

Motivation

Subseasonal (S2S) rainfall prediction suffers from rapid skill decay beyond 2-3 weeks, particularly for rainfall. Statistical post-processing techniques can improve skill but are limited in their ability to handle nonlinear, flow-dependent errors. A hybrid framework involving ML, which preserves physical relationships while correcting dynamical model errors.

Methodology

A hybrid DL-Dynamical Statistical Projection Model (DL-DSPM) improves sub-seasonal rainfall forecasts over India. Systematic biases in forecast large-scale circulation are corrected using ResNet and U-Net, trained on hindcast anomalies against ERA5 reanalysis. The bias-corrected circulation is then fed into a DSPM, which uses SVD to extract coupled circulation-rainfall relationships and reconstruct regional precipitation anomalies – combining deep-learning bias correction with statistically-derived circulation-rainfall relationships to improve rainfall forecast skill (e.g., Nie & Sun, ERL, 2024).

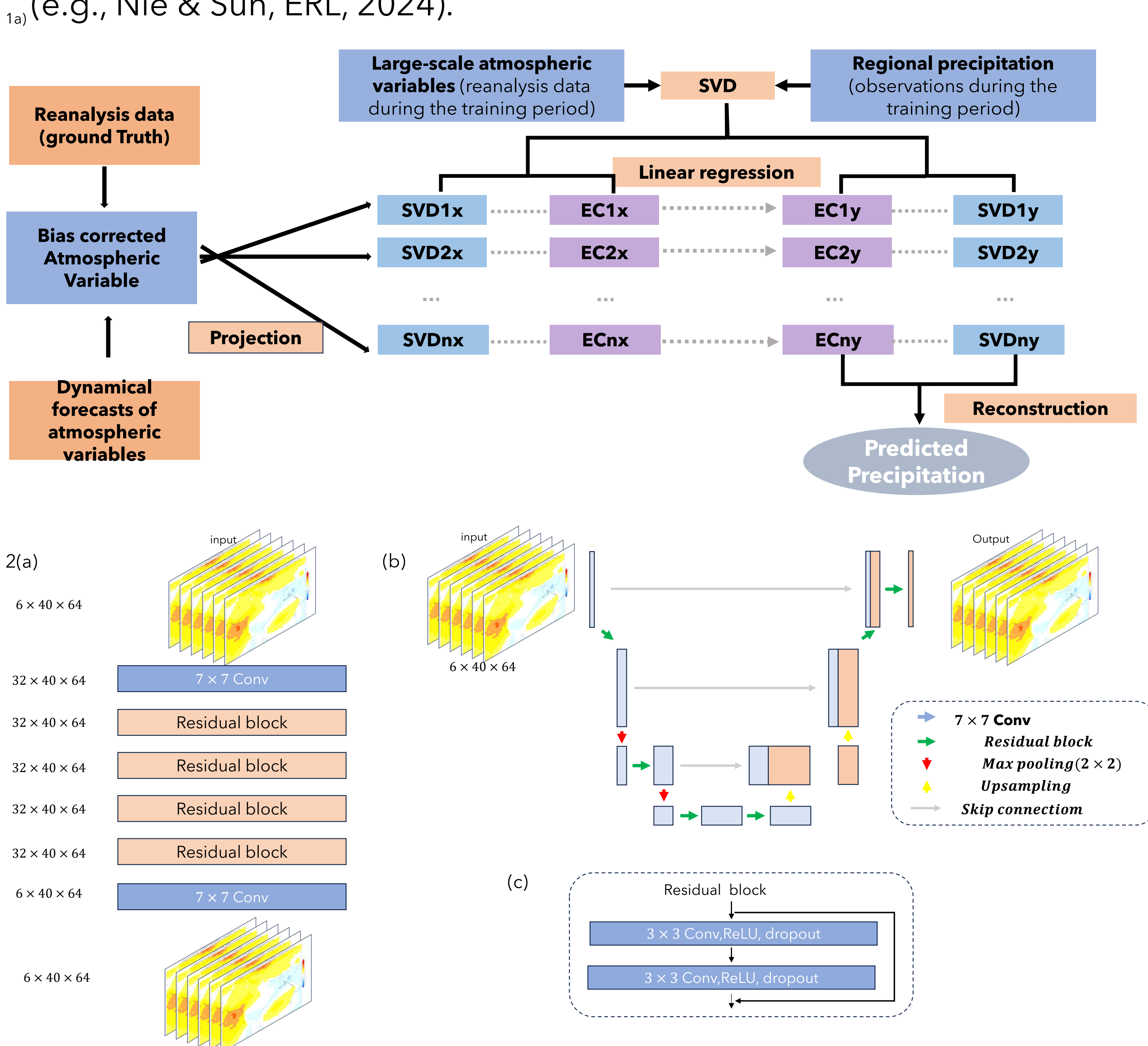


Figure 1. Schematic of the hybrid DL-DSPM framework for bias correction and sub-seasonal precipitation prediction. Figure 2. Deep learning architectures used for atmospheric variable bias correction: (a) ResNet architecture with residual blocks; (b) U-Net encoder-decoder architecture with skip connections; and (c) detailed structure of a residual block.

Objectives

- Apply k-means and other suitable clustering techniques to define sub-regions with coherent rainfall behavior.
- Use deep learning models (ResNet, U-Net) for bias correction of large-scale circulation fields.
- Apply the Dynamical Statistical Projection Model (DSPM) to derive rainfall from the bias-corrected large-scale fields.
- Enhance NCMRWF S2S rainfall forecast skills using a hybrid ML-dynamical framework.

Data & Domain

- **NCMRWF S2S Hindcast (1993-2015)**: Sub-seasonal hindcast forecasts providing the circulation variables used as predictors – Z500, U200, V200, V850, and specific humidity (SH850).
- **ERA5 (ECMWF Reanalysis v5)**: Global atmospheric reanalysis used as the ground-truth circulation field for bias-correcting the NCMRWF forecast.
- **IMD (India Meteorological Department)**: Gridded daily rainfall observations over India, used as ground truth for the precipitation prediction skill evaluation.
- **Study Area**: Indian landmass, regionalised into 6 rainfall-homogeneous clusters via K-means clustering of IMD gridded rainfall (training period only), to account for regional heterogeneity in monsoon rainfall behaviour.

Acknowledgments: The authors gratefully acknowledge Dr.V.S. Prasad, Director of NCMRWF, for his collaboration and Prof. Udgate (University of Hyderabad) for the valuable discussions and insights.

Reference: Nie et al. (2024) ,Han et al. (2021)

Result

1) Bias correction

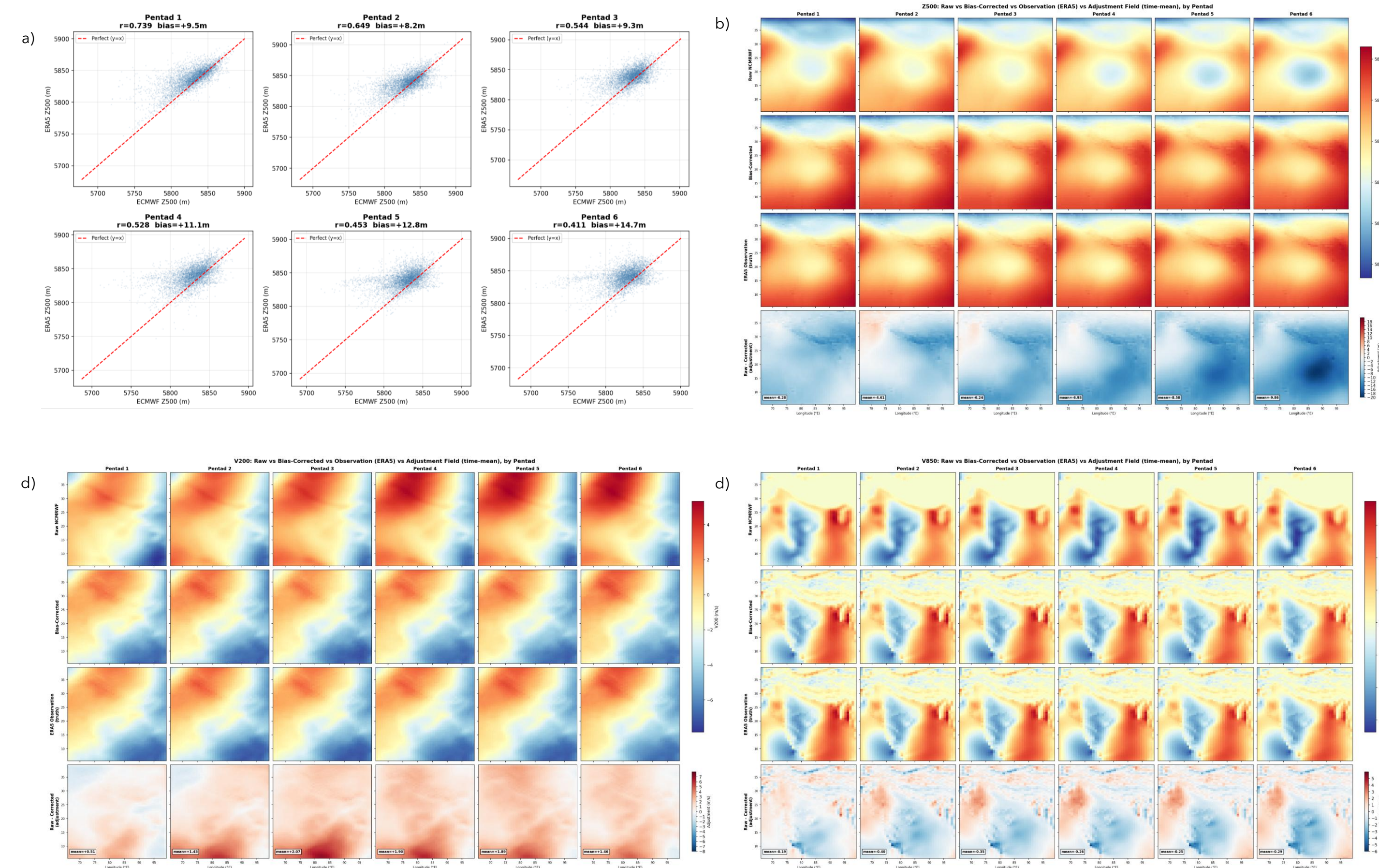


Figure 3: Z500 scatter plots showing correlation and bias between model and reanalysis, (c) V850, (d) V200, and (e) Z500 spatial distributions of raw forecasts, bias-corrected forecasts, ERA5 observations, and the corresponding bias-correction adjustment fields for six pentads.

2) ML DSPM

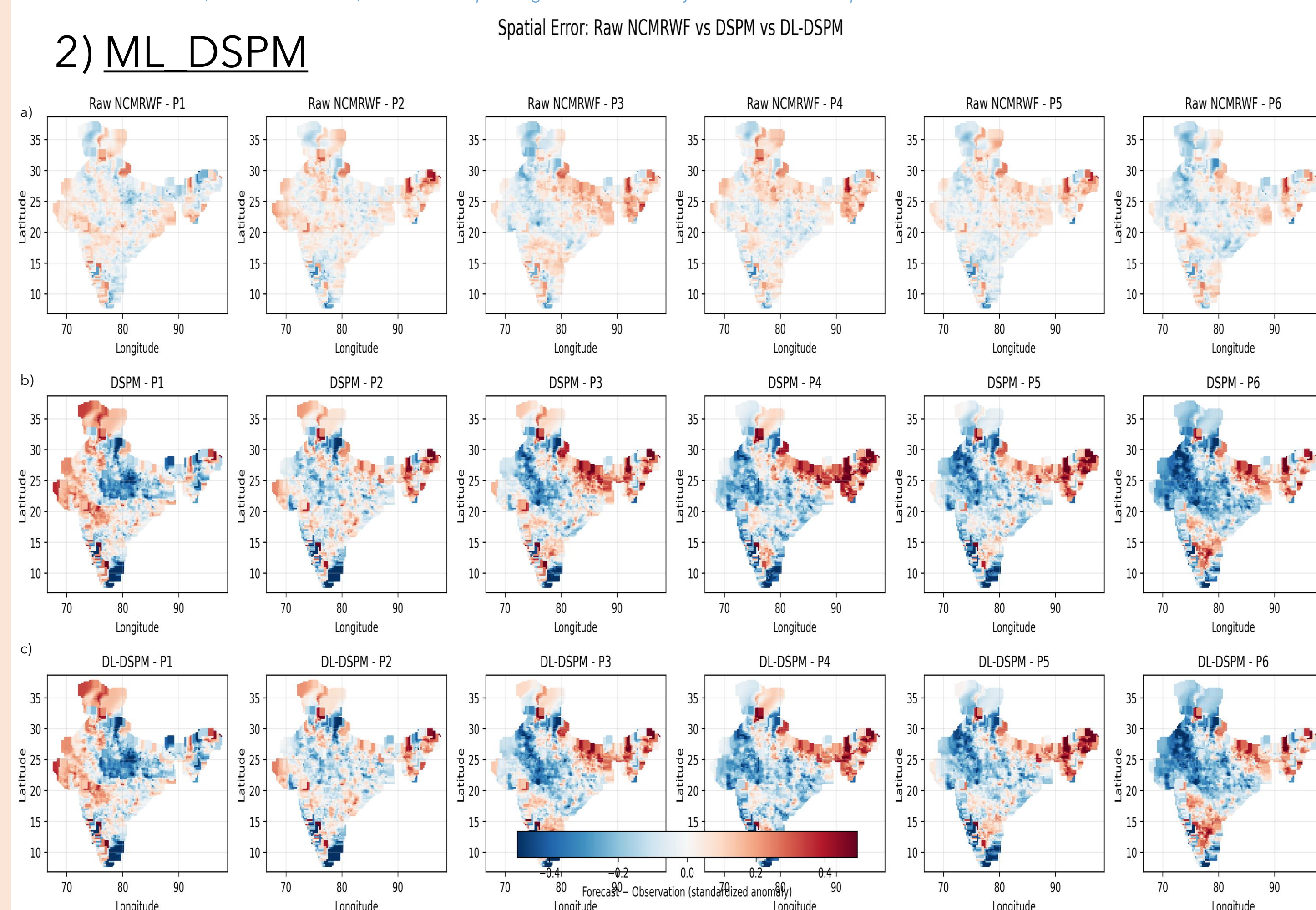


Figure 4: Pentad-wise spatial distribution of standardized forecast errors for (a) raw NCMRWF forecasts, (b) DSPM forecasts, and (c) DL-DSPM forecasts across six pentads (P1-P6), relative to observations. Blue and red shading indicate negative and positive errors, respectively.

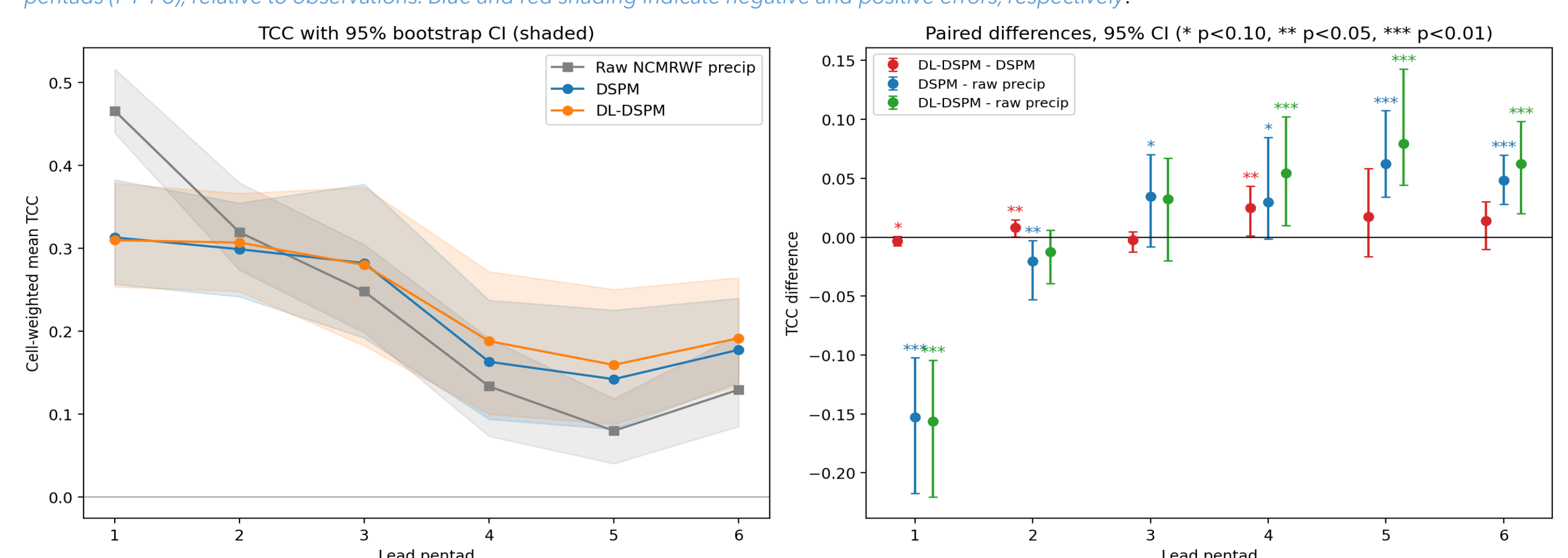


Figure 5. Lead-pentad-wise TCC comparison of raw NCMRWF, DSPM, and DL-DSPM forecasts with 95% bootstrap confidence intervals (left) and paired TCC differences with statistical significance (right).

Conclusion

- U-Net bias correction yields significant gains for circulation fields: Z500 RMSE falls 16-35% (mean $\approx 21\%$) with TCC improving significantly in 5 of 6 pentads (ΔTCC up to +0.067); U200 improves significantly at all 6 pentads (ΔTCC up to +0.076, RMSE -9 to -14%).
- DSPM/DL-DSPM precipitation skill (cell-weighted TCC) trails the raw NCMRWF forecast at pentads 1-2 (raw TCC $\approx 0.46 \rightarrow 0.32$ vs. DSPM/DL-DSPM $\approx 0.31 \rightarrow 0.30$) but significantly exceeds it from pentad 3 onward, as raw skill decays faster (pentad 5: raw ≈ 0.08 vs. DSPM/DL-DSPM $\approx 0.14-0.16$, $p<0.01$).
- The added value of bias-corrected circulation specifically (DL-DSPM vs. DSPM) is smaller and less consistent – a +23% gain at pentad 4 ($p<0.05$) but near-zero or mixed at other leads, varying by region weighting.