Supplementary Material: Bias Correction Experiment

Gemini Advanced

August 10, 2025

1 Complete Quantitative Summary

This table provides the full set of summary statistics for both the 1km bimodal and 9km normal scenarios. The metrics show the mean, standard deviation, and the percentage of the spatial domain where the "Pollutant X" value exceeds the 65 ppm threshold.

Table 1: Complete quantitative summary for all scenarios and methods.

Scenario	Method	Mean (ppm)	Std Dev (ppm)	% Area ¿ 65 ppm
1km Bimodal	Original	56.04	5.17	2.69%
	Stations (Observed)	60.42	5.95	16.44%
	Scaling	58.38	5.38	15.86%
	Delta Change	58.47	5.17	14.72%
	Variance Scaling	59.01	3.94	6.41%
	Parametric (Normal)	59.01	3.94	6.41%
	Parametric (Gamma)	58.99	3.97	6.57%
	Quantile Mapping	60.74	4.11	14.88%
	Spatial Delta	58.06	6.96	17.59%
9km Normal	Original	60.14	1.66	0.00%
	Stations (Observed)	60.42	5.95	16.44%
	Scaling	61.33	1.70	0.00%
	Delta Change	61.32	1.66	0.00%
	Variance Scaling	62.70	5.57	35.00%
	Parametric (Normal)	62.70	5.57	35.00%
	Parametric (Gamma)	55.64	28.21	46.00%
	Quantile Mapping	60.72	4.59	15.00%
	Spatial Delta	61.25	1.82	2.00%

2 Cross-Validation Results

This table provides the definitive ranking of the statistical methods based on their predictive skill. The Root Mean Square Error (RMSE) was calculated using a leave-one-out cross-validation approach on the 1km bimodal dataset. Lower RMSE indicates a better prediction.

3 Mathematical Formulations of Bias Correction Methods

Let M_{grid} be the original model data on the full grid, and M_{obs} be the model data extracted at the station locations. Let O be the full time series of station observations, and $\bar{O}$ and σ_O be the mean and standard deviation of the station observations, respectively.

Table 2: Leave-one-out cross-validation results for the 1km bimodal scenario.

Method	Average RMSE
Spatial Delta	**5.07**
Original	5.32
Variance Scaling	5.46
Parametric (Normal)	5.46
Delta Change	6.26
Scaling	6.51
Parametric (Gamma)	6.56
Quantile Mapping	8.30

3.1 Delta Change

- **Concept:** Assumes the bias is a simple additive offset, constant across the entire domain.
- **Equation:** $M'_{grid} = M_{grid} + (\bar{O} - \bar{M}_{obs})$

3.2 Scaling

- **Concept:** Assumes the bias is a multiplicative factor.
- **Equation:** $M'_{grid} = M_{grid} \times (\bar{O}/\bar{M}_{obs})$

3.3 Variance Scaling

- **Concept:** Corrects both the mean and the variance to match the observations.
- **Equation:** $M'_{grid} = (M_{grid} - \bar{M}_{obs})(\sigma_O/\sigma_{M_{obs}}) + \bar{O}$

3.4 Quantile Mapping (QM)

- **Concept:** A non-parametric method that forces the model's distribution to perfectly match the observed distribution.
- **Equation:** $M'_{grid} = \text{CDF}_O^{-1}(\text{CDF}_{M_{grid}}(M_{grid}))$

3.5 Parametric Mapping

- **Concept:** A variation of QM that assumes the data follows a specific statistical distribution.
- **Equation:** $M'_{grid} = \mathcal{F}_O^{-1}(\mathcal{F}_{M_{grid}}(M_{grid}))$ for a fitted distribution $\mathcal{F}$.

3.6 Spatial Delta

- **Concept:** Assumes the additive bias varies spatially.
- **Equation:** $M'_{grid}(x, y) = M_{grid}(x, y) + \text{Interpolate}(\bar{O}_i - M_{obs,i}, x, y)$