A Scalable Framework for Climate-Aware Satellite Precipitation Bias Correction

Yashraj Upase[1], Abhigyan Chakraborty[1] Shruti Upadhyaya[1,2],
Malarvizhi Arulraj[3], Kishalay Mitra[1]

[1]*Indian Institute of Technology, Hyderabad, India*
[2]*Advanced Radar Research Center, University of Oklahoma, USA*
[3] *Earth System Science Interdisciplinary Center, University of Maryland, College Park, Maryland, USA*

Satellite-based precipitation products like GPM-IMERG are increasingly used for hydrological and climate applications across diverse regions, yet their biases vary systematically with local climate characteristics. Traditional bias-correction approaches either operate at regional scales with limited transferability or apply uniform adjustments that ignore climate-dependent error forms. This mismatch becomes critical when scaling corrections globally, where a single model must handle everything from tropical monsoons to arid zones, each with distinct precipitation-retrieval challenges. In this work, we develop a climate-zone-aware bias-correction framework that explicitly accounts for regional climate context while maintaining scalability to larger domains. We train a convolutional neural network on all-India data to correct IMERG against gridded observations from the India Meteorological Department, but crucially, we augment the model with learned embeddings of Köppen-Geiger climate zones. These eight-dimensional embeddings allow the network to internalize climate-specific bias patterns without requiring separate region-wise models. The correction is optimised as a distributional problem, optimized through negative log-likelihood under a lognormal assumption to preserve the precipitation variability rather than just matching expected values. Validation shows that this unified model performs comparably to region-specific models trained exclusively on individual climate zones, confirming that the embeddings effectively capture zone-dependent error characteristics. Analysis of the learned embeddings reveals they encode physically meaningful distinctions- the Western Ghats, which exhibit the largest IMERG biases, are represented differently from other climate zones where biases are smaller and differently structured. By embedding climate knowledge directly into the correction architecture, the model achieves climate-aware adjustments without the overhead of maintaining separate regional models. By learning climate-specific patterns within a unified architecture, this approach enables scalable global deployment while preserving the distinct precipitation characteristics of different climate regimes.