Advancements in Neural Bayes Estimation for Spatial Processes

Jordan Richards[1], Matthew Sainsbury-Dale[1,2], Andrew Zammit-Mangion[2], and Raphaël Huser[1]

[1]King Abdullah University of Science and Technology, Saudi Arabia [2]University of Wollongong, Australia

SCAN ME

Introduction

Inference for spatial models can be **computationally troublesome** due to their reliance on **intractable** and/or **censored likelihoods**. Neural Bayes estimators (NBEs) are **likelihood-free** and **orders-of-magnitude faster** than classical likelihood-based methods. Two recent projects:

- adapt NBEs for censored input data (Richards et al., 2023),
- apply NBEs in a **peaks-over-threshold** framework,
- adapt NBEs to irregular spatial data (Sainsbury-Dale et al., 2023a),
- apply NBEs for **credible interval** estimation.

Methodology

Neural Bayes estimators are estimators that target the **Bayes risk** (Sainsbury-Dale et al., 2023b), constructed as neural networks that map model realisations $\mathbf{Z}$ to the true parameter set $\boldsymbol{\theta}$.

Censoring: helpful to **reduce bias in extremal dependence estimates**. Components of $\mathbf{Z}$ are left-censored with fixed censoring level close to one.

Irregular spatial data: previous NBEs limited to data observed on a regular grid. Constraint alleviated through use of **graph neural networks.**

Models: we consider four popular spatial extremal dependence models:

- (Inverted) Max-stable processes (I)MSPs, Gaussian processes (GPs),
- **Huser and Wadsworth (2019) (HW) random scale mixture**,

$$Z(\cdot) = R^{\delta} W(\cdot)^{1-\delta}, \quad \text{for } \delta \in [0,1],$$

with a unit-Pareto r.v. $R \geq 1$ and GP $W(\cdot)$ with unit-Pareto margins. The parameter δ **determines the extremal dependence class**, with asymptotic independence if and only if $\delta \leq 0.5$.

References

Huser, R. and Wadsworth, J. L. (2019). Modeling spatial processes with unknown extremal dependence class. *Journal of the American Statistical Association*, 114(525):434–444.

Richards, J., Sainsbury-Dale, M., Zammit-Mangion, A., and Huser, R. (2023). Neural Bayes estimators for censored peaks-over-threshold models. *arXiv preprint arXiv:2306.15642*.

Sainsbury-Dale, M., Richards, J., Zammit-Mangion, A., and Huser, R. (2023a). Neural Bayes estimators for irregular spatial data using graph neural networks. *arXiv preprint arXiv:2310.02600*.

Sainsbury-Dale, M., Zammit-Mangion, A., and Huser, R. (2023b). Likelihood-free parameter estimation with neural Bayes estimators. *The American Statistician*, (In press).

Simulation studies

Gains compared to censored pairwise likelihood (cPL): inference for 200 replicates of GP / MSP / IMSP with range $\lambda \in [2,10]$ and smoothness $\kappa \in [0.5,2]$ for correlation/power-variogram. Data are on a 16×16 grid. **Speed-up of the order of 67,500–750,000.**

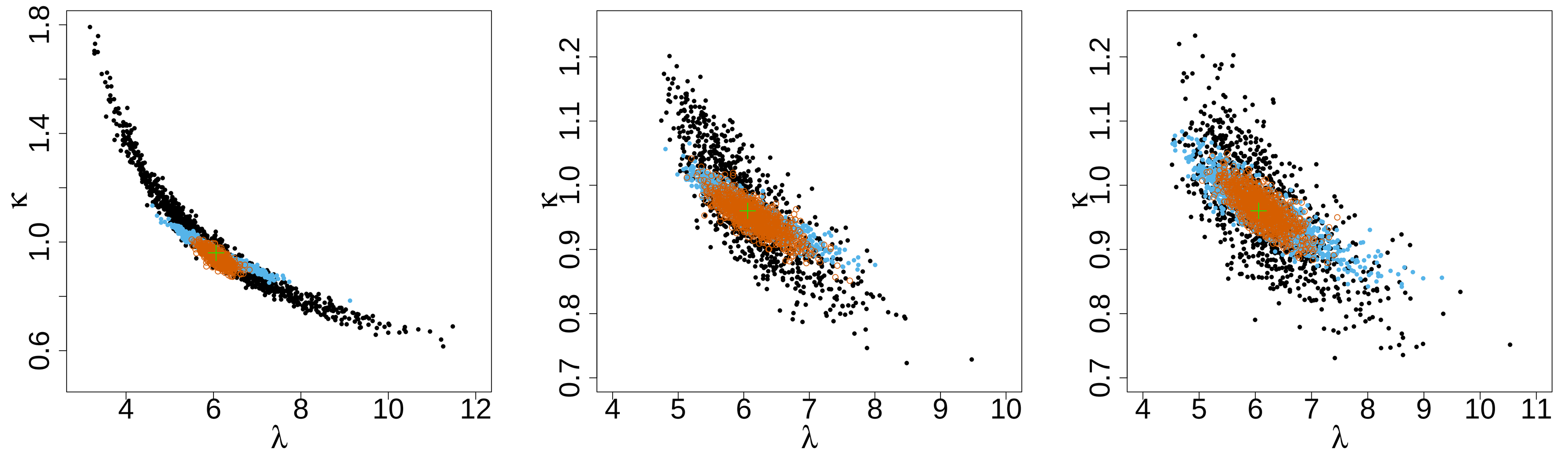

Fig. 1. Empirical joint dist. Brown: NBE. Green: true parameter set. Black: cPL (all pairs). Blue: cPL (pairs within 3 units).

Irregular spatial configurations: Matérn GP with range $\rho \in [0.05, 0.3]$, unit smoothness, and nugget $\tau \in [0.1, 1]$. Data are observed at 250 locations, whose **configuration is randomly generated during training.**

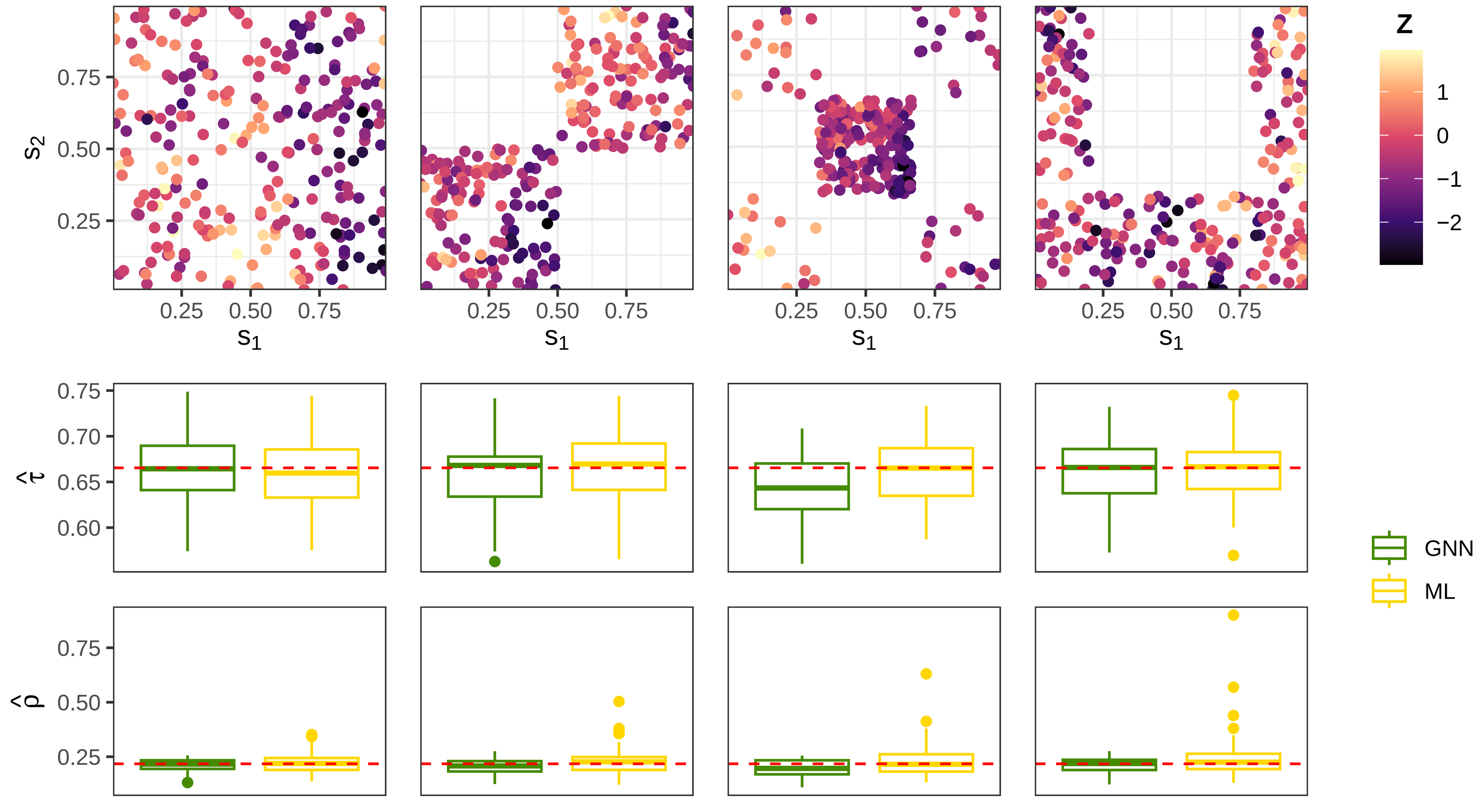

Fig. 2. Top; spatial configurations used for testing. Bottom; marginal sampling dists. for **GNN-NBE** and **maximum likelihood estimator**.

Applications

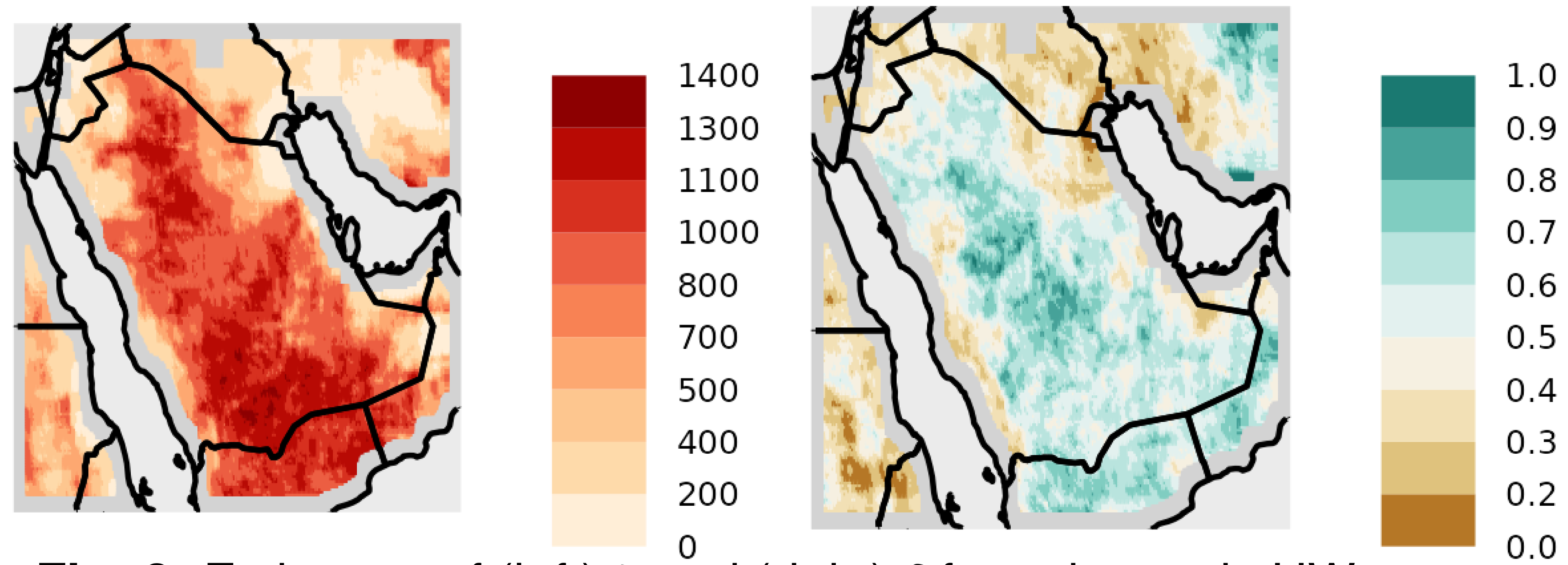

Fig. 3. Estimates of (left) λ and (right) δ for anisotropic HW process. Data are monthly mean $PM_{2.5}$ on 242×189 grid; model fitted locally on all 16×16 grids. There are **26387 estimates, each taking** $\sim 1 \times 10^{-3}$**s.**

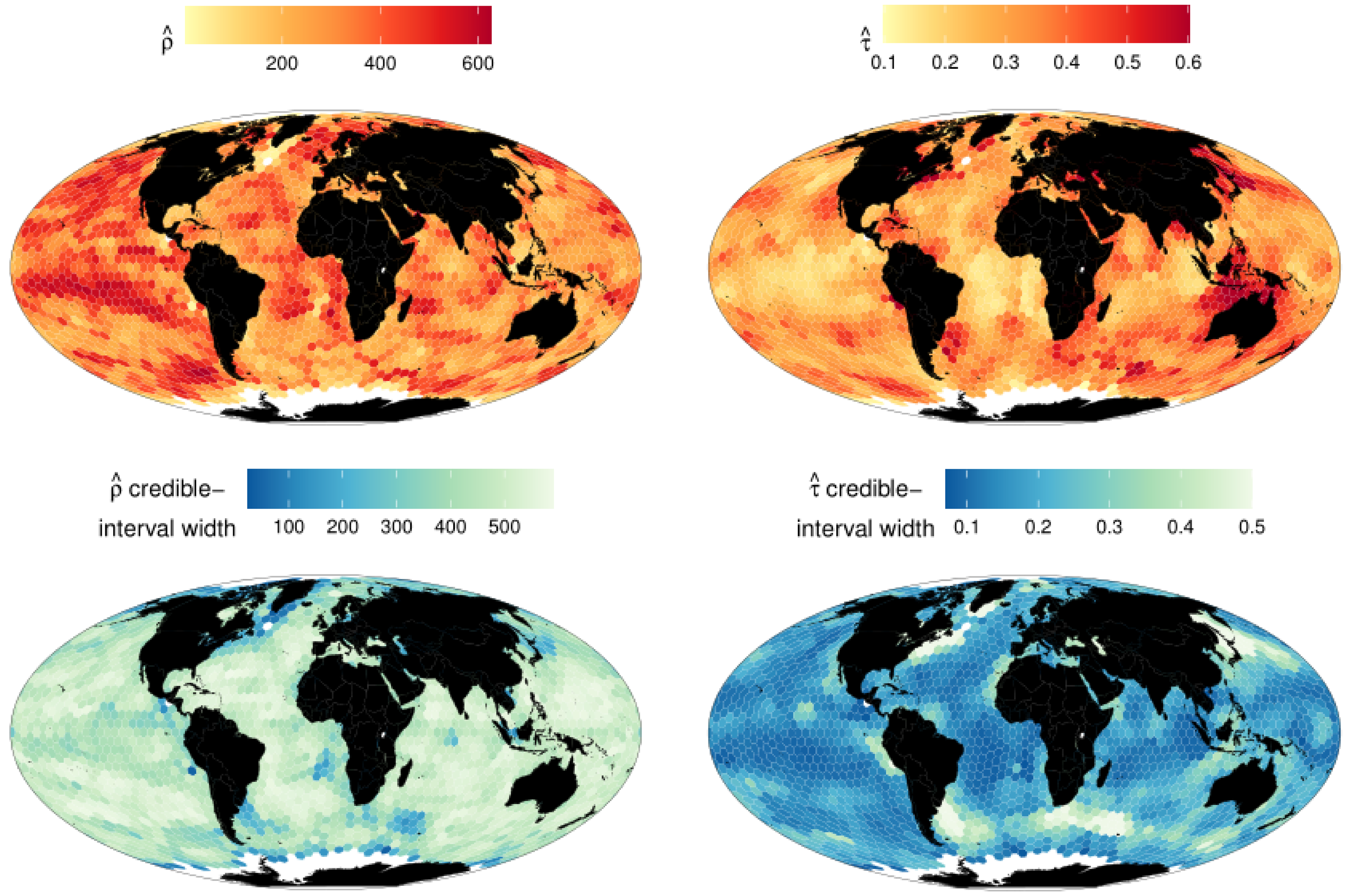

Fig. 4. Estimates of (ρ, τ) and **credible interval widths** for Matérn GP. Data are 2161 hexagonal cell clusters with SST observed at 2769 to 12591 locations. **All estimates require a single estimator and take only three minutes to compute on a single GPU.**