https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

Effectiveness of Multivariate Bias Correction in Hydrology and Agriculture: A Systematic Review

Bhuwan P. Shah[1], Ryan P. McGehee[1] and William J. Gutowski[2]

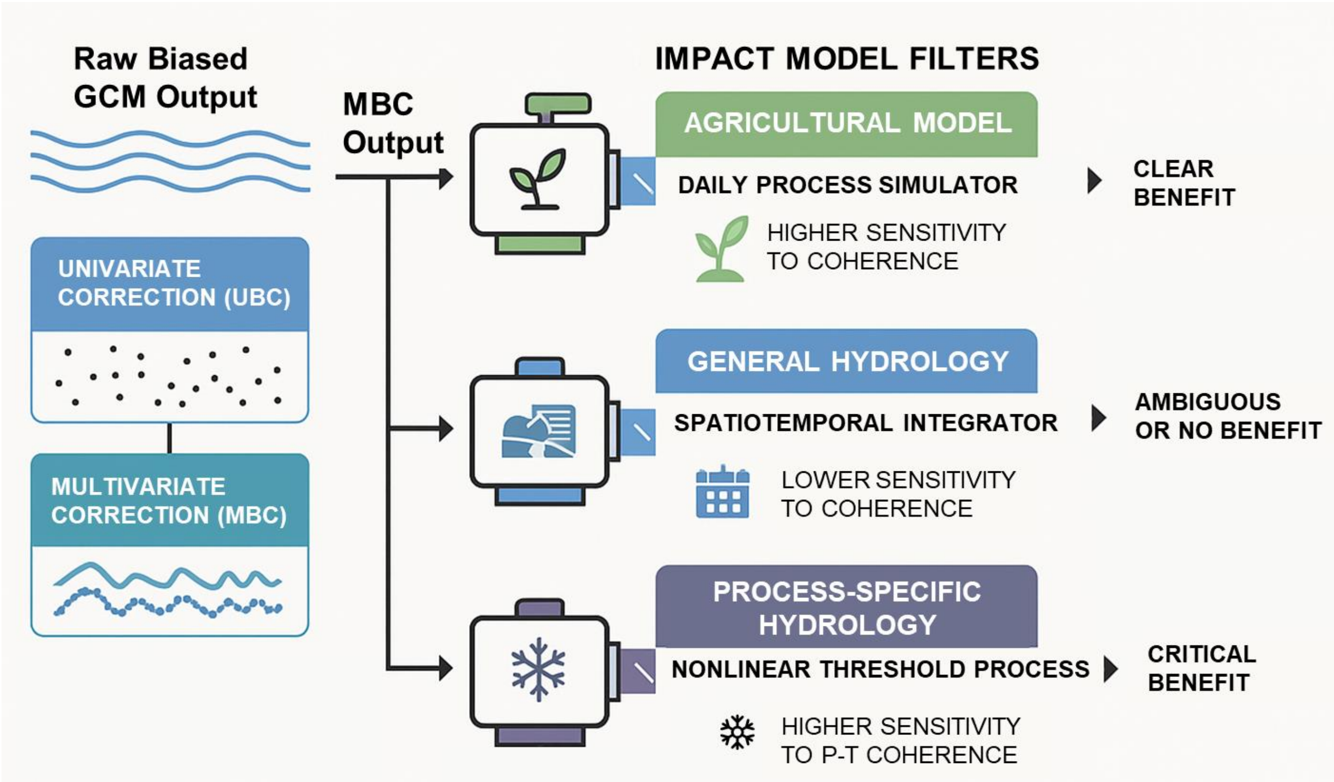
[1]Department of Agricultural and Biosystems Engineering, Iowa State University, Ames, 50011, USA
[2]Department of Agronomy, Iowa State University, Ames, 50011, USA

Correspondence to: Bhuwan P. Shah (bshah@iastate.edu)

Abstract. Climate impact assessments in hydrology and agriculture often rely heavily on outputs from Global Climate Models (GCMs). However, a fundamental scale mismatch exists between the coarse resolution of GCMs and the fine-scale, multivariate data required by impact models. While Multivariate Bias Correction (MBC) methods have emerged as a solution to restore inter-variable dependencies (e.g., the correlation between precipitation and temperature), it remains unclear whether statistical improvements in climate data translate into more accurate impact projections. This study presents a systematic review of 39 peer-reviewed articles to evaluate the added value of MBC across hydrological and agricultural domains.

Our synthesis reveals a critical "validation gap" where superior statistical performance does not consistently yield improved impact simulations. We identify a divergence in added value dictated by the characteristic response time scales of the receiving systems. Agricultural models, which are often sensitive to immediate, daily compounded extremes (e.g., heat stress during low soil moisture), demonstrate a clear benefit from MBC. In contrast, general rainfall-runoff models often function as spatiotemporal integrators, acting as low-pass filters that dampen high-frequency incoherence; consequently, simpler univariate methods frequently perform equally well for bulk streamflow simulation. Furthermore, we highlight the risks of non-stationarity, where methods calibrated to historical correlations may fail under future climate regimes. We conclude that future method development must pivot from purely statistical refinement to more process-aware, regime-dependent frameworks. The ultimate goal is to produce methods capable of addressing non-stationarity and determining when—or if—multivariate correction adds value over simpler univariate approaches.

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026



1 Introduction

The scientific consensus on anthropogenic climate change necessitates robust projections of future climate conditions to inform adaptation and mitigation strategies (Edwards, 2011). The primary tools for generating these projections are General Circulation Models (GCMs) and their dynamically downscaled counterparts, Regional Climate Models (RCMs) (François et al., 2020). While these models are skillful at simulating climate dynamics at global and continental scales, a fundamental mismatch exists between the resolution of their output and the scale at which climate change impacts are experienced and managed (Doblas-Reyes et al., 2021). This scale discrepancy renders raw GCM and RCM output largely unsuitable for the detailed, localized assessments required in critical sectors such as hydrology and agriculture (Ramirez-Villegas et al., 2013; Meyer et al., 2019). Impact applications require fine-scale data, often at 25–1 km spatial resolutions and daily to sub-hourly temporal frequencies. This creates a fundamental scale mismatch, as GCMs typically operate at coarse nominal horizontal grid spacings of 50–260 km (Flato et al., 2013). It is crucial to distinguish this grid spacing from the model's effective resolution—the scale at which physical features are robustly resolved—which is typically several times coarser than the nominal grid spacing (Skamarock, 2004). Furthermore, while Regional Climate Models (RCMs) increasingly bridge this gap with finer simulations, GCM grid spacings have not decreased linearly with available computing power; instead, resources are frequently prioritized for increasing ensemble sizes or integrating complex Earth system processes, such as the carbon cycle, dynamic

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

vegetation, or atmospheric chemistry (Haarsma et al., 2016). Consequently, these models must aggregate their internal sub-hourly calculations (100–300s) to hourly or coarser resolutions—a necessary step to manage data storage volumes that already range from terabytes to petabytes for each model's output.

1.1 Nature of Climate Model Bias and the Perils of Correction

This coarse resolution is not merely a matter of detail; it fundamentally compromises the ability of climate models to represent the physical processes that govern local and regional climate, leading to systematic biases when compared to historical observations (Cannon, 2016; François et al., 2020).

These biases arise from a few key factors, namely, resolution, parameterization, and drift. First, with respect to resolution, a model grid cell of 100–250 km averages out complex terrain, coastlines, small islands, and land cover, erroneously smoothing local climate drivers (Prein et al., 2015). Orographic precipitation and temperature dynamics, convective currents, local frontal systems, and similar dynamics all become impossible to represent at these scales. This often leads to precipitation from small, intense storms being spread across a large grid cell, resulting in the well-known "drizzle bias," where models produce rainfall too frequently but with too little intensity (Dai, 2006; Stephens et al., 2010).

Second, physical processes that occur at scales smaller than a model grid, such as cloud formation and convection, cannot be explicitly resolved and must be represented by simplified parameterizations. These necessary simplifications are a major source of uncertainty and can introduce systematic errors into model simulations over time (Rummukainen, 2010). Every model is unique, and each may opt to substitute different processes with these simplifying parameterizations. While centralized documentation initiatives such as the Earth System Documentation (ES-DOC) exist, they often focus on high-level model configurations rather than the granular algorithmic details of specific parameterizations. Because this deep technical information is frequently dispersed across technical reports or source code rather than being readily queryable, it remains difficult to comprehensively anticipate potential biases for any given model or region. This challenge is further compounded by the fact that model versions evolve over time, making the tracking of these systematic errors a significant hurdle for impact modelers.

Finally, small differences in a model's setup—such as its initial state or the specific parameters driving its processes—can cause systematic drifts in key variables, most notably ocean heat content, sea surface temperature (SST), and salinity (Sen Gupta et al., 2012; Sen Gupta et al., 2013). For example, many models exhibit "climate drift" in deep-ocean heat content and sea surface temperatures (SST), where the model simulates a warming or cooling trend even in the absence of external forcing simply because the system has not reached a state of internal equilibrium (Hobbs et al., 2016). Land-surface components also experience systematic biases and drifts in soil moisture, such as biases in soil thermal dynamics for permafrost regions (Koven et al., 2013) or underestimation of short-term soil moisture variability alongside overestimation of long-term variability (Xi et al., 2022), which can propagate into significant errors in long-term hydrological projections (IPCC, 2013). Furthermore, the observational data used for model evaluation is often temporally inhomogeneous due to the changing composition of the observing system, such as the blending of distinct satellite missions or the discontinuation of in situ stations. This lack of

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

homogeneity, combined with spatial gaps in remote regions, creates a "moving target" for bias correction efforts (Funk et al., 2015; Merchant et al., 2019).

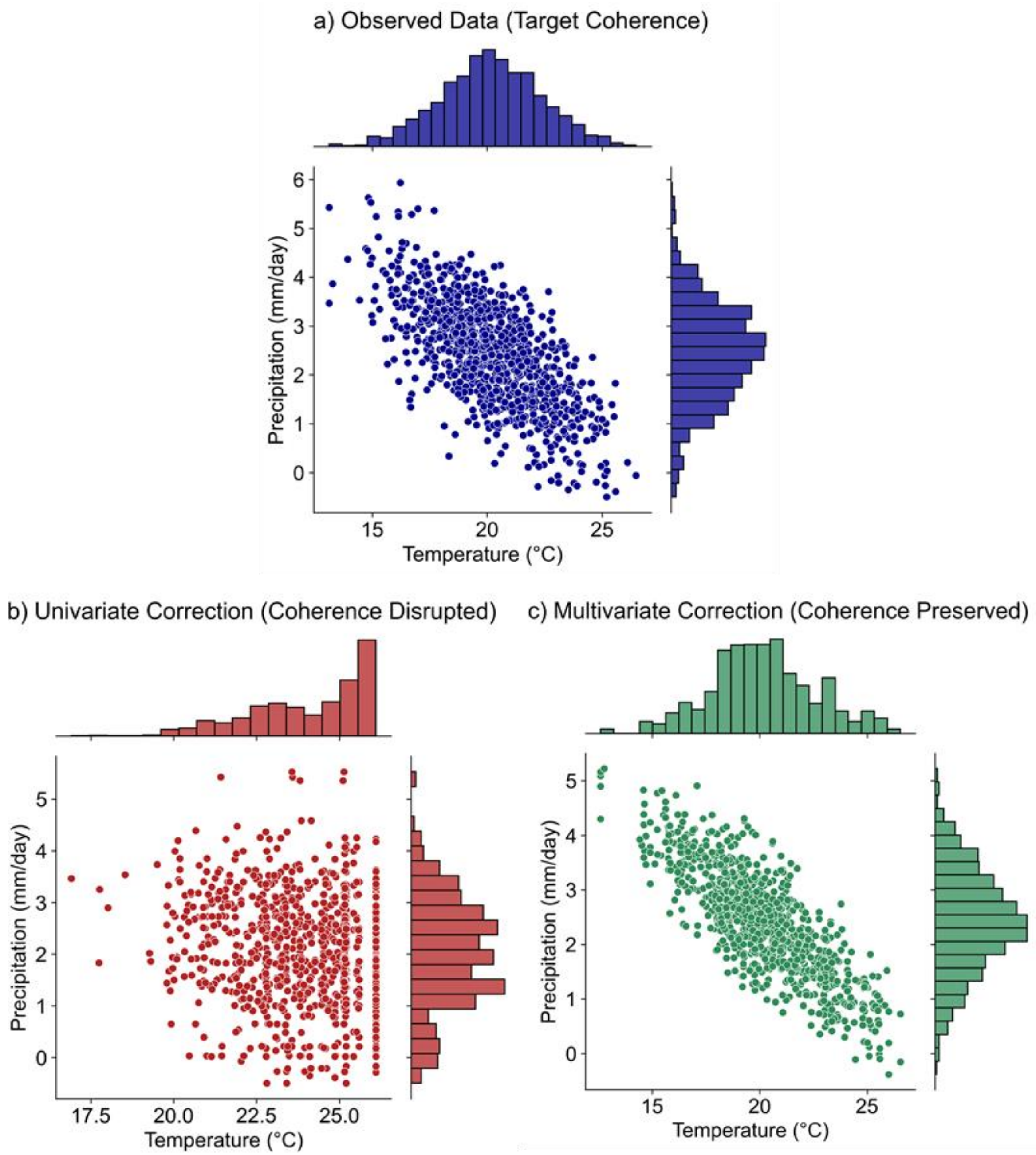
Consequently, a statistical post-processing step known as bias correction has become a standard procedure in the climate impact modeling chain (Mehrotra & Sharma, 2015). However, while powerful, bias correction is not a panacea and must be applied with caution. In many cases, the process can introduce new artifacts or distort important climate signals if applied indiscriminately. In a highly cited paper, Ehret et al. (2012) argued that while bias correction improves alignment with historical data, it may compromise the physical consistency of future projections. Almost all bias correction methods rely on the assumption of stationarity—the idea that the statistical relationship between modeled and observed climate remains constant over time. In a changing climate, this assumption may not hold, and a correction based on historical relationships could inadvertently distort long-term trends or future variability projected by the physics-based model (Christensen et al., 2008; Ehret et al., 2012). This risk is particularly acute when the magnitude of the bias approaches or exceeds the projected climate change signal itself. For instance, Boberg and Christensen (2012) demonstrated that models exhibiting systematic biases in warm, dry climates likely overestimate the regional amplification of global warming. This suggests that inherent model deficiencies can fundamentally alter simulated feedback processes, particularly those driven by enhanced drying.

1.2 Foundational Flaw of Univariate Bias Correction

For decades, the field of bias correction was dominated by univariate methods, most notably Quantile Mapping (QM) and its prominent variant, Quantile Delta Mapping (QDM), which correct the statistical distribution of each climate variable independently of all others (Räty et al., 2018; Cannon et al., 2015). This family of methods operates by aligning the cumulative distribution function (CDF) of a modeled variable (e.g., daily temperature) with the CDF of the corresponding observed variable, effectively removing biases across the full range of values, from the mean to the extremes (Thrasher et al., 2012).

Despite its effectiveness at this task, the independent treatment of each variable is a profound and critical flaw. By design, univariate methods ignore and often destroy the physically meaningful dependence structure—or coherence—that links variables like temperature, precipitation, and humidity (Sippel et al., 2016). Because each variable's time series is adjusted according to its own quantile without regard to the others, the physical relationships that bind them are broken. This can produce physically implausible weather combinations, such as intense precipitation occurring with unrealistically low relative humidity, leading to distorted and misleading results from impact models (François et al., 2020). This disruption of inter-variable coherence is illustrated in Figure 1, which visually contrasts the preserved dependence structure of a multivariate correction against the scattered, incoherent output of a univariate correction. The theoretical foundation for this critique was firmly established by Bhowmik and Sankarasubramanian (2019), who provided the mathematical proof that univariate methods are fundamentally incapable of consistently correcting inter-variable cross-correlations and, under certain conditions, can even amplify existing correlation biases.

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026



105 **Figure 1: A comparison of bias correction methods using synthetic data generated from multivariate normal distributions.** Scatter plots show Temperature vs. Precipitation for: (a) The target 'Observed' distribution (simulated with a strong negative correlation, $\rho = -0.7$); (b) The output from Univariate Quantile Mapping, which corrects the marginal distributions of temperature and precipitation independently but fails to reconstruct the dependence structure, resulting in a physically incoherent scatter; and (c) The output from Multivariate Bias Correction (MBC) using Cholesky decomposition, which corrects both the marginal distributions and the inter-variable covariance, successfully recovering the target dependence structure.

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

1.3 Evolution of Multivariate Bias Correction as a Proposed Solution

This critical limitation of univariate approaches spurred the development of multivariate bias correction (MBC) methods. These techniques represent the state-of-the-art in the field, as they aim to correct the marginal distributions of variables while simultaneously preserving or correcting their inter-variable dependence structure (Cannon, 2016). The preservation of coherence is particularly crucial for hydrological and agricultural systems, where the interplay between variables is a primary driver of system response. For instance, snow accumulation and melt are governed by the joint distribution of at least temperature and precipitation (P-T), a relationship that is fundamental to the hydrology of alpine and high-latitude regions (Meyer et al., 2019; Nury et al., 2021), and this ignores other more complex short and longwave radiative and canopy factors. Similarly, crop growth and yield are complex functions of the combined and timely effects of temperature, solar radiation, water and nutrient availability, and pest or disease presence (Galmarini et al., 2024; Zhang & Paustian, 2019).

The need for a multivariate perspective is further underscored by the growing scientific focus on compound events—high-impact events resulting from the interaction of multiple climate drivers, such as concurrent heatwaves and droughts. The assessment of such events fundamentally requires a multivariate approach to bias correction that can preserve the joint probabilities of their constituent variables (Bevacqua et al., 2017).

1.4 Rationale and Objectives of this Systematic Review

Despite the growing adoption of MBC methods and their conceptual superiority, their practical value is not always clear. A critical challenge is the potential for a "validation gap," where methods that show superior performance on direct statistical metrics (e.g., preserving inter-variable correlations) do not consistently translate into superior performance in impact-based assessments (e.g., more accurate streamflow or crop yield simulations). The reasons for this disconnect are a central focus of this review. It is unclear, for example, if this gap stems from practical barriers—such as the high implementation complexity of MBC methods, a lack of standardized software, or data limitations in CMIP archives—or from disciplinary inertia and a lack of awareness. More fundamentally, it is also possible that the conceptual claims of MBC methods do not always hold in practice, failing to achieve their stated goals in real-world applications. In short, a comprehensive and systematic synthesis of this issue is lacking. This systematic review addresses this gap by identifying, evaluating, and synthesizing the evidence from the peer-reviewed literature to answer the central research question: *How effective are multivariate bias correction methods for applications in hydrology and agriculture?* By examining the evidence base, this review aims to provide clear guidance for researchers and practitioners on the selection and application of appropriate bias correction methods, to highlight the current limitations of the state-of-the-art, and to identify the most pressing challenges for future research.

2 Systematic Review Protocol

This section details the formal protocol used to conduct this systematic review, ensuring the process is transparent, objective, and reproducible. It outlines the search strategy used to identify the initial corpus of articles, the inclusion and exclusion criteria

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

applied to select the final studies for analysis, and the data extraction and synthesis methods used to evaluate the evidence and answer the central research question.

2.1 Literature Search and Screening Process

The literature search was conducted by applying the query strings and keywords listed in Figure 2 to the Web of Science and Scopus databases, covering articles published up to June 2025. The Scopus search yielded 58 articles, of which 45 were accessible through the author's interlibrary access system, while the Web of Science search yielded 46 articles, of which 42 were accessible. After combining these results and removing duplicates, a corpus of 60 unique peer-reviewed studies was identified. These 60 studies were then screened for relevance following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) methodology, as illustrated in Figure 2. The titles and abstracts of all 60 articles were screened against the predefined inclusion and exclusion criteria outlined below. The full text of potentially eligible articles was then thoroughly reviewed. This process resulted in a final corpus of 39 unique studies that were included in the thematic synthesis. The remaining 21 studies were excluded.

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

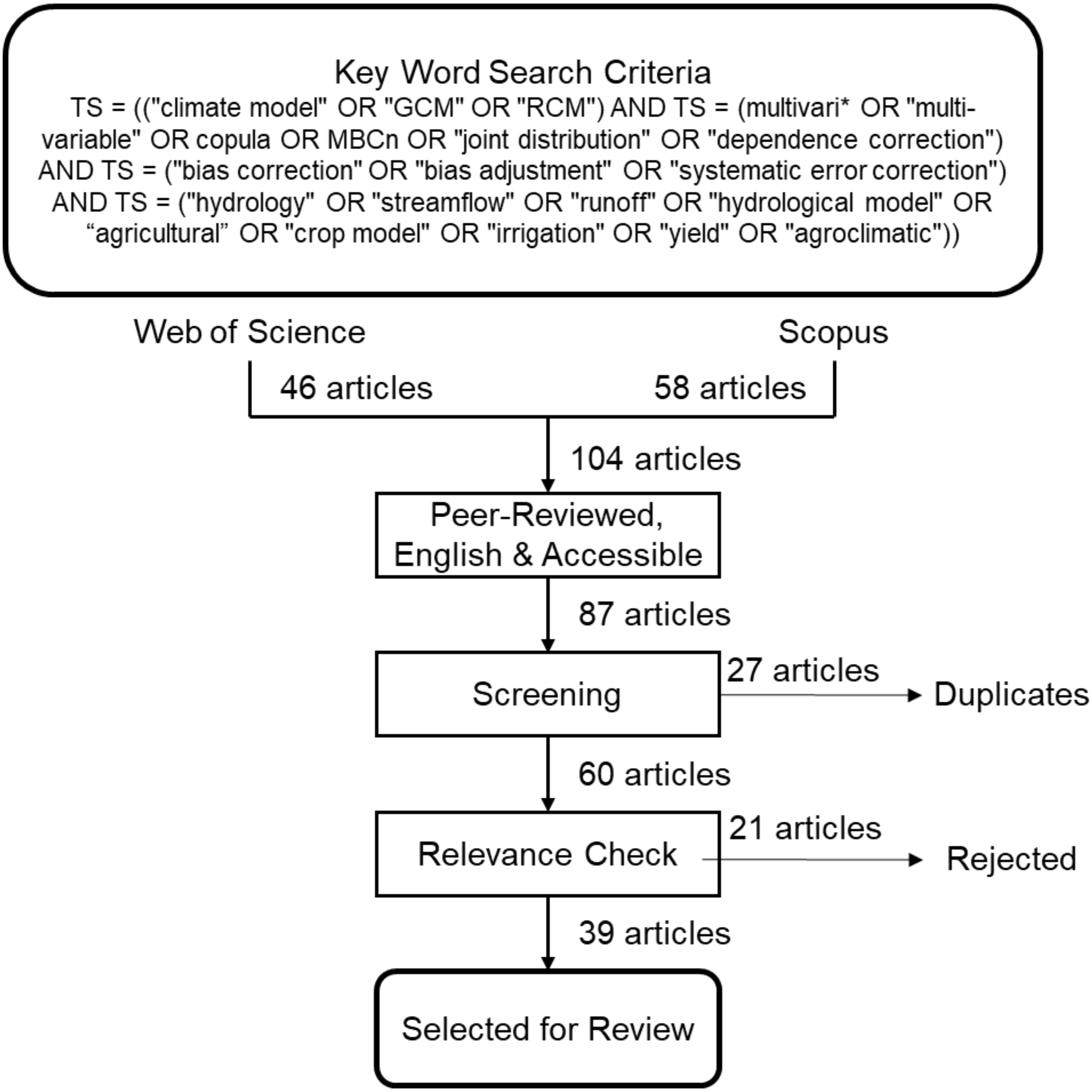


Figure 2: A flow diagram of the literature search and selection process used in this review.

155 ## 2.2 Selection Criteria

To ensure a focused and relevant synthesis, a strict set of inclusion and exclusion criteria was applied to the 63 unique studies identified.

Studies were included in the final synthesis if they met all three of the following conditions:

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

1. The study applied, developed, or performed a comparative evaluation of one or more multivariate bias correction methods.
2. The study evaluated the performance of the method(s) using either direct statistical metrics of multivariate performance (e.g., correlation matrices, energy distance) or an indirect, impact-based assessment.
3. The study's application domain was explicitly hydrology (e.g., streamflow, snowpack, or soil moisture modeling) or agriculture (e.g., crop yield, biomass, or irrigation modeling).

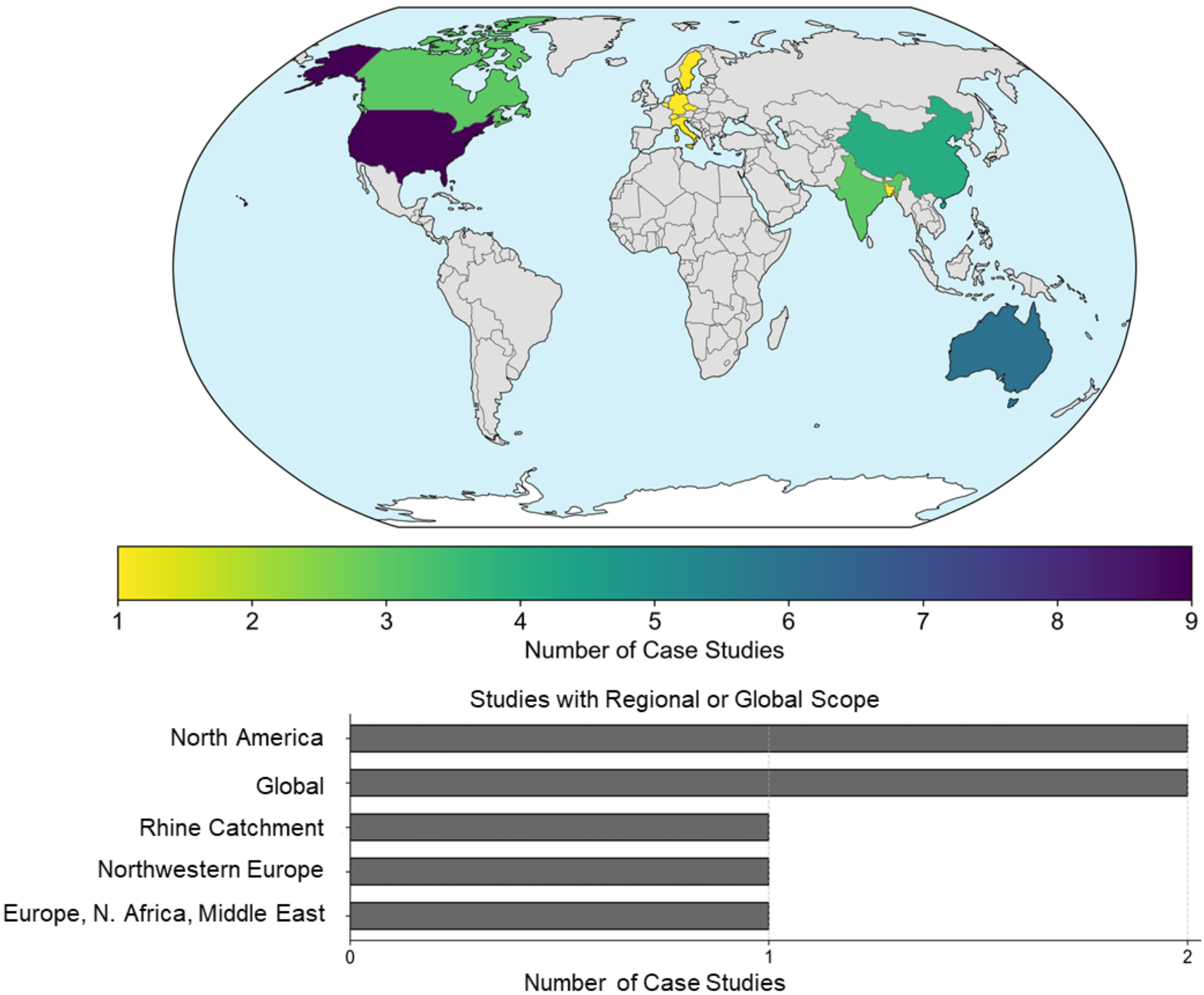
Studies were excluded from the final synthesis if they met one or more of the four following conditions:

1. The study focused exclusively on the application or comparison of univariate bias correction methods (e.g., Madani et al., 2024; Yang et al., 2022).
2. The study focused on a methodological area other than statistical bias correction, such as dynamical downscaling, GCM evaluation and selection, model ensembling, or generative AI-based downscaling (e.g., Lopez-Gomez et al., 2025; Ahmadalipour et al., 2017).
3. The study was a climate impact assessment that used pre-corrected data as an input but did not itself evaluate or compare the bias correction methodology (e.g., Singh et al., 2023; Gu et al., 2020).
4. The application of a multivariate statistical tool (e.g., a copula) was for a purpose other than bias correction of climate model variables, such as for risk analysis of compound events (e.g., Liu et al., 2019).

2.3 The Final Corpus: Overview of Included and Excluded Studies

Application of the inclusion and exclusion criteria to the 63 unique studies resulted in a final corpus of 40 studies that were included in the thematic synthesis. These 40 articles form the evidence base for this review. The geographical distribution of the case studies from these articles is shown in Figure 3, highlighting a concentration of research in North America, Europe, and parts of Asia. The key characteristics of the included articles are summarized in Table 1. The remaining 23 studies were excluded from the final synthesis. A comprehensive list of these studies and the specific rationale for their exclusion is provided in Table 2. This transparent reporting of both included and excluded studies is a cornerstone of the systematic review methodology, ensuring clarity on the scope and boundaries of the evidence synthesized.

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026



185 **Figure 3: The geographical distribution of the 39 case studies included the final synthesis. The map (top) illustrates the location of country-specific studies, while the bar chart (bottom) details studies conducted at a regional or global scale that were not mapped to a single country.**

Table 1: Characteristics of the 39 studies included in the final synthesis. Studies are sorted alphabetically by first author. Abbreviations are provided in the table footnotes.

Study	Application Domain	Primary MBC Method(s) Evaluated	Study Type	Geographical Region
Ahn et al. (2023)	Hydrology	dOTC, MBCn, MRec	CA	South Korea
Alder & Hostetler	Hydrology	MACA	CA	Western United States

(2019)				
Asong et al. (2020)	Data Provision / Climate Services	MBCn	DD	Mackenzie River Basin Canada
Bachelet et al. (2017)	Ecology / Vegetation Dynamics	MACA	AEM	Western United States
Bevacqua et al. (2017)	Hydrology, Compound Events, Methodological Evaluation	PCCs / Vine Copula	PNM	Ravenna, Italy
Bhowmik & Sankarasubramanian (2019)	Methodological Evaluation (Theoretical)	N/A (Theoretical)	OoS	Contiguous United States
Cannon (2016)	Methodological Evaluation, Hydrology	MBCp, MBCr	PNM	Global land areas, West coast of North America
Cannon et al. (2021)	Data Provision / Climate Services	MBCn	DD	North America (0.5° grid)
Chen et al. (2018)	Hydrology	JBC	CA	12 watersheds in USA, Canada, China, and Benin
Das Bhowmik et al. (2017)	Methodological Evaluation, Hydrology	ACCA	PNM	Conterminous United States (CONUS)
Dembélé et al. (2022)	Hydrology	CDF-t	AEM	Volta River basin, West Africa
Funk et al. (2025)	Methodological Evaluation	ZIV-Copula	PNM	Southern Germany
Galmarini et al. (2024)	Agriculture, Methodological Evaluation	MBCn/p/r, R2D2, ISIMIP3BASD	CA	Europe, North Africa, Middle East (21 locations)
Guo et al. (2020)	Hydrology / Methodological Evaluation	MBCn, MBCp, dOTC	CA	North America (12 catchments)
Hakala et al. (2018)	Hydrology (as an evaluation tool)	N/A (Univariate)	NEF	57 Swiss catchments
Hanggoro et al. (2020)	Agriculture / Hydrology	Copula (N-copula)	CA	Jiangxi Province China
Hnilica et al. (2017)	Methodological Evaluation	PCA + QM	PNM	Czech Republic

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

Jin et al. (2015)	Methodological Evaluation	QPM	PNM	Australia (15 sites)
Jin et al. (2022)	Agriculture	Copula	AEM (Seasonal)	Australian wheatbelt
Khatun et al. (2022)	Hydrology / Rainfall Forecasting	Copula; eKSOM	CA	Hirakud Reservoir catchment India
Lazoglou et al. (2019)	Hydrology	Copula	AEM	Mesta/Nestos river basin
Mao et al. (2015)	Methodological Evaluation	Copula	PNM	Germany
Mehrotra & Sharma (2015)	Methodological Evaluation, Hydrology	MRNBC	Proposal of new method	Sydney, Australia
Mehrotra & Sharma (2016)	Methodological Evaluation / Downscaling	MRQNBC	PNM	Sydney Australia
Meyer et al. (2019)	Hydrology	MBCn	AEM	Alpine catchments (Rhine headwaters, Switzerland)
Miralha et al. (2021)	Hydrology / Water Quality	N/A (Univariate)	AEM (New Domain)	Maumee River watershed, USA
Nury et al. (2021)	Hydrology	MRNBC	AEM	Upper Brahmaputra basin, Tibetan Plateau
Räty et al. (2018)	Hydrology	dOTC, Copula	AEM	3 catchments in Sweden and Finland
Schepen et al. (2019)	Methodological Evaluation, Agriculture, Hydrology	Schaake Shuffle	PNM	Burdekin region, Australia
Schepen et al. (2020)	Methodological Evaluation / Seasonal Forecasting	BJP; Schaake shuffle	CA	Australia
Shrestha et al. (2019)	Hydrology	MBCn	AEM	Liard basin, subarctic northwestern Canada
Sippel et al. (2016)	Climate Impact Modelling	BCC	PNM	Global (gridded land surface)
Su et al. (2020)	Hydrology	MSMBC	PNM	Gan River basin, China
Tootoonchi et al. (2022)	Hydrology	Copula, MBCn	CA	55 Swedish catchments

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

Tootoonchi et al. (2023)	Hydrology	Copula, MBCn	AEM	55 Swedish catchments
Van de Velde et al. (2022)	Methodological Evaluation / Hydrology	MRQNBC; MBCn; dOTC; R2D2	CEM	Uccle Belgium
Wen et al. (2024)	Agriculture (Drought)	VCPM	PNM	Loess Plateau, China
Yin et al. (2019)	Methodological Evaluation	MBC (Mehrotra and Sharma, 2015)	AEM	Northwestern China
Zhang & Paustian (2019)	Agriculture, Ecology	MACA, BCCA	CA	Northeastern Colorado, Southeastern North Dakota, Iowa

Study Type: AEM (Application of Existing Method); CA (Comparative Analysis); CEM (Critical Evaluation of Methods); DD (Data Descriptor); NEF (Novel Evaluation Framework); OoS (Out of Scope); PNM (Proposal of New Method).

Primary MBC Method(s) Evaluated: ACCA (Asynchronous Canonical Correlation Analysis); BCC (Multivariate bias correction and constructing storylines); BCCA (Bias-Corrected Constructed Analogs); BJP (Bayesian joint probability); CDF-t (Cumulative Distribution Function transform); dOTC (dynamical optimal transport); eKSOM (enhanced Kohonen self-organizing map); ISIMIP3BASD (ISIMIP3 Bias Adjustment method); JBC (Joint Bias Correction); MACA (Multivariate Adaptive Constructed Analogs); MBC (Multivariate Bias Correction); MBCn/p/r (Multivariate Bias Correction, n-dimensional/Pearson/rank); MRec (Matrix Recorrelation); MRNBC (Multivariate Recursive Nesting Bias Correction); MRQNBC (Multivariate Recursive Quantile Nesting Bias Correction); MSMBC (Multi-site multivariate bias correction); N/A (Not Applicable); PCCs (Pair-Copula Constructions); PCA + QM (Principal Component Analysis with Quantile Mapping); QPM (Quantile Projection Method); R2D2 (Rank Resampling for Distributions and Dependences); VCPM (Vine Copula-Based Projection Model); ZIV-Copula (Zero-inflated vine copulas).

Table 2: The rationale for exclusion of each of the 21 studies excluded from the final synthesis. Studies are sorted alphabetically by first author.

Study	Criterion ID	Justification for Exclusion
Ahmadalipour et al. (2017)	1	Proposes a framework for GCM selection, a step prior to bias correction.
Das et al. (2015)	1	Proposes a statistical model to identify drivers of extremes, not to correct model output.
Demirel & Moradkhani (2016)	1	Focuses on Bayesian Model Averaging for model ensembling, not bias correction.
Eghdamirad et al. (2019)	1	Proposes a method to propagate multivariate uncertainty, not to perform MBC.

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

Faghih et al. (2022)	4	Investigates the temporal scale (sub-daily vs. daily) of correction, not multivariate vs. univariate.
Ganguli et al. (2020)	2	A climate impact study that uses pre-processed data and does not evaluate the BC methods.
García-Valdecasas Ojeda et al. (2022)	2	Drives a hydrological model with raw RCM outputs without bias correction.
Gu et al. (2020)	2	Uses data from the ISIMIP project without evaluating the correction method.
Huang et al. (2021)	3	Proposes a sophisticated univariate method, not a multivariate one.
Janssen et al. (2021)	2	Uses univariately bias-corrected data as an input.
Ji & Ahn (2023)	1	Proposes a stochastic streamflow generator; does not apply MBC to climate models.
Khanal et al. (2019)	2	Uses univariately bias-corrected data for a compound event study.
Kim et al. (2021)	4	Uses copulas for temporal downscaling of a single variable, not inter-variable correction.
Lee & Ouarda (2018)	1	Focuses on statistical spatial downscaling, a related but distinct field.
Liu et al. (2019)	4	Uses copulas for risk analysis of compound events, not for bias correction of inputs.
Lopez-Gomez et al. (2025)	1	Focuses on generative AI-based downscaling, not statistical bias correction.
Madani et al. (2024)	3	Compares two univariate methods (Delta Change and EQM).
Qian et al. (2019)	2	Uses data corrected with a univariate trend-preserving method.
Schnur & Lettenmaier (1998)	1	An early weather-typing downscaling method, predating modern bias correction approaches.
Singh et al. (2023)	2	Uses data corrected with a univariate method (EQM) without evaluating it.
Yang et al. (2022)	3	Uses the simple univariate delta change method.

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

2.4 Data Extraction and Thematic Synthesis Framework

Following the application of the search strategy described above, the initial identification and de-duplication process yielded 60 unique articles for screening. These were subjected to a multi-stage review process. First, all 60 articles were screened by title and abstract against the inclusion/exclusion criteria, which resulted in 50 potentially relevant studies. Second, these articles underwent a full-text review, during which 11 additional studies were excluded for reasons documented in selection criteria. Finally, a standardized data extraction form was used to collect key information from each of the 39 final included studies. This form cataloged bibliographic details, study context (climate models, geographical region, variables corrected), MBC method details (name, typology), application domain, impact models used, validation metrics employed, and the authors' key findings regarding the method's effectiveness, strengths, and limitations. Following data extraction, a thematic synthesis approach was conducted. This process involved systematically coding the key findings (on effectiveness, strengths, and limitations) from each paper and grouping them based on three primary themes: the application domain (hydrology or agriculture), the validation approach (statistical or impact-based), and commonly reported limitations. This thematic grouping was used to identify recurring patterns, divergent outcomes, and overarching challenges across the body of literature.

3 Multivariate Bias Correction Methods & Validation

This section synthesizes the two key components of the reviewed literature. First, it provides a taxonomy of the prominent multivariate bias correction (MBC) methods identified in the 39-paper corpus, classifying them by their operational structure. Second, it details the different validation approaches—from direct statistical tests to indirect, impact-based assessments—used in the studies to evaluate MBC effectiveness, and it introduces the critical "validation gap" that emerges from these distinct approaches.

3.1 Prominent MBC Methods

The review of the 39 included papers identified a diverse and evolving landscape of MBC methods. These can be broadly classified into three correction categories based on their operational structure (François et al., 2020; Vrac, 2018) including: 'Marginal/Dependence', 'All-in-One', or 'Successive Conditional'. Each class of method is shown in Figure 4.

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

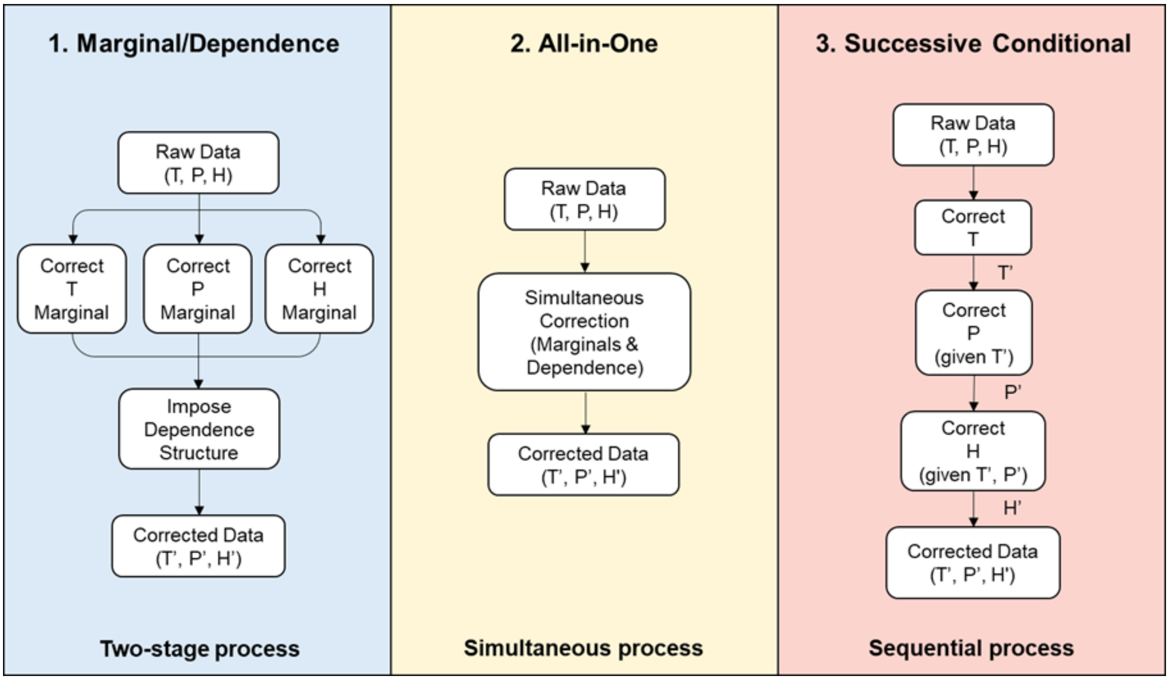


Figure 4: A conceptual diagram illustrating the different operational structures of the three main classes of Multivariate Bias Correction (MBC) methods: (1) Marginal/Dependence, (2) All-in-One, and (3) Successive Conditional. T, P, and H are used as example variables, representing Temperature, Precipitation, and Humidity, respectively. This classification follows the typology proposed by François et al. (2020) and Vrac (2018).

Marginal or dependence correction is the most common category. These methods operate in two stages: first, the marginal distribution of each variable is corrected independently using a univariate technique (like Quantile Mapping), and second, a transformation is applied to impose a target dependence structure observed in historical data. All-in-one correction methods correct both the marginal distributions and the dependence structure simultaneously in a single, integrated step. Successive conditional correction methods correct variables sequentially, where the correction of each subsequent variable is conditioned on the properties of the already-corrected variables.

Within this framework, a number of specific methods were frequently proposed, applied, and evaluated across the literature. Table 3 provides a taxonomy of the most prominent techniques identified in this review, clarifying their core principles and linking them to key references.

Table 3: A taxonomy and characteristics of prominent MBC methods, their core principles, and example references from this review. Abbreviations: multivariate bias correction (MBC), rank resampling (R2D2), dynamical optimal transport (dOTC), matrix recorrelation (MR), multi-site (MS).

Method	Class	Core Principle	Key References
MBCn, MBCp, MBCr	Marginal / Dependence	Iterative random orthogonal rotations and univariate QM to match the full multivariate distribution (MBCn) or specific correlation structures (MBCp/r).	Cannon, (2016) Miralha et al., (2021)
Copula-Based	Marginal / Dependence	Separates marginal distributions from the dependence structure (copula), with flexible modeling of complex, non-linear	Funk et al., (2025) Mao et al., (2015)

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

Methods		dependencies. Advanced forms include Vine Copulas for highly dimensional and 'zero-inflated' data.	
R2D2	Marginal / Dependence	Univariate QM, then a reordering algorithm (often the Schaake Shuffle) to impose the observed rank correlation structure.	Vrac, (2018) Dembélé et al., (2022)
dOTC	All-in-One	Optimal transport theory for finding the most efficient mapping to transform the model's joint distribution into the observed joint distribution.	Ahn et al., (2023) Räty et al., (2018)
MRec	All-in-One	Matrix transformations for simultaneously correcting marginals and the Pearson correlation matrix, often under a Gaussian assumption.	Ahn et al., (2023) François et al., (2020)
MRNBC, MRQNBC	All-in-one	Multivariate autoregressive modeling nested across daily to annual timescales to simultaneously correct biases in moments and auto/cross-correlations.	Mehrotra & Sharma, (2015, 2016)
MSMBC	All-in-one	Multivariate correction with simultaneous handling of inter-variable and inter-site (spatial) dependencies, a key advancement over single-site methods.	Su et al., (2020)

3.2 Validation Approaches: Statistical to Impact-Based

The assessment of MBC effectiveness across the literature is multifaceted, falling into two primary categories: direct statistical tests of the climate data itself, and indirect, application-oriented evaluations that use an impact model as the ultimate arbiter of performance.

Direct Statistical Assessment methods quantify how well the statistical properties of the bias-corrected climate data match those of the observational reference. The most common approach is the comparison of inter-variable and, in more advanced studies, inter-site correlation matrices, typically using Pearson or Spearman correlation coefficients (Ahn et al., 2023; Su et al., 2020). This provides a direct measure of a method's ability to reproduce the observed dependence structure. More advanced metrics include omnibus tests like the energy distance, which measures the statistical distance between entire multivariate probability distributions and is used as a formal convergence criterion in the MBCn algorithm (Cannon, 2016).

Indirect impact assessment evaluates performance within the context of a specific scientific application. While this provides a crucial test of a method's fitness-for-purpose, its role as a definitive arbiter of quality is complex. In hydrology, the predominant method is the use of hydrological signatures. A calibrated hydrological model (e.g., HBV, SWAT, VIC) is driven by different corrected climate inputs, and a suite of metrics describing the simulated streamflow regime (e.g., mean flow, high/low flow quantiles, seasonality, timing) are compared to a reference simulation driven by observed meteorology (Guo et al., 2020; Tootoonchi et al., 2023). In agriculture, a similar strategy is employed. Crop models (e.g., DSSAT, APSIM) are driven by the corrected inputs, and simulated outputs like crop yield, biomass, or Leaf Area Index (LAI) are compared to a reference run. The error, as quantified via statistical measures like Mean Bias, RMSE, Percent Bias (PBIAS; Moriasi et al., 2007), Nash-Sutcliffe Efficiency (NSE; Nash & Sutcliffe, 1970), or Kling-Gupta Efficiency (KGE; Gupta et al., 2009), relative to this reference is the primary metric of effectiveness (Galmarini et al., 2024; Hanggoro et al., 2020). Table 4 summarizes these common validation approaches.

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

Table 4: Common validation metrics for assessing inter-variable coherence

Category	Metric	Description
Direct Statistical	Correlation Matrices	Compares Pearson or Spearman correlation between variables/sites in corrected vs. observed data.
	Energy Distance	Measures the statistical distance between entire multivariate probability distributions.
Indirect Impact-Based	Hydrological Signatures	A suite of metrics describing simulated streamflow characteristics (e.g., mean flow, Q95, seasonality).
	Crop Yield/Biomass Error	The difference (e.g., RMSE) between crop model outputs simulated with corrected data vs. a reference run.
	Compound Event Analysis	Measures the frequency, magnitude, and duration of events defined by the co-occurrence of multiple climate variables.
	Skill Scores (CRPS, ES)	Evaluates the overall skill of a probabilistic forecast, used in agricultural forecasting applications.

3.3 Disconnect Between Statistical and Impact Validation

A critical and recurring theme that emerges from the synthesis of the literature is the existence of a "validation gap," particularly within hydrology. This refers to the phenomenon where methods that demonstrate superior performance on direct statistical metrics (e.g., better preservation of P-T correlations) do not consistently translate into superior performance in indirect, impact-based assessments (e.g., more accurate streamflow simulations). There are several such examples in the literature studies by Hakala et al. (2018), Räty et al. (2018), Guo et al. (2020), and Tootoonchi et al. (2023).

This disconnect suggests a potential mismatch between the objectives of statistical method developers, who often prioritize achieving statistical fidelity, and the needs of impact modelers, for whom process fidelity is paramount. The seminal comparative study by Räty et al. (2018) provided a stark illustration of this gap. They found that while the joint bias correction methods they tested did a better job of preserving inter-variable correlations in the climate data, the added value for hydrological simulations was "surprisingly small and inconsistent," with simpler univariate methods sometimes performing better for certain hydrological signatures. This finding was strongly corroborated by a large-sample study across 55 Swedish catchments by Tootoonchi et al. (2023), who concluded that the added value of multivariate methods was "not systematically reflected" in the resulting hydrological signatures and that, for a wide range of metrics, univariate methods were generally superior.

This validation gap highlights a crucial point: the choice of validation metric is not neutral. It implies that statistical metrics alone may be insufficient for judging the true "effectiveness" of a bias correction method for impact modeling. This has led to

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

the argument, put forward by Hakala et al. (2018), that a more holistic validation approach is needed, one that uses the impact model itself as a process-based evaluation tool to reveal deficiencies that are not apparent from standard statistical tests alone. This recommendation, however, raises a critical follow-up question: are all impact models equally suitable as evaluation tools? The evidence suggests they are not. An impact model is not an objective arbiter, but rather another complex system with its own structural biases. This creates the potential for a "two wrongs make a right" scenario, where a bias correction method that performs poorly on statistical metrics produces superior impact results by coincidentally canceling out a known bias in the impact model. Conversely, a method may perform exceptionally well on statistical metrics (e.g., preserving P-T correlations) but fail to improve impact results if the impact model is structurally insensitive to that specific inter-variable coherence.

For instance, simpler conceptual or empirical models—such as a statistical crop-yield model based on aggregated monthly inputs, or a hydrological model employing Horton's infiltration formulation (Horton, 1940)—tend to be structurally insensitive to the inter-variable coherence that multivariate bias correction (MBC) methods aim to preserve. In contrast, more complex process-based models—such as DSSAT (Jones et al., 2003) or those solving the Richardson-Richards equation (Richardson, 1922; Richards, 1931)—are typically more sensitive to daily inter-variable correlations and other properties that univariate (non-MBC) might inadvertently distort. Therefore, the choice of impact model is not a neutral decision; its fitness for purpose is a crucial, yet often overlooked, factor in the evaluation of bias correction methods.

4 Synthesis of Reported Effectiveness by Application Domain

The synthesis of the 39 included papers reveals a striking divergence in the reported effectiveness of MBC. The added value of correcting inter-variable dependencies is highly dependent on the application domain, with a clear consensus emerging in agriculture that contrasts sharply with a more nuanced and conditional picture in hydrology. This divergence in outcomes is summarized in Figure 5. The evidence for agricultural modeling is overwhelmingly positive, with 100% of relevant studies reporting the clear superiority of multivariate methods. In stark contrast, the benefits for hydrology are conditional. For studies focused on process-specific applications like snow or compound events, a strong majority (83%) found clear superiority for MBC. However, for studies assessing general streamflow, the results were much more mixed, with only 30% finding clear superiority, while 50% reported nuanced benefits and 20% found limited to no benefit over simpler univariate methods.

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

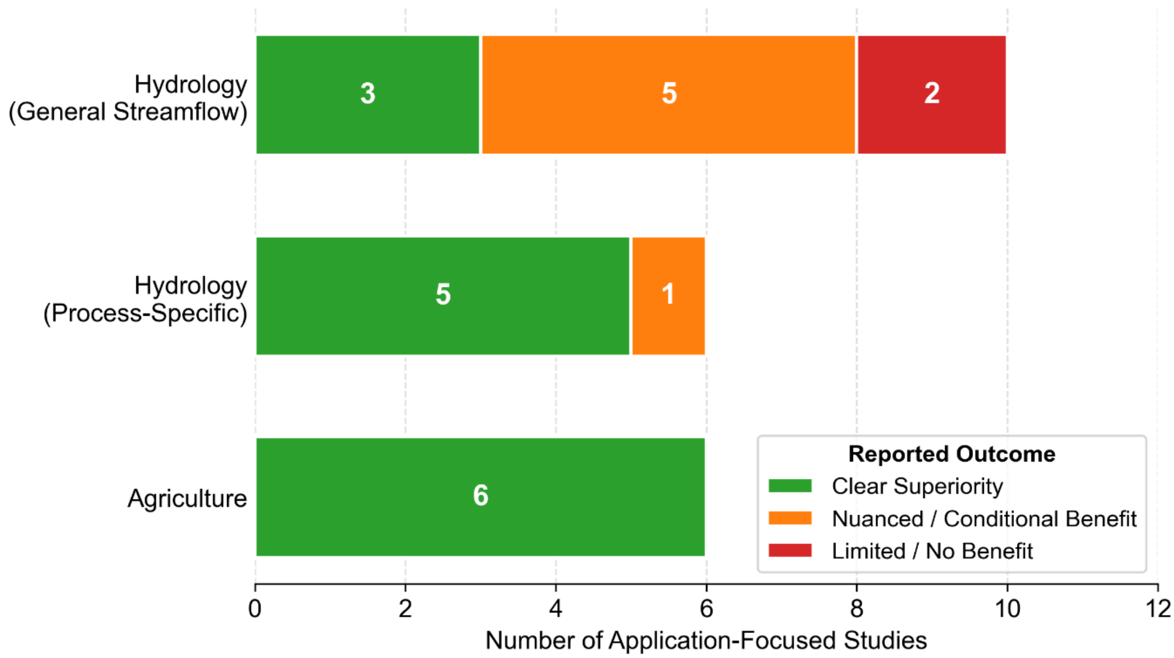


Figure 5. A synthesis of reported MBC effectiveness by application domain.

4.1 Agricultural Applications

The evidence supporting the use of MBC methods for agricultural applications is remarkably consistent and positive. The most definitive evidence comes from a large-scale intercomparison project, the Bias Adjustment for the AgMIP-Wheat project (BADJAM), reported by Galmarini et al. (2024). This comprehensive study involved 12 different AgMIP-Wheat crop models, 8 climate models, and 14 different bias adjustment methods (5 multivariate and 9 univariate) across 21 locations. The core conclusion was unambiguous: the error in crop model outputs (such as yield, biomass, and Leaf Area Index) is "systematically lower" when using multivariate bias-adjusted climate data compared to data from univariate methods (Galmarini et al., 2024). This finding was robust across a wide variety of crop models and geographical locations, leading to the clear recommendation that multivariate treatment should be preferred for driving crop models (Galmarini et al., 2024).

The underlying reason for this superiority may be attributed to the process-based nature of modern crop models. These models mechanistically simulate plant physiology on a daily time step, and their outputs are highly sensitive to the physically consistent interplay of co-occurring variables like temperature, solar radiation, and precipitation (Galmarini et al., 2024). A study by Zhang & Paustian (2019), which evaluated the sensitivity of the DayCent ecosystem model, confirmed that model outputs like crop yield are highly sensitive to the quality of the input weather data, and found that the multivariate MACA method provided better inputs than the BCCA method, which suffered from a biased representation of rainfall frequency that had cascading effects on solar radiation.

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

This consensus is reinforced by other application-specific studies. Hanggoro et al. (2020) compared univariate and multivariate methods for estimating rice irrigation water needs in China and found that the MBC method provided more realistic climate inputs by preserving inter-variable correlations, leading to more reliable estimates of water demand. Similarly, studies focusing on seasonal forecasting have shown that multivariate post-processing of climate forecasts improves the skill of early-season predictions of wheat yield (Jin et al., 2022; Schepen et al., 2019).

4.2 Hydrological Applications

In stark contrast to the clear consensus in agriculture, the evidence for the superiority of MBC in hydrology is far more nuanced, conditional, and at times contradictory. While virtually all forms of bias correction provide a substantial improvement over using raw model output, the added value of complex multivariate methods over simpler univariate ones is not consistently demonstrated for all hydrological applications.

4.2.1 Ambiguous Value for General Streamflow Characteristics

Several comprehensive comparative studies have found that for bulk hydrological signatures (i.e., metrics describing the streamflow regime) related to the overall water balance, the added value of MBC is often limited. As previously discussed in the "validation gap," the work of Räty et al. (2018) and Tootoonchi et al. (2023) found that statistical improvements from MBC were "not obvious" and "not systematically reflected" in many simulated hydrological signatures. The study by Tootoonchi et al. (2023) is particularly compelling, as it found that across 16 signatures and 55 catchments, univariate methods generally outperformed their multivariate counterparts.

Further nuance was added by Guo et al. (2020), who compared MBC methods against a univariate benchmark in 2,840 North American catchments. They found that the benefits of MBC were process-dependent, meaning more pronounced in regimes where hydrological processes are highly sensitive to inter-variable correlations (e.g., evaporation in arid/warm temperate climates, where accurate precipitation-temperature relationships are critical for modeling water loss), and limited in northern snow-characterized watersheds where marginal distributions may suffice for dominant processes like snow accumulation and melt. It is critical to note, however, that this finding applied to bulk streamflow signatures; as discussed in the next section, MBC is essential for the specific snowmelt processes that govern cold-region hydrology.

This disconnect between calibration performance and validation performance, reported in both studies, points to a critical methodological flaw. Both Guo et al. (2020) and Tootoonchi et al. (2023) used split-sample designs with 28-year and 22-year calibration periods, respectively. This raises the question of whether these data lengths are sufficient to robustly parameterize a complex and non-stationary multivariate relationship. Guo et al. (2020) explicitly confirm this weakness, noting that the advantages of MBC methods "are weakened when it comes to the validation period". They attribute this directly to the fact that the "observed intervariable correlation itself is also not invariable", having changed significantly between their calibration and validation datasets. This suggests that for general streamflow, the benefits of correcting inter-variable dependence are not only "overshadowed," a phenomenon explained by the "spatiotemporal integrator" nature of catchment models (discussed

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

further in Section 5.1), but may be fundamentally undermined by the non-stationarity of the very correlations being corrected. It is important to clarify that this non-stationarity refers to observed or simulated inter-variable dependencies (e.g., precipitation-temperature correlations) that can shift due to climate change, large-scale variability, or observational factors, while fundamental physical relationships—such as psychrometric properties governing moist air thermodynamics (e.g., the Clausius-Clapeyron relation between temperature and saturation vapor pressure)—remain invariant.

4.2.2 Critical Importance for Process-Specific Hydrology

While the benefits for general streamflow are ambiguous, the literature provides compelling evidence that MBC is essential for specific hydrological processes that are mechanistically dependent on the joint behavior of climate variables. The most prominent example is in catchments where snow and ice processes are dominant. The temperature-precipitation (P-T) relationship is physically fundamental in these environments, as it governs the phase of precipitation (rain vs. snow) and the rate of snowmelt.

A landmark study by Meyer et al. (2019) in alpine catchments demonstrated this with exceptional clarity. They showed that the simultaneous correction of P and T with the MBCn method led to more simulated snowfall compared to a univariate approach. This seemingly small statistical difference had "considerable consequences" for the simulated hydrology, producing more realistic snow cover, altering the seasonal timing of streamflow, and leading to different projections for the date of future glacier disappearance. The critical importance of preserving the P-T relationship in snow-dominated basins is confirmed by numerous other studies. Nury et al. (2021) found that the MRNBC method provided superior results for simulating streamflow in the Brahmaputra basin by preserving these critical interdependencies. Similarly, Alder & Hostetler (2019) compared a multivariate (MACA) and a univariate (BCSD) downscaling product and found that the choice of method was a major factor in hydroclimate projections, significantly affecting simulated high-elevation snowpack.

Beyond cryospheric processes, there is broad agreement that MBC improves the representation of compound events like floods and droughts by preserving the joint behavior of their driving variables (Bevacqua et al., 2017). Furthermore, the realistic simulation of soil moisture, a key state variable controlling runoff generation, depends on physically coherent climate inputs, as it is a function of the balance between precipitation and evapotranspiration (Dembélé et al., 2022).

5. Discussion

The synthesis of evidence reveals a clear paradox that can be framed through the lens of the "validation gap" introduced earlier: multivariate bias correction is systematically beneficial for agricultural models but offers only conditional benefits for hydrological models. This discussion section addresses this paradox by first exploring the structural differences in impact models that explain this divergence. It then synthesizes the pervasive limitations and "grand challenges" reported across the literature, such as bias non-stationarity and the neglect of temporal structures. Finally, it outlines the practical implications of these findings for researchers and presents a roadmap for future research.

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

5.1 Sectorial Response Differences

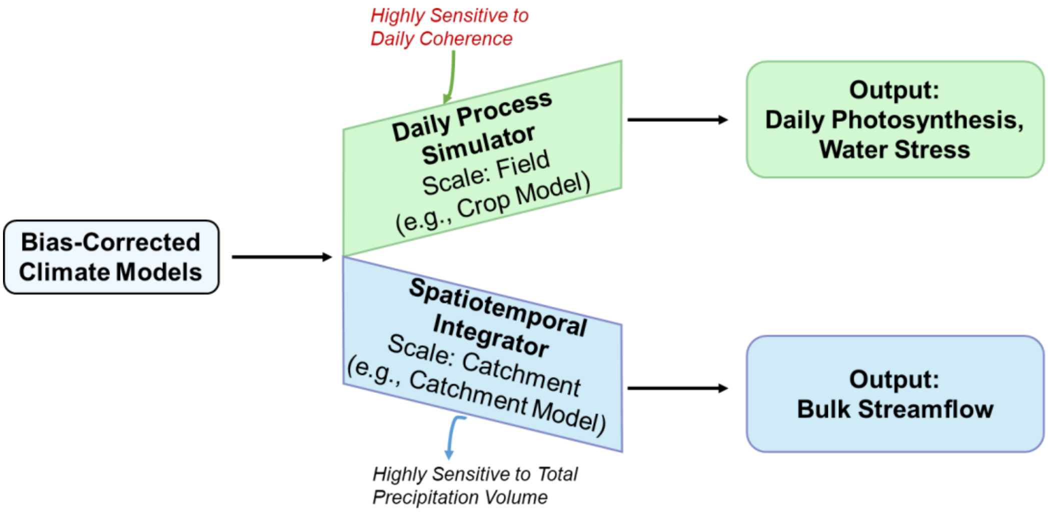
The explanation for this divergence lies not in the MBC methods themselves, but in the fundamental structural differences between the impact models. The impact model effectively acts as a filter, and its sensitivity to different properties of the climate input data determines which aspects of bias correction—and thus which validation approach(es) - are most important.

Agricultural crop models often operate at a field scale and a daily time step, typically representing upland, rain-fed landscapes. They are, in essence, daily process simulators highly sensitive to local, vertical fluxes and the physical plausibility of co-occurring weather variables. An unrealistic combination of high temperature, low solar radiation, or extreme lack or presence of precipitation can have an immediate and direct impact on the model's calculation of photosynthesis, evapotranspiration, water stress, and nutrient uptake. For these models, the coherence of the input variables on a day-to-day basis is paramount for achieving a realistic simulation of plant growth (Galmarini et al., 2024; Zhang & Paustian, 2019). A marginally late or early frost, for instance, can kill an entire crop, leading to premature decomposition, exposure of the soil surface to erosion, and the release of nutrients from decaying plant matter.

In contrast, many hydrological models act as spatiotemporal integrators, often simulating bulk streamflow at a lowland catchment outlet. This model structure naturally aggregates, or integrates, the effects of all upland processes and routes the resulting water through the catchment over periods of days, weeks, or even months. This creates a fundamental distinction in the relevant time scales of the system response. Unlike agricultural systems that often react immediately to daily compound anomalies, the catchment functions as a physical low-pass filter, where the aggregated response dampens high-frequency input fluctuations. Consequently, for simulating bulk streamflow characteristics, the primary driver is often the total volume and broad seasonal distribution of precipitation. In this context, the accuracy of the marginal distribution of precipitation may be far more important than its precise day-to-day correlation with temperature. The integrating nature of the catchment model can smooth out the effects of minor inconsistencies in the daily inputs, explaining why many studies find limited added value from MBC for general streamflow simulation.

However, this logic breaks down when the hydrological process being modeled is not a simple integrator but is instead governed by a sharp, non-linear threshold. The partitioning of precipitation into rain or snow at the 0°C threshold is the canonical example. In this case, the model's sensitivity to the joint P-T distribution returns, and the coherence of the climate inputs becomes critical once again. This conceptual model—of the impact model as a filter that is either a daily process simulator or a spatiotemporal integrator—provides a coherent physical explanation for both the ambiguity of MBC's value for general streamflow and its critical importance for snow-dominated processes. The core components of this conceptual model are illustrated in Figure 6.

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026



420 **Figure 6. Impact model sensitivity to bias-corrected climate inputs.**

5.2 Pervasive Limitations and Methodological Grand Challenges

Across the reviewed literature, a clear set of trade-offs and limitations associated with current MBC methods emerged. These represent the major methodological grand challenges that must be addressed to advance the field (summarized in Table 5).

Table 5: Synthesis of reported strengths and limitations of MBC methods

Aspect	Reported Strengths	Reported Limitations
Dependence Structure	Successfully reduces biases in inter-variable and inter-site (spatial) correlations (Su et al., 2020).	Methods like R2D2 assume a stationary dependence structure, which may not hold in the future (Vrac, 2018).
Compound Events	Broadly improves the representation of the frequency and magnitude of compound events.	Performance can vary by event type and is limited by the observational record of such events.
Temporal Structure	Some methods (e.g., MRNBC) are explicitly designed to correct auto- and cross-correlations across timescales (Mehrotra & Sharma, 2015, 2016).	A major, consistent weakness: most methods degrade or fail to correct the temporal autocorrelation of the time series (François et al., 2020).
Bias Stationarity	Performs well when biases are stationary.	Performance of all methods, especially multivariate ones, degrades significantly under non-stationary biases (Van de Velde et al., 2022).
Complexity & Stability	Conceptually more robust by representing physical linkages.	Can be computationally expensive (MBCn) or suffer from instability and poor performance in high-dimensional settings (dOTC, MRec) (Ahn et al., 2023; François et al., 2020).

425

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

5.2.1 The Unresolved Problem of Bias Non-Stationarity

Perhaps the most critical weakness of all current bias correction methods is their foundational assumption that the statistical biases in climate models are stationary—that is, the nature of the bias remains constant over time. There is growing evidence that this assumption is frequently violated, as the relationships learned during a historical training period may not remain valid under a changing climate (Dixon et al., 2016). A crucial study by Van de Velde et al. (2022) directly tested this assumption and demonstrated that the performance of all tested methods, and especially the multivariate ones, 'degrades significantly' when biases are non-stationary. This 'stationarity pitfall' is especially dangerous because an MBC method may appear to perform perfectly during a historical cross-validation (training) period, yet lead to 'highly erroneous' results when applied to future climate states that differ from the past (Lanzante et al., 2018). This provides a compelling explanation for why the advantages of MBC seen in calibration are often 'weakened' or 'disappear' in independent validation periods (Chen et al., 2018; Guo et al., 2020). Indeed, Guo et al. (2020) directly attribute this weakened performance to the fact that 'the bias of intervariable correlation... is not stationary, and the observed intervariable correlation itself is also not invariable'. The non-stationarity of biases represents an 'Achilles' heel' for the entire field of statistical post-processing and remains a primary source of uncertainty in climate change projections.

5.2.2 The Neglected Temporal Dimension

Another consistent limitation reported across the literature is the failure of most common MBC methods to correctly adjust the temporal structure of the climate time series. A conceptual review by François et al. (2020) found that many methods, while correcting inter-variable dependence, can degrade the day-to-day autocorrelation of the variables, often generating bias-corrected data with weaker temporal dependencies than are present in observations. This is a significant flaw, as the persistence of weather conditions is critical for the development of multi-day extreme events like heatwaves, droughts, and prolonged rainfall leading to floods. While some advanced 'all-in-one' methods, such as the MRNBC/MRQNBC family developed by Mehrotra & Sharma (2015, 2016), are explicitly designed to correct auto- and cross-correlations across multiple timescales, this remains a general weakness of the more widely used 'marginal/dependence' approaches.

5.2.3 Expanding to Spatiotemporal Coherence

The vast majority of MBC methods, including most of those evaluated in this review, are applied at single sites or on a grid-cell-by-grid-cell basis. This approach ignores the spatial structure of climate variables, such as the fact that precipitation at one location is highly correlated with precipitation at nearby locations. Correcting each grid cell independently can degrade this spatial coherence, leading to unrealistic "patchy" climate fields. The development of methods that can simultaneously correct inter-variable, temporal, and spatial dependencies represent the next major frontier in the field. A key advancement in this direction is the Multi-site MBC (MSMBC) method proposed by Su et al. (2020). Their work demonstrated that a method designed to handle both inter-variable and inter-site dependencies outperformed both single-site multivariate and univariate

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

methods, leading to improved hydrological simulations. This highlights the growing need for scalable methods that can address the full spatiotemporal, multivariate nature of the bias correction problem.

5.3 Implications for Practitioners and a Roadmap for Future Research

The findings of this systematic review have clear implications for practitioners in the climate impact modeling community. The choice of a bias correction method is not a simple matter of selecting the most complex or statistically comprehensive technique. Instead, the decision must be guided by a careful consideration of the specific sensitivities of the impact model being used. For agricultural modelers, the evidence strongly suggests that multivariate methods should be the default choice, as their models are highly sensitive to the daily coherence of climate inputs. For hydrologists, the choice is more nuanced. If the study focuses on snow and ice processes, compound extremes, or soil moisture dynamics, MBC is the more appropriate and effective tool. However, for assessments of general water balance or bulk streamflow in rain-dominated catchments, simpler univariate methods may be sufficient and, as some evidence suggests, may even perform better. This is because univariate methods apply the full statistical power of the dataset to optimizing a single, critical variable (e.g., precipitation), while multivariate methods must navigate a trade-off. In correcting for inter-variable dependencies, many MBC methods can degrade other statistical properties, such as the temporal structure, which univariate methods leave unaltered. For a bulk streamflow model that is insensitive to daily inter-variable coherence, this degradation of temporal properties can be more detrimental than the original coherence bias, leading to a superior result from the simpler univariate method.

The limitations identified here point toward a clear roadmap for future research in bias correction methodology. Foremost among these is the grand challenge of tackling non-stationarity by developing methods that can account for or adapt to changes in climate model biases and dependence structures in a future climate. This is closely related to the need for integrating the temporal dimension, as future method development should prioritize the simultaneous correction of multivariate dependencies and the temporal structure of time series. Furthermore, advancing the field will require scaling to spatiotemporal coherence. However, this challenge extends beyond computational scalability; it requires a process-based perspective. Because spatial dependence structures vary drastically by weather regime—for example, the localized footprint of convective precipitation contrasts sharply with the broad extent of an extratropical cyclone—future methods should prioritize regime-dependent frameworks that segregate the bias adjustment based on the dominant physical process producing the field. This approach would ensure that the imposed spatial coherence is physically consistent with the governing meteorology. Consequently, there remains a critical need for computationally efficient methods that can implement such regime-aware corrections while handling inter-variable, inter-site, and temporal dependencies simultaneously, especially for high-resolution, large-ensemble climate model datasets. Finally, promoting standardized validation through community-based frameworks and benchmark experiments would allow for more objective, transparent, and rigorous inter-comparison of methods, helping to bridge the "validation gap" between statistical fidelity and impact relevance.

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

5.4 Limitations of this Systematic Review

While this review was conducted following a rigorous and transparent protocol, several limitations should be acknowledged. First, the literature search was restricted to peer-reviewed articles published in English, which may have excluded relevant studies published in other languages or presented in conference proceedings. Second, while the search keywords were designed to be comprehensive, it is possible that some relevant studies using different terminology for multivariate bias correction were not captured. Finally, the findings are susceptible to a potential publication bias within the scientific literature itself, where studies demonstrating a clear positive or novel result may be more likely to be published than those with null or negative findings (i.e., where MBC showed no significant benefit over univariate methods).

6. Conclusion

This systematic review of 39 peer-reviewed articles confirms that multivariate bias correction methods are a necessary and critical evolution from univariate approaches, directly addressing the fundamental flaw of destroying physically meaningful relationships between climate variables. The synthesis of evidence reveals that the effectiveness of these advanced methods is context-dependent, a conclusion that challenges a simplistic narrative that ‘more complex is always better’.

The primary finding is a clear divergence between application domains. For agricultural applications, where process-based models are sensitive to the daily coherence of climate inputs, MBC provides a clear and systematic improvement over univariate methods and should be considered the standard practice. In hydrology, however, the benefits are more ambiguous. MBC proves essential for process-specific applications that depend on the joint behavior of variables, such as snow and glacier modeling, but it does not consistently demonstrate superiority over simpler methods for the simulation of general streamflow characteristics. This arises because many catchment models act as spatiotemporal integrators: they are relatively insensitive to inter-variable coherence yet highly sensitive to temporal structures, which current MBC methods often fail to preserve.

The effectiveness of MBC is further constrained by two critical limitations: (1) the degradation of performance under non-stationary biases, where the statistical relationships corrected during calibration may not hold in a changing climate, and (2) the failure to maintain realistic temporal and spatial dependencies in corrected datasets. Together, these issues underscore the need for new approaches that can jointly handle multivariate, spatial, and temporal dependencies while being robust to evolving bias characteristics.

It is therefore imperative that future research pivots from static statistical refinement to the urgent development of adaptive, scalable, and process-aware MBC frameworks. These next-generation tools must be capable of operating effectively under non-stationary conditions and across multiple scales of variability. Strengthening the dialogue between statistical developers and impact modelers is no longer optional but essential; it is the only path to ensure that the next generation of bias correction tools is not only methodologically sophisticated but also genuinely fit for purpose in delivering credible, physically consistent, and actionable projections of climate impacts.

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

7. Data availability

No original data were generated in this study. The systematic review is based on peer-reviewed articles listed in the references and summarized in Tables 1 and 2.

8. Author contributions

B.P.S. had the initial idea of the study, conducted the systematic review, prepared all figures, and wrote the first complete draft of the article. R.P.M. revised the draft. Subsequently, W.J.G. reviewed the manuscript.

9. Competing interests

The authors declare that they have no conflict of interest.

10. Acknowledgements

This work was supported by the U.S. Department of Agriculture (USDA) Natural Resources Conservation Service (NRCS) under Cooperative Agreement number NR243A750008C001. The analyses presented in this publication should not be construed to represent any official USDA or U.S. Government determination or policy. Generative AI, specifically Gemini v3.0 Pro, was used in the following capacities: identifying and removing personal bias, improving structure and grammar in post-draft edits, and enhanced search of included/excluded research articles. No AI was used for the following: design of the systematic review, all database systematic search processes, all plots/figures generation, and drafting the review manuscript.

References

Ahmadalipour, A., Rana, A., Moradkhani, H., and Sharma, A.: Multi-criteria evaluation of CMIP5 GCMs for climate change impact analysis, Theor. Appl. Climatol., 128, 241–258, https://doi.org/10.1007/s00704-015-1695-4, 2017.

Ahn, K.-H., de Padua, V. M. N., Kim, J., and Yi, J.: Impact of diverse configuration in multivariate bias correction methods on large-scale hydrological modelling under climate change, J. Hydrol., 625, 130406, https://doi.org/10.1016/j.jhydrol.2023.130406, 2023.

Alder, J. R. and Hostetler, S. W.: The dependence of hydroclimate projections in snow-dominated regions of the Western United States on the choice of statistically downscaled climate data, Water Resour. Res., 55, 7957–7976, https://doi.org/10.1029/2018WR023458, 2019.

Asong, Z. E., Elshamy, M. E., Princz, D., Wheater, H. S., Pomeroy, J. W., Pietroniro, A., and Cannon, A. J.: High-resolution meteorological forcing data for hydrological modelling and climate change impact analysis in the Mackenzie River Basin, Earth Syst. Sci. Data, 12, 629–645, https://doi.org/10.5194/essd-12-629-2020, 2020.

Bachelet, D., Sheehan, T., Ferschweiler, K., and Abatzoglou, J.: Simulating vegetation change, carbon cycling, and fire over the Western United States using CMIP5 climate projections, in: Natural Hazard Uncertainty Assessment: Modeling and Decision Support, Geophys. Monogr. Ser., vol. 223, edited by: Riley, K., Webley, P., and Thompson, M., AGU, 277–306, https://doi.org/10.1002/9781119028116.ch17, 2017.

Bevacqua, E., Maraun, D., Haff, I., Widmann, M., and Vrac, M.: Multivariate statistical modelling of compound events via pair-copula constructions: analysis of floods in Ravenna (Italy), Hydrol. Earth Syst. Sci., 21, 2701–2723, https://doi.org/10.5194/hess-21-2701-2017, 2017.

Bhowmik, R. D. and Sankarasubramanian, A.: Limitations of univariate linear bias correction in yielding cross-correlation between monthly precipitation and temperature, Int. J. Climatol., 39, 4111–4124, https://doi.org/10.1002/joc.6086, 2019.

Boberg, F. and Christensen, J. H.: Overestimation of Mediterranean summer temperature projections due to model deficiencies, Nat. Clim. Change, 2, 433–436, https://doi.org/10.1038/nclimate1454, 2012.

Cannon, A. J.: Multivariate bias correction of climate model output: matching marginal distributions and intervariable dependence structure, J. Climate, 29, 7045–7064, https://doi.org/10.1175/JCLI-D-15-0679.1, 2016.

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.

Cannon, A. J., Alford, H., Shrestha, R. R., Kirchmeier-Young, M. C., and Najafi, M. R.: Canadian Large Ensembles Adjusted Dataset version 1 (CanLEADv1): Multivariate bias-corrected climate model outputs for terrestrial modelling and attribution studies in North America, Geosci. Data J., 8, 139–150, https://doi.org/10.1002/gdj3.142, 2021.

Chen, J., Li, C., Brissette, F. P., Chen, H., Wang, M., and Essou, G. R. C.: Impacts of correcting the inter-variable correlation of climate model outputs on hydrological modeling, J. Hydrol., 560, 446–461, https://doi.org/10.1016/j.jhydrol.2018.03.040, 2018.

Christensen, J. H., Boberg, F., Christensen, O. B., and Déqué, M.: On the need for bias correction of regional climate change projections of temperature and precipitation, Geophys. Res. Lett., 35, L20709, https://doi.org/10.1029/2008GL035694, 2008.

Dai, A.: Precipitation characteristics in eighteen coupled climate models, J. Climate, 19, 4605–4630, https://doi.org/10.1175/JCLI3884.1, 2006.

Das, D., Ganguly, A. R., and Obradovic, Z.: A Bayesian Sparse Generalized Linear Model With an Application to Multiscale Covariate Discovery for Observed Rainfall Extremes Over the United States, IEEE Trans. Geosci. Remote Sens., 53, 2546–2555, https://doi.org/10.1109/TGRS.2015.2445911, 2015.

Das Bhowmik, R., Sankarasubramanian, A., Sinha, T., Patskoski, J., Mahinthakumar, G., and Kunkel, K.: Multivariate Downscaling Approach Preserving Cross Correlations across Climate Variables for Projecting Hydrologic Fluxes, J. Hydrometeorol., 18, 2247–2265, https://doi.org/10.1175/JHM-D-16-0160.1, 2017.

Dembélé, M., Vrac, M., Ceperley, N., Zwart, S. J., Larsen, J., Dadson, S. J., Mariéthoz, G., and Schaefli, B.: Contrasting changes in hydrological processes of the Volta River basin under global warming, Hydrol. Earth Syst. Sci., 26, 1481–1507, https://doi.org/10.5194/hess-26-1481-2022, 2022.

Demirel, M. C. and Moradkhani, H.: Assessing the impact of CMIP5 climate multi-modeling on estimating the precipitation seasonality and timing, Climatic Change, 135, 433–447, https://doi.org/10.1007/s10584-015-1559-z, 2016.

Doblas-Reyes, F. J., Sörensson, A. A., Almazroui, M., Dosio, A., Gutowski, W. J., Haarsma, R., et al.: Linking global to regional climate change, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Masson-Delmotte, V., et al., Cambridge University Press, 1363–1512, https://doi.org/10.1017/9781009157896.012, 2021.

Edwards, P. N.: History of climate modeling, Wiley Interdiscip. Rev. Clim. Change, 2, 128–139, https://doi.org/10.1002/wcc.95, 2011.

Eghdamirad, S., Johnson, F., Sharma, A., and Kim, J.: The influence of dependence in characterizing multi-variable uncertainty for climate change impact assessments, Hydrolog. Sci. J., 64, 696–709, https://doi.org/10.1080/02626667.2019.1602777, 2019.

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.

Faghih, M., Brissette, F., and Sabeti, P.: Impact of correcting sub-daily climate model biases for hydrological studies, Hydrol. Earth Syst. Sci., 26, 1545–1564, https://doi.org/10.5194/hess-26-1545-2022, 2022.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., et al.: Evaluation of climate models, in: 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., et al., Cambridge University Press, 741–866, https://doi.org/10.1017/CBO9781107415324.020, 2013.

François, B., Vrac, M., Cannon, A. J., Robin, Y., and Allard, D.: Multivariate bias corrections of climate simulations: which benefits for which losses?, Earth Syst. Dynam., 11, 537–562, https://dx.doi.org/10.5194/esd-11-537-2020, 2020.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., et al.: The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes, Sci. Data, 2, 1–21, https://doi.org/10.1038/sdata.2015.66, 2015.

Funk, H., Ludwig, R., Küchenhoff, H., and Nagler, T.: Towards more realistic climate model outputs: a multivariate bias correction based on zero-inflated vine copulas, J. R. Stat. Soc. C., https://doi.org/10.1093/jrsssc/qlaf044, 2025.

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

Galmarini, S., Solazzo, E., Ferrise, R., et al.: Assessing the impact on crop modelling of multi- and uni-variate climate model bias adjustments, Agric. Syst., 213, 103846, https://doi.org/10.1016/j.agsy.2023.103846, 2024.

García-Valdecasas Ojeda, M., Di Sante, F., Coppola, E., Fantini, A., Nogherotto, R., Raffaele, F., and Giorgi, F.: Climate change impact on flood hazard over Italy, J. Hydrol., 614, 128628, https://doi.org/10.1016/j.jhydrol.2022.128628, 2022.

Gu, L., Chen, J., Yin, J., Xu, C.-Y., and Zhou, J.: Responses of Precipitation and Runoff to Climate Warming and Implications for Future Drought Changes in China, Earths Future, 8, e2020EF001718, https://doi.org/10.1029/2020EF001718, 2020.

Guo, Q., Chen, J., Zhang, X. J., Xu, C.-Y., and Chen, H.: Impacts of Using State-of-the-Art Multivariate Bias Correction Methods on Hydrological Modeling Over North America, Water Resour. Res., 56, e2019WR026659, https://doi.org/10.1029/2019WR026659, 2020.

Gupta, A. S., Muir, L. C., Brown, J. N., Phipps, S. J., Durack, P. J., Monselesan, D., and Wijffels, S. E.: Climate drift in the CMIP3 models, J. Climate, 25, 4621–4640, https://doi.org/10.1175/JCLI-D-11-00312.1, 2012.

Gupta, A. S., Jourdain, N. C., Brown, J. N., and Monselesan, D.: Climate drift in the CMIP5 models, J. Climate, 26, 8597–8615, https://doi.org/10.1175/JCLI-D-12-00521.1, 2013.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, J. Hydrol., 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., Chang, P., Corti, S., Fučkar, N. S., Guemas, V., von Hardenberg, J., Hazeleger, W., Kodama, C., Koenigk, T., Leung, L. R., Lu, J., Luo, J.-J., Mao, J., Mizielinski, M. S., Mizuta, R., Nobre, P., Satoh, M., Scoccimarro, E., Semmler, T., Small, J., and von Storch, J.-S.: High Resolution Model Intercomparison Project (HighResMIP v1.0) for CMIP6, Geosci. Model Dev., 9, 4185–4208, https://doi.org/10.5194/gmd-9-4185-2016, 2016.

Hakala, K., Addor, N., and Seibert, J.: Hydrological Modeling to Evaluate Climate Model Simulations and Their Bias Correction, J. Hydrometeorol., 19, 1891–1907, https://doi.org/10.1175/JHM-D-17-0189.1, 2018.

Hanggoro, W., Yuanshu, J., Cudemus, L., and Zhihao, J.: Impact Analysis of Univariate and Multivariate Bias Correction on Rice Irrigation Water Needs in Jiangxi Province, China, Water, 12, 349, https://doi.org/10.3390/w12020381, 2020.

Hnilica, J., Hanel, M., and Puš, V.: Multisite bias correction of precipitation data from regional climate models, Int. J. Climatol., 37, 2404–2416, https://doi.org/10.1002/joc.4890, 2017.

Hobbs, W., Palmer, M. D., and Monselesan, D.: An energy conservation analysis of ocean drift in the CMIP5 global coupled models, J. Climate, 29, 1639–1653, https://doi.org/10.1175/JCLI-D-15-0477.1, 2016.

Horton, R. E.: An approach toward a physical interpretation of infiltration capacity, Soil Sci. Soc. Am. Proc., 5, 399–417, https://doi.org/10.2136/sssaj1941.036159950005000C0075x, 1940.

Huang, Z., Zhao, T., Zhang, Y., Cai, H., Hou, A., and Chen, X.: A five-parameter Gamma-Gaussian model to calibrate monthly and seasonal GCM precipitation forecasts, J. Hydrol., 601, 126893, https://doi.org/10.1016/j.jhydrol.2021.126893, 2021.

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

IPCC: 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., et al., Cambridge University Press, 1535 pp., https://doi.org/10.1017/CBO9781107415324, 2013.

Janssen, J., Radić, V., and Ameli, A.: Assessment of Future Risks of Seasonal Municipal Water Shortages Across North America, Front. Earth Sci., 9, https://doi.org/10.3389/feart.2021.730631, 2021.

Ji, S. and Ahn, K.-H.: Temperature change-informed future multisite streamflow generation to support water supply vulnerability assessments under climate change, J. Hydrol., 622, 129928, https://doi.org/10.1016/j.jhydrol.2023.129928, 2023.

Jin, H., Li, M., Hopwood, Z., Hochman, Z., and Bakar, K. S.: Improving early-season wheat yield forecasts driven by probabilistic seasonal climate forecasts, Agric. For. Meteorol., 317, 108832, https://doi.org/10.1016/j.agrformet.2022.108832, 2022.

Jones, J. W., Hoogenboom, G., Porter, C. H., Boote, K. J., Batchelor, W. D., Hunt, L. A., et al.: The DSSAT cropping system model, Eur. J. Agron., 18, 235–265, https://doi.org/10.1016/S1161-0301(02)00107-7, 2003.

Khanal, S., Ridder, N., de Vries, H., Terink, W., and van den Hurk, B.: Storm Surge and Extreme River Discharge: A Compound Event Analysis Using Ensemble Impact Modeling, Front. Earth Sci., 7, https://doi.org/10.3389/feart.2019.00224, 2019.

Khatun, A., Sahoo, B., and Chatterjee, C.: Assessment of enhanced Kohonen self-organizing map quantile mapping and copula-based bias-correction approaches for constructing basin-scale rainfall forecasts, Hydrolog. Sci. J., 67, 1831–1851, https://doi.org/10.1080/02626667.2022.2109972, 2022.

Kim, Y.-T., Kim, M. J., and Kwon, H.-H.: Development of daily spatio-temporal downscaling model with conditional Copula based bias-correction of GloSea5 monthly ensemble forecasts, J. Korea Water Resour. Assoc., 54, 1317–1329, https://doi.org/10.3741/JKWRA.2021.54.12.1317, 2021.

Lazoglou, G., Anagnostopoulou, C., Skoulikaris, C., and Tolika, K.: Bias Correction of Climate Models Precipitation Using the Copula Method and Its Application in River Basin Simulation, Water, 11, 616, https://doi.org/10.3390/w11030600, 2019.

Lee, T. and Ouarda, T. B. M. J.: Conditional stochastic simulation model for spatial downscaling for assessing the effects of climate change on hydro-meteorological variables, Climatic Change, 149, 439–453, https://doi.org/10.1007/s10584-018-2276-1, 2018.

Liu, M., et al.: Encounter Probability and Risk of Flood and Drought under Future Climate Change in the Two Tributaries of the Rao River Basin, China, Water, 12, 86, https://doi.org/10.3390/w12010104, 2019.

Lopez-Gomez, I., et al.: Dynamical-generative downscaling of climate model ensembles, Proc. Natl. Acad. Sci. U.S.A., https://doi.org/10.1073/pnas.2420288122, 2025.

Madani, A. Z., Hermassi, T., Taibi, S., Dakhlaoui, H., and Mechergui, M.: Climate change impacts on the Chiffa basin (northern Algeria) using bias-corrected RCM data, Front. Water, 6, https://doi.org/10.3389/frwa.2024.1507961, 2024.

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

Mao, G., Vogl, S., Laux, P., Wagner, S., and Kunstmann, H.: Stochastic bias correction of dynamically downscaled precipitation fields for Germany through Copula-based integration of gridded observation data, Hydrol. Earth Syst. Sci., 19, 1787–1804, https://doi.org/10.5194/hess-19-1787-2015, 2015.

Mehrotra, R. and Sharma, A.: Correcting for systematic biases in multiple raw GCM variables across a range of timescales, J. Hydrol., 520, 214–232, https://doi.org/10.1016/j.jhydrol.2014.11.037, 2015.

Mehrotra, R. and Sharma, A.: A Multivariate Quantile-Matching Bias Correction Approach with Auto- and Cross-Dependence across Multiple Time Scales: Implications for Downscaling, J. Climate, 29, 3519–3539, https://doi.org/10.1175/JCLI-D-15-0356.1, 2016.

Merchant, C. J., Embury, O., Bulgin, C. E., Block, T., Corlett, G. K., Fiedler, E., et al.: Satellite-based time-series of sea-surface temperature since 1981 for climate applications, Sci. Data, 6, 223, https://doi.org/10.1038/s41597-019-0236-x, 2019.

Meyer, J., Kohn, I., Stahl, K., Hakala, K., Seibert, J., and Cannon, A. J.: Effects of univariate and multivariate bias correction on hydrological impact projections in alpine catchments, Hydrol. Earth Syst. Sci., 23, 1339–1354, https://doi.org/10.5194/hess-23-1339-2019, 2019.

Miralha, L., Muenich, R. L., Scavia, D., Wells, K., Steiner, A. L., Kalcic, M., et al.: Bias correction of climate model outputs influences watershed model nutrient load predictions, Sci. Total Environ., 758, 143039, https://doi.org/10.1016/j.scitotenv.2020.143039, 2021.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., and Veith, T. L.: Model evaluation guidelines for systematic quantification of accuracy in watershed simulations, Trans. ASABE, 50, 885–900, https://doi.org/10.13031/2013.23153, 2007.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I—A discussion of principles, J. Hydrol., 10, 282–290, https://doi.org/10.1016/0022-1694(70)90255-6, 1970.

Nury, A. H., Sharma, A., Marshall, L., and Cordery, I.: Modelling climate change impacts on the Brahmaputra streamflow resulting from changes in snowpack attributes, J. Hydrol., 602, 126998, https://doi.org/10.1016/j.jhydrol.2021.126998, 2021.

Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., et al.: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges, Rev. Geophys., 53, 323–361, https://doi.org/10.1002/2014RG000475, 2015.

Qian, B., et al.: Climate change impacts on Canadian yields of spring wheat, canola and maize for global warming levels of 1.5 C, 2.0 C, 2.5 C, and 3.0 C, Environ. Res. Lett., 14, 075004, https://doi.org/10.1088/1748-9326/ab17fb, 2019.

Ramirez-Villegas, J., Challinor, A. J., Thornton, P. K., and Jarvis, A.: Implications of regional improvement in global climate models for agricultural impact research, Environ. Res. Lett., 8, 024018, https://doi.org/10.1088/1748-9326/8/2/024018, 2013.

Richards, L. A.: Capillary conduction of liquids through porous mediums, Physics, 1, 318–333, https://doi.org/10.1063/1.1745010, 1931.

Richardson, L. F.: Weather prediction by numerical process, Cambridge University Press, 236 pp., https://doi.org/10.1002/qj.49704820311, 1922.

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

Rummukainen, M.: State-of-the-art with regional climate models, Wiley Interdiscip. Rev. Clim. Change, 1, 82–96, https://doi.org/10.1002/wcc.8, 2010.

Räty, O., Räisänen, J., Bosshard, T., and Donnelly, C.: Intercomparison of Univariate and Joint Bias Correction Methods in Changing Climate From a Hydrological Perspective, Climate, 6, 33, https://doi.org/10.3390/cli6020033, 2018.

Schepen, A., Everingham, Y., and Wang, Q. J.: Coupling forecast calibration and data-driven downscaling for generating reliable, high-resolution, multivariate seasonal climate forecast ensembles at multiple sites, Int. J. Climatol., 40, 1587–1601, https://doi.org/10.1002/joc.6346, 2019.

Schepen, A., Everingham, Y., and Wang, Q. J.: On the Joint Calibration of Multivariate Seasonal Climate Forecasts from GCMs, Mon. Weather Rev., 148, 1105–1122, https://doi.org/10.1175/MWR-D-19-0046.1, 2020.

Schnur, R. and Lettenmaier, D. P.: A case study of statistical downscaling in Australia using weather classification by recursive partitioning, J. Hydrol., 212–213, 294–309, https://doi.org/10.1016/S0022-1694(98)00217-0, 1998.

Shrestha, R. R., Cannon, A. J., Schnorbus, M. A., and Alford, H.: Climatic Controls on Future Hydrologic Changes in a Subarctic River Basin in Canada, J. Hydrometeorol., 20, 2153–2172, https://doi.org/10.1175/JHM-D-19-0010.1, 2019.

Singh, D., Vardhan, M., Sahu, R., Chatterjee, D., Chauhan, P., and Liu, S.: Machine-learning- and deep-learning-based streamflow prediction in a hilly catchment for future scenarios using CMIP6 GCM data, Hydrol. Earth Syst. Sci., 27, 1047–1071, https://doi.org/10.5194/hess-27-1047-2023, 2023.

Sippel, S., et al.: A novel bias correction methodology for climate impact simulations, Earth Syst. Dynam., 7, 71–88, https://doi.org/10.5194/esd-7-71-2016, 2016.

Skamarock, W. C.: Evaluating mesoscale NWP models using kinetic energy spectra, Mon. Weather Rev., 132, 3019–3032, https://doi.org/10.1175/MWR2830.1, 2004.

Stephens, G. L., L'Ecuyer, T., Forbes, R., Gettelmen, A., Golaz, J. C., Bodas-Salcedo, A., et al.: Dreary state of precipitation in global models, J. Geophys. Res.-Atmos., 115, D24211, https://doi.org/10.1029/2010JD014532, 2010.

Su, T., Chen, J., Cannon, A. J., Xie, P., and Guo, Q.: Multi-site bias correction of climate model outputs for hydro-meteorological impact studies: An application over a watershed in China, Hydrol. Process., 34, 2486–2502, https://doi.org/10.1002/hyp.13750, 2020.

Thrasher, B., Maurer, E. P., McKellar, C., and Duffy, P. B.: Bias correcting climate model simulated daily temperature extremes with quantile mapping, Hydrol. Earth Syst. Sci., 16, 3309–3314, https://doi.org/10.5194/hess-16-3309-2012, 2012.

Tootoonchi, F., Haerter, J. O., Todorović, A., Räty, O., Grabs, T., and Teutschbein, C.: Uni- and multivariate bias adjustment methods in Nordic catchments: Complexity and performance in a changing climate, Sci. Total Environ., 851, 158615, https://doi.org/10.1016/j.scitotenv.2022.158615, 2022.

Tootoonchi, F., et al.: Uni- and multivariate bias adjustment of climate model simulations in Nordic catchments: Effects on hydrological signatures relevant for water resources management in a changing climate, J. Hydrol., 622, 129807, https://doi.org/10.1016/j.jhydrol.2023.129807, 2023.

https://doi.org/10.5194/egusphere-2026-529
Preprint. Discussion started: 16 March 2026

Van de Velde, J., Demuzere, M., De Baets, B., and Verhoest, N. E. C.: Impact of bias nonstationarity on the performance of uni- and multivariate bias-adjusting methods: a case study on data from Uccle, Belgium, Hydrol. Earth Syst. Sci., 26, 2319–2348, https://doi.org/10.5194/hess-26-2319-2022, 2022.

Vrac, M.: Multivariate bias adjustment of high-dimensional climate simulations: the Rank Resampling for Distributions and Dependences (R2D2) bias correction, Hydrol. Earth Syst. Sci., 22, 3175–3196, https://doi.org/10.5194/hess-22-3175-2018, 2018.

Wen, Y., Fei, Y., Fan, Y., Yang, A., Wang, B., Gao, P., and Scott, D.: Development of a probabilistic agricultural drought forecasting (PADF) framework under climate change, Agric. For. Meteorol., 348, 109965, https://doi.org/10.1016/j.agrformet.2024.109965, 2024.

Xi, X., Gentine, P., Zhuang, Q., and Kim, S.: Evaluating the variability of surface soil moisture simulated within CMIP5 using SMAP data, J. Geophys. Res.-Atmos., 127, e2021JD035363, https://doi.org/10.1029/2021JD035363, 2022.

Yang, S., Yang, D., Zhao, B., Ma, T., Lu, W., and Santisirisomboon, J.: Future Changes in High and Low Flows under the Impacts of Climate and Land Use Changes in the Jiulong River Basin of Southeast China, Atmosphere, 13, 150, https://doi.org/10.3390/atmos13020150, 2022.

Yin, Z., Feng, Q., Yang, L., Deo, R. C., Adamowski, J. F., Wen, X., et al.: Projected spatial patterns in precipitation and air temperature for China's northwest region derived from high-resolution regional climate models, Int. J. Climatol., 40, 2346–2363, https://doi.org/10.1002/joc.6435, 2019.

Zhang, Y. and Paustian, K.: Sensitivity of Predicted Agro-Ecosystem Variables to Errors in Weather Input Data, Trans. ASABE, 62, 209–222, https://doi.org/10.13031/trans.13044, 2019.