Geosci. Model Dev., 19, 11225–11242, 2026
https://doi.org/10.5194/gmd-19-11225-2026
© Author(s) 2026. This work is distributed under
the Creative Commons Attribution 4.0 License.

Geoscientific Model Development
Open Access
EGU

A differentiable framework for general circulation model precipitation bias correction

Kamlesh Sawadekar[1], **Seth McGinnis**[2], **Peijun Li**[1], **Kathryn Lawson**[1], **and Chaopeng Shen**[1]

[1]Department of Civil and Environmental Engineering, The Pennsylvania State University, University Park, Pennsylvania, 16801, USA
[2]Computational & Information Systems Laboratory, National Center for Atmospheric Research, Boulder, Colorado, 80307, USA

Correspondence: Chaopeng Shen (cshen@engr.psu.edu)

Received: 3 May 2026 – Discussion started: 20 May 2026
Revised: 7 October 2026 – Accepted: 26 October 2026 – Published: 25 November 2026

Abstract. Systematic biases in General Circulation Model (GCM) outputs limit their direct applicability in regional planning, making bias correction a technically demanding but necessary step for both short-term and long-term impact assessment. Correcting precipitation is particularly challenging due to its non-Gaussian distribution, intermittent nature, and heavy-tailed extremes. However, traditional statistical bias-correction methods have limited ability to learn systematic patterns from large datasets or generalize to new locations. While machine learning (ML) provides greater flexibility, it can produce unpredictable and difficult-to-interpret results, limiting generalization across GCMs and locations. In this study, we propose a differentiable bias-adjustment framework called δCLIMBA, or dCLIMBA, that learns a spatiotemporally adaptive parametric bias-adjustment procedure, rather than corrected precipitation directly, between historical CMIP6 model outputs and a gridded observation-based dataset, Livneh. Results demonstrate that the proposed method corrects the magnitude and distribution of extreme precipitation with particularly strong performance in the upper tail but underestimates events in the far tail. The quantile distribution of precipitation was well reproduced across diverse US cities, and spatial patterns were comparable to those from the widely used LOCA2 statistical downscaling product. In addition, the framework showed partial future trend preservation and promising attenuation of marginal biases in unseen regions. This work presents a modular and efficient bias-correction approach. The differentiable approach provides an easy-to-use option for connecting atmospheric-model outputs to on-the-ground impacts.

1 Introduction

General Circulation Models (GCMs) are essential tools for projecting future atmospheric regime change, but their outputs are subject to systematic biases that limit their direct use in regional impact assessments. These models simulate interactions between the atmosphere, hydrosphere, cryosphere, and biosphere, and are used by policymakers and scientists to develop mitigation and adaptation strategies (IPCC, 2023). Our research focuses on simulations from the Coupled Model Intercomparison Project (CMIP), which produces centennial climate projections from multiple GCMs (Eyring et al., 2016). Fundamental insights from these simulations guide policymakers in the fields of water, energy, and biodiversity, but their applicability for local-scale impact studies is hindered by their coarse spatial resolution and intrinsic biases (Palmer and Stevens, 2019). These systematic biases, which critically limit the direct use of GCMs in regional impact assessments, stem from multiple factors, including structural errors and simplifications, inaccurate parameter values, and uncertainties in forcings and boundary conditions. Approximations of sub-grid-scale processes such as cloud formation, convection, and land-atmosphere interactions create errors that may be negligible at the GCM scale (Hourdin et al., 2017), but result in systematic biases when disaggregated or interpolated to finer scales; these biases can

be particularly troublesome in regions with complex topography or heterogeneous land surface properties where coarse-scale GCMs struggle to capture local variability (Giorgi and Mearns, 2002).

To enhance the applicability of GCMs at finer scales, laborious and technically demanding bias correction techniques (also referred to as bias adjustment) have often been used to align GCM outputs with chosen observational datasets. Bias adjustment is traditionally carried out by methods like quantile mapping, linear scaling, and regression-based transfer functions (Piani et al., 2010). Though these methods are very popular, their utility is impaired by their dependence on linear assumptions and inability to capture complex, nonlinear dependencies between atmospheric variables (Ehret et al., 2012). A major limitation of these methods is that they usually depend on the assumption of a stationary relationship between model outputs and observations. As atmospheric dynamics change over time, the calibrations learned on historical data become inaccurate for future conditions, leading to errors in extreme event projections (Deser et al., 2012). Traditional methods are also applied on a point-by-point basis with no regard for spatial coherence, but maintaining realistic spatial patterns is also important, e.g., for understanding storm clustering (Hess et al., 2023; Lovejoy et al., 1987). In addition, unless special measures are taken, adjusting bias based on stationary relationships can distort the physical consistency and the climate-change signal of the adjusted fields (Buser et al., 2009; Ehret et al., 2012).

Machine learning (ML)-based bias correction has recently progressed beyond traditional quantile mapping toward architectures explicitly tailored to correcting errors (Hess et al., 2023; Pan et al., 2021; Wang et al., 2023). For example, Hess et al. (2023) used an adversarial domain adaptation framework to learn a mapping between historical GCM precipitation and observations that corrects biases while preserving spatiotemporal coherence. Pan et al. (2021) demonstrated that generative adversarial networks can improve daily precipitation fields from CMIP6-class Earth system models, reducing systematic errors in means, extremes, and spatial patterns compared to standard statistical corrections. Wang et al. (2023) developed a customized convolutional neural network that incorporates physically relevant covariates, multitask learning, and weighted loss functions, showing enhanced skill relative to conventional approaches in downscaling and bias-correcting hourly precipitation, particularly for extreme events. Together, these ML-based frameworks illustrate the potential of flexible, data-driven bias correction of GCM outputs.

Despite these successes, purely data-driven ML approaches face challenges like generalization and interpretability. Like their statistical counterparts, ML approaches are trained on historical datasets, and hence are prone to overfitting and may fail to generalize across multiple GCMs. They can produce outputs in opaque ways – it is unclear what exact transformations they apply, when they are reliable, and whether they have failed silently. ML training can result in the disruption of overall physical consistency due to the lack of physical or (in case of precipitation) hydrological constraints (Harder et al., 2022; Maraun et al., 2010). Their black-box nature makes it difficult to ensure that their outputs align with established physical principles (Harder et al., 2022). ML models also require large amounts of training data, which can be scarce or inconsistent in certain regions, reducing their robustness for global applications.

To obtain an interpretable bias corrector, we adopt differentiable modeling frameworks, which integrate prescribed equations with data-driven models. "Differentiability" refers to the computing paradigm that, with any method, rapidly and accurately produces gradients to achieve the large-scale optimization of the combined system (Shen et al., 2023). These frameworks allow for end-to-end training of neural networks that are embedded into a system that adheres to physical principles such as conservation of energy and mass (Tsai et al., 2021). Differentiable programming has already demonstrated success in hydrological modeling and ecological simulations, offering a pathway to more interpretable and generalizable models (Aboelyazeed et al., 2023; Baste et al., 2025; Bindas et al., 2024; Fang and Gentine, 2024; Feng et al., 2022; Huynh et al., 2025; Ji et al., 2025; Li et al., 2024a, b; Rahmani et al., 2023; Song et al., 2026). Such frameworks have already shown potential in understanding and mitigating some biases in reanalysis datasets through parametric fusion informed by hydrology (Sawadekar et al., 2025).

Here, we develop a Differentiable CLImate Model Bias Adjustment (δCLIMBA) model for daily precipitation simulated by GCMs. This method applies a parameterized transformation optimized over matching quantiles. This framework can be applied to any GCM and observational dataset, and the trained neural network can be easily applied in space and time. The framework also offers the potential to study the biases it corrects. The proposed framework improves bias adjustment of GCM outputs for variables including precipitation, which is one of the most difficult variables to bias correct or downscale due to its intermittent nature and highly skewed, non-Gaussian distribution (Hess and Boers, 2021). By leveraging large-scale gridded observation datasets, this framework learns spatial and temporal patterns of bias while incorporating the requisite constraints on the outputs. We also leverage the differentiable framework to bias correct precipitation in unseen regions (those without extensive ground truth data) and learn about its capability to generalize spatially since we explicitly integrate topographical attributes into our model.

To achieve this, the study focuses on the following research questions:

- Can the framework generalize across space and across GCMs while preserving trends?

– How well does the framework attenuate marginal biases in intensity, extremes, and threshold exceedances?

– Does the framework distort spatial patterns in precipitation?

2 Methods and data

We used two jointly connected temporal and spatial encoders to predict coefficients for a monotonic softplus transformation of precipitation from GCM outputs, with the GCM itself and a small set of attributes serving as model inputs (Fig. 1). This transformation is "learned" against a quantile-based loss function (along with a few additional complementary loss functions). We evaluate the overall utility of the bias-adjusted GCM output using a number of Expert Team on Climate Change Detection and Indices (ETCCDI, details in Sect. 2.4) precipitation indices.

2.1 Differentiable bias adjustment model (δCLIMBA)

The bias adjustment here uses a differentiable model framework inspired by Feng et al. (2022) and Li et al. (2025), wherein a neural network is connected to a process-based model to improve parameterization. However, unlike these differentiable models, δCLIMBA is not coupled to a process-based atmospheric or hydrological model, and neither modifies nor returns information to the parent GCM. Differentiability here refers to the property that the complete forward computational chain from the input GCM, fed and propagated through the transformation network to the loss function, is differentiable with respect to the trainable weights of neural networks utilized here. The variables with respect to which the system is differentiated are the network weights, not the climate variables themselves. The philosophy behind this methodology is to parameterize the transformation of GCM precipitation by optimizing the match between quantiles, with the aim of learning interpretable and generalizable mappings that can be ported to untrained locations.

The δCLIMBA model comprises two neural networks: NN_1 and NN_2. NN_1 is a temporal encoder which is responsible for learning biases based on temporal patterns. We evaluated several candidates for NN_1 including a simple Multi-Layer Perceptron (MLP) (Li et al., 2020; Rumelhart et al., 1986), a Long Short-Term Memory (LSTM) model (Feng et al., 2020; Hochreiter and Schmidhuber, 1997), and a one-dimensional Convolutional Neural Network (CNN-1D) (Feng et al., 2021; LeCun et al., 1989; Li et al., 2024a) that would generate parameters to facilitate the GCM's transformation. NN_2 is an attention-based (Vaswani et al., 2017) spatial encoder that learns spatial relationships between neighboring inputs. The spatial self-attention is performed independently at each time step over the nodes within a patch, where queries, keys, and values are derived from node features and augmented with a learnable, head-specific offset

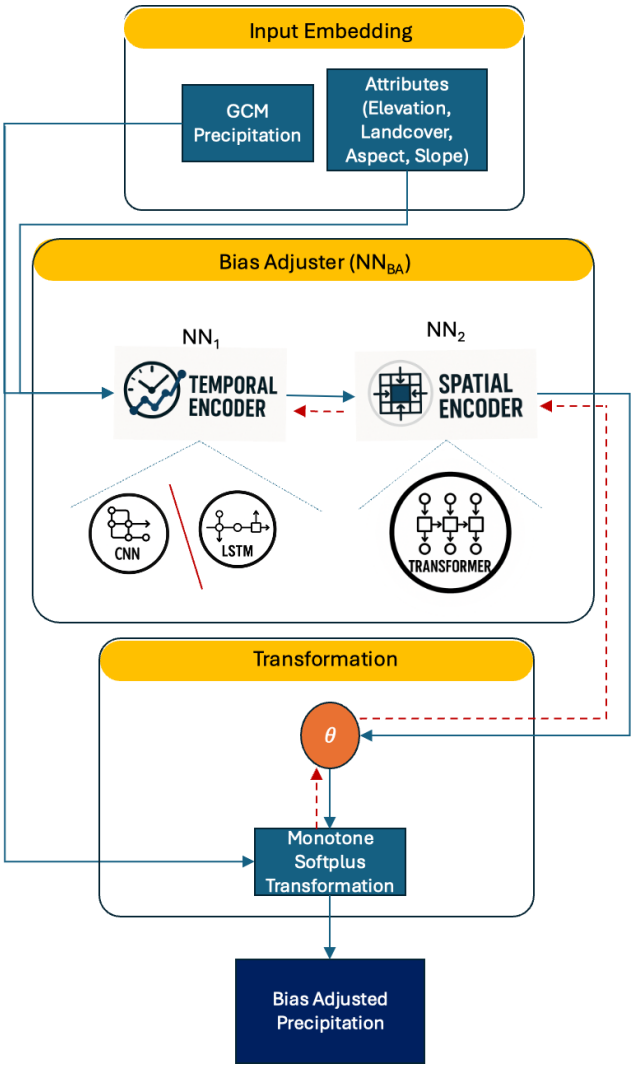

Figure 1. Schematic of the δCLIMBA framework for differentiable precipitation bias adjustment. Raw General Circulation Model (GCM) precipitation and static spatial attributes (elevation, landcover, aspect, slope) are embedded and passed to a bias-adjustment network (NN_{BA}). The network consists of temporal and spatial encoders which jointly estimate parameters θ governing a monotonic transformation ($f(x)$). The transformation module applies a learnable, differentiable monotone mapping to produce bias-adjusted precipitation based on the feedback from the composite quantile loss function. Gradients of the composite loss propagate backward through the transformation to the encoder weights; the schematic contains no process-based model component and no feedback to the parent GCM.

that encodes pairwise geodesic relationships (relative displacement, distance, and bearing). This geometry-aware formulation lets the model weight neighboring nodes according to physically meaningful spherical geometry rather than fixed or isotropic spatial kernels, akin to geometry-aware and geodesic self-attention mechanisms proposed for point clouds and images (Guo et al., 2020; Miyato et al., 2023).

For each grid cell, we chose the 16 nearest neighbors whose time series are positively correlated with that of the target grid cell. Our final model configuration uses a CNN-1D with 2 hidden layers of size 64 for NN_1, and a Transformer with a self-attention mechanism employing 2 attention heads and 64 model (hidden) dimensions for NN_2. LSTM can be used for NN_1 instead, though we went with the CNN-1D. We trained the model using the Adam optimizer (Kingma and Ba, 2017) with a learning rate of 1×10^{-4}. The batch size was chosen to be 5, and the time series length was 365 (days). This configuration was chosen based on overall performance across different precipitation indices. The full process can be mathematically described as:

$$\mathrm{NN_{BA}}\left(\mathbf{X}, \mathbf{X}_{t-1}, \mathbf{X}_{t-2}, \mathbf{X}_{t-3}, \mathbf{A}, \mathbf{I}_t\right) = [\boldsymbol{\alpha}, \mathbf{w}, \mathbf{s}, \mathbf{b}, \mathbf{c}] \equiv \boldsymbol{\theta} \quad (1)$$

$$\mathbf{X}_{\mathrm{BA}} = \boldsymbol{\alpha}\mathbf{X} + \left(\sum_{z=1}^{Z} \mathbf{w}_z \mathrm{softplus}\left(\mathbf{s}_z\left(\mathbf{X} - \mathbf{b}_z\right)\right)\right) + \mathbf{c}. \quad (2)$$

Here the bias adjustment neural network, NN_{BA}, is the jointly connected network of NN_1 and NN_2. $\mathbf{X}$ is the matrix of GCM precipitation with dimensions (batch size: 5, neighbors: 16, time: t (1–365)). $\mathbf{A}$ is the list of static attributes that could logically influence precipitation bias i.e., elevation, slope, aspect, and landcover. $\mathbf{I}_t$ is a wet day indicator where $\mathbf{I}_t = 1$ if daily precipitation $(\mathbf{X}) \geq 1$ and $\mathbf{I}_t = 0$ otherwise. The output of NN_{BA} is a set ($\boldsymbol{\theta}$) of temporally- and spatially varying parameters ($\boldsymbol{\alpha}$, $\mathbf{w}_z$, $\mathbf{s}_z$, $\mathbf{b}_z$, $\mathbf{c}$) for the "softplus monotone basis" mapping, which enforces monotonicity through softplus-based basis bumps (Zhou, 2016). We use the softplus function (Dugas et al., 2001) to ensure positivity while maintaining smooth gradients, avoiding optimization issues associated with non-differentiable or saturating alternatives. Z is the number of basis bumps (a hyperparameter, tuned to 8 here). $\mathbf{X}_{\mathrm{BA}}$ is the resulting bias-corrected GCM precipitation.

We used the following composite loss function during training. Here the loss function is the scalar objective that the training procedure minimises with respect to the trainable network weights:

$$L = p_1 Q + p_2 R + p_3 S. \quad (3)$$

Here, L (the total loss) combines three components: quantile-based loss (Q), rainy-day loss (R), and spatial correlation loss (S). These components were weighted with $p_1 = 0.99$, $p_2 = 0.01$ and $p_3 = 1$ to constrain the values within similar ranges. This formulation was designed to ensure overall distributional similarity while also paying attention to spatial structures and rainy-day occurrence during model training. A description of individual loss components follows:

$$Q = \frac{1}{K}\sum_{k=1}^{K} g\left(q_k\right) \cdot \left|Q_X\left(q_k\right) - Q_Y\left(q_k\right)\right|. \quad (4)$$

Here $Q_X(q_k)$ is the value of the empirical quantile function (inverse CDF) of data $\mathbf{X}$ at quantile level q_k; $Q_Y(q_k)$ represents the same but for reference data $\mathbf{Y}$. K is the total number of quantile levels used (1000), and q_k refers to the uniformly-spaced quantile levels. The weights $g(q_k)$ allow emphasis on specific quantiles; we used the kernel function

$$g\left(q_k\right) = \exp\left(-\left|q_k - q^*\right|\right) \quad (5)$$

where q^* is the quantile we want to emphasize. This is particularly useful for precipitation, where it is important to correct bias in the extremes as well as the bulk of the distribution. If no emphasis is specified, then $g(q_k) = 1$ (uniform). We treated q^* as a hyperparameter in our training, choosing a value of either 0.5 (bulk of the precipitation distribution) or 0.9 (tail of the distribution) for each GCM based on performance (other values were tested).

The second loss component is the rainy-day loss function, which is a soft constraint formulated to preserve the frequency of wet/dry days in the data. As a note, if this frequency shifts over time, the monotone transformation used should force the model to learn this shift based on the loss feedback.

$$R = \frac{1}{N}\sum_{i=1}^{N}\left|\sum_{t}^{T}\sigma\left(X_{it} - \tau\right) - \sum_{t}^{T}\sigma\left(Y_{it} - \tau\right)\right| \quad (6)$$

Here, X_{it} and Y_{it} are bias-adjusted and reference data, respectively, for site i and time t. σ is the sigmoid function which provides a differentiable approximation of the indicator function for rainy/wet days. τ is the wet day threshold (1 mm). N is the number of samples or sites.

The third loss component is the spatial correlation loss, which is a constraint added to preserve spatial structures in the bias-corrected simulation.

$$S = \frac{1}{BT}\sum_{b=1}^{B}\sum_{t=1}^{T}\left(1 - \mathrm{correlation}\left(\mathbf{X}_{b,t}, \mathbf{Y}_{b,t}\right)\right) \quad (7)$$

$$\mathrm{correlation}\left(\mathbf{X}_{b,t}, \mathbf{Y}_{b,t}\right) = \frac{\sum_{p=1}^{P}\left(Y_{b,p,t} - \overline{Y}_{b,t}\right)\left(X_{b,p,t} - \overline{X}_{b,t}\right)}{\sqrt{\sum_{p=1}^{P}\left(Y_{b,p,t} - \overline{Y}_{b,t}\right)^2 + \epsilon}\sqrt{\sum_{p=1}^{P}\left(X_{b,p,t} - \overline{X}_{b,t}\right)^2 + \epsilon}} \quad (8)$$

Here, $\boldsymbol{X}_{b,t}$ and $\boldsymbol{Y}_{b,t}$ are bias-adjusted and reference data, respectively, for batch b and time t. B is the total number of batches, T is the total number of time steps, and P is the number of neighbors for each of the grid cells in B, with iterator p. $\overline{X}_{b,t}$ and $\overline{Y}_{b,t}$ are the means over the P neighbours. ϵ is a small constant (i.e., 10^{-8}) to prevent division by zero.

Because every GCM has a distinctive, characteristic large-scale bias pattern (Krinner and Flanner, 2018), no single bias adjustment transformation is suitable for all GCMs. We therefore performed multiple training runs for each GCM, varying hyperparameters like q^* and the number of training epochs. For every candidate configuration, we first checked whether the quantile loss was monotonically decreasing over

the epochs and then calculated the ETCCDI indices (described in Sect. 2.4) for the validation period. Next, we computed a composite score defined as the mean percentage bias across all indices and selected the hyperparameter set with the lowest composite score for that GCM.

2.2 Datasets

We employed daily precipitation data from Coupled Model Intercomparison Project Phase 6 (CMIP6) (Eyring et al., 2016) model outputs from historical experiments and future scenarios with high emissions and stronger reliance on fossil fuels, i.e., Shared Socioeconomic Pathways (SSP5-8.5). For the current scope, the following six CMIP6 models were selected: ACCESS-CM2, GFDL-ESM4, IPSL-CM6A-LR, MIROC6, MPI-ESM1-2-LR, and MRI-ESM2-0 for bias correction of their precipitation. Each of these models was used at their native resolution (1 to 2.5°; see Table 1) within the conterminous US (CONUS) region at the daily timescale for historical simulation.

We considered using gridded Livneh (Livneh et al., 2013) data as a reference for comprehensive evaluation of our model over the entire CONUS. However, we used a revised version called unsplit-Livneh (Pierce et al., 2021) for training our model since it eliminates the problematic time adjustment applied in the original dataset and was also used for calibration of Localized Constructed Analog – 2 (LOCA2). To enable comparison at the GCM scale, the unsplit Livneh dataset was aggregated (spatially) from the original resolution of 1/16° to the coarser resolution of the respective CMIP6 members via nearest-neighbor interpolation.

The Localized Constructed Analogs (LOCA2) dataset (Pierce et al., 2014, 2021) is considered as a very important resource for downscaled products of CMIP6 (Vano et al., 2020). It presents a downscaled product with substantially improved resolution of 1/16° grid as it was calibrated using unsplit-Livneh. LOCA2 leverages spatial analog approaches to match patterns from regional- to local-scale features. Due to its ubiquity in our target CONUS region and thorough vetting (Ullrich, 2023), we decided to use it as a benchmark product against our differentiable approach along with other bias correction methods. To use this, we upscaled LOCA2 to GCM's spatial resolution to evaluate purely for systematic biases.

Furthermore, input features i.e., elevation, aspect, slope, and landcover were used to assist precipitation bias adjustment. Elevation, slope, and aspect were extracted from Digital Elevation Models at 0.01° resolution from NASA Shuttle Radar Topography Mission Global 1 arcsec V003 dataset (SRTMGL1) (NASA JPL, 2013). Landcover data was produced from Landsat at 30 m US coverage and downloaded from CEC's (https://www.cec.org/north-american-land-change-monitoring-system/, last access: 18 November 2025, NALCMS) (Pasos, 2020). All these data features were regridded to respective CMIP6 GCM grids to match the domain.

2.3 Experiment design

We trained the model in the period 1979–2000 and validated it in the period 1965–1978. We then tested bias-corrected outputs of the GCMs over the period 2001–2014. For the evaluation of future scenarios, we chose Shared Socioeconomic Pathways (SSP5-8.5) (Intergovernmental Panel On Climate Change (Ipcc), 2023) for the period 2015–2099. We interpolated the reference data (unsplit-Livneh) using nearest-neighbor interpolation to the resolution of each GCM before using it for modeling. We evaluated the model by comparing it to a set of five bias-correction methodologies (Table 2).

We performed our experiments on the Perlmutter system at the National Energy Research Scientific Computing Center (NERSC). Training our model on six GCM members across the CONUS for 100 epochs on different hyperparameter combinations distributed over 4 A100 GPUs in parallel took 7 h. After training, it only took 2 min to produce a bias-corrected simulation across the whole CONUS for the testing and future periods. This asymmetry between a single amortised training cost and a negligible inference cost is a direct consequence of learning a shared parameterisation network rather than fitting an independent correction at every grid cell, and it is what makes application to additional periods, scenarios, and unseen regions inexpensive.

These bias correction methods could in principle be recast as differentiable parameterised operators and optimised with comparable objectives, and we do not claim otherwise. However, methods like Quantile Mapping rely on conventional forms of operations that are not differentiable like sorting, empirical-CDF lookup, piecewise interpolation and hard wet-day thresholds. Any new reformulations of these operations would require smooth surrogates, for example soft sorting or smoothed indicator functions, and the resulting estimator would no longer be identical to the conventional implementation it replaces. δCLIMBA should therefore be read as one differentiable parameterisation within a broader class of bias-adjustment procedures that could be made differentiable, rather than as a claim that conventional methods are inherently unsuited to gradient-based optimisation.

2.4 Evaluation design and metrics

To evaluate our model, we focused on investigating the characteristics of bias-corrected simulations through Expert Team on Climate Change Detection and Indices (ETCCDI) for marginal biases, fractal dimension for spatial patterns/-clustering, and trend bias percentage.

To evaluate the temporal structure and marginal biases of the corrected data, we calculated the ETCCDI indices listed in Table 3. To place the different GCMs and bias adjustment

Table 1. Description of General Circulation Models (GCMs).

Short name	Full name	Institution/consortium	Native resolution
ACCESS-CM2	Australian Community Climate and Earth System Model version 2	CSIRO/Bureau of Meteorology/Australian universities (ACCESS community) (Bi et al., 2020)	$\sim 1.25° \times \sim 1.87°$
GFDL-ESM4	Geophysical Fluid Dynamics Laboratory Earth System Model 4	NOAA Geophysical Fluid Dynamics Laboratory (NOAA-GFDL) (Krasting et al., 2018)	$\sim 1.00° \times \sim 1.25°$
IPSL-CM6A-LR	Institut Pierre-Simon Laplace Climate Model version 6A, low resolution	Institut Pierre-Simon Laplace (IPSL), France (Boucher et al., 2018)	$\sim 1.26° \times \sim 2.50°$
MIROC6	Model for Interdisciplinary Research on Climate version 6	MIROC consortium: JAMSTEC, AORI (Univ. Tokyo), NIES, R-CCS Kobe, others (Mochizuki et al., 2019)	$\sim 1.40° \times \sim 1.40°$
MPI-ESM1-2-LR	Max Planck Institute Earth System Model version 1.2, low resolution	Max Planck Institute for Meteorology (MPI-M), Germany (Gutjahr et al., 2019)	$\sim 1.87° \times \sim 1.87°$
MRI-ESM2-0	Meteorological Research Institute Earth System Model version 2.0	Meteorological Research Institute (MRI), Japan (Kawai et al., 2019)	$\sim 1.12° \times \sim 1.12°$

Table 2. Summary of the bias adjustment methods used for evaluation. All methods except LOCA2 (reference comparison) were implemented using the "ibicus" framework proposed in Spuler et al. (2024).

Method	Description
Quantile Mapping (QM)	Statistical bias correction that adjusts the modeled distribution so its quantiles match those of observations (Cannon et al., 2015; Maraun, 2016).
ISIMIP (ISIMIP3BASD)	Trend-preserving bias adjustment and statistical downscaling framework developed for the Inter-Sectoral Impact Model Intercomparison Project; uses parametric/nonparametric quantile mapping tailored to each variable, with explicit preservation of trends across quantiles and options for multivariate/space–time structure (Lange, 2019).
ECDFM	Nonparametric quantile mapping where empirical CDFs are used and the correction is applied as an additive (or multiplicative) "distance" between modeled and observed quantiles (Li et al., 2010; Wang and Chen, 2014).
Quantile Delta Mapping (QDM)	Bias-correction method that corrects biases at each quantile while explicitly preserving the modeled climate-change signal at that quantile (Cannon et al., 2015; Wang and Chen, 2014).
LOCA2 (upscaled)	Updated hybrid statistical downscaling method that extends LOCA: uses an analog-based constructed-analog framework with observational training data, two-step coarse-to-fine downscaling, and explicit seasonal, intensity-dependent bias correction (Pierce et al., 2021).

methods on an equal footing during the evaluation, we first interpolated all of the bias-adjusted simulations to the resolution of the coarsest GCM (IPSL-CM6A-LR, $1.25° \times 2.5°$).

To evaluate spatial structure, we used the scale-dependent fractal dimension (FD) computed via the box-counting method (Husain et al., 2021; Lovejoy et al., 1987; Meisel et al., 1992) applied to thresholded binary precipitation fields at each quantile level. This method has been used to study how precipitation patterns change across scales (Hess et al., 2023; Lovejoy et al., 1987). For a given quantile threshold h, the precipitation field J is converted to a binary mask M_h:

$$M_h = \{1 \text{ if } h > J \text{ else } 0\}. \tag{9}$$

To measure how this thresholded pattern occupies space across scales, the mask is then partitioned using square boxes of side length ℓ. For each ℓ, we count the number of partially filled boxes, $\mathcal{N}(\ell)$, i.e., boxes that are neither empty nor completely full. If fractal scaling is assumed to be

Table 3. Summary of ETCCDI Indices for precipitation evaluation.

ETCCDI indices	Definition
R10mm	Annual count of days when precipitation ≥ 10 mm
R20mm	Annual count of days when precipitation ≥ 20 mm
Rx1day	Monthly maximum 1 d precipitation amount
Rx5day	Monthly maximum 5 d precipitation amount
Simple Daily Intensity Index (SDII)	Monthly average precipitation amount on wet days (precipitation ≥ 1 mm)
Consecutive Dry Days (CDD)	Annual maximum number of consecutive days with precipitation < 1 mm
Consecutive Wet Days(CWD)	Annual maximum number of consecutive days with precipitation ≥ 1 mm
R95pTOT	Annual total precipitation from days exceeding the 95th-percentile of daily precipitation during a reference period
R99pTOT	Annual total precipitation from days exceeding the 99th-percentile of daily precipitation during a reference period

$$N_\ell \propto \ell^{-\mathrm{FD}} \tag{10}$$

then

$$\mathrm{FD} = \frac{\log(N_\ell)}{\log(1/\ell)}. \tag{11}$$

For evaluating changes in climate signal trend, we used the framework from the "ibicus" package (Spuler et al., 2024). For a particular GCM and bias-correction method, we first computed the raw trend ($\boldsymbol{T}_{\mathrm{raw}}$) for a statistic $\boldsymbol{S}$ (e.g., mean or 95th quantile) between the future (2015–2099) and historical (1980–2014) periods:

$$\boldsymbol{T}_{\mathrm{raw}} = \boldsymbol{S}_{\mathrm{raw,future}} - \boldsymbol{S}_{\mathrm{raw,historical}}. \tag{12}$$

Then, we calculated the trend after carrying out bias adjustment:

$$\boldsymbol{T}_{\mathrm{debiased}} = \boldsymbol{S}_{\mathrm{debiased,future}} - \boldsymbol{S}_{\mathrm{debiased,historical}}. \tag{13}$$

Finally, the trend bias ($\boldsymbol{TB}$) was calculated as the percentage change between $\boldsymbol{T}_{\mathrm{raw}}$ and $\boldsymbol{T}_{\mathrm{debiased}}$:

$$\boldsymbol{TB} = 100\frac{\boldsymbol{T}_{\mathrm{debiased}} - \boldsymbol{T}_{\mathrm{raw}}}{\boldsymbol{T}_{\mathrm{raw}}}. \tag{14}$$

One of the advantages of this differentiable framework is its ability to generalize not only in time but also in space, provided we ingest data that helps it to learn spatial context (e.g., elevation, slope, and landcover). Our final evaluation was a spatial test where we trained, validated, and tested the model on different regions (Fig. 2). We trained and evaluated the model on the period 1990–2014, holding out the Upper Mississippi region (blue) for validation to select hyperparameters, and used this learned configuration to test in the adjacent Ohio region (green).

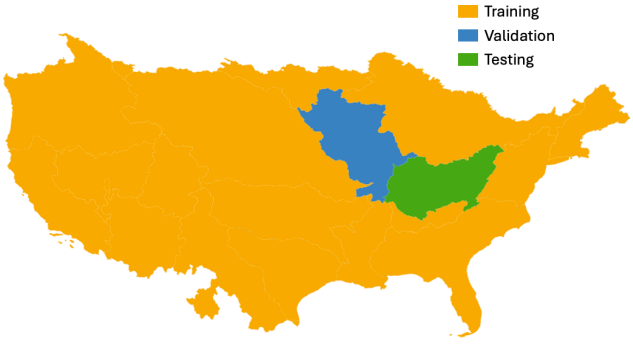

Figure 2. Spatial configuration of training, validation, and testing domains used to evaluate δCLIMBA in data-scarce settings. Validation and testing regions were entirely withheld from model training.

3 Results

δCLIMBA was evaluated using a series of experiments to test its generalizability across different GCMs at different locations, including its ability to adjust biases through precipitation indices, conserve spatial structures, and capture long-term trends.

3.1 Quantile comparison

Quantile comparisons across five climatically diverse US cities demonstrate that δCLIMBA systematically aligns the daily precipitation distributions of all six CMIP6 models with the Livneh reference across the full distributional range, including the upper tail, without inducing the non-physical amplification of extremes that is characteristic of several conventional adjustment methods (Fig. 3). The five cities, Phoenix, Yosemite, Seattle, Orlando, and Philadelphia

respectively span arid, mountainous, maritime, humid subtropical, and temperate climatic regimes.

While the raw models exhibited large spread and biases compared to observations, δCLIMBA consistently aligned model quantiles with the reference across the full distribution while preserving relative inter-model differences and avoiding artificial tail inflation (Fig. 3). The raw models' large inter-model spread and systematic biases were strongly regime-dependent: most models generally matched observations in Philadelphia, but displayed dry-region overestimation in Phoenix, exaggerated tails in mountainous Yosemite, and inconsistent heavy-rain behavior in humid Seattle and Orlando. Conventional bias-correction methods reduce mean bias but frequently distort tail behavior: Quantile Mapping, ISIMIP, and ECDFM often over-amplified upper quantiles, while Quantile Delta Mapping produced unstable and occasionally completely unrealistic tails (most evident in Phoenix and Orlando). LOCA2 improved inter-model consistency but still showed city-dependent residual biases in extremes. δCLIMBA performed well across multiple regimes and GCMs, indicating that the differentiable, spatially-conditioned formulation corrected distributional bias without inducing non-physical extreme amplification, a failure mode clearly visible in several traditional methods.

3.2 Precipitation indices

Across nearly all indices, including intensity (Rx1day, Rx5day), threshold exceedance (R10mm, R20mm), persistence (CDD, CWD), and extreme totals (R95pTOT), δCLIMBA exhibited median biases closer to zero than traditional statistical methods (Fig. 4). In contrast, raw model output and several conventional corrections showed large positive biases for extreme and heavy-precipitation indices, as well as substantial dispersion, highlighting their limited ability to jointly control mean bias and spatial consistency.

While Fig. 4 summarizes the distribution of grid-wise biases, it does not reveal whether residual errors exhibit coherent spatial structure. Figure 5 therefore examines the spatial distribution of mean percentage bias across CONUS for selected ETCCDI indices. This analysis allows assessment of whether δCLIMBA merely reduces aggregate bias or also mitigates geographically systematic error patterns.

Across all metrics except CDD and R99pTOT, the raw GCM output exhibited widespread positive bias over the CONUS (Fig. 5). While conventional bias adjustment methods such as Quantile Mapping (QM) effectively reduced this bias, their performance varied substantially across regions. For instance, QM retained a positive residual bias in the southeastern United States relative to both LOCA2 and δCLIMBA for multiple extreme metrics including R95pTOT, SDII, Rx5day, and CWD.

For heavy precipitation totals (R95pTOT), QM exhibited a pronounced positive bias across the entire region. In contrast, δCLIMBA mitigated these extremes without introducing systematic negative compensation elsewhere, indicating that its learned monotonic transformation adapted spatially. This behavior aligns with the spatial conditioning embedded in the neural architecture which enables the model to gain spatial context while attenuating bias. In addition to that, the weighted quantile loss function (assigning a higher penalty to higher quantiles) gives direct feedback to the model to de-bias precipitation at the higher end of the distribution. While LOCA2 achieved a comparable performance to δCLIMBA, it retained a moderately higher residual bias across the CONUS.

In contrast to R95pTOT, the extremely-wet-day totals index (R99pTOT) exhibits a negative bias in the raw GCM ensemble over most of the CONUS (Fig. 4), indicating that the models underestimate the annual precipitation delivered by the rarest wet days (99th percentile) even while overestimating the contribution of very wet days (95th percentile). None of the bias-adjustment methods corrected this deficit, and all of them amplified it. This behavior is consistent with a reversal in the sign of the raw bias between the 95th and 99th percentiles of wet-day precipitation: a single monotone transformation, fitted against a quantile loss in which fewer than 1 % of the quantile levels lie above the 99th percentile, learns a downward correction in the well-sampled region where the models are too wet and extrapolates that correction into the far tail where they are already too dry. R99pTOT therefore marks the limit of what the present formulation can adjust, and the degradation relative to the raw simulation is a structural consequence of the prescribed mapping rather than a failure of optimization.

Intensity and frequency-based indices (SDII and R20mm) were challenging for all approaches. For SDII, bias adjustment techniques counterintuitively increase bias relative to raw CMIP6 output rather than reducing it (Fig. 5). This degradation was most pronounced in the central United States, where QM introduced the largest bias amplification, while LOCA2 and δCLIMBA exhibited comparable residual bias. The threshold-based index R20mm presented a distinct challenge: unlike other metrics, R20mm exhibited systematic underestimation, particularly across the western mountainous region. Nevertheless, δCLIMBA and LOCA2 successfully reduced positive bias of R20mm at many locations, suggesting that the monotonic transformation constraint preferentially suppresses high-frequency precipitation events.

All approaches achieved substantial bias reduction for both temporal sequencing metrics (CDD and CWD) across the CONUS. For CWD, positive bias was effectively attenuated throughout most of the domain, with persistent residual bias limited to a few West Coast locations. Both δCLIMBA and LOCA2 outperformed QM in the high-rainfall southeastern region. For CDD, negative bias was similarly reduced across most locations, though a cluster of grid points in the western interior retained positive residual bias across all methods.

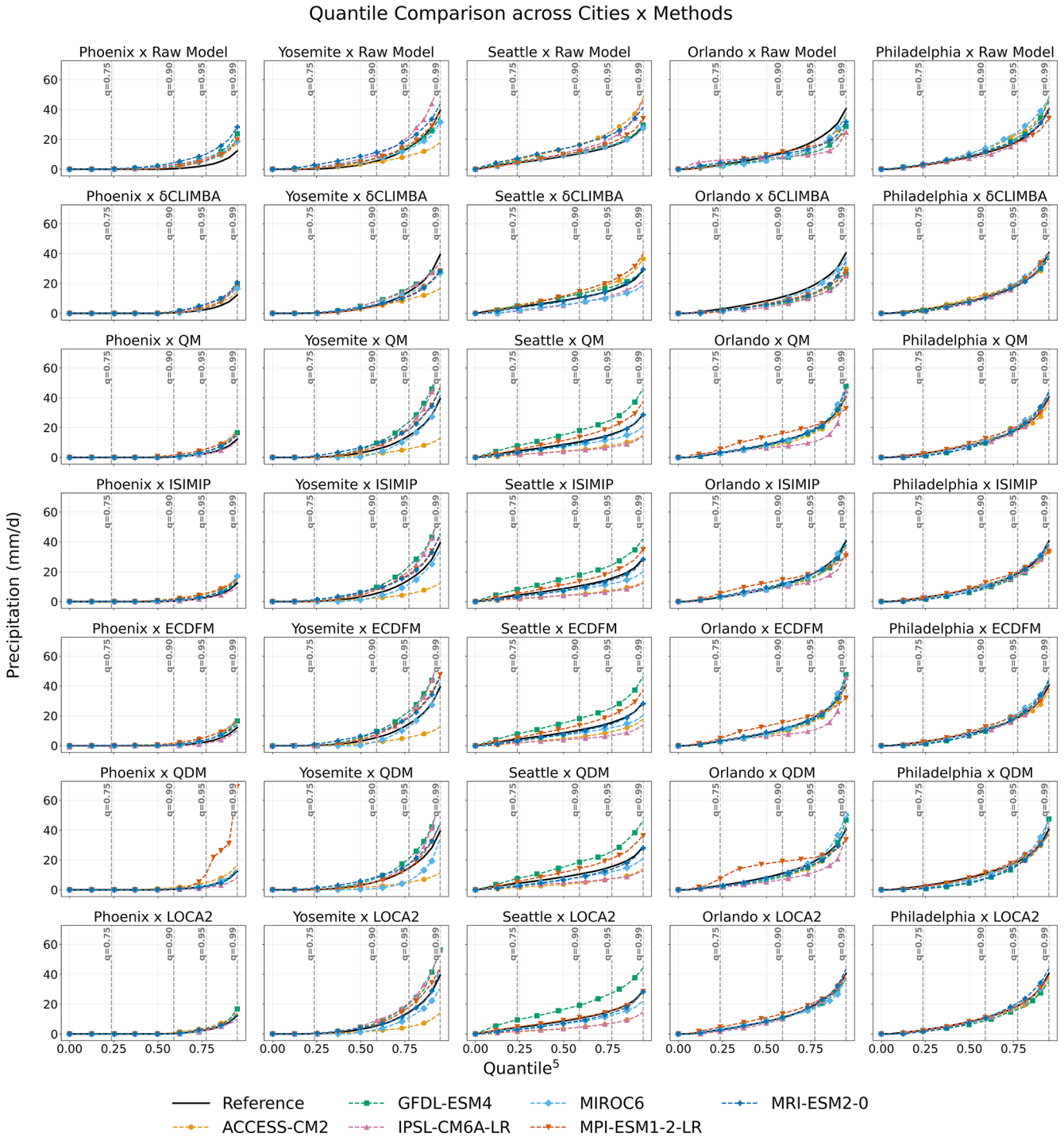

Figure 3. Quantile comparison of daily precipitation across selected climatically-diverse cities and bias adjustment methods during the testing period (2001–2014). For each city–method combination, empirical quantile curves of the six CMIP6 General Circulation Models are shown alongside the Livneh reference distribution. The x axis displays quantiles to the 5th power to expand the upper tail and emphasize extreme precipitation behavior (the original quantiles are marked using vertical dashed lines across each plot). The y axis represents precipitation (mm d^{-1}). Columns correspond with different cities and rows with bias adjustment methods.

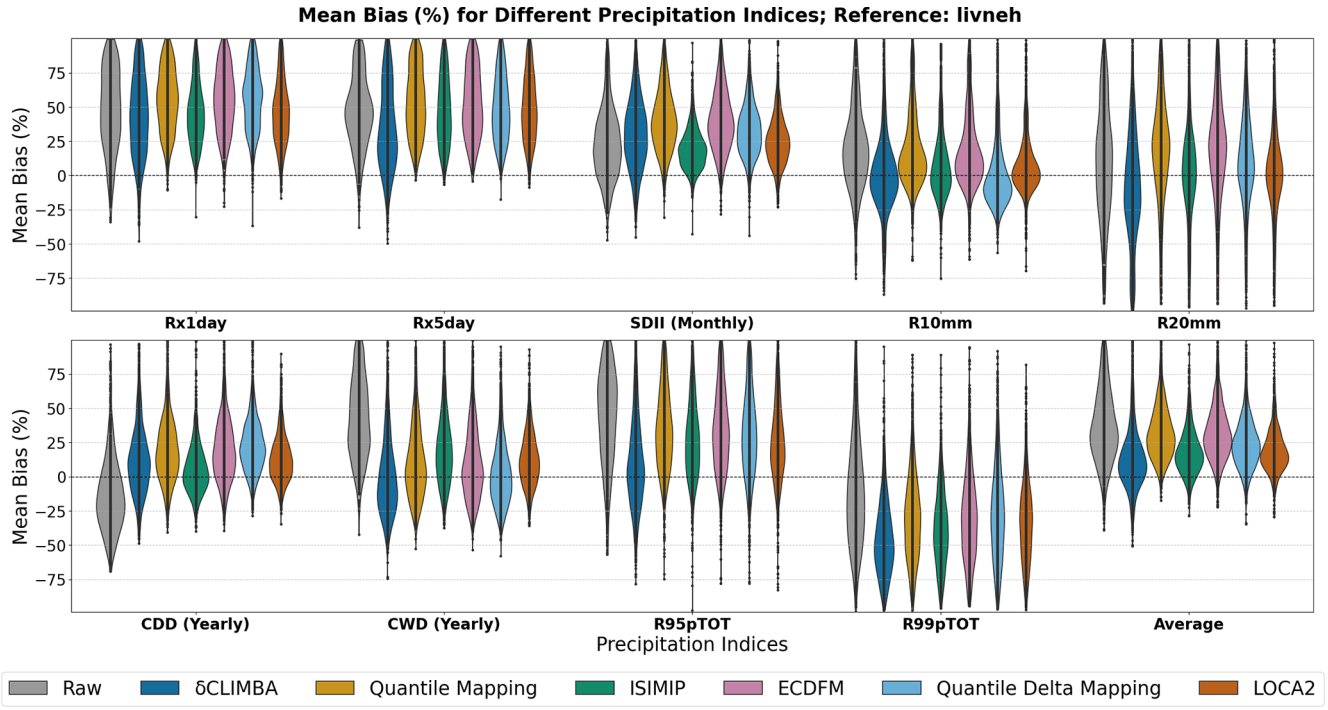

Figure 4. Spatial distribution of mean percentage bias for ETCCDI precipitation indices over the conterminous US (CONUS) during the test period (2001–2014), evaluated against the Livneh reference dataset. Results are shown for the ensemble of six CMIP6 General Circulation Models. Each violin represents the distribution of grid-point mean percentage bias across the CONUS for a given index and method. The "Average" column reports the arithmetic mean of the mean percentage bias across all ETCCDI indices for each method. Narrower distributions and medians (widest point in violin) closer to zero indicate improved bias stability and reduced spatial heterogeneity.

Collectively, these spatial diagnostics revealed heterogeneous performance of δCLIMBA across different precipitation characteristics. The method achieved consistent improvements in heavy but routine tails (R95pTOT) and temporal sequencing metrics (CDD, CWD), while matching or exceeding LOCA2's performance and substantially outperforming conventional QM but increased bias in rare extreme totals (R99pTOT). Since GCMs tend to carry systematic wet (positive) biases (Srivastava et al., 2020), the transformation in δCLIMBA seems particularly amenable to mitigating positive bias; especially in regions like the southeastern US where the precipitation distribution is highly skewed and frequency and intensity are driven by mesoscale convective systems (Rahman et al., 2023). However, intensity-based indices (SDII, R20mm) remain challenging for all bias correction approaches, highlighting the need for methods that can explicitly account for changes in precipitation event structure.

3.3 Spatial patterns

δCLIMBA substantially preserved the multiscale spatial structure of precipitation fields, achieving a mean absolute error (MAE) of 0.022 relative to the Livneh reference fractal dimension curve, a marked improvement over raw CMIP6 output (MAE = 0.030) and approaching the performance of LOCA2 (MAE = 0.017) as shown in Fig. 6. This result is notable because conventional univariate bias adjustment methods, operating independently at each grid point, lack any mechanism for learning spatial relationships and therefore tend to attenuate marginal bias at the cost of degrading spatial organization (Allard et al., 2025; Maraun et al., 2017). The fact that δCLIMBA achieves competitive spatial fidelity without an explicit analog procedure suggests that geodesic-aware spatial attention in the spatial encoder is capturing meaningful inter-node structure during training.

While LOCA2 achieved the lowest overall MAE (0.017), δCLIMBA competitively tracked the reference fractal-dimension curve across nearly the entire quantile range, with substantially reduced error relative to the raw model (MAE = 0.022 vs. 0.030). It exhibited good performance across the distribution, except for a deviation in the upper tail that is likely due to numerical instability, indicating that the differentiable framework corrects biases while retaining a physically realistic multiscale spatial structure. LOCA2's superior MAE is consistent with its design: the spatial analog method explicitly maps regional precipitation patterns to local scales and thus inherently reproduces aspects of spatial intermittency as a direct consequence of its construction, rather than as an emergent property of learned representations.

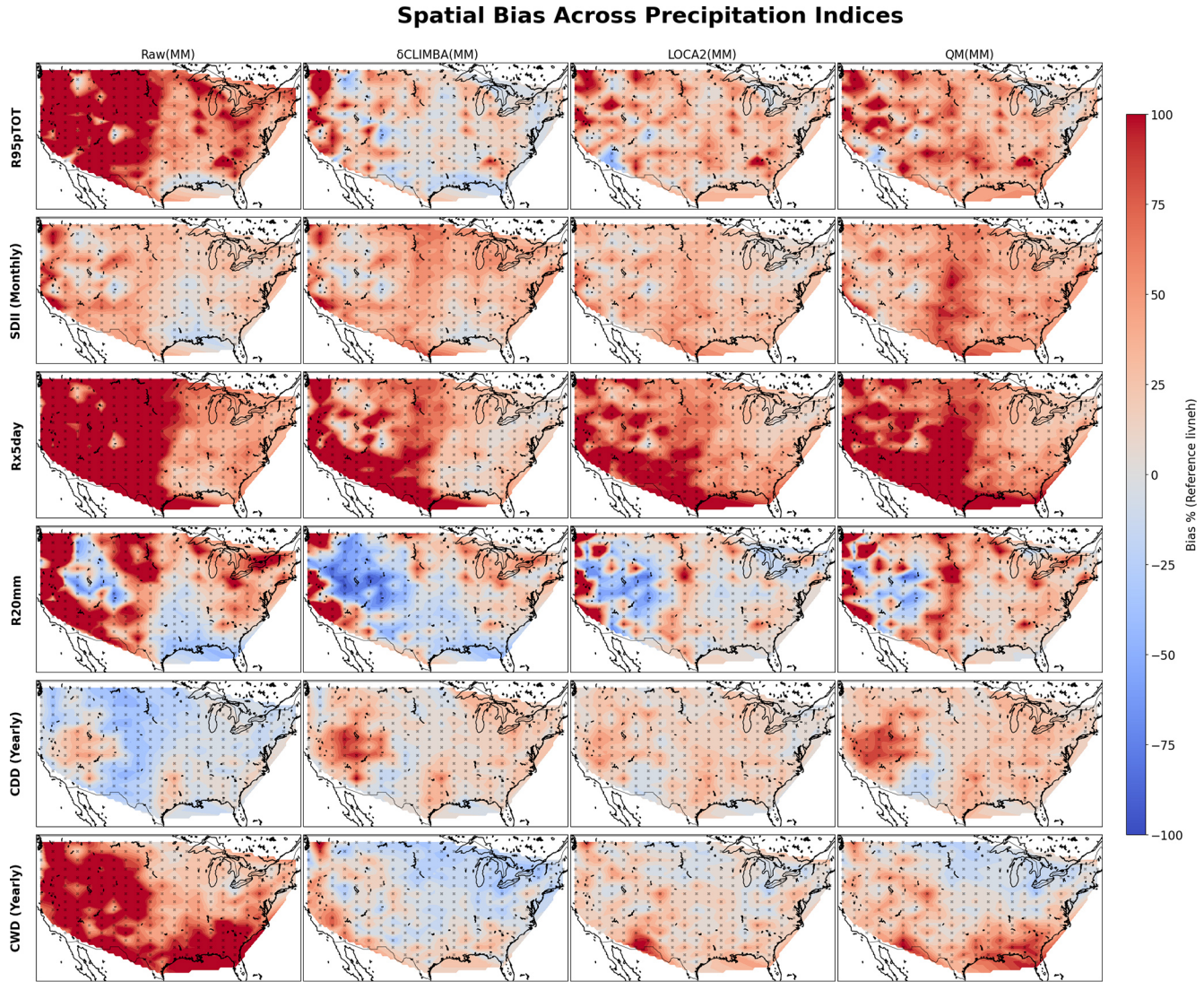

Figure 5. Spatial distribution of mean percentage bias (%) for selected ETCCDI precipitation indices over the conterminous US during the test period (2001–2014), evaluated against the Livneh reference dataset. Results are shown for the ensemble mean (model mean, MM) of six CMIP6 General Circulation Models. Rows correspond to precipitation indices, and columns correspond to bias-adjustment methods. Positive values indicate overestimation relative to Livneh, and negative values indicate underestimation. This figure highlights the spatial structure of residual bias, complementing the distributional summaries shown in Fig. 4.

3.4 Partial trend preservation

δCLIMBA partially preserved the trend in projected precipitation for GFDL-ESM4 under SSP5-8.5 without any explicit trend-preservation mechanism embedded in the framework, outperforming QDM and LOCA2 across all evaluated metrics while trailing ECDFM and ISIMIP for mean precipitation and the 95th percentile of daily precipitation. However, this behavior did not generalize consistently across GCMs, indicating that trend preservation is an emergent rather than a structural property of the current framework and warrants further investigation.

Figure 7 shows the bias in projected precipitation trends between the historical (1980–2014) and future periods (2015–2099) under SSP5-8.5 for mean precipitation, 95th percentile of daily precipitation, wet days ($> 1\,\mathrm{mm\,d^{-1}}$), and very wet days ($> 10\,\mathrm{mm\,d^{-1}}$) across multiple bias-adjustment methods. Quantile Delta Mapping (QDM) and LOCA2 (resolution upscaled) exhibited pronounced negative trend biases across all metrics, indicating systematic attenuation of trends after bias correction. Quantile Mapping performed better than its counterparts in preserving the trend of mean and 95th percentile daily precipitation, lagged behind for wet days, and performed comparably for trend in very wet days ($> 10\,\mathrm{mm}$).

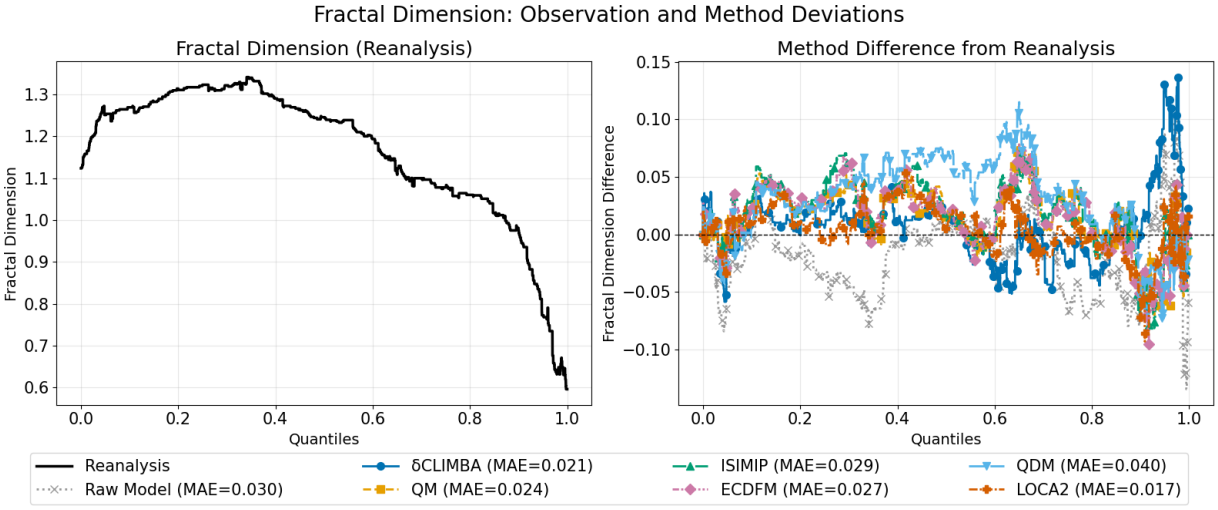

Figure 6. Fractal dimension of precipitation fields as a function of quantile over the conterminous US during the test period (2001–2014). The black curve represents the Livneh reference dataset, while colored curves correspond to ensemble model mean (MM) fields after bias adjustment (δCLIMBA, LOCA2, Quantile Delta Mapping – QDM) and raw General Circulation Model output. Mean absolute error (MAE) values reported in the legend quantify deviation from the reference fractal dimension curve across quantiles.

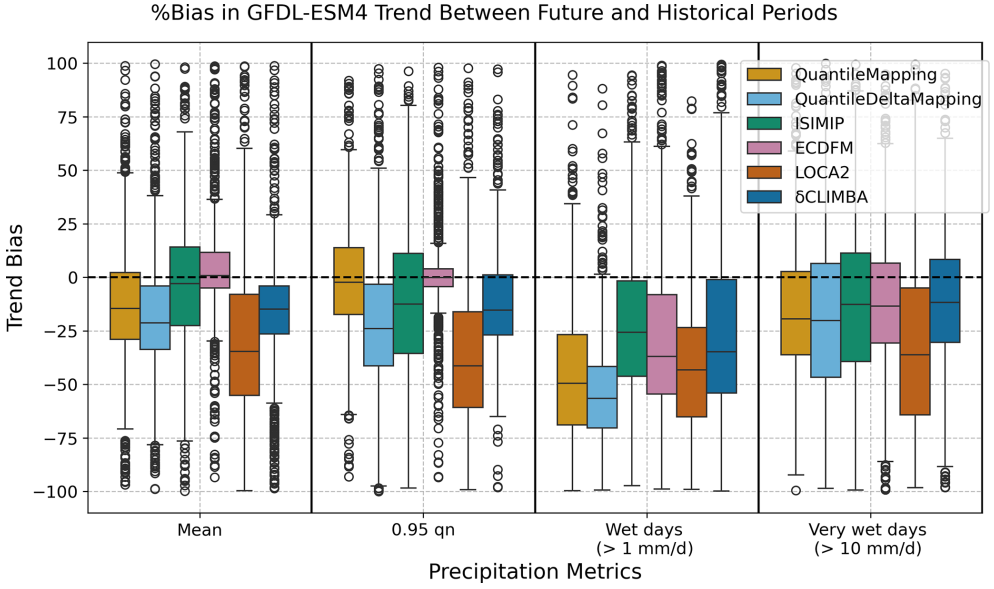

Figure 7. Percentage bias in projected climate trends of GFDL-ESM4 between future and historical periods under SSP5-8.5, evaluated across precipitation metrics over CONUS. Values closer to zero indicate improved preservation of trends following bias adjustment. The central line on each box indicates the median, the top and bottom of each box respectively indicate the 25th and 75th percentiles, and the top and bottom whiskers of each box respectively indicate the 2.5th and 97.5th percentiles. The open dots indicate outliers.

ECDFM, which is well recognized for its trend preservation (Li et al., 2010), did a particularly good job with trend bias in the mean and 95th percentile of daily precipitation, although its performance was more typical when evaluated for trends in number of wet days and very wet days. ISIMIP, another bias-correction method designed for trend-preservation, displayed similar performance to ECDFM, though with a higher inter-quartile range, implying a larger number of locations with distorted trends.

δCLIMBA has no built-in trend-preservation mechanisms, but shows acceptable performance, nonetheless. For mean and 95th percentile of daily precipitation, it performed similarly to the two quantile-mapping methods, lagging behind ISIMIP and ECDFM but doing better than LOCA2, while for wet and very wet days, its performance was competitive

with ISIMIP and ECDFM. It exhibited similar performance on some other GCMs, like ACCESS-CM2, but not all, indicating that the differentiable framework was unable to generalize the trend-preservation behavior across GCMs. Future δCLIMBA versions could investigate loss terms focusing on long-term trends.

3.5 Bias correction in data scarce regions

δCLIMBA demonstrated meaningful spatial generalization to an entirely withheld region, reducing positive bias in extreme precipitation indices including R95pTOT and Rx5day without any local calibration except for validation (Fig. 8). However, indices sensitive to daily wet-day intensity (SDII, R20mm) exhibited residual overestimation, indicating a limit on the spatial transferability of the current framework.

Improvements were most pronounced for intensity-based and percentile-based indices, including Rx1day, Rx5day, and high-percentile precipitation (R95pTOT) totals, indicating that the learned monotonic transformation effectively corrected systematic distributional distortions in unseen regions. Similar to the temporal test results in Sect. 3.2 (Fig. 4), rare event precipitation index (R99pTOT) bias was difficult to attenuate using the current transformation scheme as the model underestimates rare wet day events. Some indices, such as SDII, CDD and CWD, retained wider bias distributions, suggesting that duration-based metrics remain more challenging to correct under spatial holdout. However, the longer interquartile range in the violin plots for all indices shows that many of the grid cells in the test region tend to end up with biases even larger than with the raw simulation, such as SDII, R10mm, and CDD.

This experiment constitutes a particularly stringent test of the framework, as conventional bias-correction methods, whether distribution-based, moment-scaling, or spatial analog methods, are explicitly calibrated against local observations and lack a principled mechanism for spatial generalization; their performance inherently degrades outside the calibration domain, making a like-for-like spatial comparison uninformative. In addition, LOCA2 is a bespoke product (Pierce et al., 2014, 2021), so there is no way to test its ability to generalize spatially.

Overall, these results demonstrate that there is promise in using δCLIMBA to generalize spatially. Since the grid cells are moving towards zero bias, the differentiable framework has the potential to maintain stable bias-adjustment behavior when transferred to an unseen but adjacent region, reducing distributional biases without access to local calibration data. Yet, more effort is needed to better regularize the system.

4 Conclusions

Bias correction or bias adjustment is a domain that cannot be addressed or optimized by a single metric. Every GCM output carries different kinds of biases, and therefore an exhaustive evaluation of not only marginal bias but also their temporal, spatial, and scale-dependent inherent characteristics is important. With its composite quantile loss alone, δCLIMBA excels in adjusting the distributions of GCMs across diverse cities, surpassing all traditional methods while being on par with LOCA2. This similar or even better performance is also depicted while attenuating marginal biases of extreme precipitation indices like R95pTOT. The corrected indices, however, separate sharply at the far tail. While marginal biases in very wet day totals (R95pTOT) were attenuated, extremely wet day totals (R99pTOT) were not: the raw ensemble already underestimates this index, and every adjustment method examined here, δCLIMBA included, deepened rather than reduced that deficit. We attribute this to the structure of the correction rather than to its optimization, since a single monotone mapping fitted against a quantile loss dominated by the bulk of the distribution learns a downward adjustment where the models are too wet and carries that adjustment into the far tail where they are already too dry. Correcting this regime would require a formulation that treats the far tail separately, for example a tail-anchored basis, an explicit constraint above a chosen quantile, or a loss weighted on exceedance totals rather than on quantile levels. We identify it as the clearest structural limitation of the present framework.

Since precipitation is a multiscale phenomenon, δCLIMBA uses a transformer-based spatial encoder to learn spatial structures from neighboring grid cells. Though the current framework can learn only short-range dependencies through each grid cell's 16 closest correlated neighbors, δCLIMBA outputs were able to imitate spatial patterns (based on fractal dimension) second closest to the reference dataset, after LOCA2. This framework could be a gateway for more work in adding mechanisms that learn long-range dependencies to output more realistic precipitation simulations.

Even though δCLIMBA performed well on reducing distributional biases and conserving spatial patterns, one of the biggest concerns from every bias correction method is about preservation of trends between historical and future simulations. Unlike ECDFM and ISIMIP, δCLIMBA does not include an explicit trend-preservation mechanism, but still exhibited better behavior than purely quantile-based methods and LOCA2 in the case of GFDL-ESM4. However, this behavior was not repeated with other GCMs, indicating the need for a data-driven mechanism in δCLIMBA to preserve trends. One of the advantages of machine learning based approaches (including differentiable models) is that the trained models are applicable to completely unseen regions and timeframes. Unlike all the traditional methods including LOCA2, δCLIMBA utilizes information from inputs like elevation and landcover to mitigate biases, which allow it to extrapolate those learned patterns in unseen but similar regions. Marginal biases in the testing region such as R95pTOT and Rx5day were demonstrated to be attenuated with their

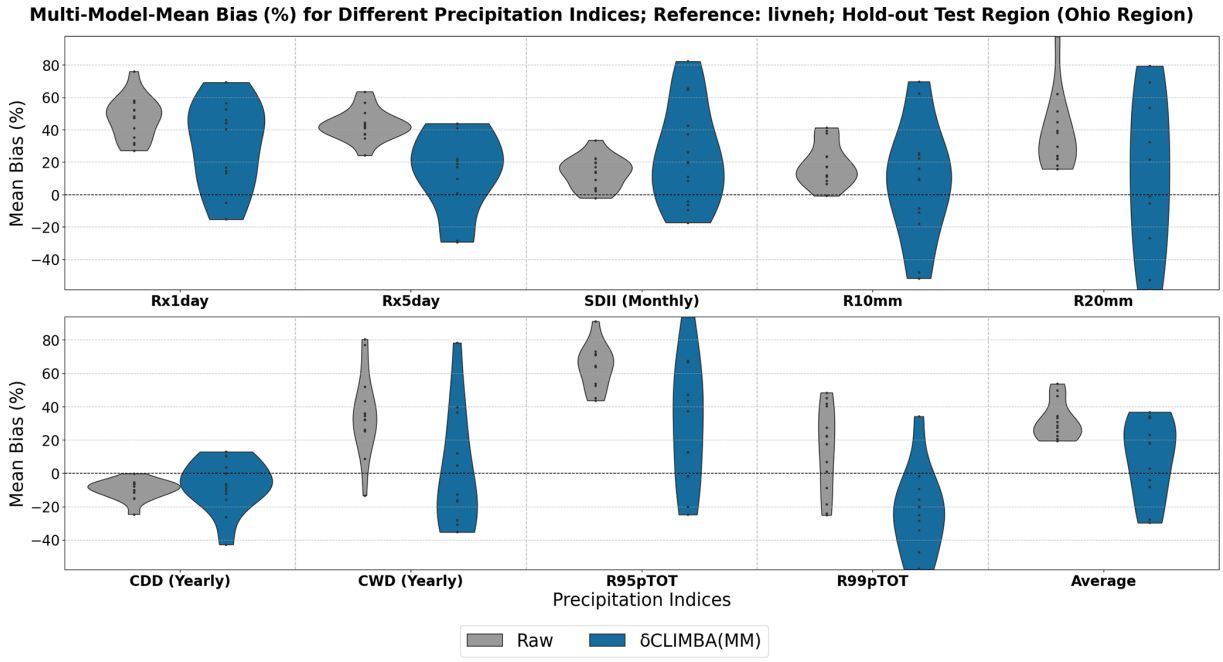

Figure 8. Spatial distribution of mean percentage bias for ETCCDI precipitation indices over the testing region (Ohio region), evaluated against the Livneh reference dataset. Results are shown for the ensemble mean (model mean, MM) of six CMIP6 General Circulation Models. Each violin represents the distribution of grid-point mean percentage bias across the testing region for a given index and method. The "Average" column reports the arithmetic mean of mean percentage bias across all ETCCDI indices. Narrower distributions and medians closer to zero indicate improved bias stability and reduced spatial heterogeneity.

positive bias distribution moving towards zero. However, the larger interquartile range and biases exacerbated in indices like SDII and R20mm indicated overestimation of precipitation per day, limiting the model's reliability to gauge intensity and frequency of events.

Classically, there could be three main advantages of an end-to-end differentiable pipeline. First, no observed value of the transformation coefficients exists anywhere, yet they are learned entirely from the gradient of a distributional loss; the framework therefore requires no intermediate supervision. Second, because the objective is a single global constraint defined over the entire training domain rather than a collection of independent per-grid-cell calibrations, one parameterisation network is shared across all locations and remains defined at locations outside the calibration domain, which point-wise quantile-mapping methods cannot be by construction. Third, objectives of different character like a weighted quantile loss, a soft wet-day frequency constraint, and a spatial-correlation term can be combined in a single scalar objective and balanced through gradients, instead of being imposed as a sequence of hand-tuned post-processing steps whose interactions are not controlled.

It is worth stating plainly what differentiability did and did not contribute in this study, since the term admits broader readings than the one used here. δCLIMBA is not a coupled physics–machine learning model, and none of the results above depend on differentiating a process-based model. What differentiability provided is the ability to fit a prescribed, monotone, interpretable transformation end to end against a composite objective that mixes distributional, frequency-based, and spatial terms, with no intermediate supervision and a single parameterisation shared across the entire domain. The transparency and the physical admissibility of the correction come from the prescribed mapping, the ability to transfer to unseen regions comes from conditioning that mapping on static attributes and neighbouring grid cells, and differentiability is what allows those two elements to be optimised jointly rather than in sequence. The ablation experiments reported in the Supplement (Sect. S1) are intended to make this attribution empirical rather than asserted.

Overall, the differentiable model is learning the aspects of bias correction on which it is optimized. The bias adjustor paired with a composite quantile loss function makes sure that the bias correction distribution and spatial patterns remain closer to the observations than do simulations from traditional methods. Though the framework is generalizable, it is not a universal solution designed to attenuate bias in any GCM, but rather provides a framework to train and tune our choice of GCM. Extending this framework using a foundational model approach for one-shot bias correction is one possible way to make it truly generalizable. This differentiable framework could also be leveraged to understand the mitigation of biases directly by analyzing the coefficients generated from neural networks. All in all, the dif-

ferentiable framework has the potential to isolate and attenuate bias based on the mechanisms implemented inside it. It provides a pathway, not exercised in this study, for coupling the adjustment to downstream models. Because the bias-adjusted field is produced by a differentiable operator, a differentiable downstream model, for example a differentiable hydrological model (Feng et al., 2022; Rahmani et al., 2023), could be appended so that an impact-relevant objective such as streamflow error or water-balance closure informs the bias-adjustment coefficients directly through backpropagation, without requiring gridded precipitation to be treated as an intermediate calibration target. We identify this as a future opportunity. Its flexibility makes it suitable for integration into Earth system post-processing pipelines and impact workflows, as it is a modular, computationally efficient, and extensible bias correction approach that could be physically informed, scalable, and compatible with both historical and future applications.

Code and data availability. The source code for δCLIMBA and data processing can be found here: https://doi.org/10.5281/zenodo.23159628 (Sawadekar, 2026). All datasets used are publicly available and referenced accordingly. The CMIP6 data was accessed via Copernicus Climate Data Store under Copernicus license: https://doi.org/10.24381/cds.c866074c (Copernicus Climate Change Service, Climate Data Store, 2021). The unsplit-Livneh was downloaded from: https://cirrus.ucsd.edu/~pierce/nonsplit_precip/ (last access: 6 February 2025). The land cover was downloaded from: https://www.cec.org/publications/nalcms/ (last access: 18 November 2025). The elevation was extracted and processed from SRTM: https://www.earthdata.nasa.gov/data/catalog/lpcloud-srtmgl1-003 (last access: 19 November 2025). Slope and aspect were derived from elevation using ArcGIS.

Supplement. The supplement related to this article is available online at https://doi.org/10.5194/gmd-19-11225-2026-supplement.

Author contributions. CS conceived the initial study design. KS and CS designed the experiment. KS developed the model code and performed the analysis. KS, SM, PL, KL, and CS analysed the results and helped prepare the manuscript.

Competing interests. Chaopeng Shen and Kathryn Lawson have financial interests in HydroSapient, Inc., a company which could potentially benefit from the results of this research. These interests have been reviewed by The Pennsylvania State University in accordance with its individual conflict of interest policy, for the purpose of maintaining the objectivity and the integrity of research at The Pennsylvania State University.

Disclaimer. Publisher's note: 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. The authors bear the ultimate responsibility for providing appropriate place names. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

Acknowledgements. The authors acknowledge the National Energy Research Scientific Computing Center (NERSC), particularly the Perlmutter system, for providing computational resources used in this study. We appreciate Institute for Computational and Data Sciences (ICDS) at Penn State for parts of the data storage. We also thank the developers and providers of the CMIP6, Livneh, and LOCA2 datasets for making these resources available to the research community.

Financial support. This work was supported by the US Department of Energy RGCM awards DOE DE-SC0016605 and DE-AC52-07NA27344. This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977.

Review statement. This paper was edited by Stefan Rahimi-Esfarjani and reviewed by two anonymous referees.

References

Aboelyazeed, D., Xu, C., Hoffman, F. M., Liu, J., Jones, A. W., Rackauckas, C., Lawson, K., and Shen, C.: A differentiable, physics-informed ecosystem modeling and learning framework for large-scale inverse problems: demonstration with photosynthesis simulations, Biogeosciences, 20, 2671–2692, https://doi.org/10.5194/bg-20-2671-2023, 2023.

Allard, D., Vrac, M., François, B., and García De Cortázar-Atauri, I.: Assessing multivariate bias corrections of climate simulations on various impact models under climate change, Hydrol. Earth Syst. Sci., 29, 4711–4738, https://doi.org/10.5194/hess-29-4711-2025, 2025.

Baste, S., Klotz, D., Acuña Espinoza, E., Bardossy, A., and Loritz, R.: Unveiling the limits of deep learning models in hydrological extrapolation tasks, Hydrol. Earth Syst. Sci., 29, 5871–5891, https://doi.org/10.5194/hess-29-5871-2025, 2025.

Bi, D., Dix, M., Marsland, S. J., O'Farrell, S., Sullivan, A. P., Bodman, R., Law, R. J., Harman, I. N., Srbinovsky, J., Rashid, H., Dobrohotoff, P., Mackallah, C., Yan, H., Hirst, A. C., Savita, A., Dias, F. S., Woodhouse, M. T., Fiedler, R., and Heerdegen, A.: Configuration and spin-up of ACCESS-CM2, the new generation Australian Community Climate and Earth System Simulator Coupled Model, J. South. Hemis. Earth Syst. Sci., 70, 225–251, https://doi.org/10.1071/es19040, 2020.

Bindas, T., Tsai, W. P., Liu, J., Rahmani, F., Feng, D., Bian, Y., Lawson, K., and Shen, C.: Improving River Routing Using a Differentiable Muskingum-Cunge Model and

Physics-Informed Machine Learning, Water Resour. Res., 60, https://doi.org/10.1029/2023wr035337, 2024.

Boucher, O., Denvil, S., Levavasseur, G., Cozic, A., Caubel, A., Foujols, M.-A., Meurdesoif, Y., Cadule, P., Devilliers, M., Ghattas, J., Lebas, N., Lurton, T., Mellul, L., Musat, I., Mignot, J., and Cheruy, F.: IPSL IPSL-CM6A-LR model output prepared for CMIP6 CMIP (Version 20250528) [Application/x-netcdf], Earth System Grid Federation, https://doi.org/10.22033/esgf/cmip6.1534, 2018.

Buser, C. M., Künsch, H. R., Lüthi, D., Wild, M., and Schär, C.: Bayesian multi-model projection of climate: bias assumptions and interannual variability, Clim. Dynam., 33, 849–868, https://doi.org/10.1007/s00382-009-0588-6, 2009.

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. Climate, 28, 6938–6959, https://doi.org/10.1175/JCLI-D-14-00754.1, 2015.

Copernicus Climate Change Service, Climate Data Store: CMIP6 climate projections, Copernicus Climate Change Service (C3S) Climate Data Store (CDS) [data set], https://doi.org/10.24381/cds.c866074c, 2021.

Deser, C., Phillips, A., Bourdette, V., and Teng, H.: Uncertainty in climate change projections: the role of internal variability, Clim. Dynam., 38, 527–546, https://doi.org/10.1007/s00382-010-0977-x, 2012.

Dugas, C., Bengio, Y., Bélisle, F., Nadeau, C., and Garcia, R.: Incorporating second-order functional knowledge for better option pricing, in: Advances in Neural Information Processing Systemsm, Vol. 13, edited by: Leen, T., Dietterich, T., and Tresp, V., MIT Press, ISBN 9780262122412, 2001.

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., and Liebert, J.: HESS Opinions "Should we apply bias correction to global and regional climate model data?", Hydrol. Earth Syst. Sci., 16, 3391–3404, https://doi.org/10.5194/hess-16-3391-2012, 2012.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Fang, J. and Gentine, P.: Exploring Optimal Complexity for Water Stress Representation in Terrestrial Carbon Models: A Hybrid-Machine Learning Model Approach, J. Adv. Model. Earth Syst., 16, e2024MS004308, https://doi.org/10.1029/2024MS004308, 2024.

Feng, D., Fang, K., and Shen, C.: Enhancing streamflow forecast and extracting insights using long-short term memory networks with data integration at continental scales, Water Resour. Res., 56, e2019WR026793, https://doi.org/10.1029/2019WR026793, 2020.

Feng, D., Lawson, K., and Shen, C.: Mitigating prediction error of deep learning streamflow models in large data-sparse regions with ensemble modeling and soft data, Geophys. Res. Lett., 48, e2021GL092999, https://doi.org/10.1029/2021GL092999, 2021.

Feng, D., Liu, J., Lawson, K., and Shen, C.: Differentiable, learnable, regionalized process-based models with multiphysical outputs can approach state-of-the-art hydrologic prediction accuracy, Water Resour. Res., 58, e2022WR032404, https://doi.org/10.1029/2022WR032404, 2022.

Giorgi, F. and Mearns, L. O.: Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the "reliability ensemble averaging" (REA) method, J. Climate, 15, 1141–1158, https://doi.org/10.1175/1520-0442(2002)015<1141:COAURA>2.0.CO;2, 2002.

Guo, L., Liu, J., Zhu, X., Yao, P., Lu, S., and Lu, H.: Normalized and Geometry-Aware Self-Attention Network for Image Captioning, in: IEEE 2020/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 10324–10333, https://doi.org/10.1109/CVPR42600.2020.01034, 2020.

Gutjahr, O., Putrasahan, D., Lohmann, K., Jungclaus, J. H., von Storch, J.-S., Brüggemann, N., Haak, H., and Stössel, A.: Max Planck Institute Earth System Model (MPI-ESM1.2) for the High-Resolution Model Intercomparison Project (HighResMIP), Geosci. Model Dev., 12, 3241–3281, https://doi.org/10.5194/gmd-12-3241-2019, 2019.

Harder, P., Yang, Q., Ramesh, V., Sattigeri, P., Hernandez-Garcia, A., Watson, C., Szwarcman, D., and Rolnick, D.: Physics-Constrained Deep Learning for Climate Downscaling (Version 9), arXiv [preprint], https://doi.org/10.48550/arxiv.2208.05424, 2022.

Hess, P. and Boers, N.: Deep learning for improving numerical weather prediction of rainfall extremes, ESS Open Archive, https://doi.org/10.1002/essoar.10507827.1, 2021.

Hess, P., Lange, S., Schötz, C., and Boers, N.: Deep learning for bias-correcting CMIP6-class earth system models, Earth's Future, 11, e2023EF004002, https://doi.org/10.1029/2023EF004002, 2023.

Hochreiter, S. and Schmidhuber, J.: Long Short-Term Memory, Neural Comput., 9, 1735–1780, https://doi.org/10.1162/neco.1997.9.8.1735, 1997.

Hourdin, F., Mauritsen, T., Gettelman, A., Golaz, J.-C., Balaji, V., Duan, Q., Folini, D., Ji, D., Klocke, D., Qian, Y., Rauser, F., Rio, C., Tomassini, L., Watanabe, M., and Williamson, D.: The Art and Science of Climate Model Tuning, B. Am. Meteorol. Soc., 98, 589–602, https://doi.org/10.1175/bams-d-15-00135.1, 2017.

Husain, A., Reddy, J., Bisht, D., and Sajid, M.: Fractal dimension of coastline of Australia, Sci. Rep., 11, 6304, https://doi.org/10.1038/s41598-021-85405-0, 2021.

Huynh, N. N. T., Garambois, P.-A., Renard, B., Colleoni, F., Monnier, J., and Roux, H.: A distributed hybrid physics – AI framework for learning corrections of internal hydrological fluxes and enhancing high-resolution regionalized flood modeling, Hydrol. Earth Syst. Sci., 29, 3589–3613, https://doi.org/10.5194/hess-29-3589-2025, 2025.

IPCC – Intergovernmental Panel On Climate Change: Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, in: 1st Edn., Cambridge University Press, https://doi.org/10.1017/9781009157896, 2023.

Ji, H., Song, Y., Bindas, T., Shen, C., Yang, Y., Pan, M., Liu, J., Rahmani, F., Abbas, A., Beck, H., Lawson, K., and Wada, Y.: Distinct hydrologic response patterns and trends worldwide revealed by physics-embedded learning, Nat. Commun., 16, 9169, https://doi.org/10.1038/s41467-025-64367-1, 2025.

Kawai, H., Yukimoto, S., Koshiro, T., Oshima, N., Tanaka, T., Yoshimura, H., and Nagasawa, R.: Significant improvement of cloud representation in the global climate

model MRI-ESM2, Geosci. Model Dev., 12, 2875–2897, https://doi.org/10.5194/gmd-12-2875-2019, 2019.

Kingma, D. P. and Ba, J. A.: A method for stochastic optimization, arXiv [preprint], https://doi.org/10.48550/arXiv.1412.6980, 2017.

Krasting, J. P., John, J. G., Blanton, C., McHugh, C., Nikonov, S., Radhakrishnan, A., Rand, K., Zadeh, N. T., Balaji, V., Durachta, J., Dupuis, C., Menzel, R., Robinson, T., Underwood, S., Vahlenkamp, H., Dunne, K. A., Gauthier, P. P., Ginoux, P., Griffies, S. M., Hallberg, R., Harrison, M., Hurlin, W., Malyshev, S., Naik, V., Fabien, P., Paynter, D. J., Ploshay, J., Reichl, B. G., Schwarzkopf, D. M., Seman, C. J., Silvers, L., Wyman, B., Zeng, Y., Adcroft, A., Dunne, J. P., Dussin, R., Huan, G., Jian, H., Held, I. M., Horowitz, L. W., Pu, L., Milly, P. C. D, Shevliakova, E., Stock, C., Winton, M., Wittenberg, A. T., Xie, Y., and Zhao, M.: NOAA-GFDL GFDLESM4 model output prepared for CMIP6 CMIP (Version 20250528) [Application/x-netcdf], Earth System Grid Federation, https://doi.org/10.22033/ESGF/CMIP6.1407, 2018.

Krinner, G. and Flanner, M. G.: Striking stationarity of large-scale climate model bias patterns under strong climate change, P. Natl. Acad. Sci. USA, 115, 9462–9466, https://doi.org/10.1073/pnas.1807912115, 2018.

Lange, S.: Trend-preserving bias adjustment and statistical downscaling with ISIMIP3BASD (v1.0), Geosci. Model Dev., 12, 3055–3070, https://doi.org/10.5194/gmd-12-3055-2019, 2019.

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., and Jackel, L. D.: Backpropagation applied to handwritten zip code recognition, Neural Comput., 1, 541–551, https://doi.org/10.1162/neco.1989.1.4.541, 1989.

Li, H., Sheffield, J., and Wood, E. F.: Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching, J. Geophys. Res.-Atmos., 115, 2009JD012882, https://doi.org/10.1029/2009JD012882, 2010.

Li, P., Zha, Y., Shi, L., Tso, C. H. M., Zhang, Y., and Zeng, W.: Comparison of the use of a physical-based model with data assimilation and machine learning methods for simulating soil water dynamics, J. Hydrol., 584, 124692, https://doi.org/10.1016/j.jhydrol.2020.124692, 2020.

Li, P., Zha, Y., Zhang, Y., Michael Tso, C.-H., Attinger, S., Samaniego, L., and Peng, J.: Deep learning integrating scale conversion and pedo-transfer function to avoid potential errors in cross-scale transfer, Water Resour. Res., 60, e2023WR035543, https://doi.org/10.1029/2023WR035543, 2024a.

Li, P., Shen, C., Liu, J., Rahmani, F., and Lawson, K. E.: Structural bias should be addressed before effective parameter learning – Insights from SMAP soil moisture simulations using differentiable process-based models, ESS Open Archive, https://doi.org/10.22541/essoar.175855507.72741303/v1, 2025.

Li, Z., Yuan, Q., Yang, Q., Li, J., and Zhao, T.: Differentiable modeling for soil moisture retrieval by unifying deep neural networks and water cloud model, Remote Sensing of Environment, 311, 114281, 2024b.

Livneh, B., Rosenberg, E. A., Lin, C., Nijssen, B., Mishra, V., Andreadis, K. M., Maurer, E. P., and Lettenmaier, D. P.: A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States: Update and Extensions, J. Climate, 26, 9384–9392, https://doi.org/10.1175/jcli-d-12-00508.1, 2013.

Lovejoy, S., Schertzer, D., and Tsonis, A. A.: Functional box-counting and multiple elliptical dimensions in rain, Science, 235, 1036–1038, https://doi.org/10.1126/science.235.4792.1036, 1987.

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

Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themeßl, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M., and Thiele-Eich, I.: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, Rev. Geophys., 48, https://doi.org/10.1029/2009rg000314, 2010.

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. Change, 7, 764–773, https://doi.org/10.1038/nclimate3418, 2017.

Meisel, L. V., Johnson, M., and Cote, P. J.: Box-counting multifractal analysis, Phys. Rev. A, 45, 6989–6996, https://doi.org/10.1103/PhysRevA.45.6989, 1992.

Miyato, T., Jaeger, B., Welling, M., and Geiger, A.: GTA: A Geometry-Aware Attention Mechanism for Multi-View Transformers (Version 3), arXiv [preprint], https://doi.org/10.48550/ARXIV.2310.10375, 2023.

Mochizuki, T., Tatebe, H., Kataoka, T., and Koyama, H.: MIROC MIROC6 model output prepared for CMIP6 DCPP (Version 20250527) [Application/x-netcdf], Earth System Grid Federation, https://doi.org/10.22033/ESGF/CMIP6.890, 2019.

NASA JPL: NASA Shuttle Radar Topography Mission Global 1 arc second [Data set], NASA Land Processes Distributed Active Archive Center [data set], https://doi.org/10.5067/MEASURES/SRTM/SRTMGL1.003, 2013.

Palmer, T. and Stevens, B.: The scientific challenge of understanding and estimating climate change, P. Natl. Acad. Sci. USA, 116, 24390–24395, https://doi.org/10.1073/pnas.1906691116, 2019.

Pan, B., Anderson, G. E., Gonçalves, A. R., Lucas, D., Bonfils, C., Lee, J., Tian, Y., and Ma, H.-P.: Learning to correct climate projection biases, J. Adv. Model. Earth Syst., 13, e2021MS002509, https://doi.org/10.1029/2021MS002509, 2021.

Pasos, M.: North American Land Cover, 2020 (Landsat, 30m) [Dataset], https://www.cec.org/north-american-environmental-atlas/land-cover-30m-2020/ (last access: 25 March 2026), 2020.

Piani, C., Weedon, G. P., Best, M., Gomes, S. M., Viterbo, P., Hagemann, S., and Haerter, J. O.: Statistical bias correction of global simulated daily precipitation and temperature for the application of hydrological models, J. Hydrol., 395, 199–215, https://doi.org/10.1016/j.jhydrol.2010.10.024, 2010.

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

Pierce, D. W., Su, L., Cayan, D. R., Risser, M. D., Livneh, B., and Lettenmaier, D. P.: An extreme-preserving long-term gridded

daily precipitation data set for the conterminous United States, J. Hydrometeorol., 22, 1883–1895, https://doi.org/10.1175/JHM-D-20-0212.1, 2021.

Rahman, M. S., Senkbeil, J. C., and Keellings, D. J.: Spatial and temporal variability of extreme precipitation events in the southeastern United States, Atmosphere, 14, 1301, https://doi.org/10.3390/atmos14081301, 2023.

Rahmani, F., Appling, A., Feng, D., Lawson, K., and Shen, C.: Identifying structural priors in a hybrid differentiable model for stream water temperature modeling, Water Resour. Res., 59, e2023WR034420, https://doi.org/10.1029/2023WR034420, 2023.

Rumelhart, D. E., Hinton, G., and Williams, R. J.: Learning representations by back-propagating errors, Nature, 323, 533–536, https://doi.org/10.1038/323533a0, 1986.

Sawadekar, K.: δCLIMBA: A Differentiable Framework for GCM Precipitation Bias Correction – code and processed data (manuscript version) [Computer software], Zenodo [code and data set], https://doi.org/10.5281/zenodo.23159628, 2026.

Sawadekar, K., Song, Y., Pan, M., Beck, H., McCrary, R., Ullrich, P., Lawson, K., and Shen, C.: Improving differentiable hydrologic modeling with interpretable forcing fusion, J. Hydrol., 659, 133320, https://doi.org/10.1016/j.jhydrol.2025.133320, 2025.

Shen, C., Appling, A. P., Gentine, P., Bandai, T., Gupta, H., Tartakovsky, A., Baity-Jesi, M., Fenicia, F., Kifer, D., Li, L., Liu, X., Ren, W., Zheng, Y., Harman, C. J., Clark, M., Farthing, M., Feng, D., Kumar, P., Aboelyazeed, D., Rahmani, F., Song, Y., Beck, H. E., Bindas, T., Dwivedi, D., Fang, K., Höge, M., Rackauckas, C., Mohanty, B., Roy, T., Xu, C., and Lawson, K.: Differentiable modelling to unify machine learning and physical models for geosciences, Nat. Rev. Earth Environ., 4, 552–567, https://doi.org/10.1038/s43017-023-00450-9, 2023.

Song, Y., Sawadekar, K., Frame, J. M., Pan, M., Clark, M. P., Knoben, W. J. M., Wood, A. W., Lawson, K. E., Patel, T., and Shen, C.: Physics-Informed, Differentiable Hydrologic Models for Capturing Unseen Extreme Events, Water Resour. Res., 62, e2025WR040414, https://doi.org/10.1029/2025WR040414, 2026.

Spuler, F. R., Wessel, J. B., Comyn-Platt, E., Varndell, J., and Cagnazzo, C.: ibicus: A new open-source Python package and comprehensive interface for statistical bias adjustment and evaluation in climate modelling (v1.0.1), Geosci. Model Dev., 17, 1249–1269, https://doi.org/10.5194/gmd-17-1249-2024, 2024.

Srivastava, A. K., Grotjahn, R., and Ullrich, P. A.: Evaluation of historical CMIP6 model simulations of extreme precipitation over contiguous US regions, Weather Clim. Ext., 29, 100268, https://doi.org/10.1016/j.wace.2020.100268, 2020.

Tsai, W.-P., Feng, D., Pan, M., Beck, H. E., Lawson, K., Yang, Y., Liu, J., and Shen, C.: From calibration to parameter learning: Harnessing the scaling effects of big data in geoscientific modeling, Nat. Commun., 12, 5988, https://doi.org/10.1038/s41467-021-26107-z, 2021.

Ullrich, P.: Validation of LOCA2 and STAR-ESDM Statistically Downscaled Products, OSTI OAI, US Department of Energy Office of Scientific and Technical Information, https://doi.org/10.2172/2202926, 2023.

Vano, J., Hamman, J., Gutmann, E., Wood, A., Mizukami, N., Clark, M., Pierce, D., Cayan, D., Wobus, C., Nowak, K., and Arnold, J.: Comparing downscaled LOCA and BCSD CMIP5 climate and hydrology projections – Release of downscaled LOCA CMIP5 hydrology, US Bureau of Reclamation, https://gdo-dcp.ucllnl.org/downscaled_cmip_projections/ (last access: 9 November 2026, 2020.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I.: Attention is all you need, arXiv [preprint], https://doi.org/10.48550/arXiv.1706.03762, 2017.

Wang, F., Tian, D., and Carroll, M.: Customized deep learning for precipitation bias correction and downscaling, Geosci. Model Dev., 16, 535–556, https://doi.org/10.5194/gmd-16-535-2023, 2023.

Wang, L. and Chen, W.: Equiratio cumulative distribution function matching as an improvement to the equidistant approach in bias correction of precipitation, Atmos. Sci. Lett., 15, 1–6, https://doi.org/10.1002/asl2.454, 2014.

Zhou, M.: Softplus regressions and convex polytopes, arXiv [preprint], https://doi.org/10.48550/arXiv.1608.06383, 2016.