Non-stationary Gaussian fields and Penalized Complexity Priors

Liam Llamazares

with Finn Lindgren and Jonas Latz

April 30, 2026

Outline

Outline

Problem: Where to go during the holidays?

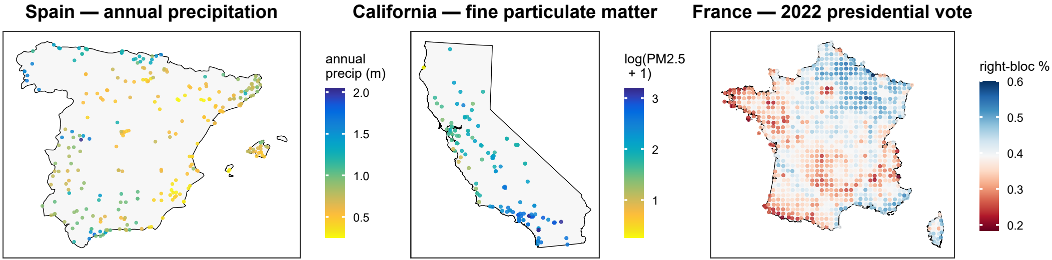
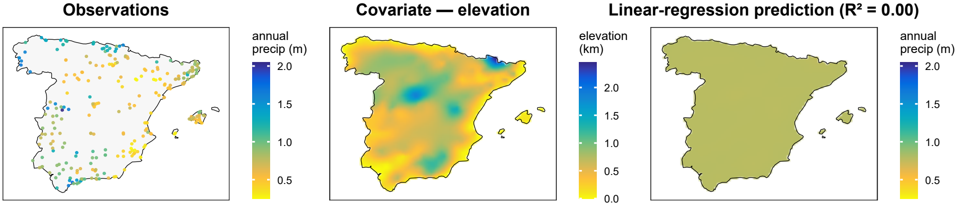


Figure: Spain, California, or France?

Building a model

Basic linear regression: $y(\boldsymbol{x}_i) = \beta_0 + \beta_1 h(\boldsymbol{x}_i) + \varepsilon_i$



Spatial field

$$y(x_i) = \beta_0 + \beta_1 h(\boldsymbol{x}_i) + u(\boldsymbol{x}_i) + \varepsilon_i$$

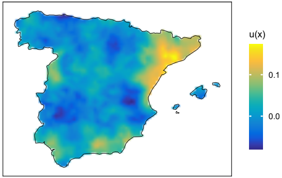


Figure: Example spatial field $u(\boldsymbol{x})$

Requirements on u: Physical, probabilistic (Bayesian), computable.

Outline

Physics through PDEs

This is a PDE

$$(\zeta - \Delta)u(\boldsymbol{x}) = f(\boldsymbol{x}).$$

The solution u models some quantity of interest

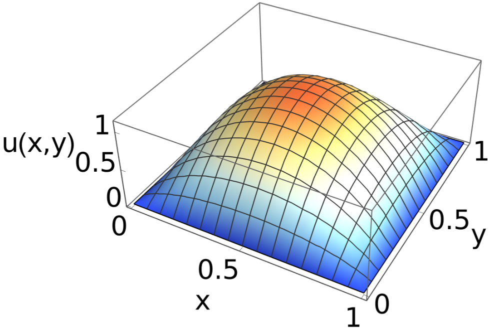


Figure: Solution u of the PDE.

Physics & uncertainty through SPDEs

This is a SPDE

$$(\zeta - \Delta)u(\boldsymbol{x}) = \mathcal{W}(\boldsymbol{x}).$$

The solution u models some quantity we are uncertain about.

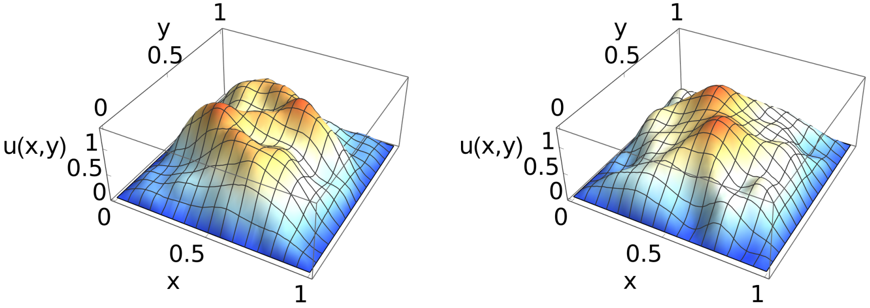


Figure: Two realizations of solution u of the SPDE.

Gaussian fields

The solution u to an SPDE $\mathcal{L}u = \mathcal{W}$ is a Gaussian field

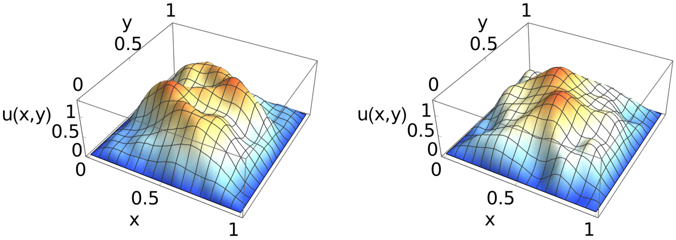


Figure: Gaussian field

Definition (Gaussian field)

A *Gaussian field* on $\mathcal{D} \subseteq \mathbb{R}^d$ is a collection of random variables $\{u(\boldsymbol{x}) : \boldsymbol{x} \in \mathcal{D}\}$, such that, for all $\{\boldsymbol{x}_i\}_{i=1}^n \subset \mathcal{D}$, $(u(\boldsymbol{x}_1), \ldots, u(\boldsymbol{x}_n))$, is a Gaussian vector.

Introduction: Covariance and precision

Notation (Covariance and precision)

Let u be Gaussian with mean zero, the **covariance** *of u is*

$$K(\boldsymbol{x}, \boldsymbol{y}) := \mathbb{E}[u(\boldsymbol{x})u(\boldsymbol{y})], \quad \forall x, y \in \mathcal{D}.$$

We call $Q = K^{-1}$ the precision and write $u \sim \mathcal{N}\left(0, Q^{-1}\right)$.

We say that u is *stationary* if $K(\boldsymbol{x}, \boldsymbol{y}) = r(\boldsymbol{x} - \boldsymbol{y})$, it is isotropic if $K(\boldsymbol{x}, \boldsymbol{y}) = r(|\boldsymbol{x} - \boldsymbol{y}|)$.

Stationary and non-stationary Matérn fields

Definition (Isotropic *Matérn* field)

The *isotropic* solution u with parameters $\zeta > 0$ and $\sigma > 0$ to

$$(\zeta - \Delta)u(\boldsymbol{x}) = \sigma \mathcal{W}(\boldsymbol{x}).$$

Matérn 1947; Whittle 1963; Stein 1999

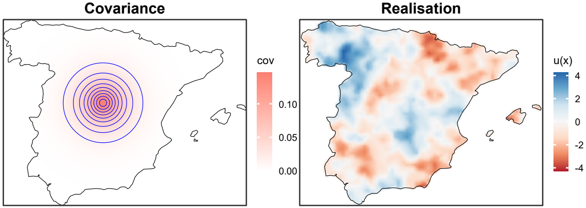


Figure: Isotropic Matérn field over Spain

Stationary Matérn fields

Definition (Anisotropic *Matérn* field (Fuglstad et al. 2015))

The *stationary* solution u with diffusion matrix $\boldsymbol{H} \in \mathbb{R}^{2\times 2}$.

$$(\zeta - \nabla \boldsymbol{H} \nabla)\frac{u(\boldsymbol{x})}{\sigma} = \mathcal{W}(\boldsymbol{x}).$$

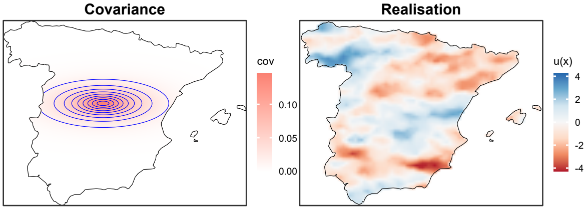


Figure: Anisotropic Matérn field over Spain

Stationary Matern fields

$\boldsymbol{H} \equiv \boldsymbol{H}_{\boldsymbol{v}}$ can be parametrized invertibly and smoothly using a vector $\boldsymbol{v} \in \mathbb{R}^2$ (Llamazares-Elias, Latz, and Lindgren 2024).

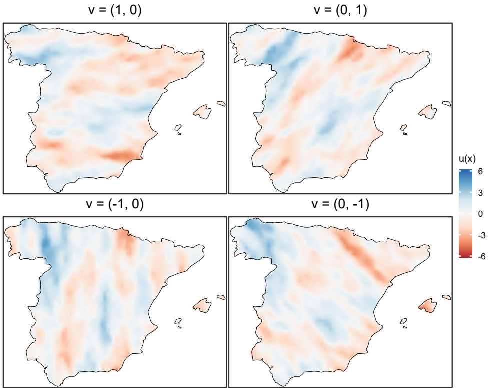


Figure: Anisotropic Matérn field for different values of $\boldsymbol{v}$

Non-stationary Matérn fields

Definition (*Non-stationary* Matérn field .)

The *non-stationary* solution u with diffusion matrix $\boldsymbol{H}_{\boldsymbol{v}(x)} \in \mathbb{R}^{2\times 2}$

$$(\zeta(\boldsymbol{x}) - \nabla \boldsymbol{H}_{\boldsymbol{v}(\boldsymbol{x})} \nabla)\frac{u(\boldsymbol{x})}{\sigma(\boldsymbol{x})} = \mathcal{W}(\boldsymbol{x}).$$

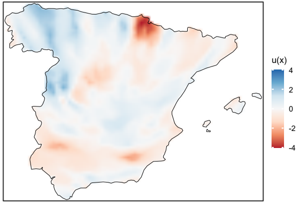


Figure: Non-stationary Matérn field

Outline

PC priors: Bayesian inference

Problem setting: Given an observation $\boldsymbol{u} = (u(x_1), \dots, u(x_n))$ of u how likely is a given $\boldsymbol{\theta} = (\zeta, \boldsymbol{v}, \sigma)$?

$$p(\boldsymbol{\theta}|\boldsymbol{u}) \propto p(\boldsymbol{u}|\boldsymbol{\theta})p(\boldsymbol{\theta}).$$

- The likelihood $p(\boldsymbol{u}|\boldsymbol{\theta}) \sim \mathcal{N}\left(0, \boldsymbol{Q}_{\boldsymbol{\theta}}^{-1}\right)$ comes from solving the SPDE (FEM).
- How should we define the prior $p(\boldsymbol{\theta})$? (crucial)

PC priors: Motivation

“*Penalised complexity* (PC) priors” [Simpson et al. 2014].

Main idea

Occam’s razor: *Simple models are preferred.*

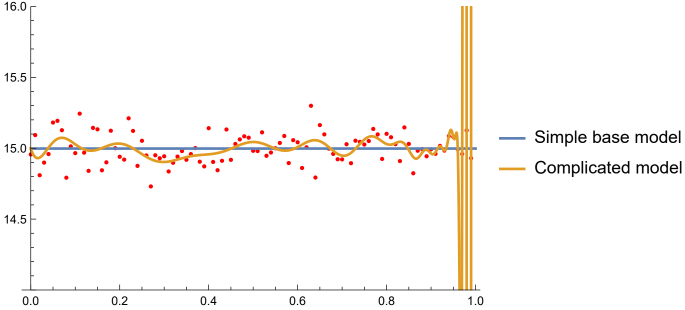


Figure: Which model is better?

PC priors 101

Implementation

1. Select a simple base model $\boldsymbol{\theta}_{\mathbf{0}}$.
2. Define a distance to the base model

$$m(\boldsymbol{\theta}) := d(u_{\boldsymbol{\theta}}, u_{\boldsymbol{\theta}_{\mathbf{0}}}).$$

3. **Penalization of complexity:** Impose a distribution on $\boldsymbol{\theta}$ so that $m(\boldsymbol{\theta}) \sim \mathrm{Exp}(\lambda)$. That is,

$$p(m) = \lambda \exp(-\lambda m).$$

Stationary spectral theory

Given a stationary field u with covariance $r(x-y) = K(x,y)$ we define the *spectral density of u* as

$$S(\boldsymbol{\xi}) = \int_{\mathbb{R}^2} r(\boldsymbol{x}) e^{-2\pi i \boldsymbol{\xi}\cdot\boldsymbol{x}} \, \mathrm{d}\boldsymbol{\xi}$$

The spectral density encodes information about u

Theorem (Spectral theorem)

Let u be stationary, then there exists a (generalized) field $\mathrm{d}Z$ such that

$$u(\boldsymbol{x}) = \int_{\mathbb{R}^2} e^{2\pi i \boldsymbol{\xi}\cdot\boldsymbol{x}} \, \mathrm{d}Z(\boldsymbol{\xi}), \quad \textit{where}, \quad \mathbb{E}[\,\mathrm{d}Z(\boldsymbol{\xi})\,\mathrm{d}Z(\boldsymbol{\omega})] = S(\boldsymbol{\xi})\delta(\boldsymbol{\xi}-\boldsymbol{\omega})$$

Stationary PC priors: A spectral distance

For constant θ (stationary model) (Llamazares-Elias, Latz, and Lindgren 2024)

- L^2 distance of spectral densities relative to base model.

$$d(u_{\boldsymbol{\theta}}, u_{\boldsymbol{\theta_0}}) := \left(\int_{\mathbb{R}^2} S_{\boldsymbol{\theta}_0}^{-1}(\boldsymbol{\xi}) \left(S_{\boldsymbol{\theta}}(\boldsymbol{\xi}) - S_{\boldsymbol{\theta_0}}(\boldsymbol{\xi}) \right)^2 \, \mathrm{d}\boldsymbol{\xi} \right)^{\frac{1}{2}} .$$

- We set $\boldsymbol{\theta_0} = \mathbf{0}$ and **calculate** $m(\boldsymbol{\theta})$ **exactly**.
- We obtain a PC prior $p_{\text{stat}}(\boldsymbol{\theta})$.

Non-stationary fields

Problem: The spectral density for non-stationary fields is undefined.

Proposition (Bochner's isometry)

Given stationary field u *with covariance* $r(\boldsymbol{x}-\boldsymbol{y}) = K(\boldsymbol{x},\boldsymbol{y})$,

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} r(\boldsymbol{x}-\boldsymbol{y}) f(\boldsymbol{y})\overline{g(\boldsymbol{x})} = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} S(\boldsymbol{\xi})\widehat{f}(\boldsymbol{\xi})\overline{\widehat{g}(\boldsymbol{\xi})}.$$

Definition (Non-stationary spectral density)

We define the *spectral density of a non-stationary field* u with covariance K as

$$S(\boldsymbol{\xi},\boldsymbol{\omega}) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K(\boldsymbol{x},\boldsymbol{y}) e^{-2\pi i\boldsymbol{\xi}\cdot\boldsymbol{x}} e^{2\pi i\boldsymbol{\omega}\cdot\boldsymbol{y}}\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{y}.$$

Non-stationary spectral theory

Proposition (Bochner's isometry)

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} K(\boldsymbol{x},\boldsymbol{y}) f(\boldsymbol{y})\overline{g(\boldsymbol{x})} = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} S(\boldsymbol{\xi},\boldsymbol{\omega})\widehat{f}(\boldsymbol{\omega})\overline{\widehat{g}(\boldsymbol{\xi})}.$$

Theorem (Non-starionary spectral theorem)

There exists a unique generalized field $\mathrm{d}Z$ *such that*

$$u(\boldsymbol{x}) = \int_{\mathbb{R}^d} e^{2\pi i \boldsymbol{x}\cdot\boldsymbol{\xi}}\,\mathrm{d}Z(\boldsymbol{\xi}), \quad \forall \boldsymbol{x} \in \mathbb{R}^d. \tag{1}$$

Furthermore, $\mathbb{E}[\,\mathrm{d}Z(\boldsymbol{\xi})\,\overline{\mathrm{d}Z(\boldsymbol{\omega})}] = S(\boldsymbol{\xi},\boldsymbol{\omega})$*.*

Distance for non-stationary field

Definition

We define the distance from *non-stationary* $u_{\boldsymbol{\theta}}$ to *stationary* $u_{\boldsymbol{\theta}_0}$ as

$$D(u_{\boldsymbol{\theta}}, u_{\boldsymbol{\theta}_0}) := \left(\int_{\mathbb{R}^2} \int_{\mathbb{R}^2} S_{\boldsymbol{\theta}_0}(\boldsymbol{\xi})^{-1} \left| S_{\boldsymbol{\theta}}(\boldsymbol{\xi}, \boldsymbol{\omega}) - S_{\boldsymbol{\theta}_0}(\boldsymbol{\xi}) \right|^2 \mathrm{d}\boldsymbol{\omega} \, \mathrm{d}\boldsymbol{\xi} \right)^{\frac{1}{2}}$$

Problem: No closed form expression for $S_{\boldsymbol{\theta}}(\boldsymbol{\xi}, \boldsymbol{\omega})$. How to calculate the distance?

Taylor approximation

Proposition

Write $\boldsymbol{\theta} = \boldsymbol{\theta}_0 + \boldsymbol{\theta}_\Delta$ *we exactly calculate the operator* $\boldsymbol{Q}_{\theta_0}$ *such that*

$$\begin{aligned} D(u_{\boldsymbol{\theta}}, u_{\boldsymbol{\theta}_0})^2 &= \|K_{\boldsymbol{\theta}} - K_{\boldsymbol{\theta}_0}\|^2_{\mathrm{HS}(L^2(\mathbb{R}^2)\to H^2(\mathbb{R}^2))} \\ &= \langle \boldsymbol{Q}_{\boldsymbol{\theta}_0}\boldsymbol{\theta}_\Delta, \boldsymbol{\theta}_\Delta \rangle_{L^2(\mathbb{R}^2)} + \mathcal{O}\left(\|\boldsymbol{\theta}_\Delta\|^4\right). \end{aligned}$$

We set $p(\boldsymbol{\theta}_{\boldsymbol{\Delta}}) \propto \exp\left(-\frac{1}{2}\langle \boldsymbol{Q}_{\boldsymbol{\theta}_0}\boldsymbol{\theta}_\Delta, \boldsymbol{\theta}_\Delta \rangle\right)$ and combine with stationary PC-priors

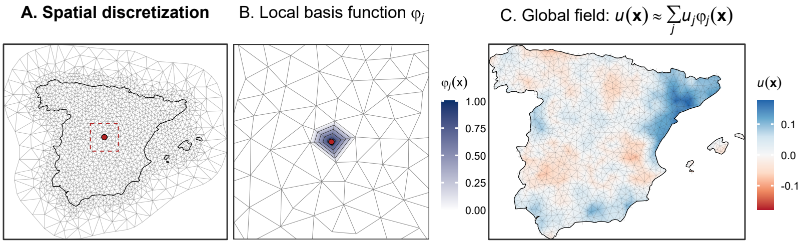
Theorem (Non-stationary PC-priors)

PC priors for the non-stationary model are given by setting

$$\boldsymbol{\theta} = \boldsymbol{\theta}_0 + \boldsymbol{\theta}_\Delta, \quad \boldsymbol{\theta}_0 \sim p_{\text{stat}}(\boldsymbol{\theta}_0), \quad \boldsymbol{\theta}_\Delta | \boldsymbol{\theta}_0 \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{Q}^{-1}_{\boldsymbol{\theta}_0}\right).$$

Outline

The likelihood



Writing $\boldsymbol{u} = \sum_{j=1}^{n} u_j \phi_j$ and imposing

$$\langle \mathcal{L}\boldsymbol{u}, \phi_j \rangle = \langle \mathcal{W}, \phi_j \rangle \,, \quad \forall j = 1, \ldots, n$$

gives for sparse $\boldsymbol{Q}_{\boldsymbol{u}|\boldsymbol{\theta}}$

$$\boldsymbol{u}|\boldsymbol{\theta} \sim \mathcal{N}(0, \boldsymbol{Q}_{\boldsymbol{u}|\boldsymbol{\theta}}^{-1})$$

Can we learn the physics?

We fix a true smooth $\boldsymbol{\theta} = (\zeta, \boldsymbol{v}, \sigma)$ and observe r times

$$\boldsymbol{y}^j = \boldsymbol{u}_{\boldsymbol{\theta}}^j + \boldsymbol{\varepsilon}^j, \quad j = 1, \dots, r.$$

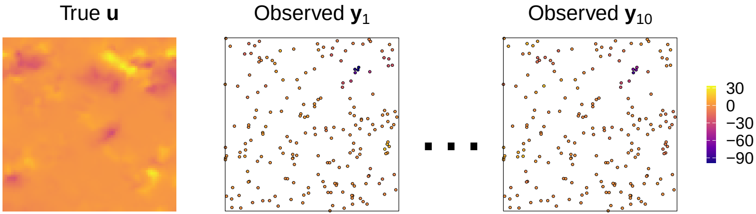


Figure: Noisy observations $\boldsymbol{y}_j$ of true $\boldsymbol{u}_{\boldsymbol{\theta}}$

Can we learn the physics?

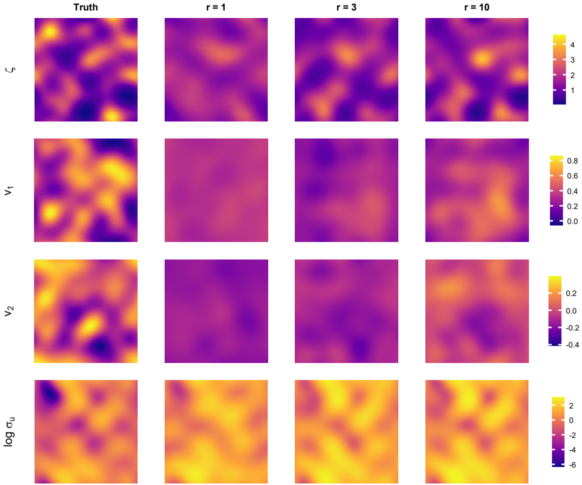


Figure: Recovery of ζ, v_1, v_2, σ as we observe more realizations $r = 1, ..., 10$.

Can we learn the physics ?

We fix a true *discontinuous* $\boldsymbol{\theta} = (\zeta, \boldsymbol{v}, \sigma)$ and observe r times

$$\boldsymbol{y}^j = \boldsymbol{u}_{\boldsymbol{\theta}}^j + \boldsymbol{\varepsilon}^j, \quad j = 1, \dots, r.$$

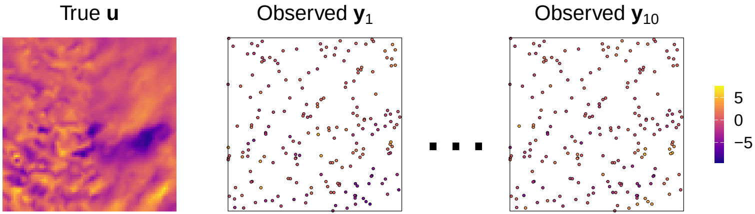


Figure: Noisy observations $\boldsymbol{y}_j$ of true $\boldsymbol{u}_{\boldsymbol{\theta}}$

Can we learn the physics?

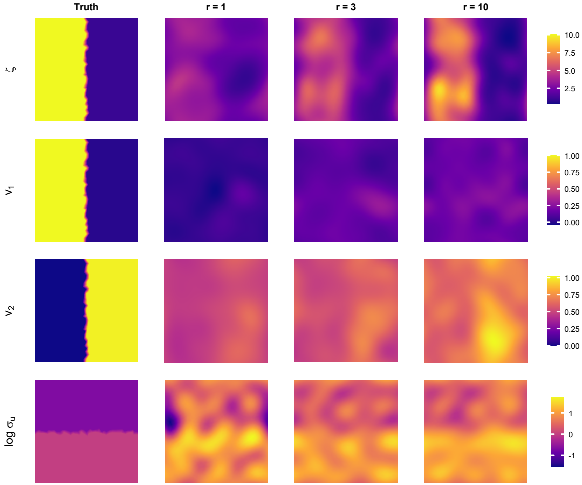
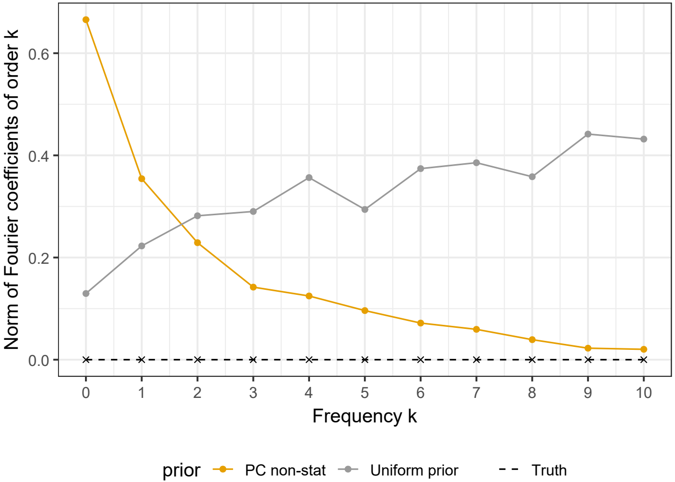


Figure: Recovery of ζ, v_1, v_2, σ as we observe more realizations $r = 1, ..., 10$.

Do we need priors?

If we don’t penalize through prior,s we get white noise



Spanish precipitation: Predictions

Plot side-by-side predictions

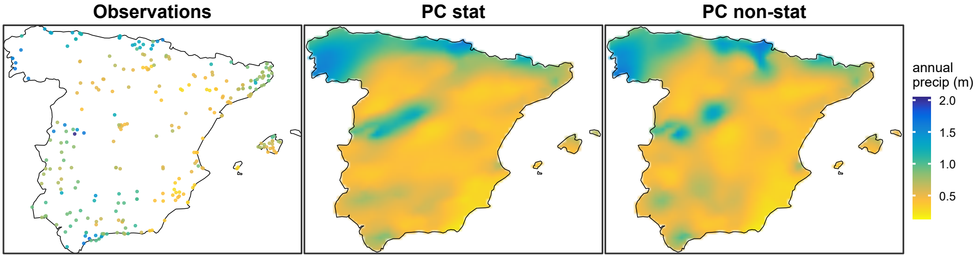
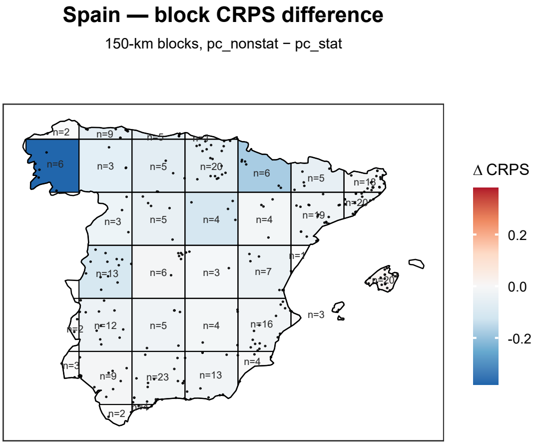


Figure: Predicted precipitation in Spain under a stationary and non-stationary model

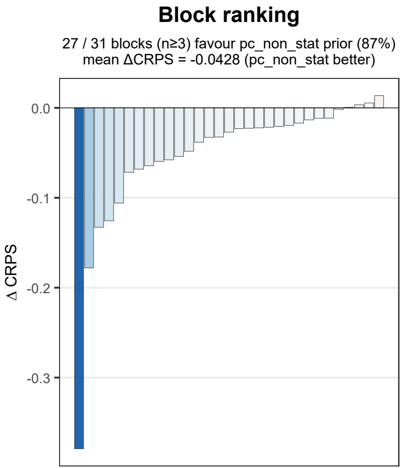
Model performance: Block CRPS

How well does the model generalize?



Spain — block CRPS difference

150-km blocks, pc_nonstat − pc_stat



Block ranking

27 / 31 blocks (n≥3) favour pc_non_stat prior (87%)
mean ΔCRPS = -0.0428 (pc_non_stat better)

Recovered physics

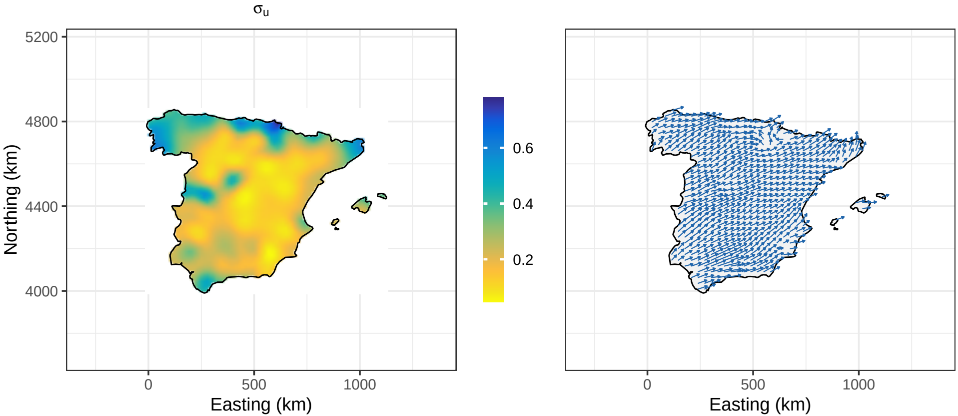


Figure: Variance of u left, and diffusion direction $\widetilde{v}$

Air pollution California

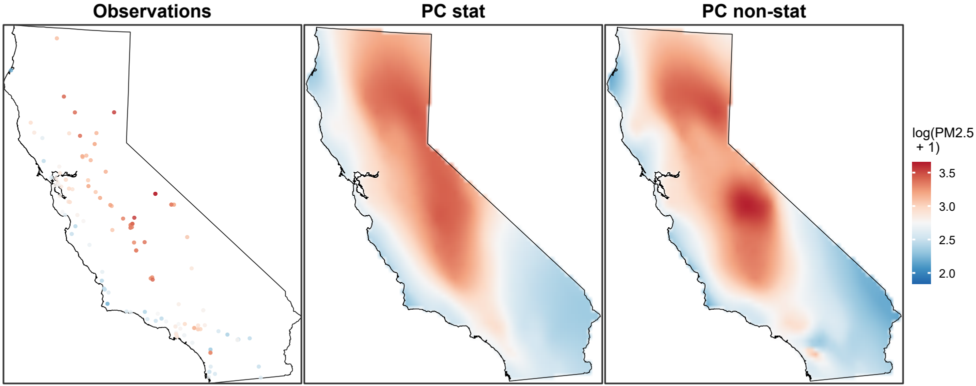


Figure: Comparison of prediction of stationary and non-stationary models

Recovered physics

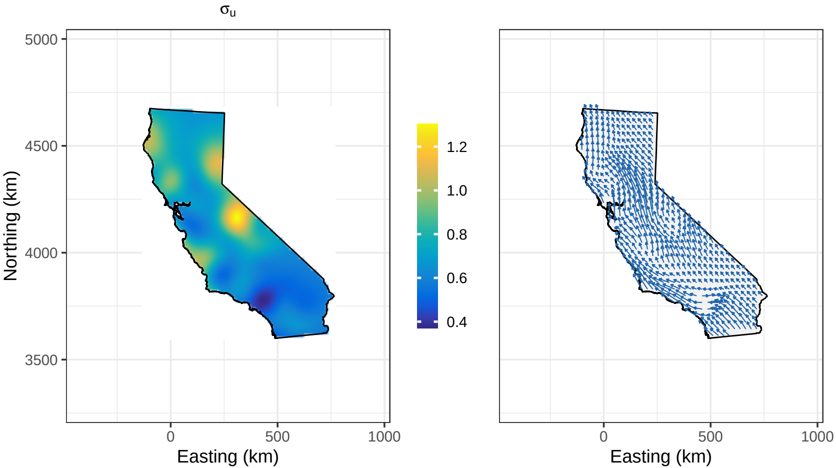
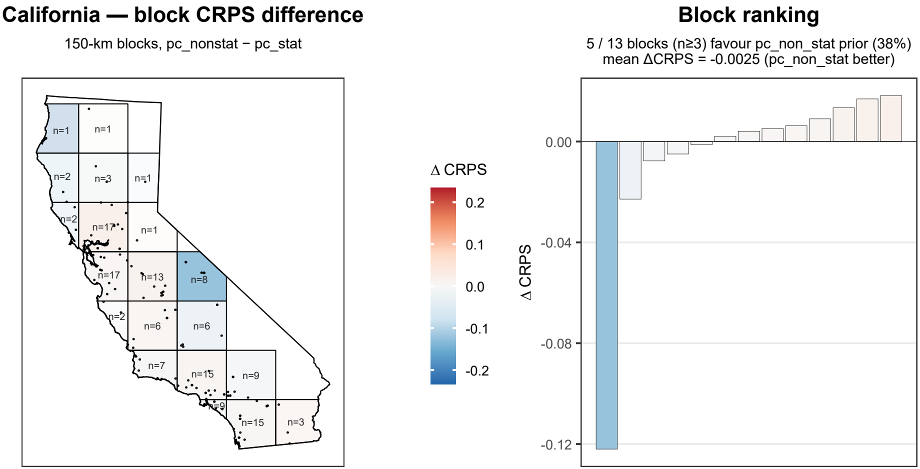


Figure: Variance of u left, correlation range ρ and diffusion direction $\boldsymbol{v}$

Model performance



Figure

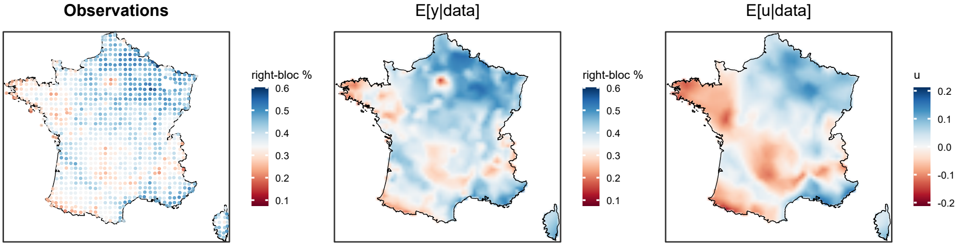
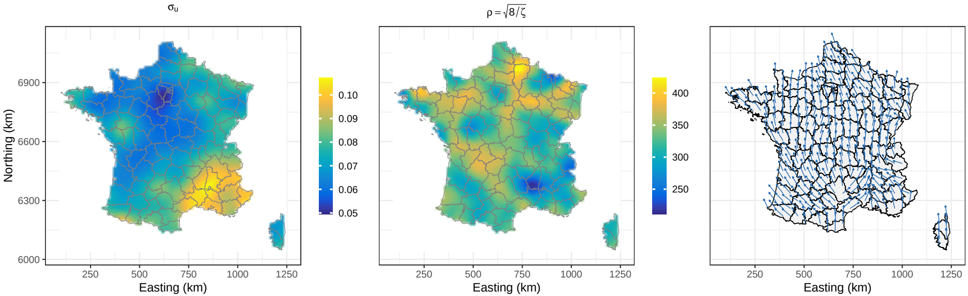


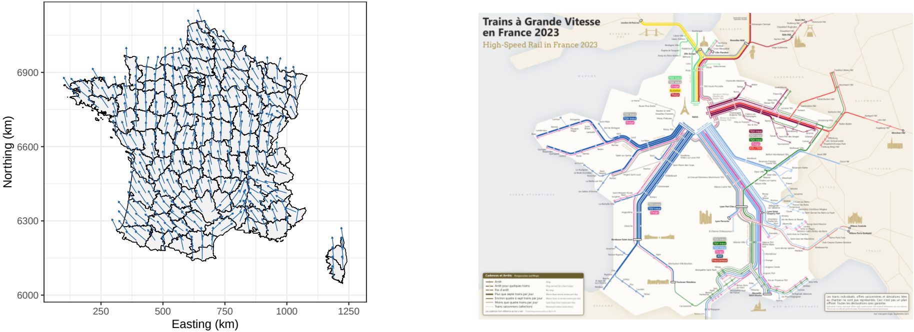
Figure: Observed vote, predicted vote and field

Learned physics



Figure

Learned physics vs. rail network



(b) French rail network.

(a) Posterior diffusion field $\widetilde{v}$.

Figure: Recovered diffusion directions align with the principal rail corridors.

Conclusion and future work

- We considered a **non-stationary** model

$$(\zeta(\boldsymbol{x}) - \nabla \cdot \boldsymbol{H}_{\boldsymbol{v}(\boldsymbol{x})}\nabla)\frac{u(\boldsymbol{x})}{\sigma(\boldsymbol{x})} = \mathcal{W}(\boldsymbol{x}).$$

- We obtained **penalized complexity priors** for $\boldsymbol{\theta}(\cdot) = (\zeta(\cdot), \boldsymbol{v}(\cdot), \sigma(\cdot))$.
- We showed how the model reveals physical structure in precipitation, particle matter and electoral data.
- **Future work:**
  1. Extending to higher smoothness of u.
  2. Extendinf to spatio-temporal fields.
  3. Going on vacation.

References I

- Matérn, Bertil (1947). "Metoder att uppskatta noggrannheten vid linje-och provytetaxering". *Meddelanden från statens skogsforskningsinstitut.*
- Whittle, Peter (1963). "Stochastic-processes in several dimensions". *Bulletin of the International Statistical Institute* 40.2, pp. 974–994.
- Stein, Michael L (1999). *Interpolation of spatial data: Some theory for kriging.* Springer Science & Business Media.
- Fuglstad, Geir-Arne et al. (2015). "Exploring a new class of non-stationary spatial Gaussian random fields with varying local anisotropy". *Statistica Sinica*, pp. 115–133.
- Llamazares-Elias, Liam, Jonas Latz, and Finn Lindgren (2024). "A parameterization of anisotropic Gaussian fields with penalized complexity priors". *arXiv preprint arXiv:2409.02331.*
- Simpson, Daniel P. et al. (2014). "Penalising Model Component Complexity: A Principled, Practical Approach to Constructing Priors". *Statistical Science* 32, pp. 1–28.
- Lindgren, Finn, Håvard Rue, and Johan Lindström (2011). "An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach". *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 73.
- Lindgren, Finn, David Bolin, and Håvard Rue (2022). "The SPDE approach for Gaussian and non-Gaussian fields: 10 years and still running". *Spatial Statistics*, p. 100599.
- Lototsky, Sergey V, Boris L Rozovsky, et al. (2017). *Stochastic partial differential equations.* Springer.

References II

Matérn, Bertil (1960). "Spatial variation. Stochastic models and their application to some problems in forest surveys and other sampling investigations.".

Llamazares-Elias, Liam (Oct. 2023). *SPDEaniso*. Version 0.0.0.9000. URL: `https://github.com/inlabru-org/SPDEaniso`.