

Bias Correction and Downscaling

CEVE 543 - Fall 2025

Dr. James Doss-Gollin

2025-10-20

Outline (45 minutes)

- Motivation and GCM biases (minutes 0-4)
- Weather vs climate distinction + supervised learning (minutes 4-9)
- Delta method (minutes 9-21)
- QQ-mapping (minutes 21-33)
- Quantile delta mapping (QDM) (minutes 33-41)
- Uncertainty and critical considerations (minutes 41-44)
- Wrap-up and looking ahead (minutes 44-45)

1 Motivation: Why Bias Correction?

Minutes 0-4

- Climate models have systematic biases that directly affect impact assessments.
- Impacts depend on distributions (frequency, intensity, duration) — not just means or totals.

1.1 Two Types of GCM Biases

- Drizzle bias (pervasive):
 - Too many light rainfall events with intensities too low relative to stations
 - Root causes: spatial averaging within large grid cells ($\approx 100\text{--}300$ km) and sub-grid parameterizations
 - Impacts: distorts frequency/intensity; affects crop growth, soil moisture dynamics, infiltration/runoff, and floods even when monthly totals match
- Dynamical biases (circulation errors):
 - Examples: ENSO amplitude/frequency, monsoon onset/timing, midlatitude storm tracks
 - Impacts: shifts climatology and interannual variability; affects predictability and teleconnections

Discussion Question

Pick one impact domain (agriculture, floods, urban drainage, or energy). Which bias is more damaging in your domain—drizzle bias or dynamical bias—and why? What observable metrics would you compute to diagnose and quantify the bias?

Suggested format: 2 min pair + 2 min share.

1.2 Weather vs Climate: A Critical Distinction

Minutes 4-9

- Weather forecasting: initial value problem with paired forecast–observation data → supervised learning works well.
- Climate projection: boundary value problem with no paired future observations → compare distributions under like-forcing; different statistical tools.

Aspect	Weather forecasting	Climate projection
Mathematical framing	Initial value problem	Boundary value problem
Training data	Paired forecasts and observations	No pairs for the future; rely on historical distributions under same forcings
Target	Next state/field at high resolution	Changes in distributions, moments, quantiles, extremes
Horizon	Hours–weeks	Decades–centuries
Typical methods	Supervised ML, data assimilation, super-resolution	Bias correction, quantile mapping, QDM, emulators respecting physics
Validation	Skill vs paired observations	Hindcast distributional match under historical forcings
Failure modes	Error growth, chaos	Nonstationarity, regime shifts, process changes
Outputs	Deterministic/probabilistic fields	Corrected time series, deltas, distributional adjustments

i Note: Rapid progress in supervised downscaling

- Why progress is fast: large paired forecast–observation datasets enable powerful supervised models (e.g., diffusion models like CorrDiff).
- Why climate downscalers are harder: no paired future data, large distribution shift/nonstationarity, long-term physical constraints.
- Implication: higher risk; requires rigorous validation and uncertainty quantification.

When we do have paired data (predictor–predictand), supervised learning is natural:

- Super-resolution: coarse → fine fields (deterministic or probabilistic ensembles)
- Learned precipitation modules: use T, q, winds → estimate precip at fine scales; often patch-based, newer methods improve cross-patch consistency
- Simple bias correction for forecast systems: regress out mean/variance errors (e.g., +2°C warm bias)
- Key advantage: train directly on paired observations with standard ML losses

💡 Discussion Question

If a supervised diffusion model like CorrDiff excels at weather-scale downscaling, what specific obstacles arise when trying to train a similar model to downscale future climate projections? Which of those can you mitigate, and how?

Suggested format: 2 min pair + 2 min share.

1.3 Climate: The Hard Case

- No future observations → train/fit on historical distributions only (under same forcings)
- Core assumption: model and observed describe the same climate distribution historically
- Workflow: match GCM-hist to Obs-hist → derive transfer function → apply to GCM-fut
- Caveat: may fail under regime shifts/nonstationarity or novel process changes

2 Delta Method

Minutes 9-21

The delta method is simple but widely used [1].

2.0.a Delta method (at a glance)

- Inputs: Observed historical series; GCM historical and future series (or statistics) by month/season
- Use when: You trust the modeled change signal more than model absolute levels
- Outputs: Corrected future series/statistics anchored to observations, preserving modeled change
- Pitfalls: Stationary-bias assumption; no frequency/intensity correction; deterministic
- Quick steps: compute Δ (additive for T, multiplicative for P) by month → apply Δ to observed baseline

2.1 Assumption and When to Use

- GCM can have systematic absolute biases (e.g., radiation/clouds, ocean heat transport, topography)
- Assume the modeled change signal (trend/delta) is credible
- Example: model is +2°C warm bias but simulates +3°C future warming → apply +3°C to observed baseline

2.2 Mathematical Framework

Notation:

- $X_{\text{hist}}^{\text{obs}}$ = observed variable in historical period
- $X_{\text{hist}}^{\text{GCM}}$ = GCM-simulated variable in historical period
- $X_{\text{fut}}^{\text{GCM}}$ = GCM-simulated variable in future period
- $\bar{X}$ = temporal average (e.g., monthly mean)

Compute the change (delta):

$$\Delta = \bar{X}_{\text{fut}}^{\text{GCM}} - \bar{X}_{\text{hist}}^{\text{GCM}}$$

Apply to observations:

$$X_{\text{fut}}^{\text{corr}} = X_{\text{hist}}^{\text{obs}} + \Delta$$

Note on notation: in the delta method, we apply the modeled change signal (the delta) to an observed baseline series. That's why the baseline term is $X_{\text{hist}}^{\text{obs}}$ (not GCM). By contrast, QQ-mapping and QDM directly transform the GCM daily values to corrected values.

2.3 Variants

For temperature, we typically use an additive delta:

$$T^{\text{corr}}(d, m, y) = T^{\text{obs}}(d, m, y_{\text{hist}}) + [\bar{T}_{\text{fut}}^{\text{GCM}}(m) - \bar{T}_{\text{hist}}^{\text{GCM}}(m)]$$

Here d is the day of month, m is the calendar month, y is the year, and y_{hist} is a year from the historical period. We take a day from the historical record and add the monthly-mean temperature change.

For precipitation, we use a multiplicative delta:

$$P^{\text{corr}}(d, m, y) = P^{\text{obs}}(d, m, y_{\text{hist}}) \times \frac{\bar{P}_{\text{fut}}^{\text{GCM}}(m)}{\bar{P}_{\text{hist}}^{\text{GCM}}(m)}$$

2.3.a Applying delta to derived statistics (quantiles, IDF curves)

The same idea applies when the target is a derived statistic rather than the raw daily series. Two common cases:

- Quantiles (e.g., monthly temperature quantiles), additive form:

$$Q_X^{\text{corr}}(p, m) = Q_{X, \text{hist}}^{\text{obs}}(p, m) + [Q_{X, \text{fut}}^{\text{GCM}}(p, m) - Q_{X, \text{hist}}^{\text{GCM}}(p, m)]$$

- IDF curves (precipitation intensity–duration–frequency), multiplicative form:

$$\text{IDF}^{\text{corr}}(D, r) = \text{IDF}_{\text{hist}}^{\text{obs}}(D, r) \times \frac{\text{IDF}_{\text{fut}}^{\text{GCM}}(D, r)}{\text{IDF}_{\text{hist}}^{\text{GCM}}(D, r)}$$

Here p is a percentile, m a month, D a duration, and r a return period (or equivalently a high quantile). This “delta-on-statistics” approach preserves the modeled change in the statistic of interest while anchoring to observed historical levels.

Practical note: to reduce noise and month-boundary discontinuities, practitioners sometimes compute deltas on moving windows (e.g., 15–31 days) or seasonal aggregates; we use monthly deltas here for clarity and speed in class.

Discussion Question

Why use additive correction for temperature but multiplicative for precipitation? What would go wrong if you did it the other way?

Hint: Think about what happens if temperature drops by 5°C vs if precipitation drops by 50%.

2.4 Example Calculation

Suppose for July we have:

- Observed historical mean: $\bar{T}_{\text{hist}}^{\text{obs}} = 25^\circ\text{C}$
- GCM historical mean: $\bar{T}_{\text{hist}}^{\text{GCM}} = 27^\circ\text{C}$ (2°C warm bias)

- GCM future mean: $\bar{T}_{\text{fut}}^{\text{GCM}} = 30^{\circ}\text{C}$

The delta is $\Delta = 30 - 27 = 3^{\circ}\text{C}$.

The corrected future temperature is $T_{\text{fut}}^{\text{corr}} = 25 + 3 = 28^{\circ}\text{C}$.

We preserve the GCM's 3°C warming signal while removing the 2°C warm bias.

2.5 Advantages and Limitations

- Advantages:
 - Simple, transparent; preserves modeled change signal
 - Works for many statistics (means, quantiles, variances)
 - Seasonal specificity: apply deltas by month/season
- Limitations:
 - Deterministic; no explicit uncertainty
 - Assumes stationary bias structure across climate states
 - Doesn't correct frequency/intensity structure directly

Discussion Question

The delta method assumes the bias is additive (for temperature) or multiplicative (for precipitation) and constant across climate states.

What physical processes might violate these assumptions? For example:

1. What if precipitation shifts from frontal to convective?
2. What if cloud feedbacks change with warming?
3. What if the seasonal cycle shifts?

Think about: process-level changes, nonlinear responses, threshold effects.

3 Quantile-Quantile (QQ) Mapping

Minutes 21-33

QQ-mapping is more sophisticated and directly addresses the drizzle bias [2].

3.0.a QQ-mapping (at a glance)

- Inputs: Daily/ sub-daily GCM and observed series (historical), by month/season
- Use when: Large-scale dynamics credible; need to fix wet-day frequency and intensity distribution
- Outputs: Bias-corrected series matching observed distribution (hist), applied to future with assumed stationarity
- Pitfalls: Can distort modeled change signal when applied to future; deterministic; tail extrapolation
- Quick steps: adjust wet-day frequency via threshold $\rightarrow$ map intensities via CDF-to-CDF transform (by month)

3.1 Assumption and When to Use

- Assumes large-scale dynamics are credible but frequency/intensity at sub-grid scales are biased
- Historical GCM–Obs distributional relationship is assumed to hold in the future
- Reasonable when GCM gets large-scale patterns right; risky if key processes are missing

3.2 Motivation

Precipitation is the product of frequency and intensity:

$$\bar{X}_m = \pi_m \times \mu_{I,m}$$

where π_m is the wet-day frequency in calendar month m (dimensionless) and $\mu_{I,m}$ is the mean wet-day intensity in month m (mm/day).

Correcting only the monthly mean $\bar{X}_m$ doesn't fix the frequency/intensity distortion.

💡 Discussion Question

A GCM simulates 20 days with 5 mm/day rain (monthly total = 100 mm). Observations show 10 days with 10 mm/day rain (also 100 mm total).

If you only correct the monthly mean, what problems might arise for:

1. A soil infiltration model?
2. A crop water stress model?

Think about: infiltration capacity, runoff generation, dry spell length, water stress timing.

3.3 Visual Illustration

The two figures below illustrate how QQ-mapping works:

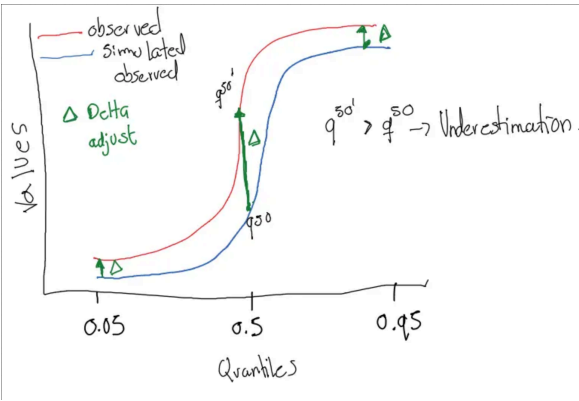


Figure 1: Schematic of quantile mapping: mapping GCM CDF (model) to observed CDF. Credit: Gutierrez (2021)

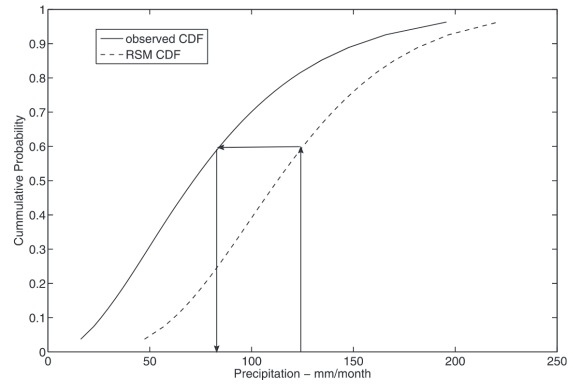


Figure 2: Example application to monthly precipitation: observed CDF (solid) vs Regional Spectral Model (RSM) CDF (dashed). Arrows show the mapping. P. J. Block, F. A. S. Filho, L. Sun, and H.-H. Kwon [3]

3.4 Two-Step Process

Notation:

- $P^{\text{GCM}}(d, m, y)$ = daily GCM precipitation on day d of month m in year y
- $P^{\text{obs}}(d, m, y)$ = daily observed precipitation
- $F(\cdot)$ = cumulative distribution function (CDF)
- $F^{-1}(\cdot)$ = inverse CDF (quantile function)

- $\tilde{x}$ = threshold for defining “wet day” (typically 0.1 mm)

Step 1: Frequency correction

Find threshold $\tilde{x}^{\text{GCM}}$ such that GCM wet-day frequency matches observed:

$$F^{\text{GCM}}(\tilde{x}^{\text{GCM}}) = F^{\text{obs}}(\tilde{x}^{\text{obs}})$$

where $\tilde{x}^{\text{obs}} = 0.1$ mm is the observed wet-day threshold.

Invert to get:

$$\tilde{x}^{\text{GCM}} = (F^{\text{GCM}})^{-1}(F^{\text{obs}}(\tilde{x}^{\text{obs}}))$$

Discard all GCM values below $\tilde{x}^{\text{GCM}}$. This removes excess drizzle events.

Step 2: Intensity correction

For each wet day i with $x_i \geq \tilde{x}^{\text{GCM}}$, map from GCM intensity distribution to observed:

$$P_i^{\text{corr}} = (F_I^{\text{obs}})^{-1}(F_I^{\text{GCM}}(P_i^{\text{GCM}}))$$

where F_I^{GCM} is the CDF of GCM intensities above threshold $\tilde{x}^{\text{GCM}}$ and F_I^{obs} is the CDF of observed intensities above threshold $\tilde{x}^{\text{obs}}$.

The intuition is to find where P_i^{GCM} falls in the GCM distribution, then sample the same percentile from the observed distribution.

3.5 Distribution Choices

For observed intensities (F_I^{obs}), we typically use a gamma distribution:

$$f(x; \alpha, \beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right), \quad x > 0$$

The shape parameter α is dimensionless and the scale parameter β has the same units as x . Both are estimated by maximum likelihood estimation (MLE).

For GCM intensities (F_I^{GCM}), two approaches are common. Gamma-Gamma (GG) fits a gamma distribution to both GCM and observed data. Empirical-Gamma (EG) uses the empirical CDF for GCM data but gamma for observed. A. V. M. Ines and J. W. Hansen [2] found EG slightly better for their crop modeling application.

3.6 Why This Matters

- Impact models are sensitive to:
 - Wet/dry timing $\rightarrow$ dry spell lengths
 - Event intensities $\rightarrow$ infiltration/runoff generation
 - Temporal clustering $\rightarrow$ water stress and extreme flows
- Monthly-total-only correction can mislead impacts; frequency/intensity corrections matter (A. V. M. Ines and J. W. Hansen [2])

3.7 Limitations

- Deterministic; no explicit uncertainty

- Assumes quantile relationship is time-invariant; process changes can break it
- Plain QM can distort modeled climate-change signals when applied to future data → motivation for QDM [4]
- Terminology: “quantile mapping (QM)” $\approx$ “QQ-mapping” in common use
- QDM preserves quantile-specific change by construction [4]

4 Quantile Delta Mapping (QDM)

Minutes 33-41

QDM blends the strengths of delta change and quantile mapping [4]. It corrects historical bias while preserving the GCM’s projected change at each quantile. This avoids “erasing” the modeled climate-change signal during bias correction.

4.1 Assumptions and Idea

- Historical quantile relation (model $\leftrightarrow$ obs) is informative
- Modeled change signal by quantile is credible
- Procedure: quantile-map using historical relation → adjust by modeled change at the same percentile

4.2 Notation

- $F_{\text{hist}}^{\text{mod}}, F_{\text{hist}}^{\text{obs}}$: historical CDFs (model, observed)
- $Q_{\text{hist}}^{\text{mod}}(p), Q_{\text{fut}}^{\text{mod}}(p), Q_{\text{hist}}^{\text{obs}}(p)$: corresponding quantile functions
- For a modeled future value y_i (from the GCM future series), define its modeled future percentile $p_i = F_{\text{fut}}^{\text{mod}}(y_i)$

4.3 QDM formula

- Additive variables (e.g., temperature):

$$x_{\text{fut},i}^{\text{QDM}} = Q_{\text{hist}}^{\text{obs}}(p_i) + \left[Q_{\text{fut}}^{\text{mod}}(p_i) - Q_{\text{hist}}^{\text{mod}}(p_i) \right]$$

- Multiplicative variables (e.g., precipitation):

$$x_{\text{fut},i}^{\text{QDM}} = Q_{\text{hist}}^{\text{obs}}(p_i) \times \frac{Q_{\text{fut}}^{\text{mod}}(p_i)}{Q_{\text{hist}}^{\text{mod}}(p_i)}$$

Intuition: map the modeled future value to a percentile p_i , take the observed historical value at that percentile, and then apply the model’s projected change at that percentile (absolute for temperature, relative for precipitation) [4].

4.4 Pros and caveats

- QDM (at a glance)
 - Inputs: Historical GCM/Obs distributions and modeled future distribution
 - Use when: You need bias correction that preserves quantile-specific modeled change
 - Outputs: Corrected future series with preserved change at each percentile
 - Pitfalls: Assumes historical quantile relation persists; zero-inflation care; tail handling
 - Quick steps: find percentile of modeled future value → map to observed-hist quantile → apply modeled change at that percentile (additive/multiplicative)
- Preserves modeled change signal across the distribution; mitigates “delta erosion.”
- Handles non-uniform changes (e.g., heavier tails) better than plain delta.

- Still assumes historical quantile relationships carry forward; tail extrapolation remains tricky.
- Requires care for zero-inflation and frequency changes for precipitation.

4.5 Method comparison (cheat sheet)

Feature		Delta	QQ-mapping (QM)	QDM
Preserves modeled change signal		Yes (mean/target stat)	No (can distort)	Yes (by quantile)
Corrects wet-day frequency	fre-	No	Yes (with threshold)	Yes (if combined with frequency step)
Corrects intensity distribution	dis-	No (only target stat)	Yes	Yes
Assumes stationarity of bias		Yes	Yes (quantile relation)	Yes (quantile relation + change factors)
Handles non-uniform change across quantiles		No	Not guaranteed	Yes
Uncertainty representation	represen-	Deterministic	Deterministic	Deterministic
Data requirements		Means/quantiles by month	Full distribution by month	Full distribution by month (hist + fut percentiles)

💡 Discussion Question

When would you prefer QDM over plain QQ-mapping or the basic delta method? Give one example where preserving quantile-specific change matters for impacts, and one case where QDM's assumptions might be risky.

Suggested format: 1 min think + 2 min pair + 2 min share.

5 Uncertainty in Bias Correction

Minutes 41-43

Bias correction is not a panacea and introduces additional uncertainty.

5.1 Sources of Uncertainty

- Sampling: limited records → few extremes → uncertain tail parameters
- Structural: method choice (delta/qq/QDM), distributional form (gamma/empirical), temporal aggregation (daily/monthly), wet-day thresholds
- Extrapolation: does the transfer function hold in future climates with no historical analog? New circulation regimes, unprecedented T-q combinations, process changes

Discussion Question

Suppose you calibrate QQ-mapping using 1980-2010 data, when the wettest month had 250 mm. In 2080, the GCM simulates a month with 400 mm.

1. What will QQ-mapping do?
2. What assumptions are you making?
3. When might this fail?

Think about: extrapolation methods, physical realism, process changes.

5.2 Methods Can Disagree Strongly

- L. E. Hay, R. L. Wilby, and G. H. Leavesley [1]: delta vs statistical downscaling for three US basins → markedly different runoff projections
- Differences largest in complex terrain; “better” method depends on validation context

Bias Correction Can ADD Uncertainty

Applying bias correction doesn't necessarily improve impact predictions. It transforms one type of uncertainty (GCM bias) into another (correction uncertainty). Multiple methods should always be compared.

5.3 Quantifying Uncertainty

- Use multiple methods and multiple GCMs; generate ensembles; validate on independent data when available
- Validation is hardest where it matters most (future conditions)

6 Critical Considerations

Minutes 43-44

6.1 Extrapolation

- All methods must extrapolate beyond calibration range
 - QM: often uses constant (cap at max obs) or linear tail extrapolation → risk of unrealistic extremes or violations (e.g., negative P)
 - Delta: assumes constant absolute (T) or relative (P) bias → may break if processes change (e.g., frontal → convective)
- No-analog futures: new circulation regimes, unprecedented T-q combos, shifts in precipitation mechanisms → transfer functions may fail

6.2 Missing Processes

Bias Correction Isn't Magic

You cannot correct for processes the GCM doesn't simulate. Garbage in, garbage out still applies.

Consider tropical cyclones as an example. Many global GCMs don't resolve tropical cyclones, which require roughly 50 km horizontal resolution. Observations clearly show TCs in coastal regions like Houston and Miami, but the GCM shows smooth, large-scale precipitation instead.

What happens with QQ-mapping? - It will create intense events to match the observed distribution, but:
- Timing wrong (seasonally smeared vs clustered TC events) - Location wrong (tracks absent/shifted) - Dynamics wrong (no rotation/RI/structure)

Discussion Question

Your GCM simulates no tropical cyclones but observations have them. You apply QQ-mapping for a coastal location.

1. What happens to the bias-corrected rainfall?
2. Is the result useful for impact assessment?
3. What alternatives might you consider?

Think about: statistical properties versus physical realism, temporal clustering, spatial structure.

6.3 Credibility Check and Scale Considerations

The GCM must credibly simulate the target variable at some level. - Good uses: correct drizzle bias when large-scale patterns are realistic; adjust temperature when circulation is credible; tweak timing/magnitude when seasonal cycle is captured - Poor uses: "fixing" missing processes; correcting where circulation is unrealistic; extrapolating far beyond historical range

Discussion Question

You're studying flood risk in a monsoon region. The GCM simulates monsoon onset 2 weeks late and underestimates peak rainfall by 30%.

1. Can bias correction fix these problems?
2. Which aspects can be corrected and which cannot?
3. What diagnostic checks would you perform before trusting the bias-corrected output?

Think about: timing versus magnitude errors, process representation, validation strategies.

Sometimes a hierarchical approach is better. We can bias-correct the GCM at coarse scale for large-scale precipitation, then run a specialized process model like a TC model or convective model, and finally combine outputs for local impacts. This preserves GCM information where it is credible while using process models for sub-grid phenomena.

7 Looking Ahead

Minutes 44-45

7.1 Beyond Basic Methods

The delta method and QQ-mapping are starting points. More sophisticated methods exist and will be covered later in the course.

- Stochastic weather generators: multiple realizations conditioned on GCM; represent sub-monthly variability and uncertainty (Module 2)
- Weather typing, HMMs: classify states and link to local distributions; naturally probabilistic (next two weeks)
- Multivariate bias correction: adjust multiple variables jointly while preserving correlations (active research)
- Spatial consistency: univariate QM can break spatial coherence → rank-based approaches, optimal transport, spatial QDM

7.2 Preparing for Wednesday

On Wednesday we will discuss A. V. M. Ines and J. W. Hansen [2] in detail.

Key questions to consider while reading:

1. How well does bias correction improve crop yield simulations?
2. What limitations remain even after bias correction?
3. Why are yields under-predicted even with “perfect” rainfall distributions?

The paper provides a concrete case study of the methods we discussed today and highlights remaining challenges.

8 References

Bibliography

- [1] L. E. Hay, R. L. Wilby, and G. H. Leavesley, “A Comparison of Delta Change and Downscaled Gcm Scenarios for Three Mountainous Basins in the United States,” *JAWRA Journal of the American Water Resources Association*, vol. 36, no. 2, pp. 387–397, 2000, doi: 10.1111/j.1752-1688.2000.tb04276.x.
- [2] A. V. M. Ines and J. W. Hansen, “Bias Correction of Daily GCM Rainfall for Crop Simulation Studies,” *Agricultural and Forest Meteorology*, vol. 138, no. 1, pp. 44–53, Aug. 2006, doi: 10.1016/j.agrformet.2006.03.009.
- [3] P. J. Block, F. A. S. Filho, L. Sun, and H.-H. Kwon, “A Streamflow Forecasting Framework Using Multiple Climate and Hydrological Models,” *JAWRA Journal of the American Water Resources Association*, vol. 45, no. 4, pp. 828–843, 2009, doi: 10.1111/j.1752-1688.2009.00327.x.
- [4] A. J. Cannon, S. R. Sobie, and T. Q. Murdock, “Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes?,” *Journal of Climate*, vol. 28, no. 17, pp. 6938–6959, Sep. 2015, doi: 10.1175/JCLI-D-14-00754.1.