

# Part II: Bayesian Spatial Statistics using INLA and inlabru

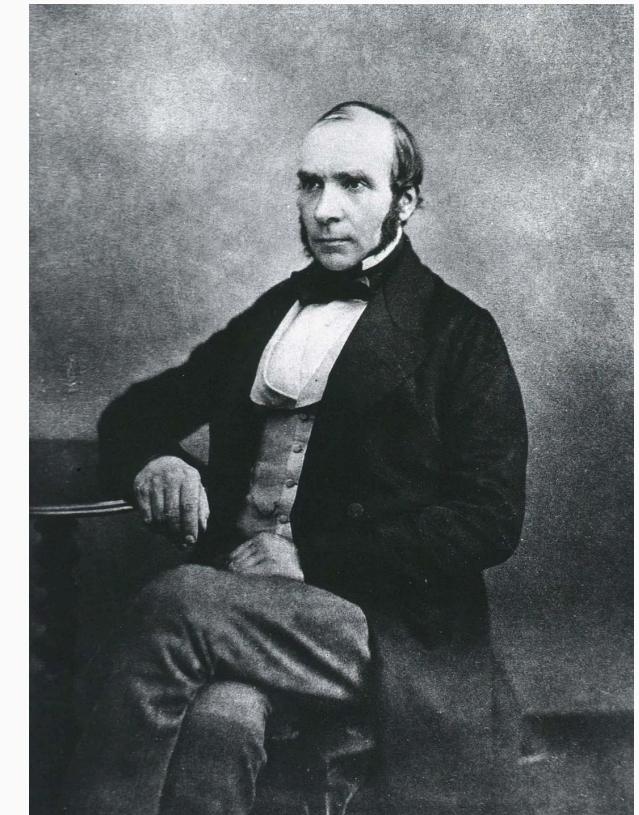
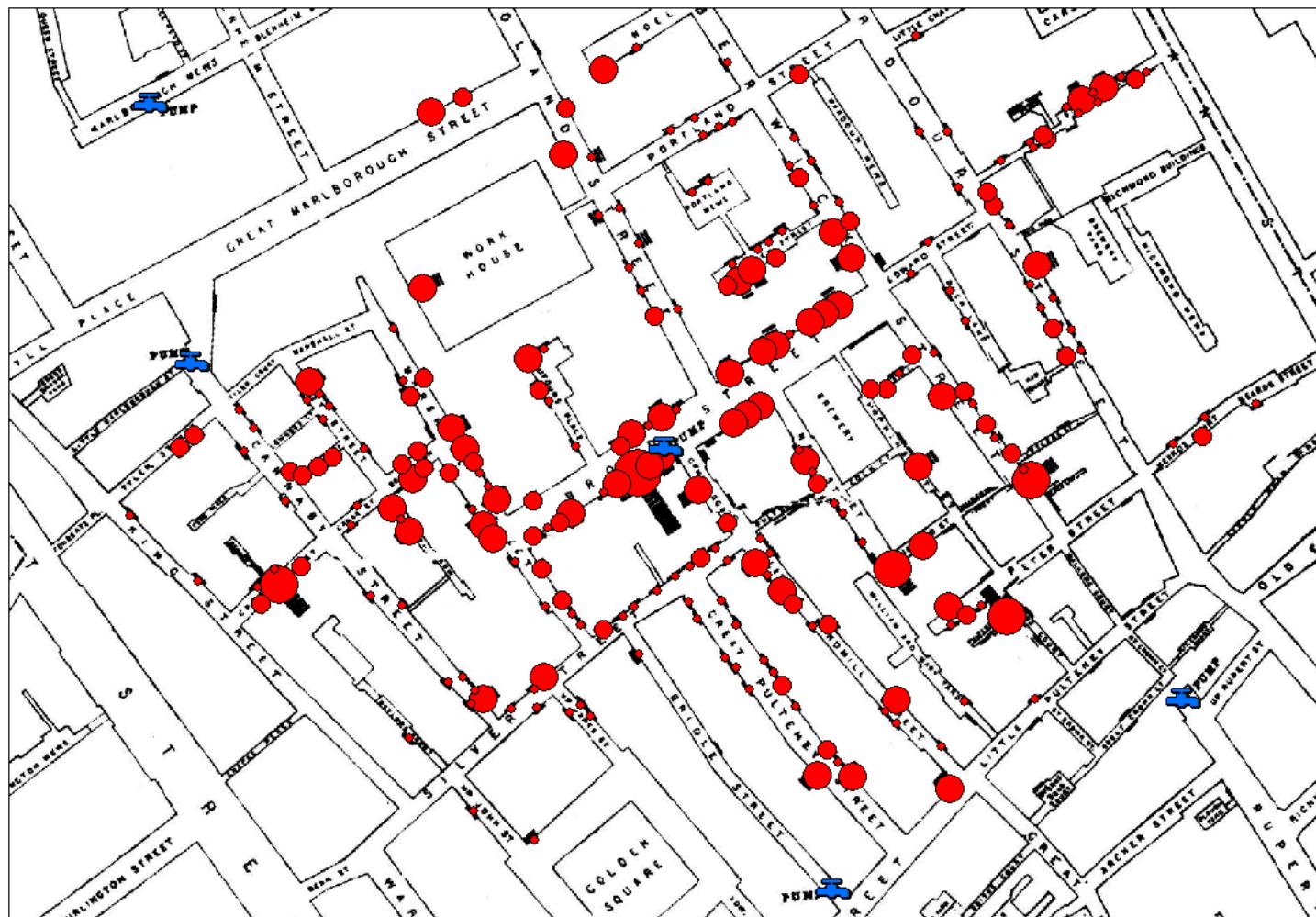
Joaquín Martínez-Minaya, November 12, 2025

VAlcencia BAyesian Research Group  
Statistical Modeling Ecology Group  
Grupo de Ingeniería Estadística Multivariante  
[jmarmin@eio.upv.es](mailto:jmarmin@eio.upv.es)



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

# John Snow's Cholera Map in London 1854'



# Outline

1. Spatial statistics. Types of spatial data
2. Disease mapping
3. Geostatistics
4. Penalized complexity priors (PC-priors)
5. References

# 1. Spatial statistics. Types of spatial data

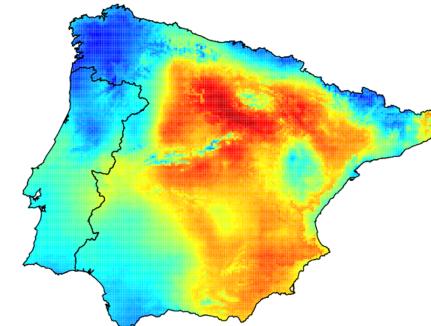
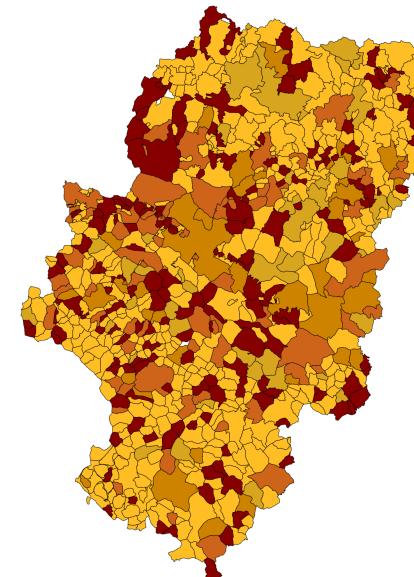


UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

# Spatial statistics. Types of spatial data

**Spatial statistics** is defined as the part of statistics which deal with spatial data and study spatial patterns.

- **Lattice or areal data:** observations are taken at a finite number of sites whose whole constitutes the entire study region (discrete space), e.g. number of sick people by provinces.
- **Point pattern:** the interest is study the process which generates the points. e.g. distribution of trees in a mountain.
- **Geostatistical data:** consist of a collection of data in a fixed set locations over a continuous spatial field, e.g. amount of fish in the ocean or presence/absence of a plant in a country.



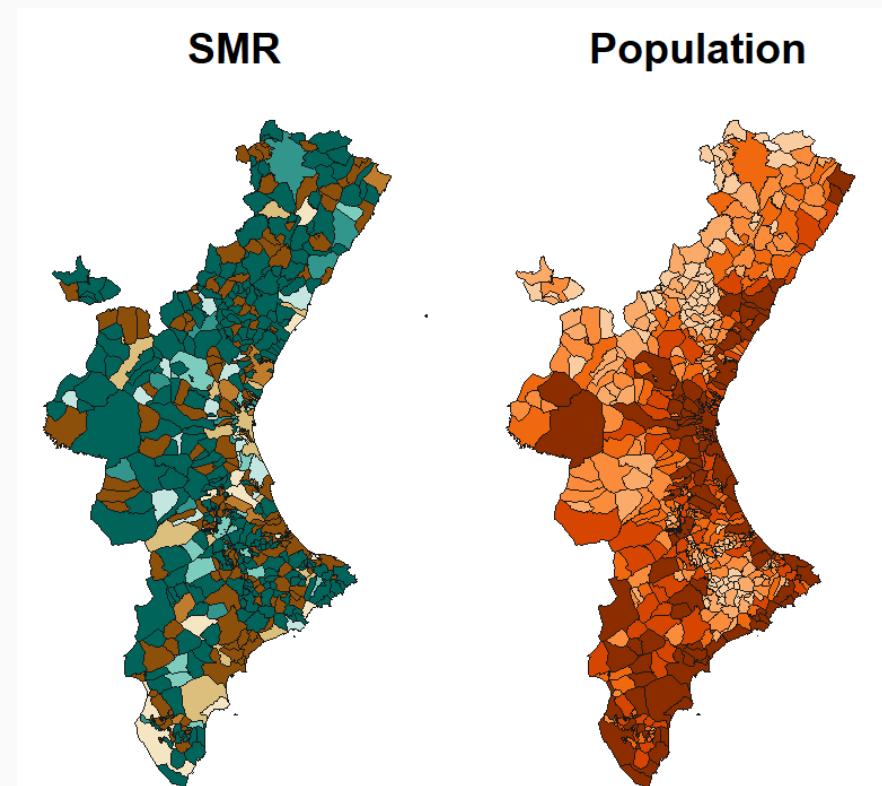
## 2. Disease mapping



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

# Oral Cancer mortality in Valencian Region

- In this analysis, we study **oral cancer mortality in the municipalities of the Valencian Region** using a disease mapping model. The aim is to understand the spatial distribution of risks and identify high-risk areas while accounting for variability due to population size and random noise. The variables are:
- **Obs**: the number of observed deaths from oral cancer in the study period.
- **Exp**: the number of expected deaths, based on population size and age-specific rates.
- **SMR**: the standardized mortality ratio, calculated as  $\text{SMR} = \frac{\text{Obs}}{\text{Exp}} \cdot 100$ 
  - **SMR = 100**: Risk is equivalent to the standard population.
  - **SMR > 100**: **excess risk**.
  - **SMR < 100**: **reduced risk**.



# The model

- A conditional independent **Poisson** likelihood function is assumed:

$$y_i \sim \text{Poisson}(\lambda_i), \quad \lambda_i = E_i \rho_i, \quad \log(\rho_i) = \eta_i, \quad i = 1, \dots, 32$$

- We assume that  $\eta_i = \beta_0 + u_i + v_i$ , being  $\mathbf{u}$  the **independent random effect** and  $\mathbf{v}$  the **spatially structured random effect**:

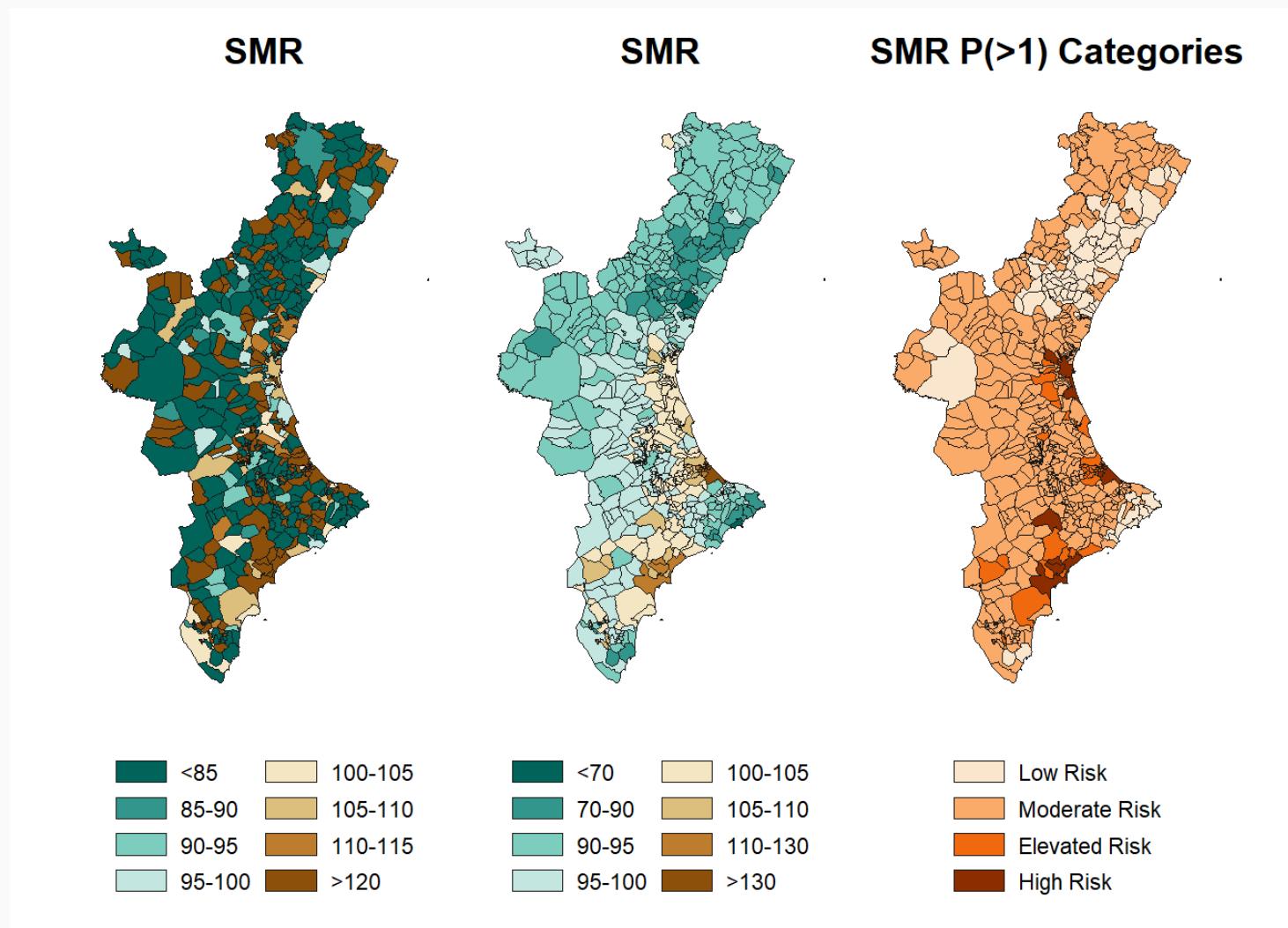
$$u_i \sim \mathcal{N}(0, \tau_{\mathbf{u}}^{-1}), \quad v_i | \mathbf{v}_{-i} \sim \mathcal{N}\left(\frac{1}{n_i} \sum_{j \sim i} v_j, \frac{1}{n_i \tau_{\mathbf{v}}}\right).$$

In this case  $\boldsymbol{\theta} = (v_1, \dots, v_{32}, u_1, \dots, u_{32})$ , and  $\boldsymbol{\theta} | \boldsymbol{\psi}$  is Gaussian distributed.

- **Hyperpriors** for the standard deviation parameters  $\sigma_u$  and  $\sigma_v$  follow uniform priors:

$$\sigma_u, \sigma_v \sim \text{Uniform}(0, \infty)$$

# Predicting Risk in Valencia Region



# 3. Geostatistics

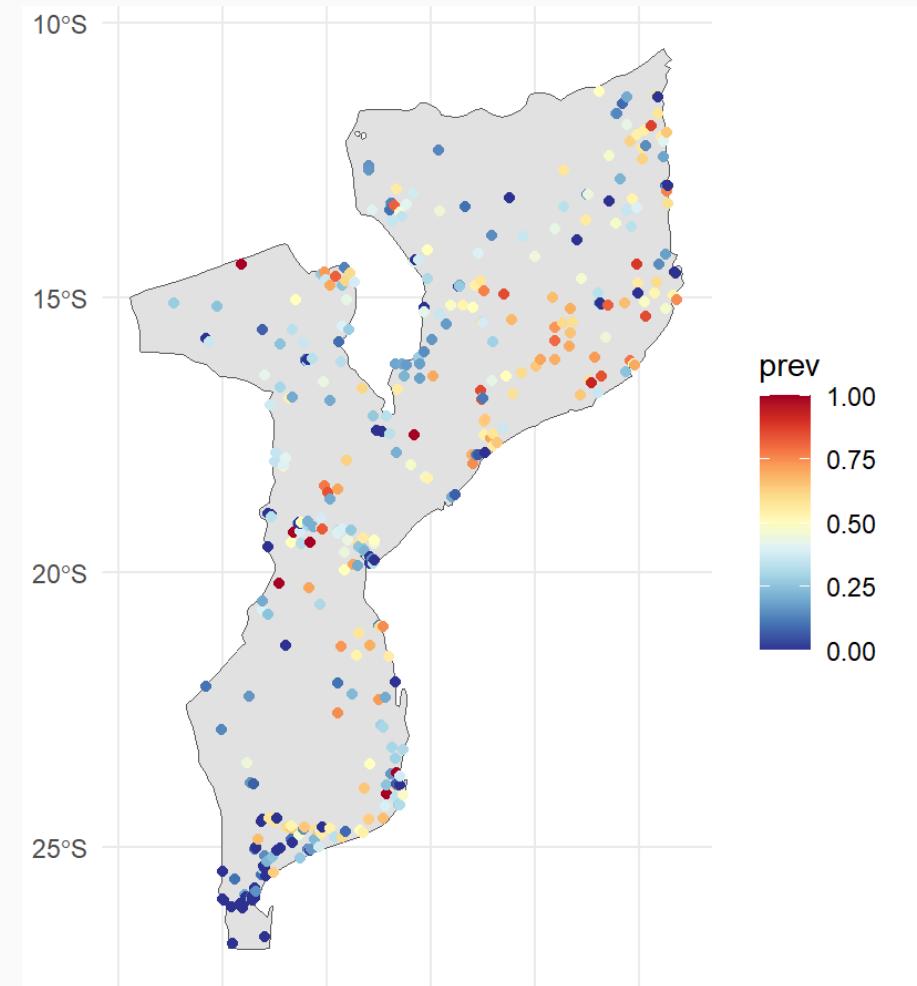


UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

# Continuous spaces

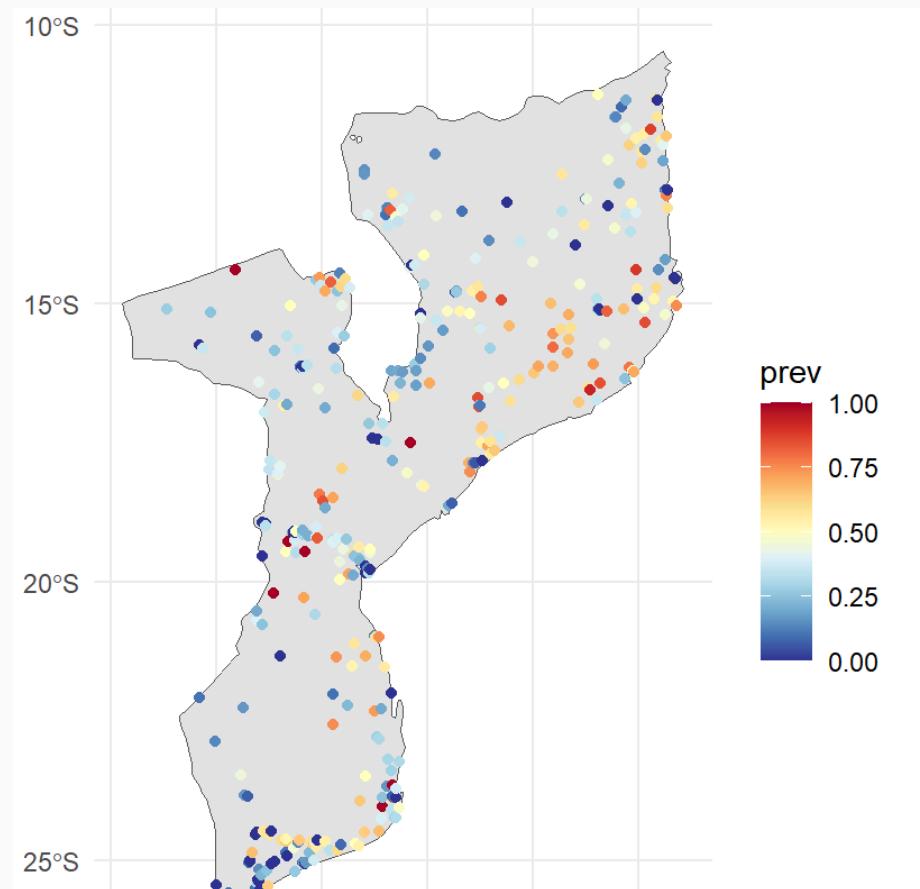
- Sometimes, the assumption that the observations have been collected over **discrete time** points have to be removed.
- The same happen in **space**.
  - If we are studying the **presence of a disease**, **pollution in an area** or the temperature of a country, the locations where the phenomenon of interest is measured are not frequently allocated in a lattice.
- Then, we are dealing with **continuous spaces** in 1D and 2D

Malaria Prevalence in Mozambique



# Malaria Prevalence in Mozambique

- This analysis studies **malaria prevalence in Mozambique** using a spatial Bayesian model. The goal is to predict malaria risk and evaluate the effects of environmental and demographic covariates.
- **Examined**: the number of individuals examined for malaria.
- **Positive**: the number of individuals testing positive for malaria.
- **Covariates**:
  - **Altitude**: Elevation of the study location (in meters).
  - **Temperature**: Average temperature (in °C).
  - **Proximity to water bodies**: Distance to the nearest water source (in kilometers).



# Geostatistics. Basis

- Geostatistical models assume that the observations are correlated.
- They are based on the following principle

**Everything is related to everything else, but near things are more related than distant things**

- So, two close locations tend to **co-vary** more than those far from each other.

# Let's be a bit more formal

- A random spatial effect  $w(s)$  at a location  $s \in \mathcal{D}$  can be considered as a **stochastic process** characterized by a spatial index  $s$  which varies continuously in the fixed domain  $\mathcal{D}$ , where  $\mathcal{D}$  is a fixed subset of  $r$ -dimensional Euclidean space.
- The spatial process  $w(s)$  is Gaussian if for any  $n \geq 1$  and any set of sites  $s = \{s_1, \dots, s_n\}$ ,  $w = \{w(s_1), \dots, w(s_n)\}$  has a multivariate normal distribution with mean  $\mu = E(w(s))$  and a structured covariance matrix  $\Sigma$ . Usually  $\mu$  is assumed to be  $\mathbf{0}$ . In the literature, this process is widely known as a **Gaussian field (GF)**.
- The key issue in spatial statistics is the covariance function  $\mathcal{C}$ , which determines the covariance between random variables in two different points. If  $s_i$  and  $s_j$  are two locations in space, then the **covariance function** is defined as

$$\mathcal{C}(w(s_i), w(s_j)) = Cov(w(s_i), w(s_j))$$

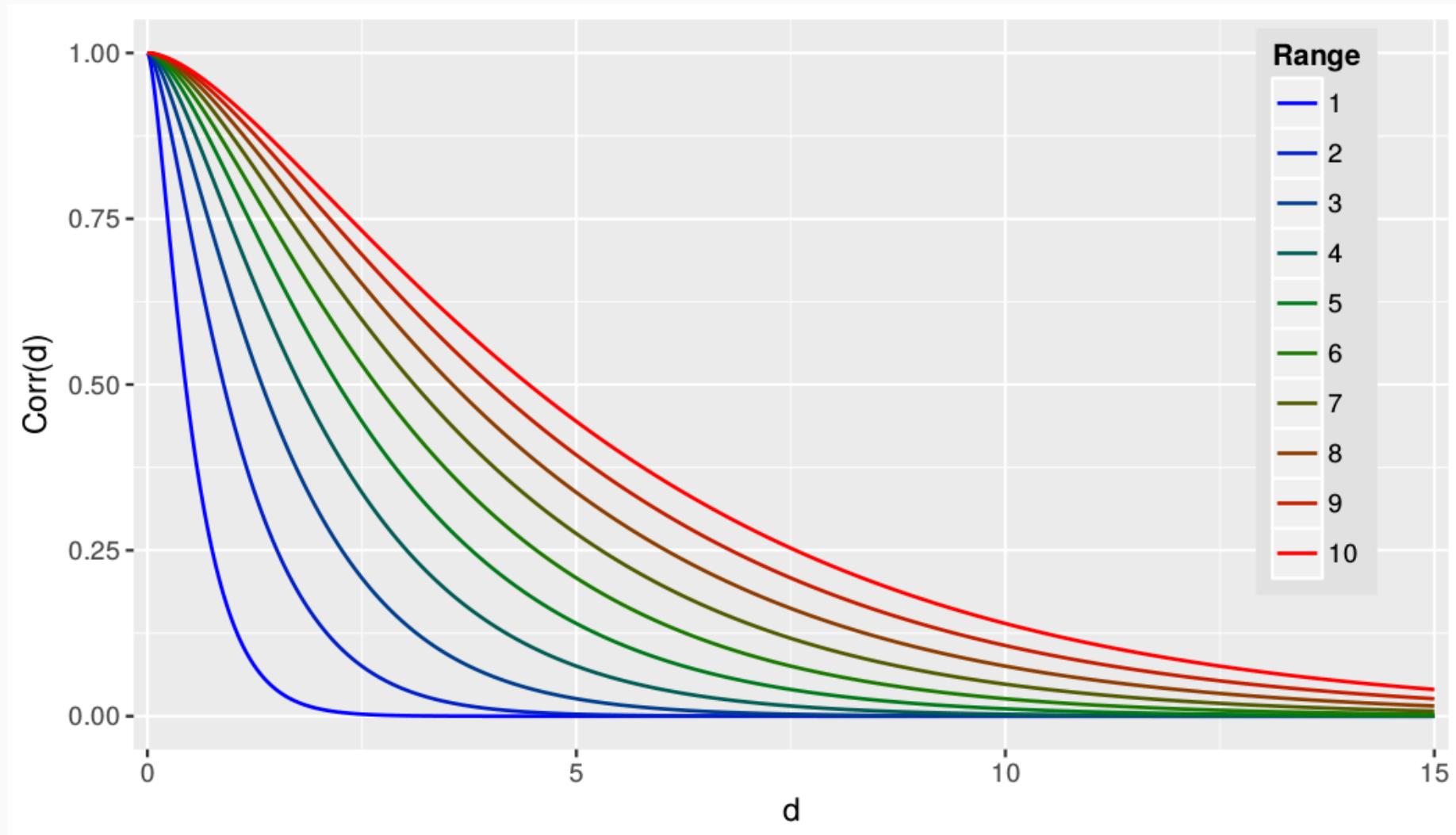
- It defines the covariance matrix  $\Sigma$  of the GF. Each element of the matrix  $\Sigma_{ij}$  is defined as:

$$\Sigma_{ij} = \mathcal{C}(w(s_i), w(s_j))$$

- **Stationarity.** We say that the GF is second-order stationary if  $\mu(s) = \mu$  and  $Cov(w(s), w(s + h)) = C(h)$  for all  $h \in \mathcal{R}$  such that  $s$  and  $s + h$  lie within  $\mathcal{D}$ . The covariance function in two different locations depends on the distance vector between these two locations.
  - An example could be the spread of a pathogen in plants. If there is a road close to the crop, maybe this pathogen could spread faster along the road in cars or trucks than in the crop, it would depend on the direction.
- **Isotropy.** We say that the GF is isotropic if the covariance function depends only on the Euclidean distance between points, i.e.,  $Cov(w(s), w(s + h)) = C(||h||)$ .
  - For instance, if we think again in the spread of a pathogen in a crop, it would mean that the spread does not depend on the direction, just on the distance.
- **Matérn correlation** function is very common.

$$C(||h||) = \sigma_w^2 \left( \frac{\sqrt{8}}{\phi} ||h|| \right) K_1 \left( \frac{\sqrt{8}}{\phi} ||h|| \right)$$

# Matérn correlation function



# Geostatistics in the context of LGMs

## Likelihood

- A conditional independent **Binomial likelihood** function is assumed:

$$y_i \mid \pi_i \sim \text{Binomial}(n_i, \pi_i), \eta_i = \text{logit}(\pi_i) = \beta_0 + \beta_1 \text{Temp} + w_i, i = 1, \dots, 447$$

## Latent Gaussian field

$$\mathbf{w} \sim \mathcal{N}(0, \Sigma(\sigma_w, \phi)), \beta_j \sim \mathcal{N}(0, \tau = 0.001)$$

$\boldsymbol{\theta} = (\beta_0, \beta_1, w_1, \dots, w_{447})$ , and  $\boldsymbol{\theta} \mid \psi$  is Gaussian distributed.

- $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \Sigma(\sigma_w, \phi))$ , i.e., the spatial effect is assumed to be a **continuous Gaussian field (GF)** with Matérn covariance structure, where:
- $\Sigma(\sigma_w, \phi)$  is a **covariance matrix** depending on the distance between locations,  $\sigma_w$  is the **variance** of the spatial effect, and  $\phi$  is the **range** of the spatial effect.

**Hyperparameters**  $\psi = (\sigma_w, \phi)$

**Problem: INLA can not fit continuous GFs**

**Solution: approximate the continuous GFs using the Stochastic Partial Differential Equation approach (SPDE)**

# SPDE. The key idea

## Likelihood

$$y_i \mid \pi_i \sim \text{Ber}(\pi_i)$$

$$\text{logit}(\pi_i) = \beta_0 + \beta_1 \text{Temp} + w_i$$

## Latent Gaussian field

$$\boldsymbol{\beta} \sim \mathcal{N}(\mathbf{0}, \tau = 0.0001)$$

$$\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \Sigma(\sigma_w, \phi))$$

## Hyperparameters

$$p(\sigma_w, \phi)$$

## Likelihood

$$y_i \mid \pi_i \sim \text{Ber}(\pi_i)$$

$$\text{logit}(\pi_i) = \beta_0 + \beta_1 \text{Temp} + w_i$$

## Latent Gaussian field

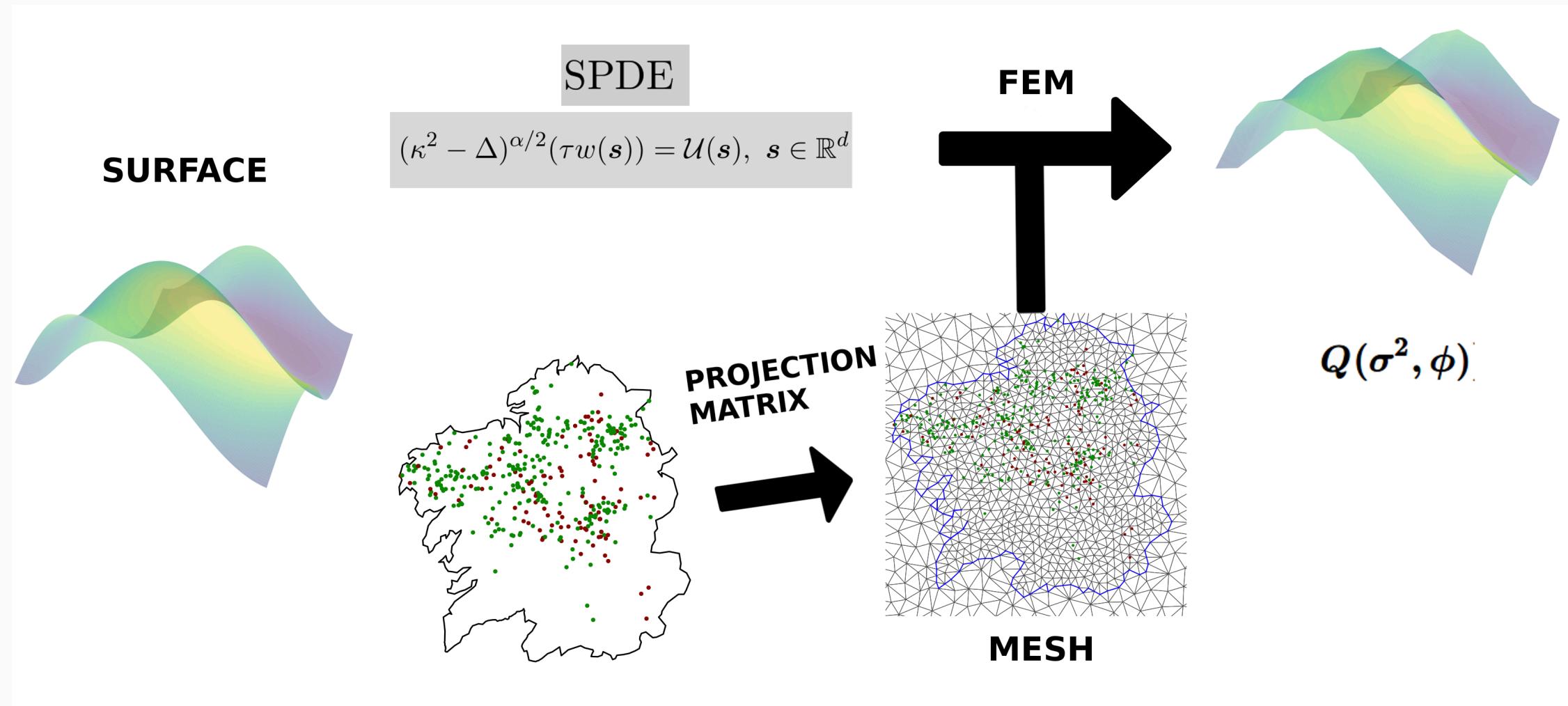
$$\boldsymbol{\beta} \sim \mathcal{N}(\mathbf{0}, \tau = 0.0001)$$

$$\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}^{-1}(\sigma_w, \phi))$$

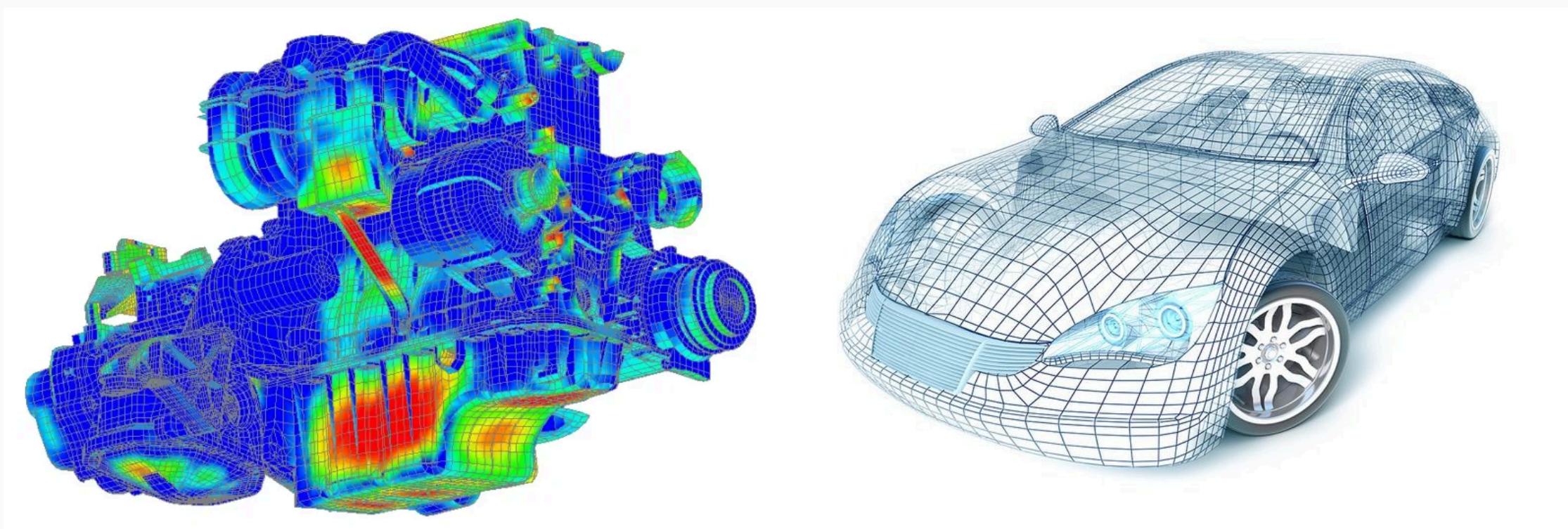
## Hyperparameters

$$p(\sigma_w, \phi)$$

# SPDE. How is the approximation conducted?

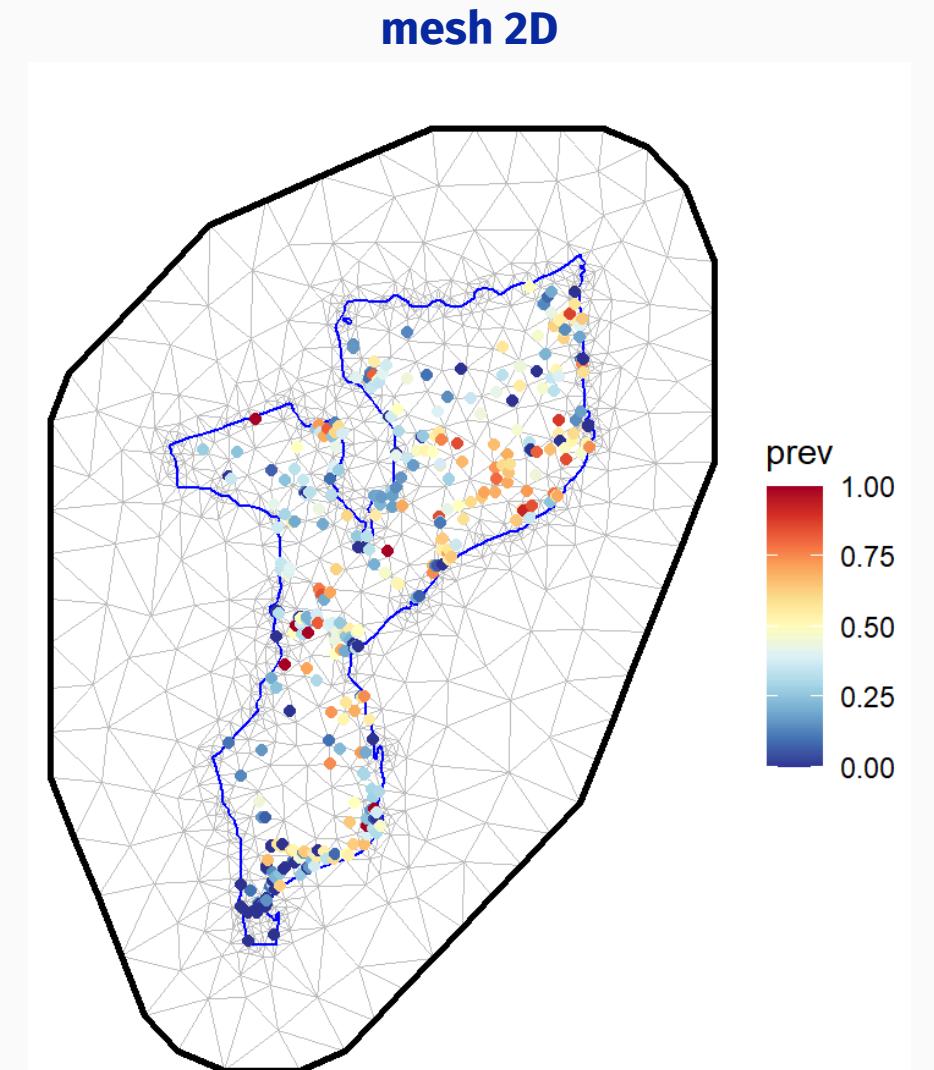


# FEM. General idea



# The mesh

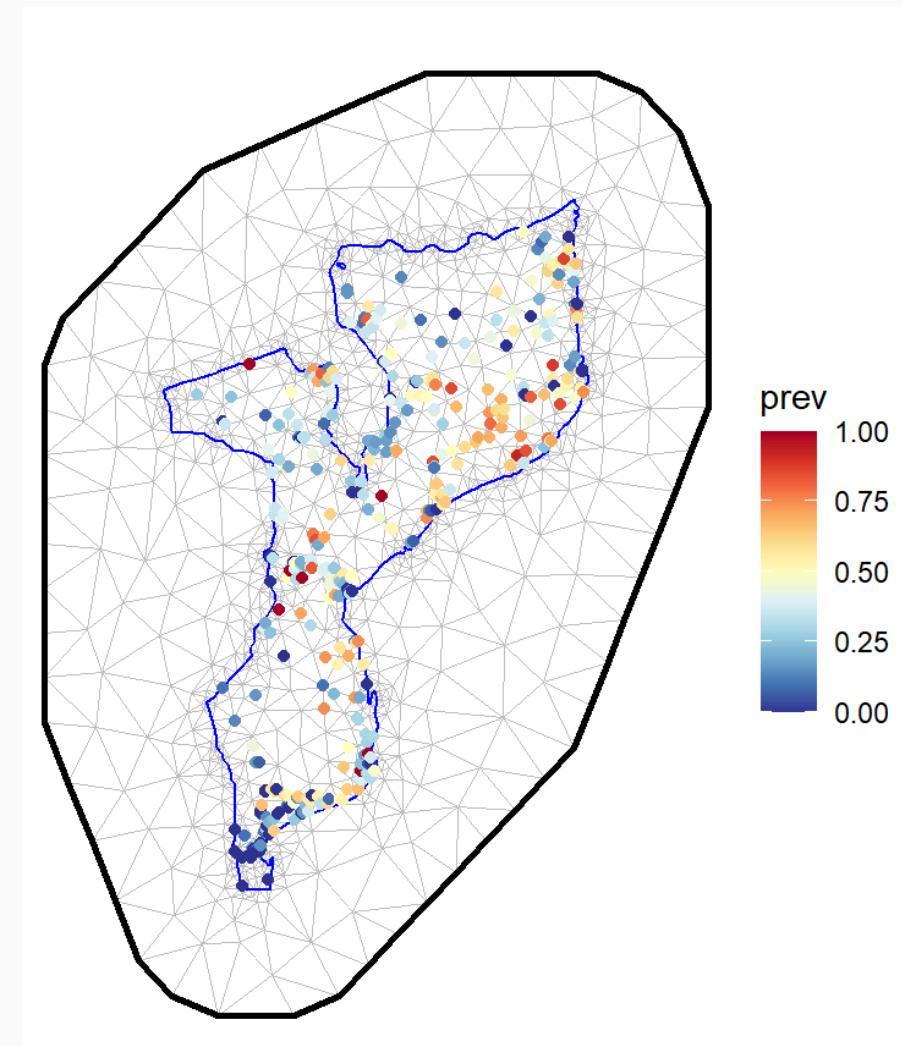
- The mesh is created using a **Delaunay triangulation**.
- The model is evaluated at the **intersecting nodes**.
- We want triangles with **similar size** over the study area.  
**Sample locations in mesh nodes not necessary.**
- The greater is the number of triangles, the bigger is the computational cost.
- Too big triangles imply that part of the small scale variability could be masked in the intrinsic variance of the process under study  **$y$** .



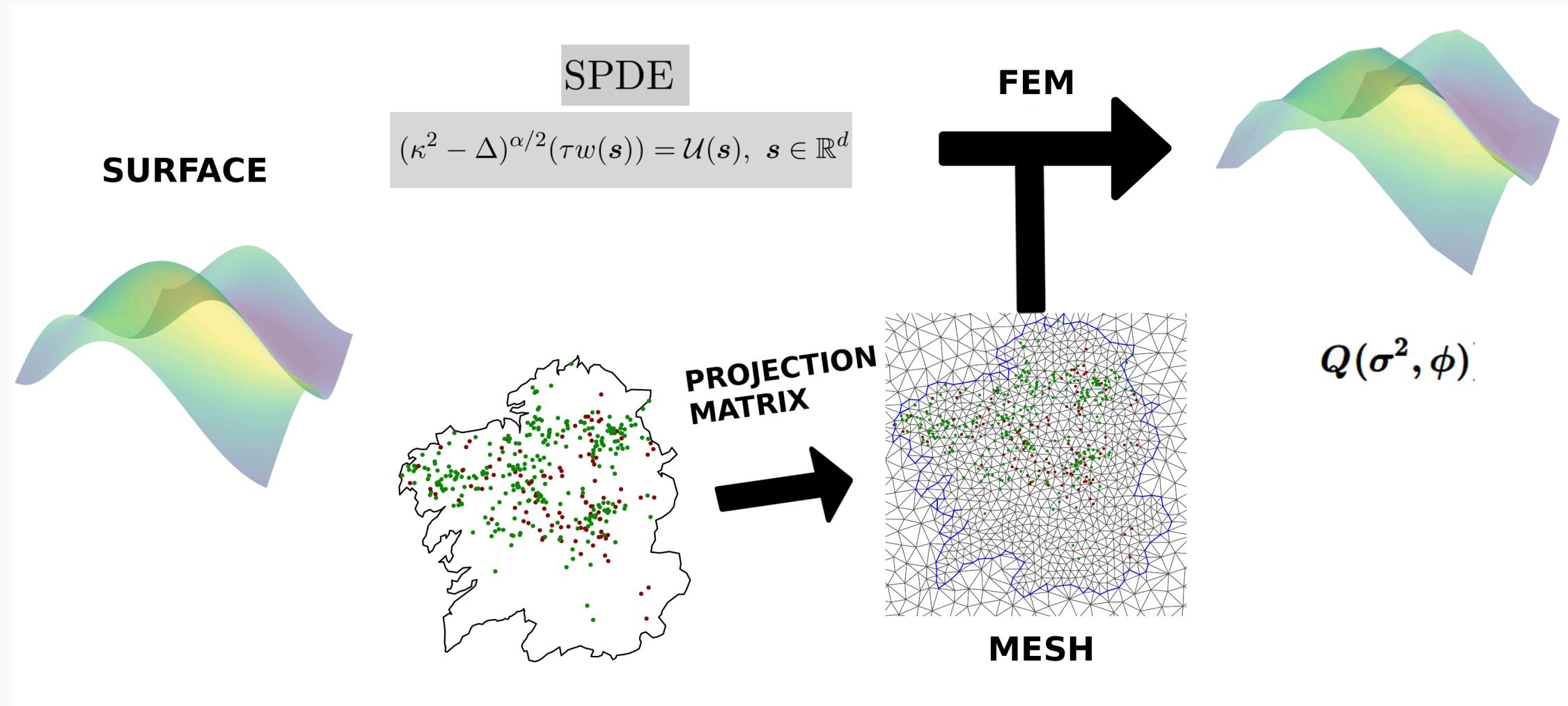
# Mesh in R-INLA

```
fm_mesh_2d_inla(boundary = mozambique_sf,  
                 max.edge = c(1, 2),  
                 cutoff = 0.05,  
                 offset = c(1, 3),  
                 crs      = 4326)
```

- How to do a bad mesh. By Haakon Bakka



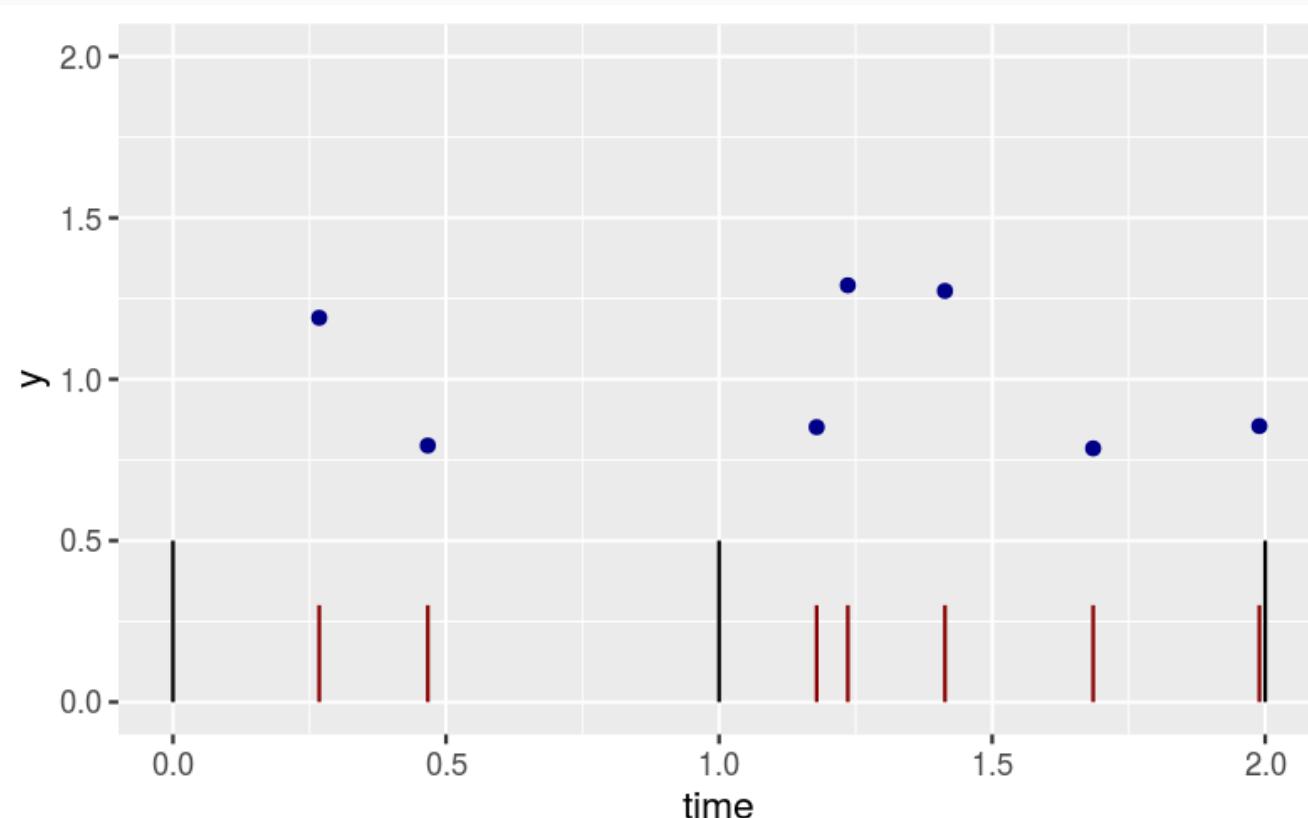
# SPDE. How the approximation is conducted?



# Connection between mesh and data: 1D

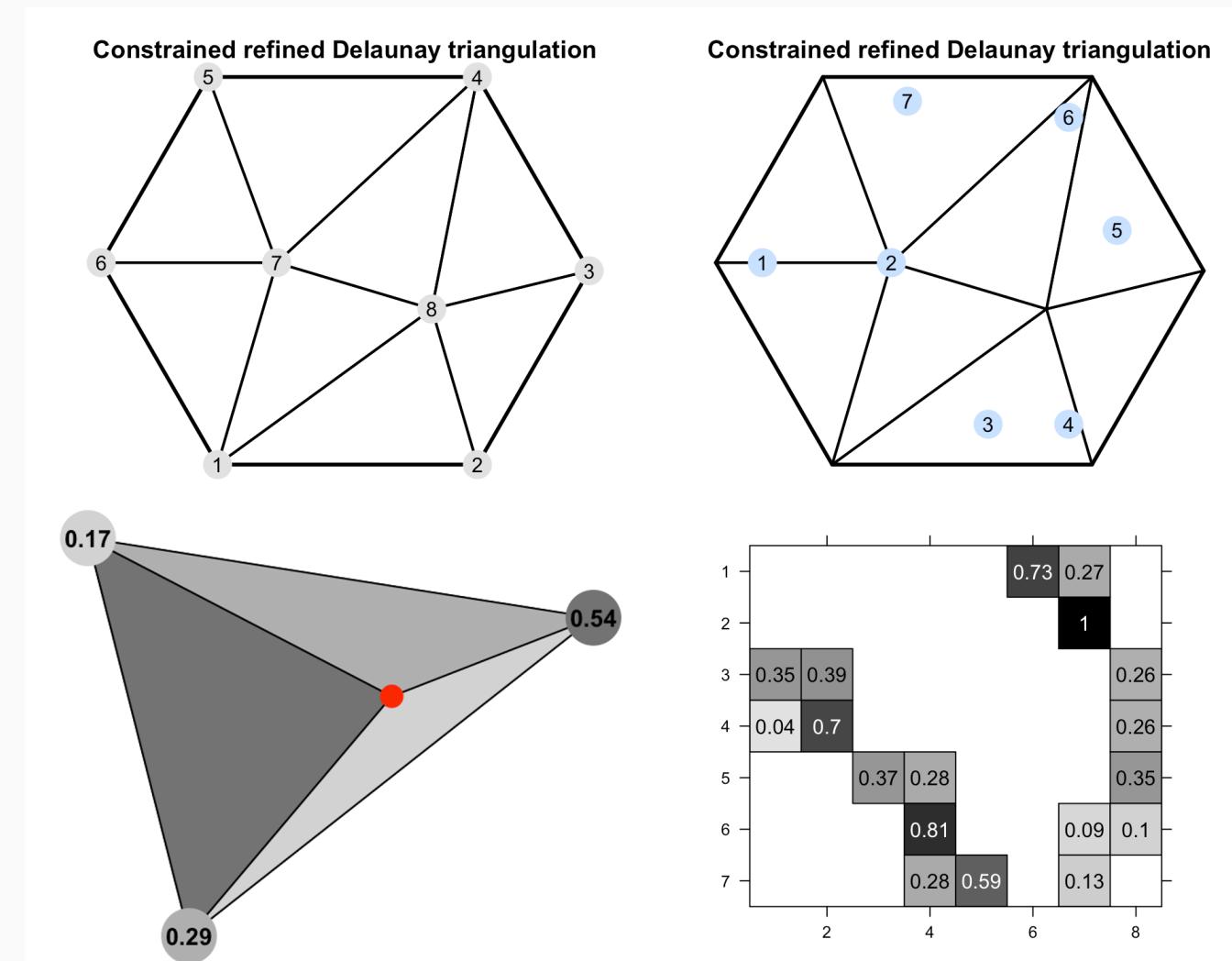
```
mesh ← inla.mesh.2d(boundary = ... ,  
                     max.edge = ... )
```

```
inla.spde.make.A(mesh = ... ,  
                  loc = ... )
```

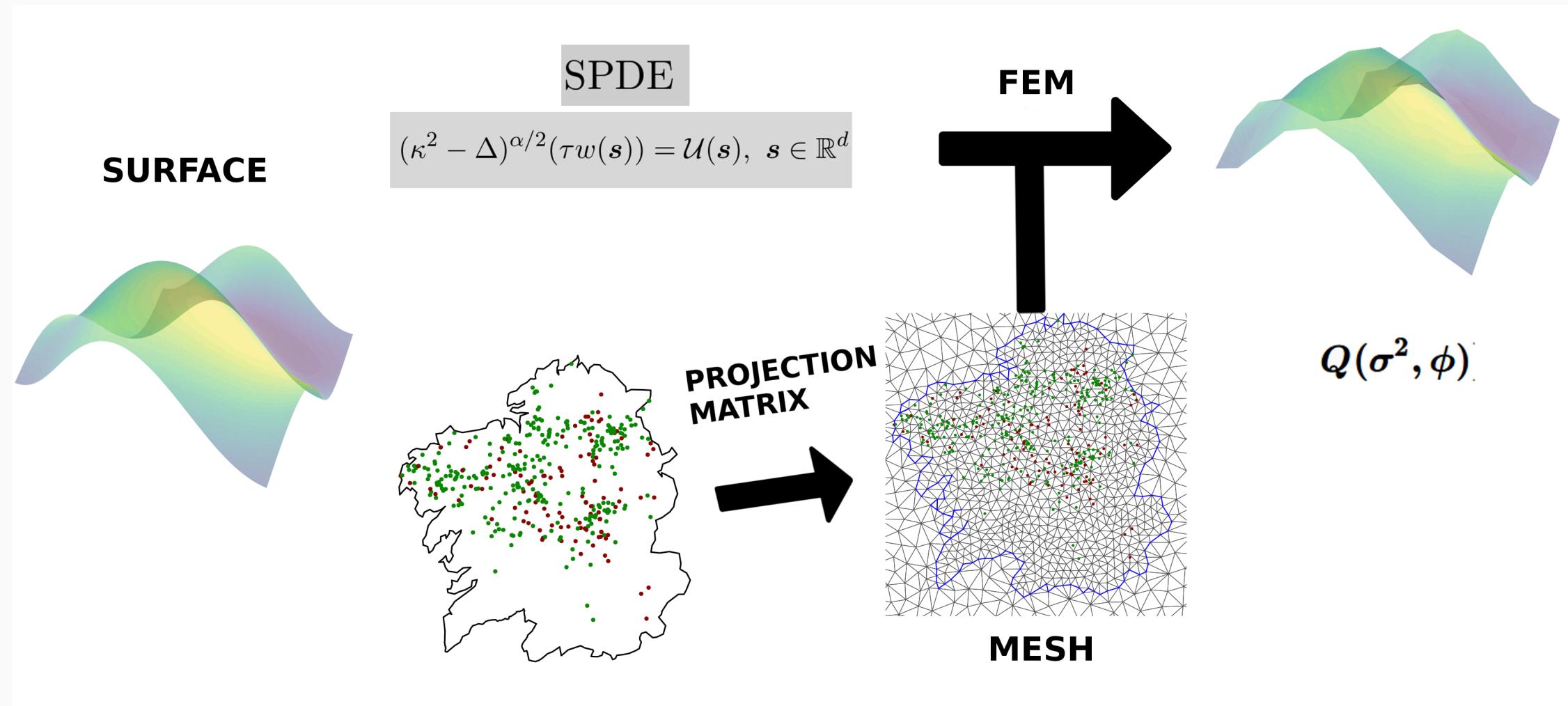


<b>knot 1</b>	<b>knot 2</b>	<b>knot 3</b>
0.7456	0.2544	0.0000
0.5567	0.4433	0.0000
0.0000	0.8802	0.1198
0.0000	0.8261	0.1739
0.0000	0.6571	0.3429
0.0000	0.3993	0.6007
0.0000	0.1101	0.8899

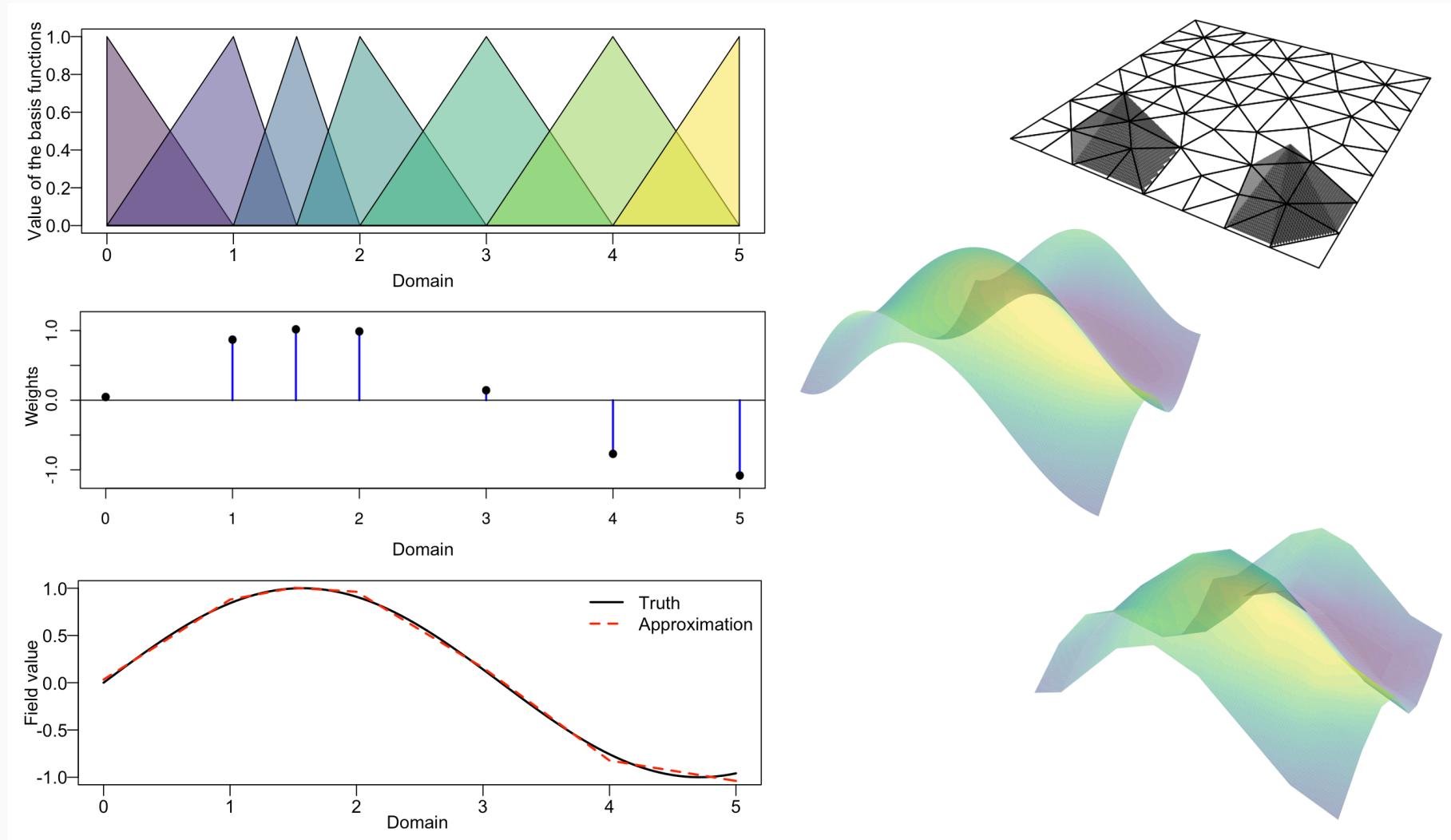
# Connection between mesh and data: 2D



# SPDE. How the approximation is conducted?



# Solving SPDE using FEM: 1D and 2D

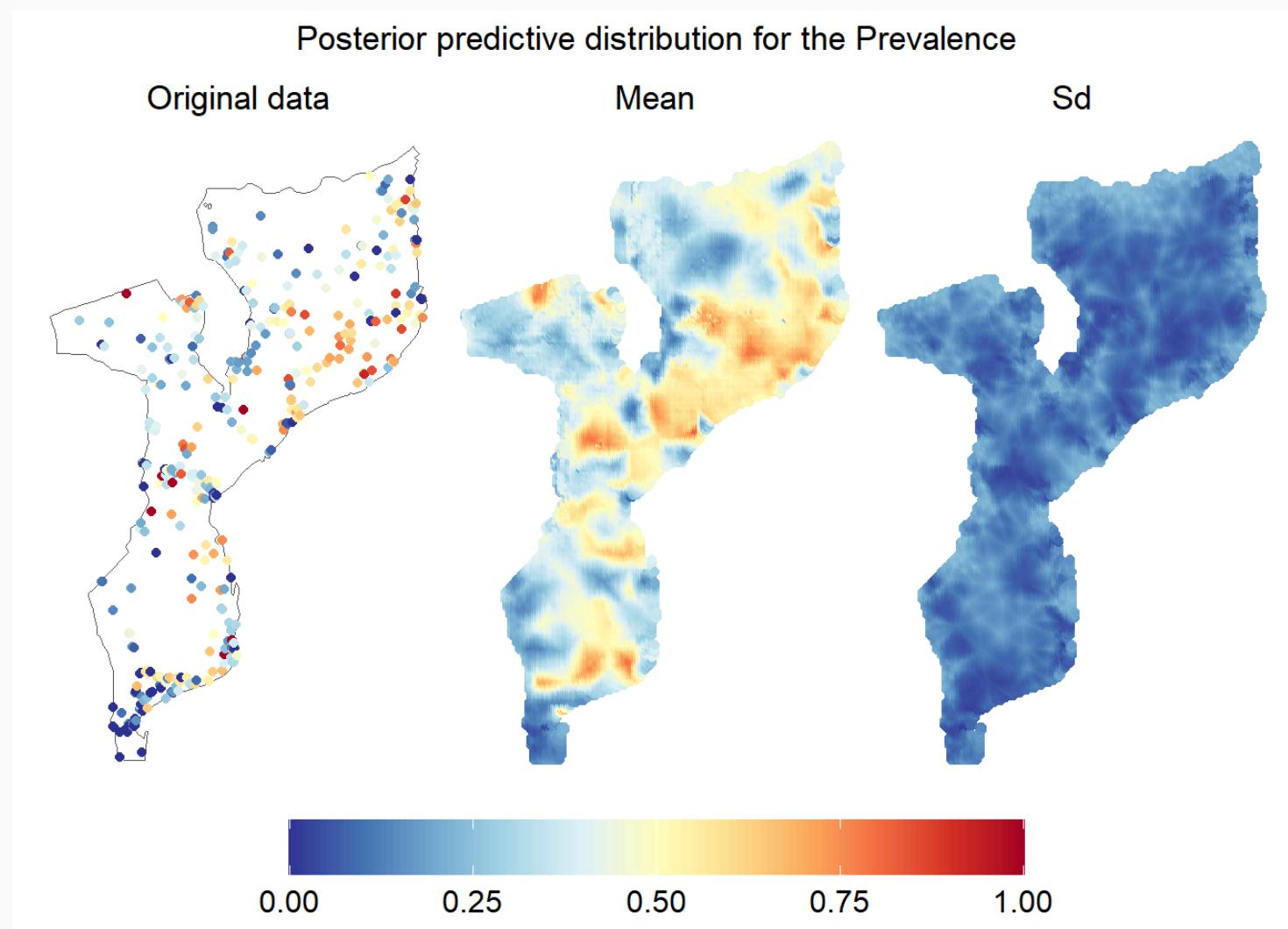


# Defining SPDE

- In order to **define the SPDE**, the following command have to be used:

```
spde ← inla.spde2.pcmatern(  
  # Mesh  
  mesh = mesh,  
  #  $P(practic.range < range0) = 0.5$   
  prior.range = c(range0, 0.5),  
  #  $P(sigma > 1) = 0.01$   
  prior.sigma = c(1, 0.01))
```

# Malaria Prevalence in Mozambique



# 4. Penalized complexity priors (PC-priors)



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

# Penalizing departure from the base model

- Simpson et al. (2017) propose priors that penalize departure from a base model and for this reason they are called **Penalized Complexity (PC) priors**.
- The prior favors the base model unless evidence is provided against it, following the principle of parsimony.
- Distance from the base model is measured using the **Kullback-Leibler** distance, and penalization from the base model is done at a **constant rate on the distance**.
- Finally, the PC prior is defined using **probability statements** on the model parameters in the appropriate scale.

# Hyperpriors for the standard deviation in an iid

- The **PC-prior for the precision**  $\tau$  has density:

$$p(\tau) = \frac{\lambda}{2} \tau^{-3/2} \exp(-\lambda \tau^{-1/2}), \quad \tau > 0,$$

where

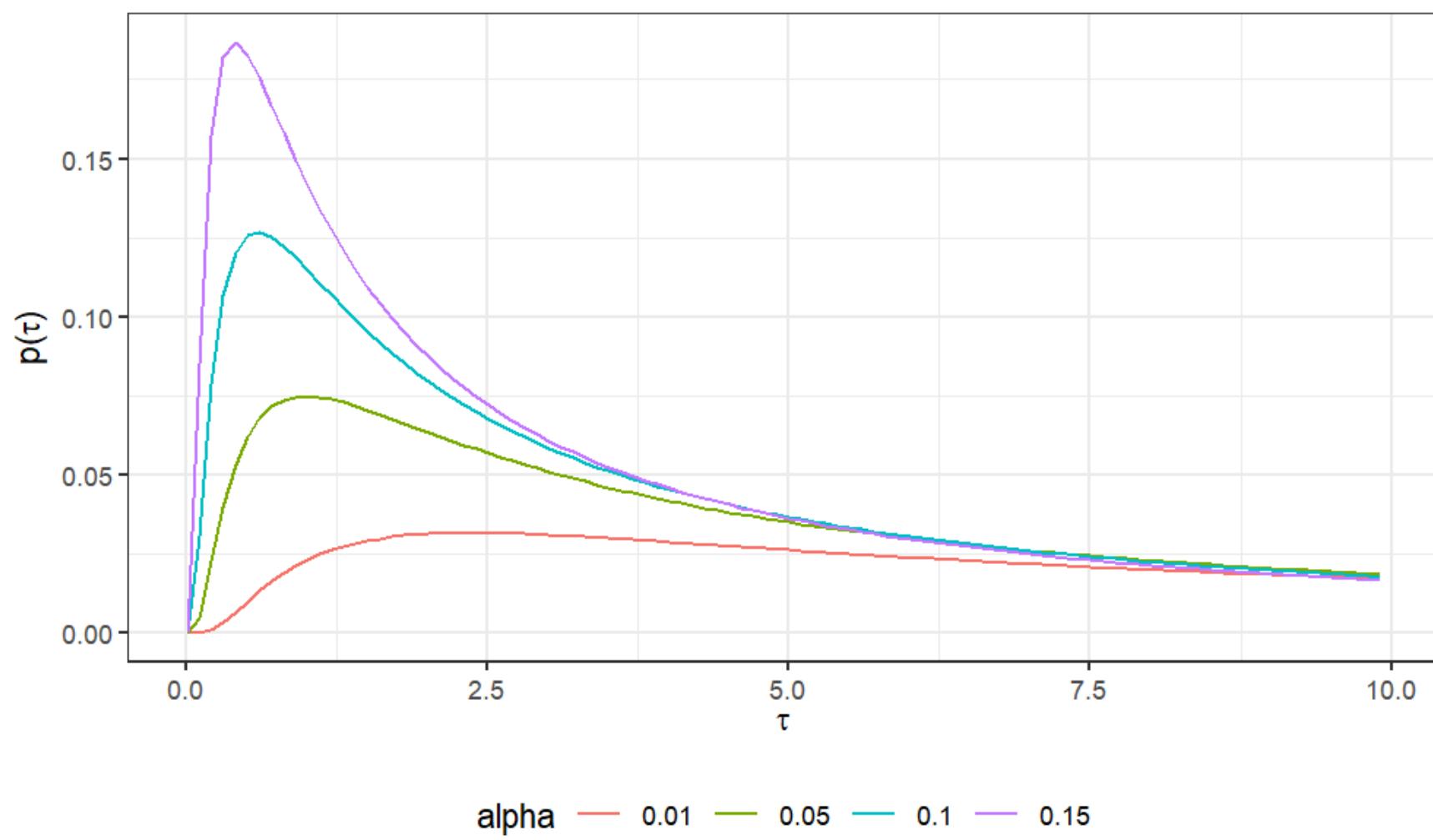
$$\lambda = -\frac{\ln(\alpha)}{u},$$

and  $(u, \alpha)$  are the parameters to this prior. The interpretation of  $(u, \alpha)$  is that:

$$\text{Prob}(\sigma > u) = \alpha, \quad u > 0, \quad 0 < \alpha < 1.$$

- Functions `inla.pc.{d,p,q,r}.prec` allow us to **deal with this priors**.
- If we want to plot the prior in terms of the **standard deviation**  $\sigma$ , remember that using function `inla.tmarginal` we can go from the  $\tau$  parameter to  $\sigma$  parameter.

# Hyperpriors for the standard deviation in an iid.



# Spatial effect: priors

- The PC-prior for the **range** is defined in terms of  $\phi_0$  and  $p_1$  so that

$$Prob(\phi < \phi_0) = p_1$$

- The PC-prior for the **standard deviation** is defined in terms of  $\sigma_0$  and  $p_2$  so that

$$Prob(\sigma_w > \sigma_0) = p_2$$

- In order to define the SPDE using PC-priors, the following command have to be used:

```
spde ← inla.spde2.pcmatern(  
mesh = ... ,  
prior.range = c(phi0, p1),  
prior.sigma = c(sigma0, p2))
```

## 5. References



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

# This material has been constructed based on:

- Moraga, P., Dean, C., Inoue, J., Morawiecki, P., Noureen, S. R., & Wang, F. (2021). Bayesian spatial modelling of geostatistical data using INLA and SPDE methods: A case study predicting malaria risk in Mozambique. *Spatial and Spatio-temporal Epidemiology*, 39, 100440.
- Blangiardo, M., & Cameletti, M. (2015). Spatial and spatio-temporal Bayesian models with R-INLA. John Wiley & Sons.
- Fuglstad, G. A., Simpson, D., Lindgren, F., & Rue, H. (2019). Constructing priors that penalize the complexity of Gaussian random fields. *Journal of the American Statistical Association*, 114(525), 445-452.
- INLA tutorials
- INLA book by Virgilio Gómez-Rúbio
- INLA book by Paula Moraga
- SPDE book by Krainski et al.

# Part II: Bayesian Spatial Statistics using INLA and inlabru

Joaquín Martínez-Minaya, November 12, 2025

VAlencia BAyesian Research Group

Statistical Modeling Ecology Group

Grupo de Ingeniería Estadística Multivariante

jmarmin@eio.upv.es



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA