Spatial modeling with INLA

Janet van Niekerk

September 24, 2025

Outline

1. Overview
2. Spatial domains
3. Areal modeling
   - Besag and BYM
   - Non-stationary Besag model
4. Geostatistics
   - Kriging
   - Matérn field and extensions
   - SPDE approach
   - Barrier model - non-stationary Matern field
5. Spatial modeling on different manifolds 3D...
   - Dementia study
6. Point processes - LGCP
   - LGCP and SPDE
7. Extensions to non-stationarity and non-seperable spatio-temporal models

Review

- Exact Bayesian inference requires computing posterior:

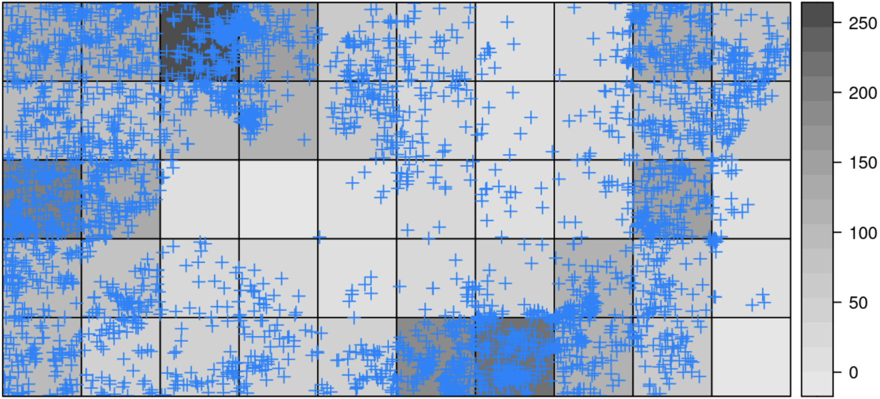
$$p(\theta \mid y) = \frac{p(y \mid \theta)p(\theta)}{p(y)}.$$

- Problem: marginal likelihood

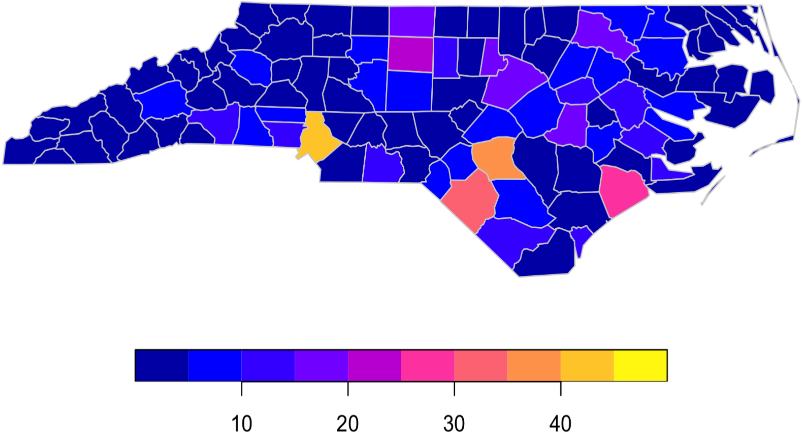
$$p(y) = \int p(y \mid \theta)p(\theta)\, d\theta$$

  often intractable.
- Laplace approximation ✓
- Variational Bayes ✓
- **INLA**

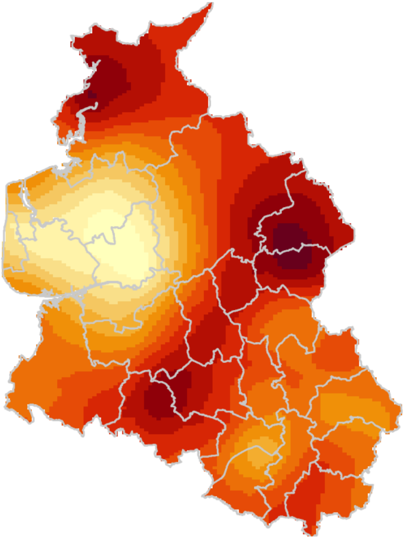
Lattice type



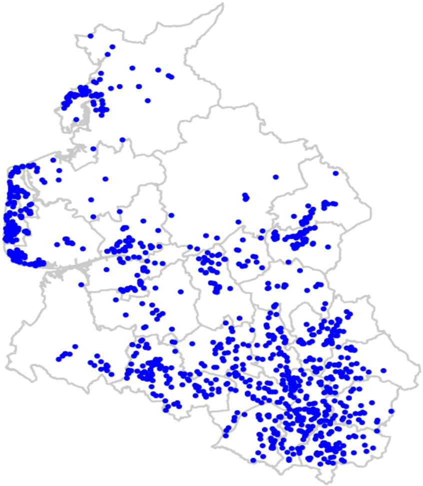
Irregular lattice - areal data



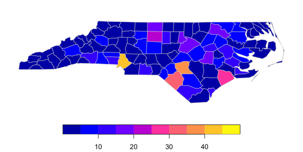
Continuous domain - geostatistics



Point process



Besag and BYM



Besag model is a "smoother" over space.

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

BYM (Besag + iid) parameterized for intrepretable 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.[1]

Malaria and G6PD R markdown file - Example 3 and 4

[1] Alahmadi, H., Van Niekerk, J., Padellini, T. and Rue, H., 2024. Joint quantile disease mapping with application to malaria and G6PD deficiency. Royal Society Open Science, 11(1), p.230851.

Flexible Besag model[2]

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

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

and

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

[2] Abdul-Fattah, E., Krainski, E., Van Niekerk, J. and Rue, H., 2024. Non-stationary Bayesian spatial model for disease mapping based on sub-regions. Statistical Methods in Medical Research, p.09622802241244613.

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 τ from the Besag model, as follows

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

This prior contracts

$$\tau_1, \tau_2, ..., \tau_P \quad \rightarrow \quad \tau$$

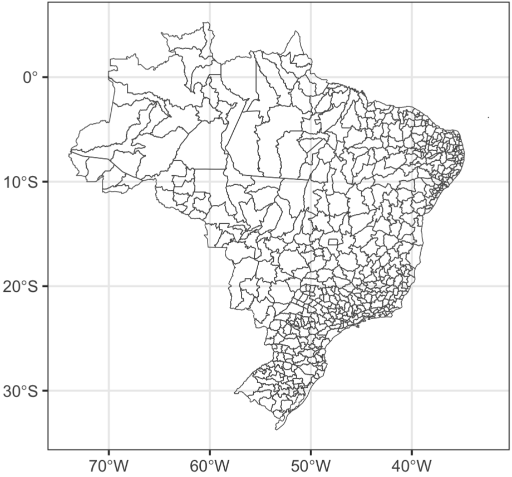
Dengue risk in Brazil

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,

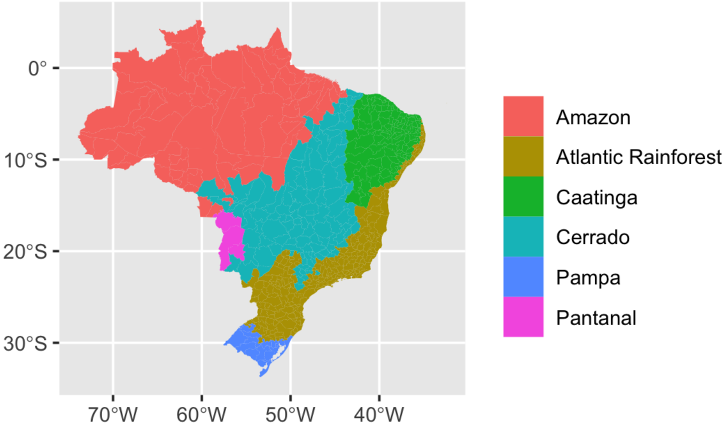
$$\boldsymbol{y} \sim \text{Poisson}(Ee^{\boldsymbol{\eta}}), \quad \boldsymbol{\eta} = \mathbf{1}^T \mu + \boldsymbol{\alpha}$$

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

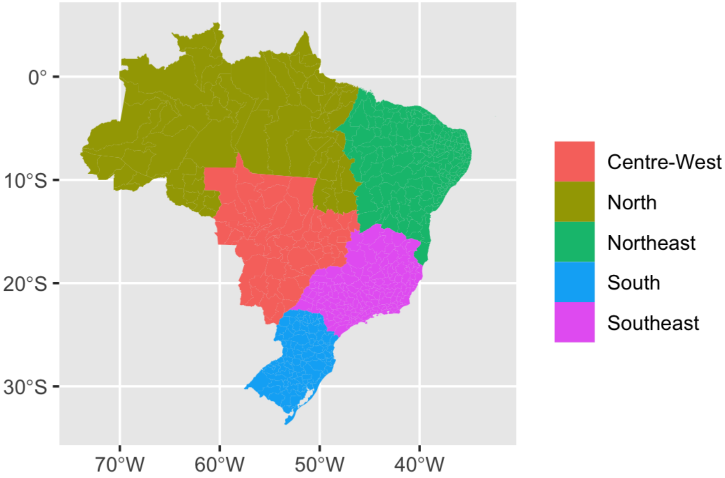
Dengue risk in Brazil



Dengue risk in Brazil

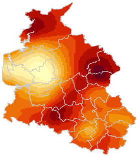


Dengue risk in Brazil



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

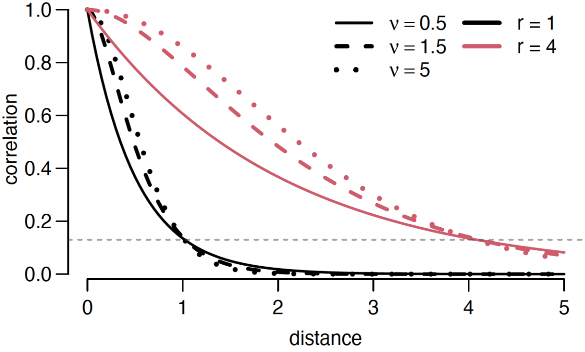
Consider a set of locations $\boldsymbol{s}$, then the spatial field $\boldsymbol{u}$ defined at $\boldsymbol{s}$ is multivariate Gaussian with the Matérn covariance function for the elements of $\Sigma(\theta)$,

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

Matérn covariance model

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

If $d = 2$ and $\nu = 1$: Whittle (1954)



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

The Matérn's SPDE

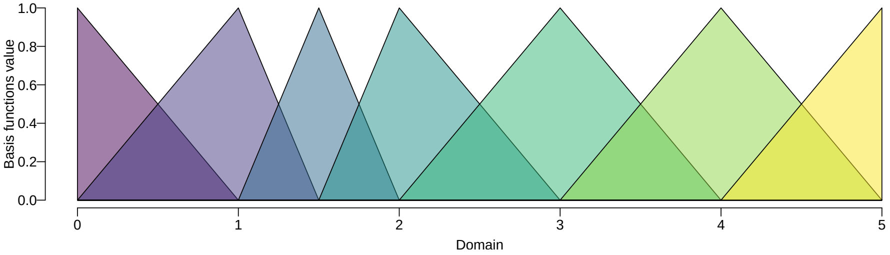
- Whittle (1954), Whittle (1963):
  - Fields with Matérn covariance are solutions to the following Stochastic Partial Differential Equation (SPDE)

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

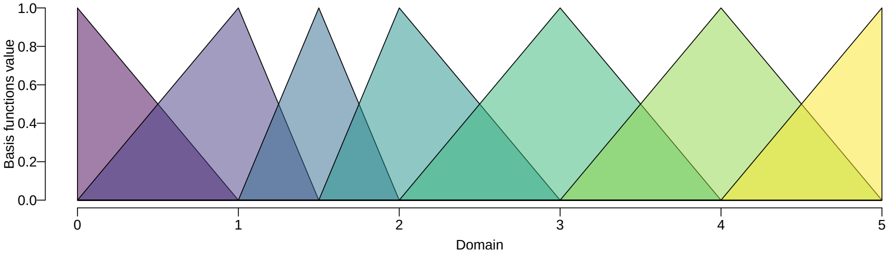
  - $\kappa > 0$: scale parameter
  - $\alpha = \nu + d/2$: smoothness
  - Δ is the Laplacian

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

How to solve the SPDE? FEM



How to solve the SPDE? FEM



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

- Discretization
  - sparse precision matrix:
  - $\mathbf{Q}_\alpha(\tau, \kappa)$, for $\alpha \in \{1, 2, ...\}$.
- α
  - $\alpha = 1$: $\tau \mathbf{K}_1(\kappa) = \tau(\kappa^2 \mathbf{C} + \mathbf{G})$
  - $\alpha = 2$: $\tau(\kappa^4 \mathbf{C} + 2\kappa^2 \mathbf{G} + \mathbf{G}\mathbf{C}^{-1}\mathbf{G})$
  - $\alpha = 2, 3, 4, ...$: $\tau \mathbf{K}_1(\kappa)\mathbf{C}^{-1}\mathbf{K}_{\alpha-2}(\kappa)\mathbf{C}^{-1}\mathbf{K}_1(\kappa)$

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)[4] II

- d=1, $u_1, u_2, ..., u_n$, two neighbours

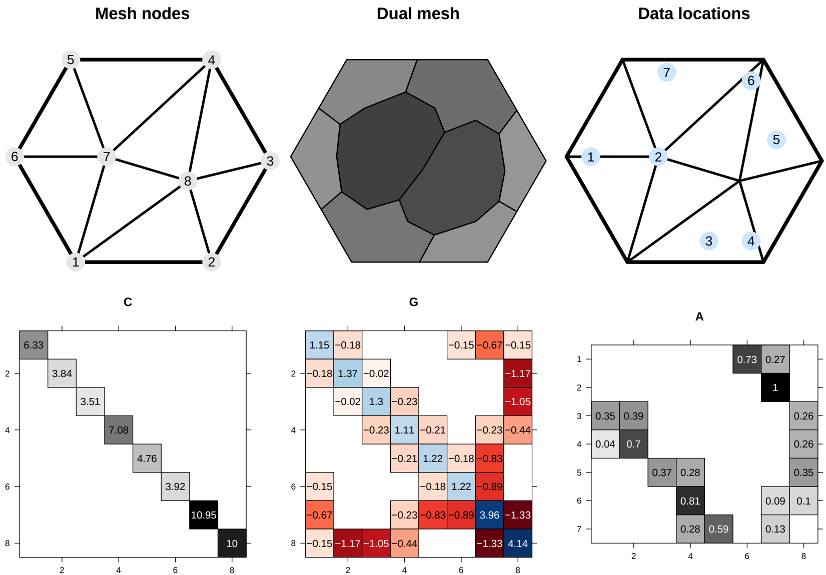
$$\tau^2 \begin{bmatrix} 1+\kappa^2 & -1 & & & \\ -1 & 2+\kappa^2 & -1 & & \\ & & \ddots & & \\ & & -1 & 2+\kappa^2 & -1 \\ & & & -1 & 1+\kappa^2 \end{bmatrix}$$

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

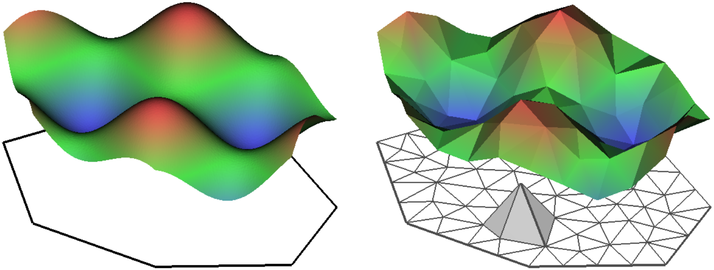
[3] Lindgren, F., Rue, H. and Lindström, J., 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. Journal of the Royal Statistical Society Series B: Statistical Methodology, 73(4), pp.423-498.

[4] Lindgren, F., Rue, H. and Lindström, J., 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(4), pp.423-498.

Piecewise linear basis, FEM matrices



FEM in 3D



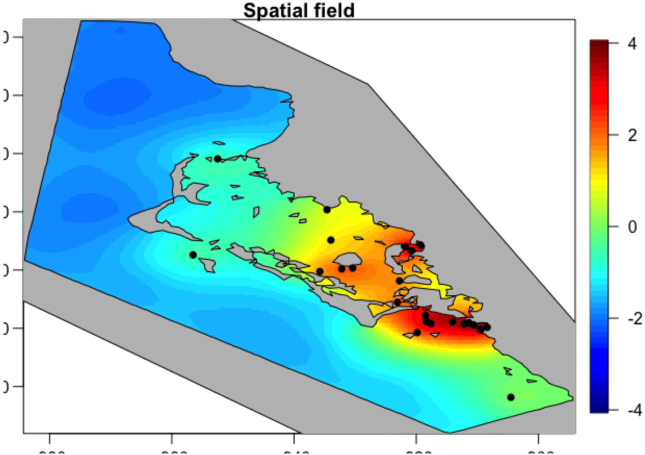
Geostatistical survival analysis

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

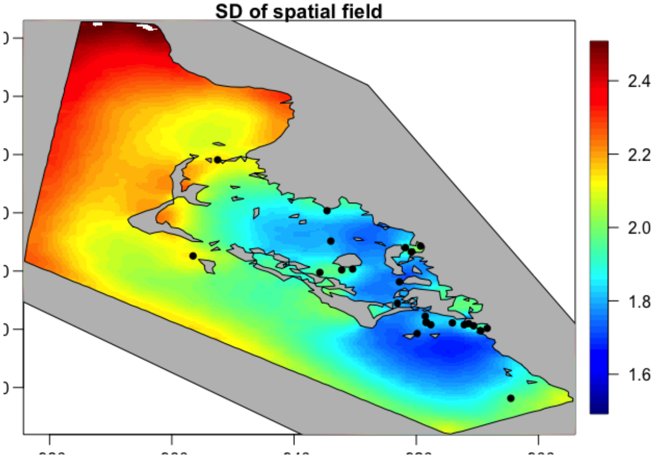
Leukemia mortality example in R

Non-stationary Matern field based on physical barriers

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



Non-stationary Matern field based on physical barriers



cs-fMRI model[5]

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 $\boldsymbol{y}_{TN\times 1}$ with the latent Gaussian model as follows:

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

where we have K task signals and J nuisance signals.

[5] Van Niekerk, J., Krainski, E., Rustand, D. and Rue, H., 2023. A new avenue for Bayesian inference with INLA. Computational Statistics & Data Analysis, 181, p.107692.

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 $\boldsymbol{y}$ 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

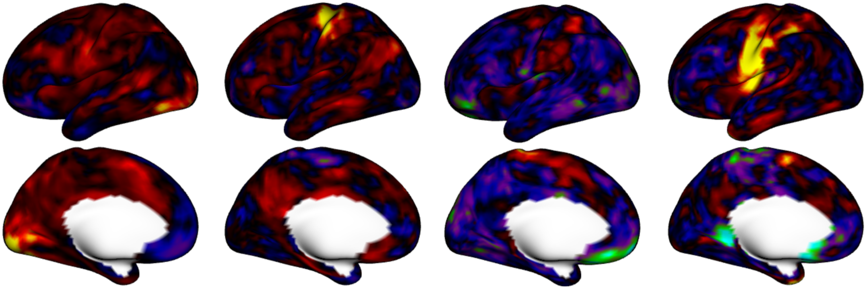


Figure: Activation areas for the different tasks in the left hemisphere - visual cue, right hand motor, right foot motor, tongue motor task (from left to right)

LGCP

- Given a set of locations on a domain $\mathcal{D}$
- One interest is to estimate the intensity function
  - $\lambda(\mathbf{l})$, $\lambda(\mathbf{l}) \geq 0$, $\mathbf{l} \in \mathcal{D}$.

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 $\mathcal{R} \subset \mathcal{D}$: $y_{\mathcal{R}} \sim \text{Poisson}(n_{\mathcal{R}})$
  - $n_{\mathcal{R}} = \int_{\mathcal{R}} \lambda(\mathbf{l}) \partial \mathbf{l}$
- Cox process (CP): $\lambda(.)$ is assumed to be a random function
  - $\lambda(\mathbf{l})$ is a random variable
- Log Gaussian Cox Process (LGCP)
  - $\log(\lambda(.)) = \text{u}(.)$ is a Gaussian process - GP, Møller, Syversveen, and Waagepetersen (1998)
  - $u(.|\theta)$, θ are GP parameters

LGCP inference

- The log-likelihood function:

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

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[6].
Details here

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

[6] Lindgren, F., Bakka, H., Bolin, D., Krainski, E. and Rue, H., 2024. A diffusion-based spatio-temporal extension of Gaussian Matérn fields. SORT-Statistics and Operations Research Transactions, pp.3-66.

Examples

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

α_t	α_s	α_e	Type	ν_t	ν_s
α_t	α_s	α_e	General	$\min\left[\alpha_t - \frac{1}{2}, \frac{\nu_s}{\alpha_s}\right]$	$\alpha_e + \alpha_s(\alpha_t - \frac{1}{2}) - \frac{d}{2}$
α_t	0	α_e	Separable	$\alpha_t - \frac{1}{2}$	$\alpha_e - \frac{d}{2}$
α_t	α_s	$\frac{d}{2}$	Critical	$\alpha_t - \frac{1}{2}$	$\alpha_s(\alpha_t - \frac{1}{2})$
α_t	α_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	2	$\frac{d}{2} - 1 < \alpha_e < \frac{d}{2}$	Super-critical diffusion	$\nu_s/2$	$\alpha_e + 1 - \frac{d}{2}$
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	$\frac{3}{2} - \frac{d}{4}$	$3 - \frac{d}{2}$