

# Spatial modeling with INLA

King Abdullah University of  
Science and Technology



June 2024

The repo for this workshop is [here](#)  
<https://tinyurl.com/SYMSTAT24W2>



## QR code for repository



<https://tinyurl.com/SYMSTAT24W2>

Please run the Libraries.R file to ensure you have the relevant libraries for this workshop



# Outline I

- 1 Spatial domains
- 2 Latent Gaussian models
- 3 The INLA methodology
  - Introduction
  - Posterior inference with INLA
- 4 Areal modeling
  - Besag and BYM
  - Non-stationary Besag model
- 5 Geostatistics
  - Kriging
  - Matérn field and extensions
  - SPDE approach
  - Barrier model - non-stationary Matern field
- 6 Spatial modeling on different manifolds 3D...



# Outline II

- Dementia study
  
- 7 Point processes - LGCP
  - LGCP and SPDE
  
- 8 Extensions to non-stationarity and non-separable spatio-temporal models
  
- 9 Discussion

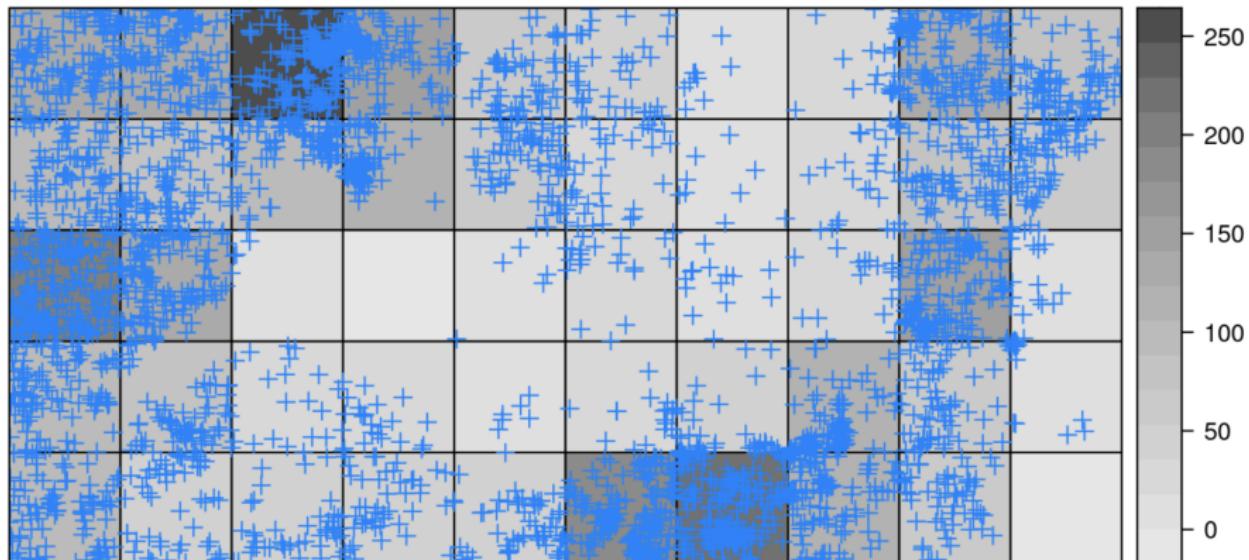


# BayesComp group at KAUST



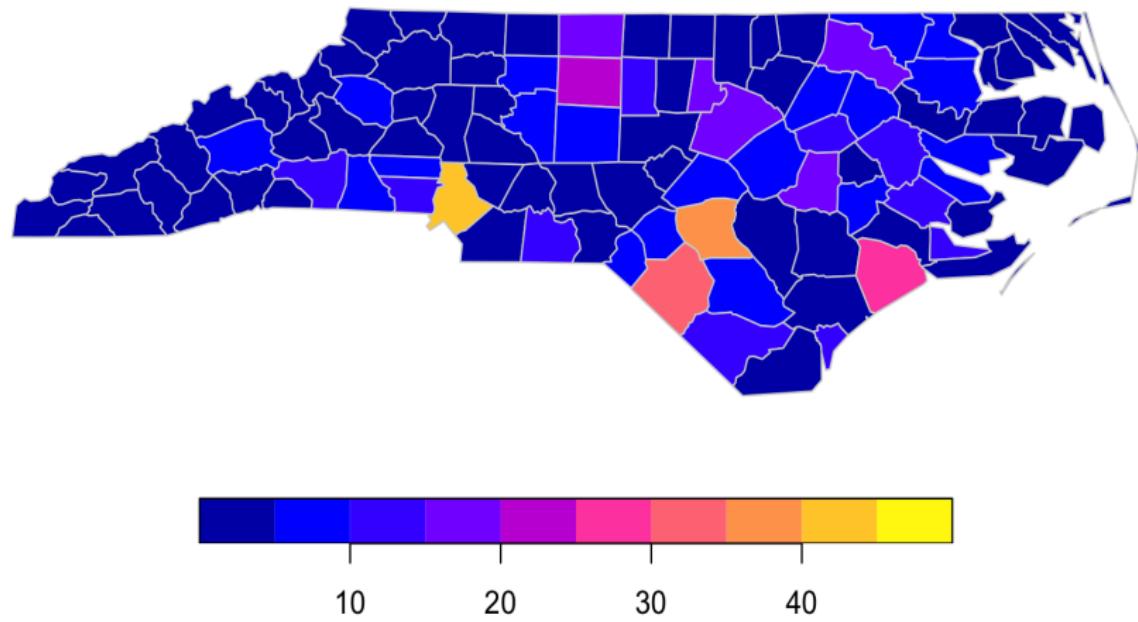


# Lattice type



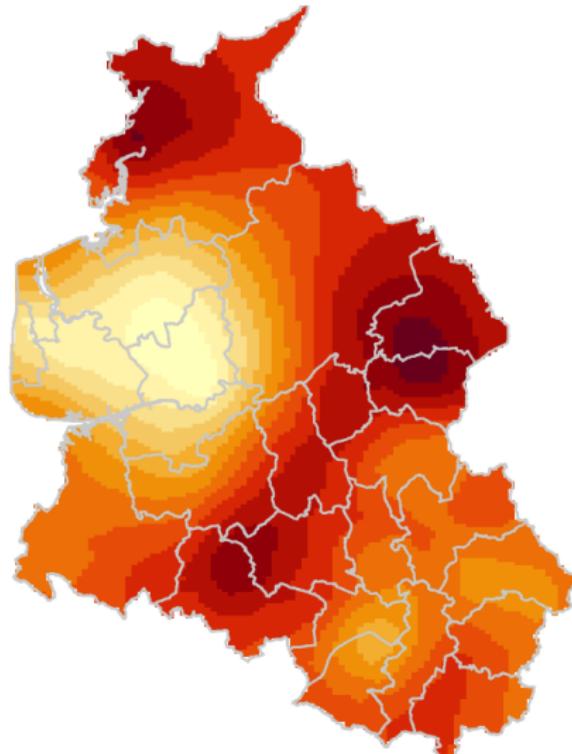


# Irregular lattice - areal data



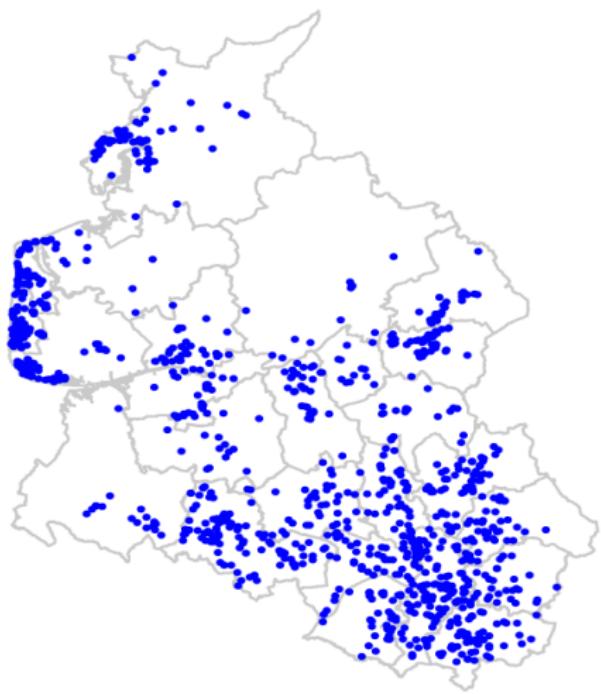


# Continuous domain - geostatistics





# Point process





# Model definition - GAMM

Suppose we have response data  $\mathbf{y}_{n \times 1}$  (conditionally independent) with density function  $\pi(y|\mathbf{X}, \boldsymbol{\theta})$  and link function  $h(\cdot)$ , that is linked to some covariates  $\mathbf{Z}$  through linear predictors

$$\boldsymbol{\eta}_n = \beta_0 + \mathbf{Z}_\beta \boldsymbol{\beta} + \sum f^k(\mathbf{Z}_f) = \mathbf{A}\mathbf{X}$$

The inferential aim is to estimate the latent field  $\mathbf{X}_m = \{\beta_0, \boldsymbol{\beta}, \mathbf{f}\}$ , and  $\boldsymbol{\theta}$ .



# GAMM → LGM

Assume

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

where  $\boldsymbol{Q}(\boldsymbol{\theta})$  is a sparse matrix ( $\boldsymbol{X}$  is a GMRF).

$p(\boldsymbol{X}, \boldsymbol{\theta}) = p(\boldsymbol{X}|\boldsymbol{\theta})p(\boldsymbol{\theta})$  and  $p(\boldsymbol{\theta})$  can be non-Gaussian.



# How common are sparse $Q(\theta)$ ?

Consider an AR(1) model..

$$x_i | \mathbf{x}_{-i} \sim N(\phi x_{i-1}, \tau^{-1}), x_1 \sim N(0, \tau^{-1})$$



# AR(1) example



# AR(1) example



# Website for R-INLA library

<https://www.r-inla.org/>



# Bayesian inference

Data  $\mathbf{y}$  (with covariates  $\mathbf{Z}$ ), depend on  $\mathbf{X}$  and  $\boldsymbol{\theta}$  such that,  $E[Y] = h(\mathbf{A}(\mathbf{Z})\mathbf{X})$ .

Bayes' theorem:

$$\begin{aligned} q(\mathbf{X}, \boldsymbol{\theta} | \mathbf{y}) &\propto L(\mathbf{y} | \mathbf{X}, \boldsymbol{\theta}) p(\mathbf{X}, \boldsymbol{\theta}) \\ \text{Posterior} &\propto \text{Likelihood} \times \text{Prior} \end{aligned}$$



# Computational aspects

- Analytical methods - conjugacy (pre-computer era)
- Approximate methods - Laplace (can be inaccurate)
- Exact methods - MCMC (very slow for complex models or large data)

Now, due to computing resources approximate methods are gaining popularity - INLA, VB, EP etc.

INLA - 2009 [Rue et al., 2009]

2021+ [Van Niekerk et al., 2023]

HPC [Gaedke-Merzhäuser et al., 2023]



# Why is INLA so accurate and so fast?

- LGM structure
- Sparse precision matrix
- Specialized matrix algebra for sparse matrices
- NEW: VB (low-rank) correction [van Niekerk and Rue, 2024]

Use precision matrix instead of covariance matrix → natural occurrence



# Posterior approximations by INLA

$$\begin{aligned}\pi(\boldsymbol{X}, \boldsymbol{\theta}, \boldsymbol{y}) &= \pi(\boldsymbol{\theta})\pi(\boldsymbol{X}|\boldsymbol{\theta}) \prod_{i=1}^n \pi(y_i | (\boldsymbol{AX})_i, \boldsymbol{\theta}) \\ \tilde{\pi}(\boldsymbol{\theta}|\boldsymbol{y}) &\propto \frac{\pi(\boldsymbol{X}, \boldsymbol{\theta}, \boldsymbol{y})}{\pi_G(\boldsymbol{X}|\boldsymbol{\theta}, \boldsymbol{y})} \Big|_{\boldsymbol{X}=\mu(\boldsymbol{\theta})} \\ \tilde{\pi}(\theta_j|\boldsymbol{y}) &= \int \tilde{\pi}(\boldsymbol{\theta}|\boldsymbol{y}) d\boldsymbol{\theta}_{-j} \\ \tilde{\pi}(\boldsymbol{X}_j|\boldsymbol{y}) &= \int \tilde{\pi}(\boldsymbol{X}_j|\boldsymbol{\theta}, \boldsymbol{y}) \tilde{\pi}(\boldsymbol{\theta}|\boldsymbol{y}) d\boldsymbol{\theta},\end{aligned}$$

$\tilde{\pi}(\boldsymbol{X}_j|\boldsymbol{\theta}, \boldsymbol{y})$  depends on the approximation used, for Gaussian it is straightforward for the Laplace approximation we do another Gaussian approximation to  $\tilde{\pi}(\boldsymbol{X}_{-j}|\boldsymbol{\theta}, \boldsymbol{y})$ .



# Modern INLA

The Gaussian approximation  $\pi_G(\mathbf{X}|\boldsymbol{\theta}, \mathbf{y})$  to  $\pi(\mathbf{X}|\boldsymbol{\theta}, \mathbf{y})$  is calculated from a second order expansion of the likelihood around the mode of  $\pi(\mathbf{X}|\boldsymbol{\theta}, \mathbf{y})$ ,  $\mu(\boldsymbol{\theta})$  as follows

$$\begin{aligned}\log(\pi(\mathbf{X}|\boldsymbol{\theta}, \mathbf{y})) &\propto -\frac{1}{2}\mathbf{X}^\top \mathbf{Q}(\boldsymbol{\theta})\mathbf{X} + \sum_{i=1}^n \left( b_i(\mathbf{AX})_i - \frac{1}{2}c_i(\mathbf{AX})_i^2 \right) \\ &= -\frac{1}{2}\mathbf{X}^\top (\mathbf{Q}(\boldsymbol{\theta}) + \mathbf{A}^\top \mathbf{D}\mathbf{A})\mathbf{X} - \mathbf{b}^\top \mathbf{A}\mathbf{X}\end{aligned}$$

where  $\mathbf{b}$  is an  $n$ -dimensional vector with entries  $\{b_i\}$  and  $\mathbf{D}$  is a diagonal matrix with  $n$  entries  $\{c_i\}$ . Note that both  $\mathbf{b}$  and  $\mathbf{D}$  depend on  $\boldsymbol{\theta}$ , so the Gaussian approximation is for a fixed  $\boldsymbol{\theta}$ .



# Modern INLA

The process is iterated to find  $\boldsymbol{b}$  and  $\boldsymbol{D}$  that gives the Gaussian approximation at the mode,  $\mu(\boldsymbol{\theta})$ , so that

$$\boldsymbol{X}|\boldsymbol{\theta}, \boldsymbol{y} \sim N\left(\boldsymbol{\mu}(\boldsymbol{\theta}), \boldsymbol{Q}_{\boldsymbol{X}}^{-1}(\boldsymbol{\theta})\right).$$

The graph of the Gaussian approximation consists of two components,

- ①  $\mathcal{G}_p$ : the graph obtained from the prior of the latent field through  $\boldsymbol{Q}(\boldsymbol{\theta})$
- ②  $\mathcal{G}_d$ : the graph obtained from the data based on the non-zero entries of  $\boldsymbol{A}^\top \boldsymbol{A}$



# Modern INLA

Next, the marginal conditional posteriors of the elements of  $\boldsymbol{X}$  is calculated from the joint Gaussian approximation as

$$\boldsymbol{X}_j | \boldsymbol{\theta}, \boldsymbol{y} \sim N \left( (\boldsymbol{\mu}(\boldsymbol{\theta}))_j, (\boldsymbol{Q}_{\boldsymbol{X}}^{-1}(\boldsymbol{\theta}))_{jj} \right).$$

but with a low-rank VB correction

$$\boldsymbol{Q}_{\boldsymbol{X}} \boldsymbol{\mu}(\boldsymbol{\theta}) = \boldsymbol{b} + \boldsymbol{\delta}$$

then

$$\boldsymbol{\mu}^*(\boldsymbol{\theta}) = \boldsymbol{Q}_{\boldsymbol{X}}^{-1} \boldsymbol{\delta} + \boldsymbol{\mu}(\boldsymbol{\theta})$$

and the marginals

$$\tilde{\pi}(\boldsymbol{X}_j | \boldsymbol{y}) = \int \pi_G(\boldsymbol{X}_j | \boldsymbol{\theta}, \boldsymbol{y}) \tilde{\pi}(\boldsymbol{\theta} | \boldsymbol{y}) d\boldsymbol{\theta} \approx \sum_{k=1}^K \pi_G(\boldsymbol{X}_j | \boldsymbol{\theta}_k, \boldsymbol{y}) \tilde{\pi}(\boldsymbol{\theta}_k | \boldsymbol{y}) \delta_k.$$

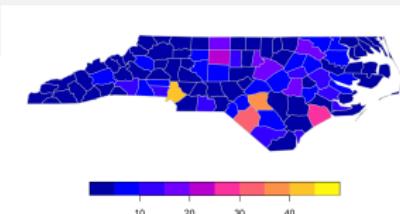


# Universal tools in INLA - also for spatial models

- Model selection metrics - WAIC, DIC
- Cross validation (1 and group) and model-based clustering
- Prediction of unobserved areas or new profiles
- Mean or quantile models
- Joint models
- Multiple imputation
- Coregionalization models
- etc..... ask at <https://groups.google.com/g/r-inla-discussion-group?pli=1> or  
e-mail [help@r-inla.org](mailto:help@r-inla.org)



# Besag and BYM



Besag model is a "smoother" over space.

$$x_i | \mathbf{x}_{-i} \sim N \left( \frac{1}{n_i} \sum_{i \sim j} x_j, \frac{1}{n_i \tau} \right)$$

BYM (Besag + iid) parameterized for interpretable parameters.

$$x_i = \frac{1}{\sqrt{\tau}} \left( \sqrt{\phi} u_i + \sqrt{1 - \phi} v_i \right).$$



# North Carolina SIDS example

NC sids R markdown file - Example 1 and 2



# Malaria and G6PD example on joint spatial modeling

We can do joint modeling and quantile models as well.

[Malaria and G6PD R markdown file - Example 3 and 4](#)



# Flexible Besag model

Instead of one precision for the entire area, we define multiple precision parameters,  $\tau_1, \tau_2, \dots, \tau_P$ , to account for covariance non-stationarity. The conditional density for the spatial effect of area  $i$  is

$$x_i | \mathbf{x}_{-i}, \tau_1, \dots, \tau_P \sim N\left(\frac{1}{2} \sum_{\substack{i \text{ in sub-region } k \\ j \text{ in sub-region } l \\ i \sim j}} (\tau_l + \tau_k) \tau_{x_i}^{-1} x_j, \tau_{x_i}^{-1}\right),$$

and

$$\tau_{x_i} = \frac{1}{2} \left( n_i \tau_k + \sum_l n_{il} \tau_l \right).$$



# Contraction prior: Non-stationary $\rightarrow$ stationary

The joint PC prior for  $\boldsymbol{\theta} = \log \boldsymbol{\tau}$  can be derived as a convolution of the PC prior for  $\boldsymbol{\tau}$  from the Besag model, as follows

$$\pi(\boldsymbol{\theta}) = 2^{-(P+2)/2} \pi^{-P/2} \lambda \sigma^{-P} \exp\left(-\frac{1}{2}(\boldsymbol{\theta} - \bar{\boldsymbol{\theta}})^T \tilde{\boldsymbol{\Sigma}}^{-1} (\boldsymbol{\theta} - \bar{\boldsymbol{\theta}}) - \bar{\theta}/2 - \lambda e^{-\bar{\theta}/2}\right),$$

This prior contracts

$$\tau_1, \tau_2, \dots, \tau_P \quad \rightarrow \quad \boldsymbol{\tau}$$



# Dengue risk in Brazil

We analyze the effects of hydrometeorological hazards on dengue risk in Brazil. To test the spatial variations in the spread of the virus in different sub-regions of Brazil, we fit dengue counts with a Poisson regression model as follows,

$$\mathbf{y} \sim \text{Poisson}(Ee^{\boldsymbol{\eta}}), \quad \boldsymbol{\eta} = \mathbf{1}^T \boldsymbol{\mu} + \boldsymbol{\alpha}$$

where  $\mathbf{y}$  is the observed counts in November of dengue cases,  $E$  is the expected number of counts,  $\boldsymbol{\eta}$  is the linear predictor,  $\boldsymbol{\mu}$  is the overall intercept, and  $\boldsymbol{\alpha}$  is the Besag or flexible Besag model over space.

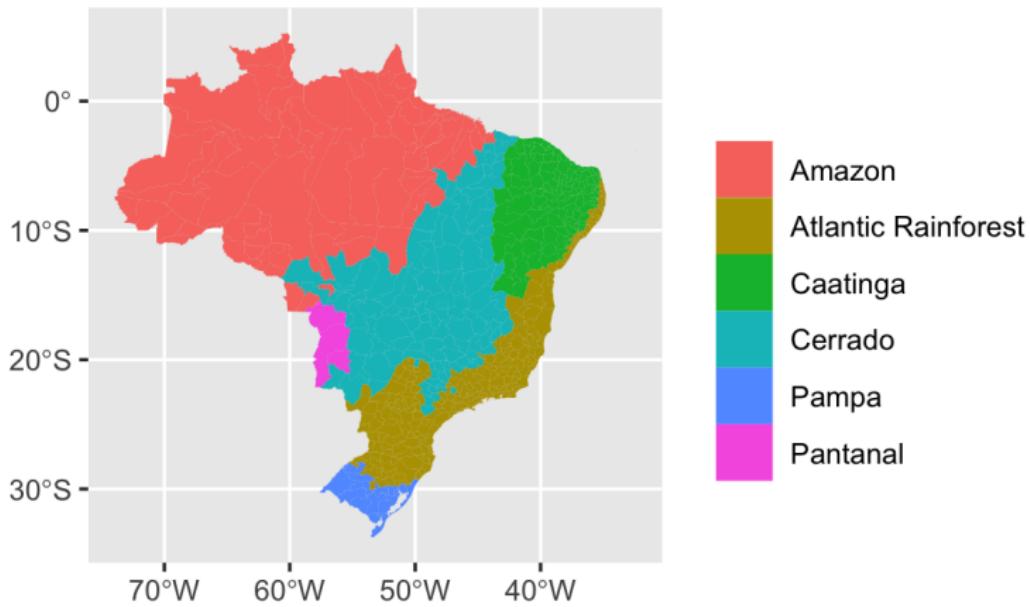


# Dengue risk in Brazil



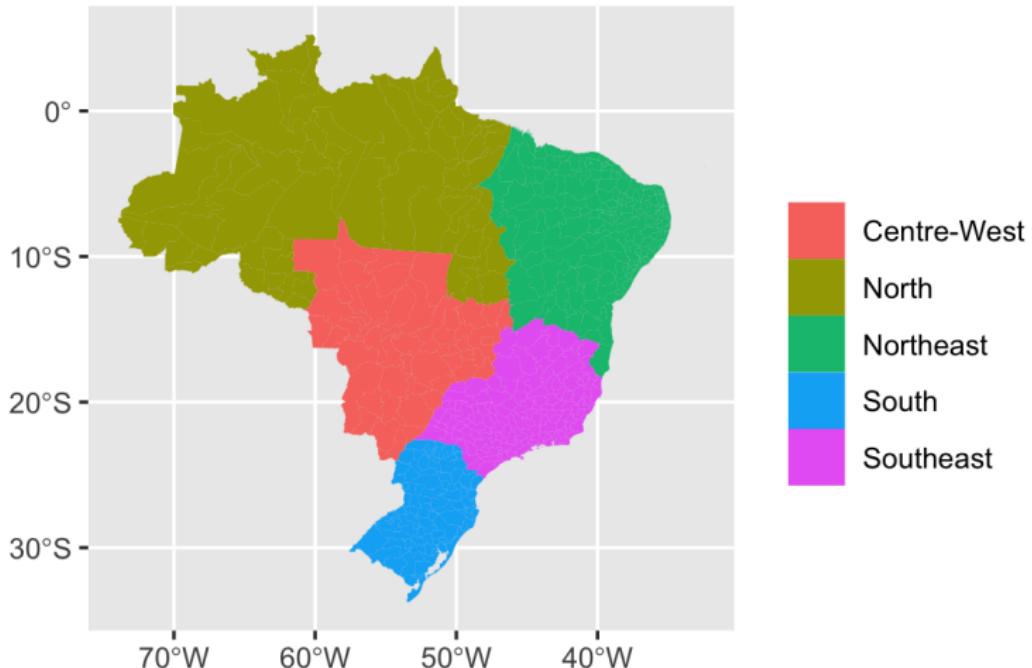


# Dengue risk in Brazil





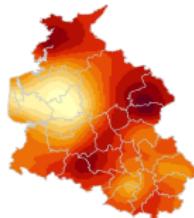
# Dengue risk in Brazil





# Dengue risk in Brazil

Dengue risk in Brazil R markdown file



Kriging provides conditional expectations of the spatial field based on covariance parameters.

With INLA we estimate "covariance" parameters in a Bayesian way and provide the marginal expectation of the spatial field.

So INLA also does "Kriging" and more - Kriging is a model, not a method.



# Matérn field

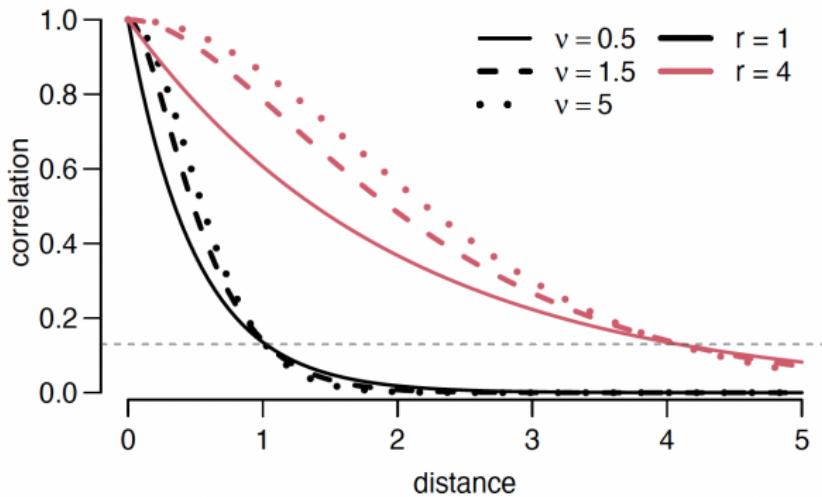
Consider a set of locations  $\mathbf{s}$ , then the spatial field  $\mathbf{u}$  defined at  $\mathbf{s}$  is multivariate Gaussian with the Matérn covariance function for the elements of  $\Sigma(\theta)$ ,

$$\pi(\mathbf{u}|\theta) = (2\pi)^{-n/2} |\Sigma(\theta)|^{-1/2} \exp\left(-\frac{1}{2}\mathbf{u}^T \Sigma(\theta)^{-1} \mathbf{u}\right)$$



# Matérn covariance model

Matérn(1960):  $\Sigma_{ij} = \frac{\sigma^2(\kappa\|\mathbf{s}_i - \mathbf{s}_j\|)^\nu K_\nu(\kappa\|\mathbf{s}_i - \mathbf{s}_j\|)}{\Gamma(\nu+d/2)(4\pi)^{d/2}\kappa^{2\nu}2^{\nu-1}}$   
 If  $d = 2$  and  $\nu = 1$ : Whittle (1954)



$$\text{practical range} = r = \sqrt{8\nu}/\kappa, \text{ corr}(r) \approx 0.13$$



# The Matérn's SPDE

- Whittle (1954), Whittle (1963):
  - Fields with Matérn covariance are solutions to the following Stochastic Partial Differential Equation (SPDE)

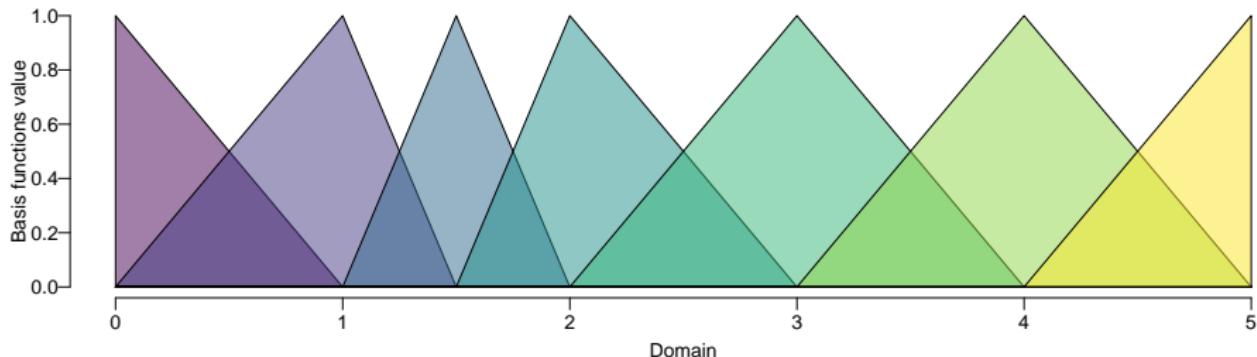
$$\tau(\kappa^2 - \Delta)^{\alpha/2} u(\mathbf{s}) = \mathcal{W}(\mathbf{s})$$

- $\kappa > 0$ : scale parameter
- $\alpha = \nu + d/2$ : smoothness
- $\Delta$  is the Laplacian

$$\Delta = \sum_{i=1}^d \frac{\partial^2}{\partial s_i^2}$$



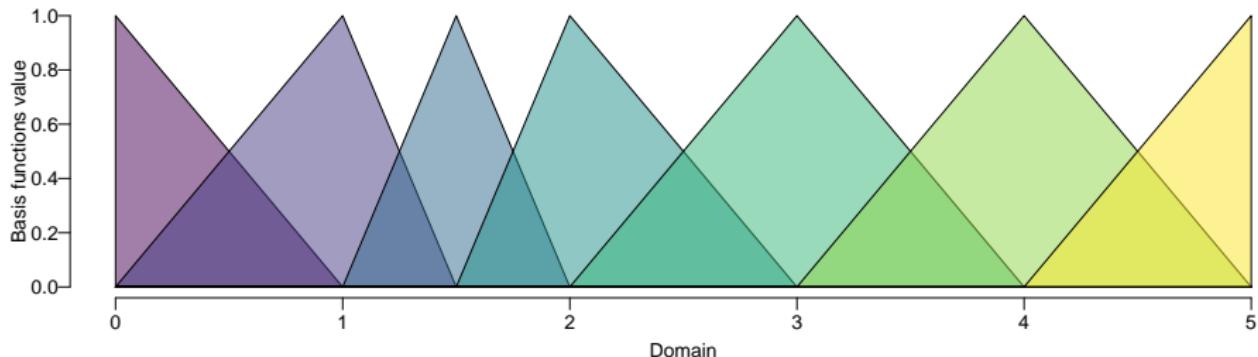
# How to solve the SPDE? FEM



- $u(\mathbf{s}) \approx \sum_{k=1}^m \psi_k(\mathbf{s}) u_k = \mathbf{A}(\mathbf{s}, \mathbf{s}_0) \mathbf{u}(\mathbf{s}_0)$ ,
  - $\psi_k$ : basis functions evaluated at data locations  $\mathbf{s}$
  - $u_k$ : the process at the discretization points  $\mathbf{s}_0$



# How to solve the SPDE? FEM



- $u(\mathbf{s}) \approx \sum_{k=1}^m \psi_k(\mathbf{s}) u_k = \mathbf{A}(\mathbf{s}, \mathbf{s}_0) u(\mathbf{s}_0)$ ,
  - $\psi_k$ : basis functions evaluated at data locations  $\mathbf{s}$
  - $u_k$ : the process at the discretization points  $\mathbf{s}_0$



# Lindgren, Rue, and Lindström (2011) I

- Discretization

- sparse precision matrix:
- $\mathbf{Q}_\alpha(\tau, \kappa)$ , for  $\alpha \in \{1, 2, \dots\}$ .

- $\alpha$

- $\alpha = 1: \tau \mathbf{K}_1(\kappa) = \tau(\kappa^2 \mathbf{C} + \mathbf{G})$
- $\alpha = 2: \tau(\kappa^4 \mathbf{C} + 2\kappa^2 \mathbf{G} + \mathbf{G}\mathbf{C}^{-1}\mathbf{G})$
- $\alpha = 2, 3, 4, \dots: \tau \mathbf{K}_1(\kappa) \mathbf{C}^{-1} \mathbf{K}_{\alpha-2}(\kappa) \mathbf{C}^{-1} \mathbf{K}_1(\kappa)$

Equivalent models (discretized): Whittle (1954), Besag (1974), Besag (1981), Besag and Kooperberg (1995) and Besag and Mondal (2005).

- $\alpha = 1: \tau \mathbf{K}_1(\kappa) = \tau(\kappa^2 \mathbf{C} + \mathbf{G})$



# Lindgren, Rue, and Lindström (2011) II

- $d=1, u_1, u_2, \dots, u_n$ , two neighbours

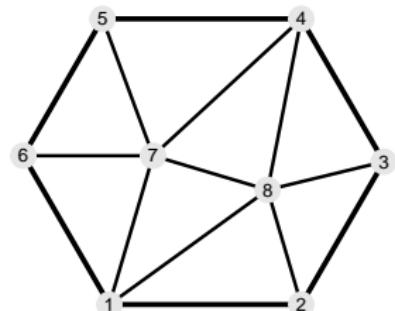
$$\tau^2 \begin{bmatrix} 1 + \kappa^2 & -1 & & \\ -1 & 2 + \kappa^2 & -1 & \\ & \ddots & & \\ -1 & 2 + \kappa^2 & -1 & \\ & -1 & 1 + \kappa^2 & \end{bmatrix}$$

- $d = 2, \mathbf{C} = \mathbf{I}, \mathbf{G} = \text{Laplacian}$  (4 neighbours)

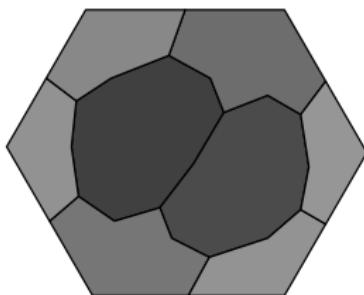


# Piecewise linear basis, FEM matrices

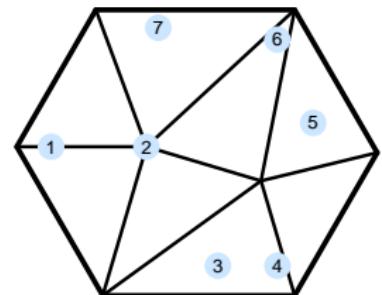
Mesh nodes



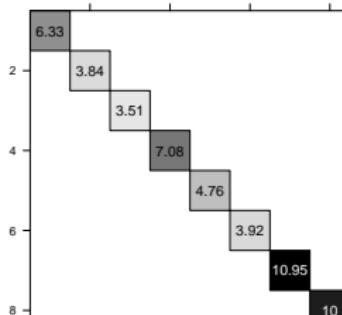
Dual mesh



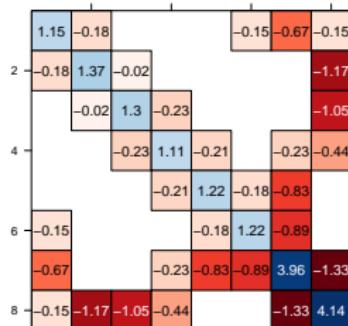
Data locations



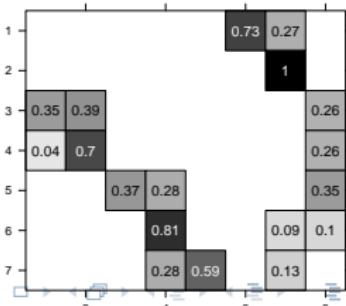
C



G

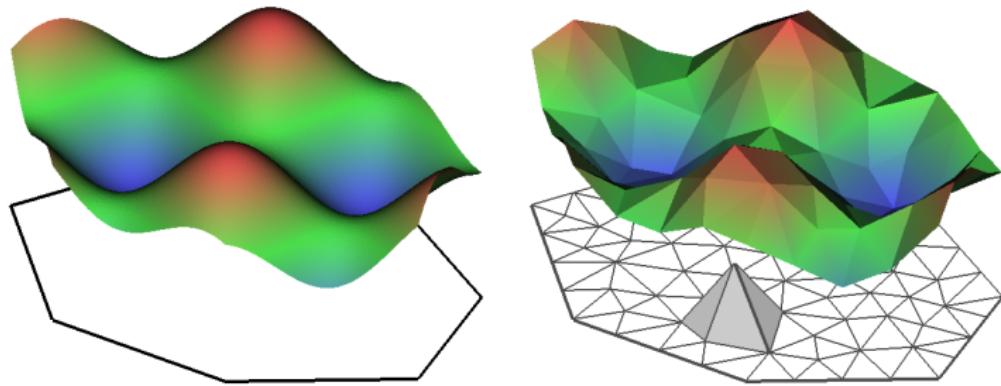


A





# FEM in 3D





# Geostatistical survival analysis

In this example we are studying the spatial distribution of leukemia mortality to inform public health policies, to gain insights for unmeasured covariates.

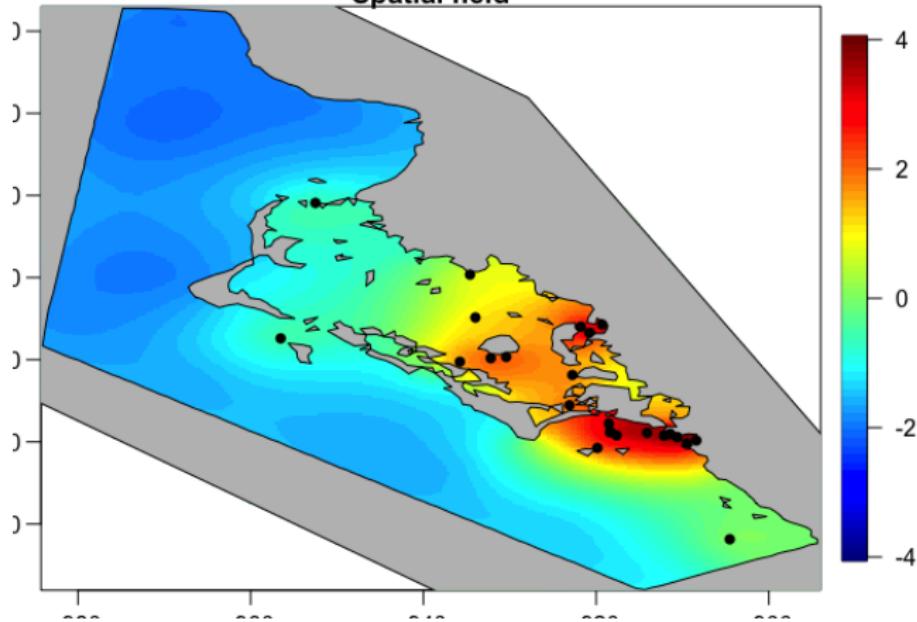
[Leukemia mortality example in R](#)



# Non-stationary Matern field based on physical barriers

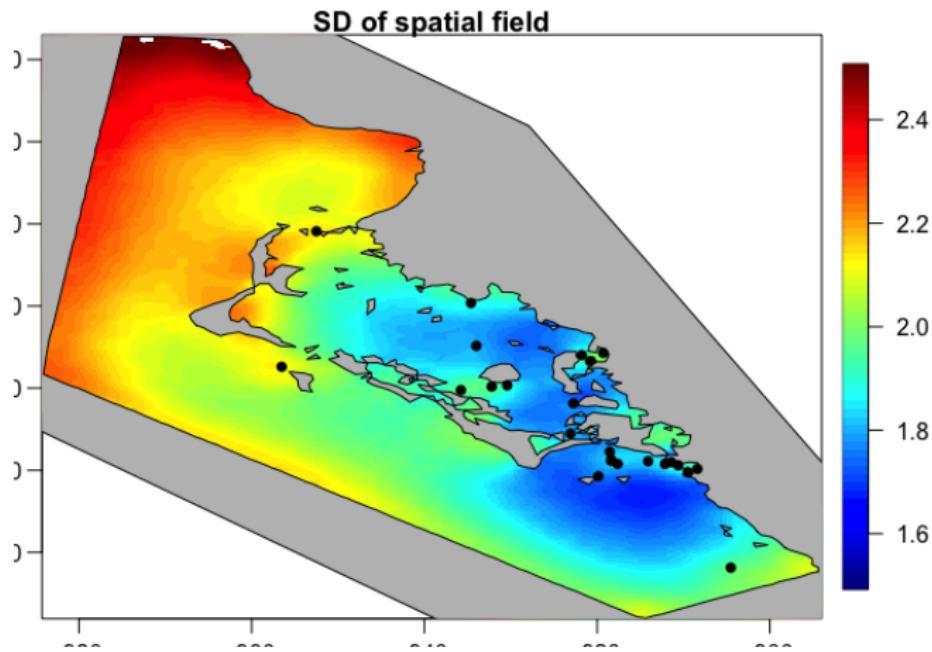
Now the distance is not Euclidean... Construct the mesh with boundaries

Spatial field





# Non-stationary Matern field based on physical barriers





## cs-fMRI model

Functional magnetic resonance imaging (fMRI) is a noninvasive neuro-imaging technique used to localize regions of specific brain activity during certain tasks.

For  $T$  timepoints and  $N$  vertices per hemisphere resulting in data  $\mathbf{y}_{TN \times 1}$  with the latent Gaussian model as follows:

$$\begin{aligned}
 \mathbf{y} | \boldsymbol{\beta}, \mathbf{b}, \boldsymbol{\theta} &\sim N(\boldsymbol{\mu}_y, \mathbf{V}), \quad \boldsymbol{\mu}_y = \sum_{k=0}^K \mathbf{x}_k \boldsymbol{\beta}_k + \sum_{j=1}^J \mathbf{z}_j \mathbf{b}_j \\
 \boldsymbol{\beta}_k &= \boldsymbol{\Psi}_k \mathbf{w}_k \quad (\text{SPDE prior on } \boldsymbol{\beta}_k) \\
 \mathbf{w}_k | \boldsymbol{\theta} &\sim N(\mathbf{0}, \mathbf{Q}_{\tau_k, \kappa_k}^{-1}) \\
 \mathbf{b}_j &\sim N(\mathbf{0}, \delta \mathbf{I}) \quad (\text{Diffuse priors for } \mathbf{b}_j) \\
 \boldsymbol{\theta} &\sim \pi(\boldsymbol{\theta}),
 \end{aligned}$$

where we have  $K$  task signals and  $J$  nuisance signals.



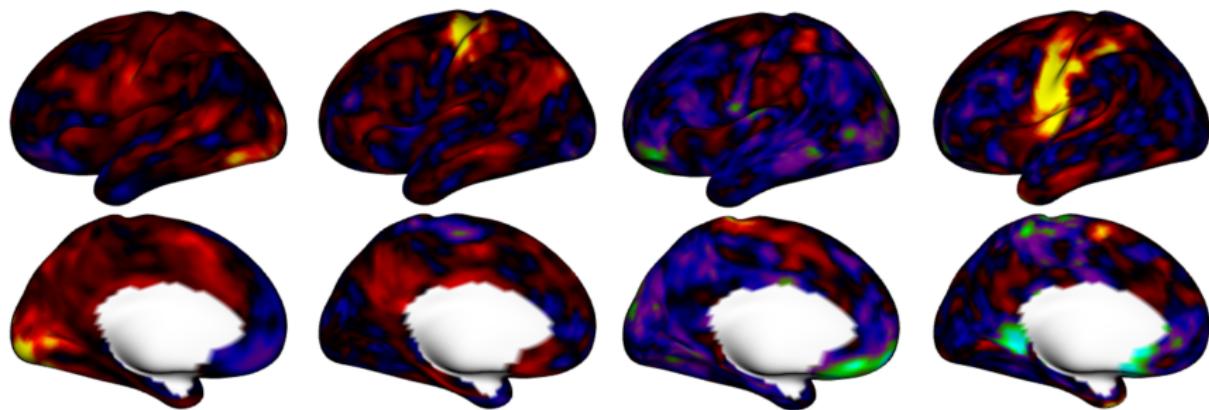
## cs-fMRI model

The data consists of a 3.5-min fMRI for each subject, consisting of 284 volumes, where each subject performs 5 different motor tasks interceded with a 3 second visual cue. Each hemisphere of the brain contained 32492 surface vertices. From these, 5000 are resampled to use for the analysis. This results in a response data vector  $\mathbf{y}$  of size **2 523 624**, with an SPDE model defined on a mesh with 8795 triangles.

The inference based on the modern formulation of INLA was computed in 148 seconds.



# cs-fMRI model



**Figure:** Activation areas for the different tasks in the left hemisphere - visual cue, right hand motor, right foot motor, tongue motor task (from left to right)



# LGCP

- Given a set of locations on a domain  $\mathcal{D}$
- One interest is to estimate the intensity function
  - $\lambda(\mathbf{l}), \lambda(\mathbf{l}) \geq 0, \mathbf{l} \in \mathcal{D}$ .
  - number of events in  $\mathcal{R} \subset \mathcal{D}$ :  $y_{\mathcal{R}} \sim \text{Poisson}(n_{\mathcal{R}})$
  - $n_{\mathcal{R}} = \int_{\mathcal{R}} \lambda(\mathbf{l}) d\mathbf{l}$
- Cox process (CP):  $\lambda(\cdot)$  is assumed to be a random function
  - $\lambda(\mathbf{l})$  is a random variable
- Log Gaussian Cox Process (LGCP)
  - $\log(\lambda(\cdot)) = u(\cdot)$  is a Gaussian process - GP, Møller, Syversveen, and Waagepetersen (1998)
  - $u(\cdot | \theta), \theta$  are GP parameters



## LGCP

- Given a set of locations on a domain  $\mathcal{D}$
- One interest is to estimate the intensity function
  - $\lambda(\mathbf{l}), \lambda(\mathbf{l}) \geq 0, \mathbf{l} \in \mathcal{D}$ .
  - number of events in  $\mathcal{R} \subset \mathcal{D}: y_{\mathcal{R}} \sim \text{Poisson}(n_{\mathcal{R}})$
  - $n_{\mathcal{R}} = \int_{\mathcal{R}} \lambda(\mathbf{l}) d\mathbf{l}$
- Cox process (CP):  $\lambda(\cdot)$  is assumed to be a random function
  - $\lambda(\mathbf{l})$  is a random variable
- Log Gaussian Cox Process (LGCP)
  - $\log(\lambda(\cdot)) = u(\cdot)$  is a Gaussian process - GP, Møller, Syversveen, and Waagepetersen (1998)
  - $u(\cdot | \theta), \theta$  are GP parameters



# LGCP inference

- The log-likelihood function:

$$l(\Lambda, \theta | \mathcal{Y}) = c - \int_{\mathcal{D}} \lambda(l) \partial l + \sum_{i=1}^n \log(\lambda(\mathbf{l}_i))$$

- The log-likelihood function direct approximation

$$\begin{aligned} l(\Lambda, \theta | \mathcal{Y}) &\approx c - \sum_{j=1}^m w_j \lambda(l) + \sum_{i=1}^n \log(\lambda(\mathbf{l}_i)) \\ &= c - \sum_{j=1}^m w_j \exp(\eta(l)) + \sum_{i=1}^n \eta(\mathbf{l}_i) \end{aligned}$$

approximated with  $m$  integration points.

- SPDE approach for easier computations, D. P. Simpson et al. (2016)
- more complex Point Process models using INLA in inlabru



# LGCP inference

- The log-likelihood function:

$$l(\Lambda, \theta | \mathcal{Y}) = c - \int_{\mathcal{D}} \lambda(l) \partial l + \sum_{i=1}^n \log(\lambda(\mathbf{l}_i))$$

- The log-likelihood function direct approximation

$$\begin{aligned} l(\Lambda, \theta | \mathcal{Y}) &\approx c - \sum_{j=1}^m w_j \lambda(l) + \sum_{i=1}^n \log(\lambda(\mathbf{l}_i)) \\ &= c - \sum_{j=1}^m w_j \exp(\eta(l)) + \sum_{i=1}^n \eta(\mathbf{l}_i) \end{aligned}$$

approximated with  $m$  integration points.

- SPDE approach for easier computations, D. P. Simpson et al. (2016)
- more complex Point Process models using INLA in inlabru



# Burkitt's lymphoma example

Locations of cases of Burkitt's lymphoma in the Western Nile district of Uganda 1960-1975. We model this data using a LGCP with INLA.

[Burkitt's lymphoma example in R](#)



# Ecological modeling under preferential sampling

This is a bit outside the scope of this workshop but I include it for your reference.

Distance sampling example with LGCP in R



# Introduction

Spatial fields sometimes exhibit covariance non-stationarity - but how to deal with this?

The SPDE approach provides the fundamental vehicle for extensions with good mathematical grounding and theoretical guarantees.

[Details here](#)

$$\left( -\gamma_t^2 \frac{d^2}{dt^2} + L_s^{\alpha_s} \right)^{\alpha_t/2} u(\mathbf{s}, t) = d\mathcal{E}_Q(\mathbf{s}, t), \quad (\mathbf{s}, t) \in \mathcal{D} \times \mathbb{R}.$$



# Examples

**Table 1.** Summary of the smoothness properties of the solutions  $u(s, t)$  for different values of the parameters  $\alpha_t, \alpha_s, \alpha_e$ , together with some examples. Here  $v_t$  and  $v_s$  respectively denote the temporal and spatial smoothnesses of the process.

$\alpha_t$	$\alpha_s$	$\alpha_e$	Type	$v_t$	$v_s$
$\alpha_t$	$\alpha_s$	$\alpha_e$	General	$\min \left[ \alpha_t - \frac{1}{2}, \frac{v_s}{\alpha_s} \right]$	$\alpha_e + \alpha_s(\alpha_t - \frac{1}{2}) - \frac{d}{2}$
$\alpha_t$	0	$\alpha_e$	Separable	$\alpha_t - \frac{1}{2}$	$\alpha_e - \frac{d}{2}$
$\alpha_t$	$\frac{d}{2}$		Critical	$\alpha_t - \frac{1}{2}$	$\alpha_s(\alpha_t - \frac{1}{2})$
$\alpha_t$	$\alpha_s$	0	Fully non-separable	$\alpha_t - \frac{1}{2} - \frac{d}{2\alpha_s}$	$\alpha_s(\alpha_t - \frac{1}{2}) - \frac{d}{2}$
1	2	$\alpha_e > \frac{d}{2}$	Sub-critical diffusion	1/2	$\alpha_e + 1 - \frac{d}{2}$
1	2	$\frac{d}{2}$	Critical diffusion	1/2	1
1	$\frac{d}{2} - 1 < \alpha_e < \frac{d}{2}$		Super-critical diffusion	$v_s/2$	$\alpha_e + 1 - \frac{d}{2}$
1	0	2	Separable	1/2	$2 - \frac{d}{2}$
3/2	2	0	Fractional diffusion	$1 - \frac{d}{4}$	$2 - \frac{d}{2}$
2	2	0	Iterated diffusion	$\frac{3}{2} - \frac{d}{4}$	$3 - \frac{d}{2}$



## Further details

[www.r-inla.org](http://www.r-inla.org)

New default setting in INLA (VB) (previously `inla.mode = "experimental"`)

- INLA can fit many different statistical models and complex models can be built using multiple "building blocks"/random effects.
- Remove the linear predictors from the latent field → accurate posterior inference with VB correction (I - VB - LA)
- New applications that aren't feasible with INLA 1.0

-  **Gaedke-Merzhäuser, L., van Niekerk, J., Schenk, O., and Rue, H. (2023).**  
Parallelized integrated nested Laplace approximations for fast Bayesian inference.  
*Statistics and Computing*, 33(1):25.
-  **Rue, H., Martino, S., and Chopin, N. (2009).**  
Approximate Bayesian inference for latent Gaussian models by using integrated nested laplace approximations.  
*Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(2):319–392.
-  **Van Niekerk, J., Krainski, E., Rustand, D., and Rue, H. (2023).**  
A new avenue for Bayesian inference with INLA.  
*Computational Statistics & Data Analysis*, 181:107692.
-  **van Niekerk, J. and Rue, H. (2024).**  
Low-rank variational Bayes correction to the Laplace method.  
*Journal of Machine Learning Research*, 25(62):1–25.



شكراً • Thank you



جامعة الملك عبد الله  
للغالوم والتكنولوجيا  
King Abdullah University of  
Science and Technology