simulations - A Critical Review, Curr. Clim. Chang. Reports, 2, 211–220, https://doi.org/10.1007/s40641-016-0050-x, 2016.
Maraun, D., Shepherd, T. G., Widmann, M., Zappa, G., Walton, D., Gutiérrez, J. M., Hagemann, S., Richter, I., Soares, P. M. M., Hall, A., and Mearns, L. O.: Towards process-informed bias correction of climate change simulations, Nat. Clim. Chang., 7, 764–773, https://doi.org/10.1038/nclimate3418, 2017.
Merks, J., Photiadou, C., Ludwig, F., and Arheimer, B.: Comparison of Open Access Global Climate Services for Hydrological Data, Hydrol. Sci. J., 67, 2369–2385, https://doi.org/10.1080/02626667.2020.1820012, 2020.
Mirza, M. and Osindero, S.: Conditional generative adversarial nets, arXiv Prepr. arXiv1411.1784, 2014.
Nishant, N., Virgilio, G. D., Ji, F., Tam, E., Beyer, K., and Riley, M.: Evaluation of Present-Day CMIP6 Model Simulations of Extreme Precipitation and Temperature Over the Australian Continent, Atmosphere (Basel)., 13, 1478, https://doi.org/10.3390/atmos13091478, 2022.
Oyama, N., Ishizaki, N. N., Koide, S., and Yoshida, H.: Deep generative model super-resolves spatially correlated multiregional climate data, Sci. Rep., 13, 5992, https://doi.org/10.1038/s41598-023-32947-0, 2023.
Pandey, V., Srivastava, P. K., Singh, S. K., Petropoulos, G. P., and Mall, R. K.: Drought Identification and Trend Analysis Using Long-Term CHIRPS Satellite Precipitation Product in Bundelkhand, India, Sustainability, 13, 1042, https://doi.org/10.3390/su13031042, 2021.
Pierce, D. W., Cayan, D. R., Maurer, E. P., Abatzoglou, J. T., and Hegewisch, K. C.: Improved Bias Correction Techniques for Hydrological Simulations of Climate Change, J. Hydrometeorol., 16, 2421–2442, https://doi.org/https://doi.org/10.1175/JHM-D-14-0236.1, 2015.
Qing, P., Wang, R., Jiang, Y., Li, C., and Guo, W.: The Change of Hydrological Variables and Its Effects on Vegetation in Central Asia, Theor. Appl. Climatol., 146, 741–753, https://doi.org/10.1007/s00704-021-03730-w, 2021.
Rajulapati, C. R. and Papalexiou, S.: Precipitation Bias Correction: A Novel Semi-parametric Quantile Mapping

https://doi.org/10.5194/egusphere-2026-963
Preprint. Discussion started: 19 May 2026

Method, Earth Sp. Sci., 10, https://doi.org/10.1029/2023ea002823, 2023.
Rampal, N., Gibson, P. B., Sherwood, S., and Abramowitz, G.: On the Extrapolation of Generative Adversarial Networks for Downscaling Precipitation Extremes in Warmer Climates, Geophys. Res. Lett., 51, e2024GL112492, https://doi.org/https://doi.org/10.1029/2024GL112492, 2024.
Reddy, P. J., Matear, R., Taylor, J., Thatcher, M., and Grose, M.: A precipitation downscaling method using a super-resolution deconvolution neural network with step orography, Environ. Data Sci., 2, e17, https://doi.org/DOI: 10.1017/eds.2023.18, 2023.
Running, S., Mu, Q., Zhao, M., and Moreno, A.: MODIS/Terra Net Evapotranspiration Gap-Filled 8-Day L4 Global 500m SIN Grid V061, https://doi.org/10.5067/MODIS/MOD16A2GF.061, 6 May 2021.
Stengel, K., Glaws, A., Hettinger, D., and King, R. N.: Adversarial super-resolution of climatological wind and solar data, Proc. Natl. Acad. Sci., 117, 16805–16815, https://doi.org/10.1073/pnas.1918964117, 2020.
Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Hanna, S., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., Von Salzen, K., Yang, D., and Winter, B.: The Canadian Earth System Model version 5 (CanESM5.0.3), Geosci. Model Dev., 12, 4823–4873, https://doi.org/10.5194/gmd-12-4823-2019, 2019.
Teng, J., Potter, N. J., Chiew, F. H. S., Zhang, L., Wang, B., Vaze, J., and Evans, J. P.: How does bias correction of regional climate model precipitation affect modelled runoff?, Hydrol. Earth Syst. Sci., 19, 711–728, https://doi.org/10.5194/hess-19-711-2015, 2015.
Thao, S., Garvik, M., Mariethoz, G., and Vrac, M.: Combining Global Climate Models Using Graph Cuts, Clim. Dyn., 59, 2345–2361, https://doi.org/10.1007/s00382-022-06213-4, 2022.
Tong, Y., Gao, X., Han, Z., Xu, Y., Xu, Y., and Giorgi, F.: Bias correction of temperature and precipitation over China for RCM simulations using the QM and QDM methods, Clim. Dyn., 57, 1425–1443, https://doi.org/10.1007/s00382-020-05447-4, 2021.
Ugolotti, A., Anders, T., Lanssens, B., Hickler, T., François, L., and Tölle, M.: Impact of Bias Correction on Climate Change Signals Over Central Europe and the Iberian Peninsula, Front. Environ. Sci., 11, https://doi.org/10.3389/fenvs.2023.1116429, 2023.
Wada, I. M., Usman, H. S., Nwankwegu, A. S., Usman, M. N., and Gebresellase, S. H.: Selection and downscaling of CMIP6 climate models in Northern Nigeria, Theor. Appl. Climatol., 153, 1157–1175, https://doi.org/10.1007/s00704-023-04534-w, 2023.
Wang, F. and Tian, D.: On Deep Learning-Based Bias Correction and Downscaling of Multiple Climate Models Simulations, Clim. Dyn., 59, 3451–3468, https://doi.org/10.1007/s00382-022-06277-2, 2022.
Wang, X., Yi, J., Guo, J., Song, Y., Lyu, J., Xu, J., Yan, W., Zhao, J., Cai, Q., and Min, H.: A Review of Image Super-Resolution Approaches Based on Deep Learning and Applications in Remote Sensing, https://doi.org/10.3390/rs14215423, 2022.
Wang, Y.: Prediction of Future Urban Rainfall and Waterlogging Scenarios Based on CMIP6: A Case Study of Beijing Urban Area, Water, 15, 2045, https://doi.org/10.3390/w15112045, 2023.
Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P.: Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process., 13, 600–612, https://doi.org/10.1109/TIP.2003.819861, 2004.
Yu, H., Chen, J., Xu, C., Xiong, L., and Chen, H.: Bias Nonstationarity of Global Climate Model Outputs: The Role of Internal Climate Variability and Climate Model Sensitivity, Int. J. Climatol., 39, 2278–2294, https://doi.org/10.1002/joc.5950, 2018.