Enhancing Bayesian Distributed Learning with Spatially Adjusted Predictive Distributions

Luca Presicce

Ph.D. student in Statistics

University of Milan-Bicocca, Department of Economics, Management & Statistics

SIS 2025 - Statistics for Innovation, June 16-18, 2025, Genoa (Italy)

Issues in Geostatistical Models

- ☞ Growing availability of large geographically referenced datasets (e.g., climate sciences, environmental monitoring, epidemiology).
- ☞ Customary Bayesian geostatistical models based on Gaussian processes face computational limitations with moderate n.
- ☞ Emerges the need for scalable methods delivering rapid inference, and spatial predictions.

Distributed learning offers scalable approximate solutions.

Motivation and Key limitations of Distributed Learning in Spatial modeling

Distributed approaches often show partition dependence when spatial dependency get involved.

- ☞ Loss of the original dependence structure in spatially regular distributed inference.
- ☞ Converting in poor performance on regular grid partitions, notwithstanding the spontaneity of spatial blocks.
- ☞ Emerges the need for more spatially coherent aggregation methods for Bayesian distributed models.

Spatial adjustment scheme for distributed models

Bayesian distributed learning yields posterior inferences as

$$\hat{p}(\,\cdot \mid \mathscr{D}\,) = \sum_{k=1}^{K} w_k \; p_k(\,\cdot \mid \mathscr{D}_k\,) \qquad (1)$$

- $\boldsymbol{w} = \{w_1, \ldots, w_K\}$ posterior weights
- $p_k(\,\cdot \mid \mathscr{D}_k)$ local posterior distribution.

For $u_j \in \mathcal{U}$ unobserved location

$$\widetilde{w}_{u_j,k} = \frac{w_k \cdot \gamma(\mathrm{dist}(u_j, c_k))}{\sum_{k=1}^{K} w_k \cdot \gamma(\mathrm{dist}(u_j, c_k))} \qquad (2)$$

- $\gamma(\cdot)$ is a decreasing function

- Centroid-based adjustment of posterior weights
- Incorporates spatial distance between prediction locations and partition centroids

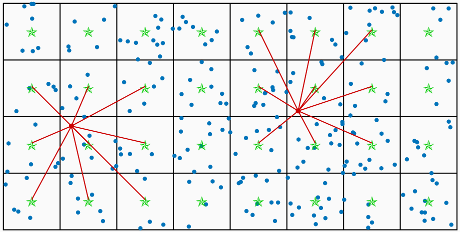

Figure 1: Spatial weights adjustment scheme.

Key contributions

Practical advantages introduced by trivial adjustment, with minimal computational overhead

- ☞ Balance for partition scheme dependence, helping to recover predictive performances in spatially distributed settings.
- ☞ Maintains scalability of distributed approaches: overhead enters in predictive sampling.
- ☞ Adjustment flexibility selecting weighting function $\gamma(\cdot)$, and compatibility with existing distributed Bayesian frameworks.

Sea Surface Temperature Data Analysis

Dimensions:

- ☞ $\approx 480,000$ train locations
- ☞ ≈ 2000 partitions (5×5 regular grid)

Achievements:

- ☞ 45% reduction RMSPE.
- ☞ Preservation of global patterns.

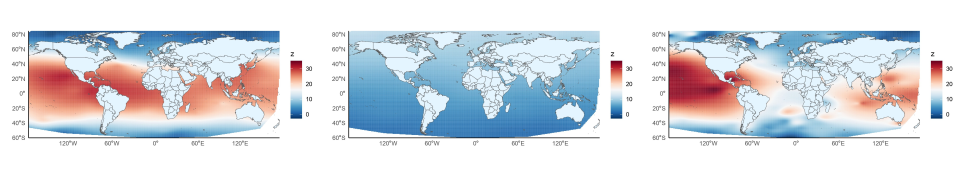

Figure 2: From left to right: test data interpolation, not-adjusted prediction, and adjusted prediction.

Bayesian distributed model fitted using spBPS R package: **L. Presicce, S. Banerjee** (2025+) **"Bayesian Transfer Learning for Artificially Intelligent Geospatial Systems: A Predictive Stacking Approach"**, Under Review.

Wrapping up

- Effective and Easy-to-use solution for partition dependence in spatial Bayesian distributed models.
- Maintains computational efficiency of distributed approaches affecting only predictive sampling.

Available in `spBPS` R package (sooonly)!

Check it out on GitHub

Thanks for your attention!