# An overview of the SPDE approach

Elias Teixeira Krainski

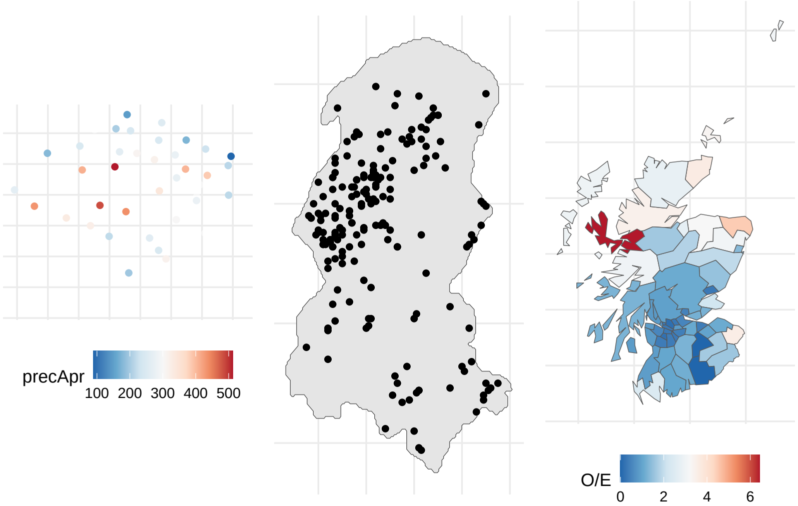
King Abdullah University of Science and Technology (KAUST)

May 2024

# Outline

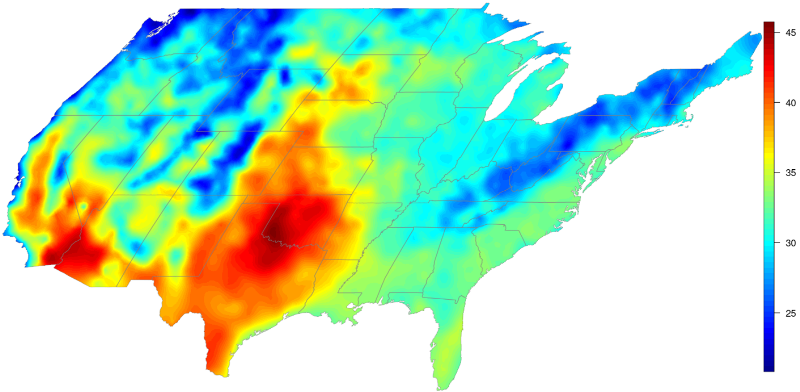
# Introduction

# Spatial statistics data



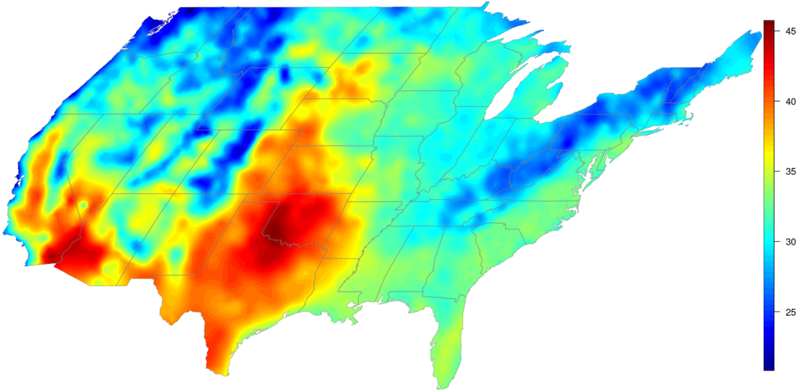
# A process over space

- (est.) Maximum temperature in US mainland, in 2022-07-20



# A process over space

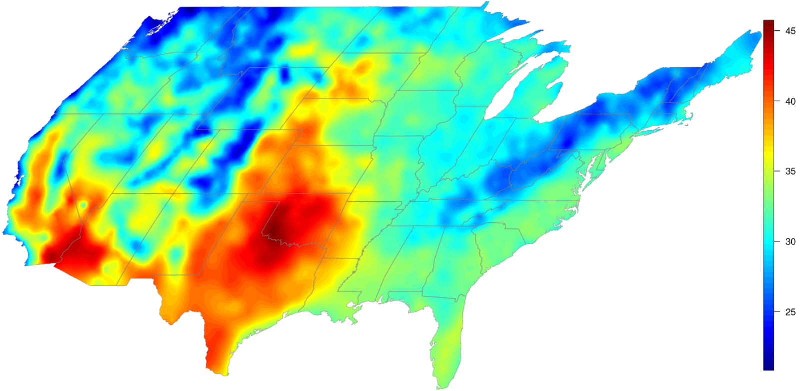
- (est.) Maximum temperature in US mainland, in 2022-07-20




- The spatial domain, S, is continuous.
  - E.g. $S \in \mathbb{R}^2$
  - E.g. $S \in \mathbb{S}^2$ (sphere)

# A process over space

- (est.) Maximum temperature in US mainland, in 2022-07-20




- The spatial domain, S, is continuous.
  - E.g. $S \in \mathbb{R}^2$
  - E.g. $S \in \mathbb{S}^2$ (sphere)

- $u(\mathbf{s})$ is a stochastic process, $\mathbf{s} \in S$

# Real world data: at a finite number of locations

At a finite set of $n$ locations

$u(\mathbf{s}_1), \ldots, u(\mathbf{s}_n)$

# Stochastic process $u(\mathbf{l})$

- Covariance

$$V(\mathbf{l}, \mathbf{l}') = \mathrm{Cov}(u(\mathbf{l}), u(\mathbf{l}'))$$

- Spectral density

$$u(\mathbf{l}) = \int_{-\infty}^{\infty} \mathrm{e}^{i\mathbf{w}\mathbf{l}} du(\mathbf{w})$$

- Kernel convolution

$$u(\mathbf{l}) = \int k(\mathbf{l} - \mathbf{u}) \mathcal{W}(\mathbf{l}) d\mathbf{u}$$

- Stochastic Partial Differential Equation - SPDE

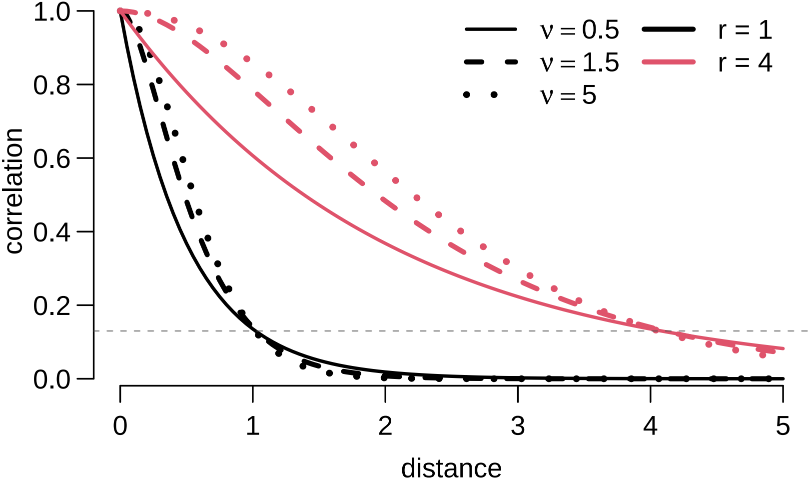
$$\mathcal{L}_1 u(\mathbf{l}) = \mathcal{L}_2 \mathcal{W}(\mathbf{l})$$

- Conditional distributions ( DISCRETE DOMAIN !!! )

$$u_{\mathbf{l}} | u_{\mathbf{neighbourhood}} \sim \mathbb{P}(.)$$

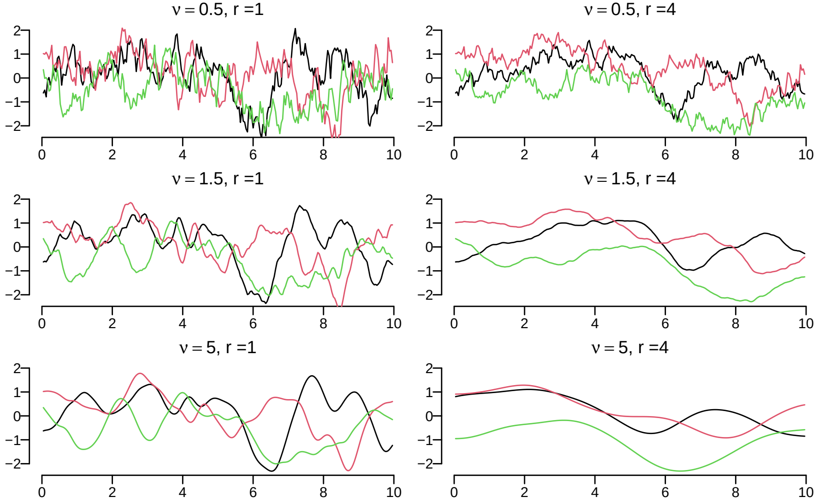
# Whittle-Matérn covariance, Matérn (1960)

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

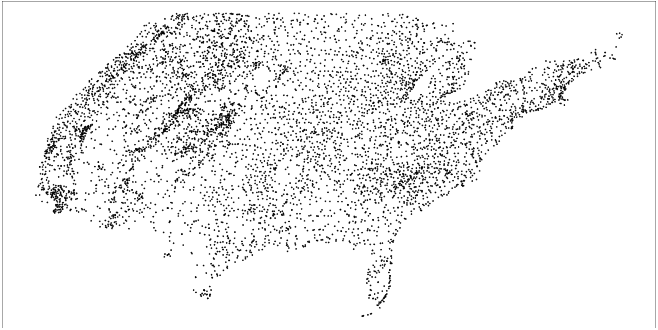



practical range $=$ r $= \sqrt{8\nu}/\kappa$, corr(r) $\approx 0.13$

# Simulations, 1D, $\sigma^2 = 1$

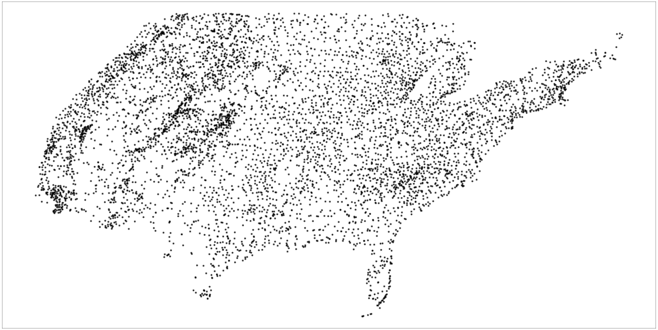


# Consider a set of locations



- Consider the distribution at these locations
  - $\pi(u(\mathbf{s}_1), ..., u(\mathbf{s}_n)|\theta) = \pi(u_1, ..., u_n|\theta) = \pi(\mathbf{u}|\theta)$

# Consider a set of locations

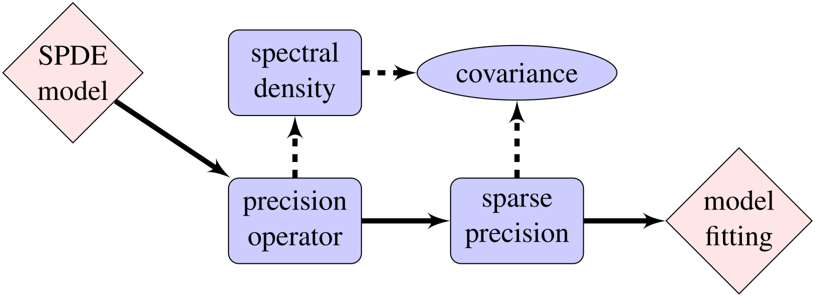


- Consider the distribution at these locations
  - $\pi(u(\mathbf{s}_1), ..., u(\mathbf{s}_n)|\theta) = \pi(u_1, ..., u_n|\theta) = \pi(\mathbf{u}|\theta)$
- If it is Gaussian with some covariance $\Sigma(\theta)$

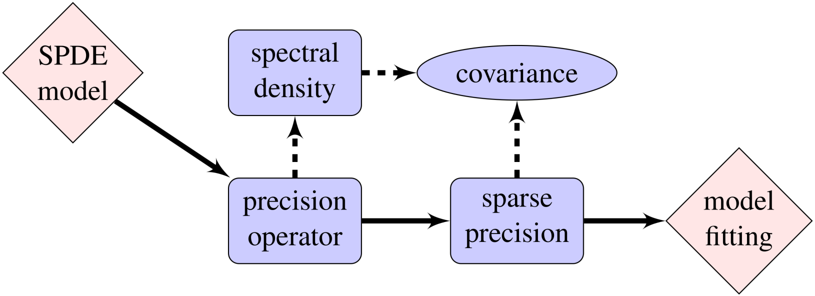
$$\pi(\mathbf{u}|\theta) = (2\pi)^{-n/2}|\Sigma(\theta)|^{-1/2} \exp\left(-\frac{1}{2}\mathbf{u}^T\Sigma(\theta)^{-1}\mathbf{u}\right)$$

# The SPDE approach

# SPDE framework



# SPDE framework




- It avoids specifying covariance!
  - D. Simpson, Lindgren, and Rue (2011)
  - D. Simpson, Lindgren, and Rue (2012)

# 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

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

# Lindgren, Rue, and Lindström (2011)

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

# Lindgren, Rue, and Lindström (2011)

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

# Lindgren, Rue, and Lindström (2011)

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

# Lindgren, Rue, and Lindström (2011)

- Discretization
  - sparse precision matrix:
  - $\mathbf{Q}_\alpha(\tau, \kappa)$, for $\alpha \in \{1, 2, ...\}$.
- $\alpha$
  - $\alpha = 1$: $\tau\mathbf{K}_1(\kappa) = \tau(\kappa^2\mathbf{C} + \mathbf{G})$
  - $\alpha = 2$: $\tau(\kappa^4\mathbf{C} + 2\kappa^2\mathbf{G} + \mathbf{G}\mathbf{C}^{-1}\mathbf{G})$
  - $\alpha = 2, 3, 4, ...$: $\tau\mathbf{K}_1(\kappa)\mathbf{C}^{-1}\mathbf{K}_{\alpha-2}(\kappa)\mathbf{C}^{-1}\mathbf{K}_1(\kappa)$
- Equivalent models (discretized): Whittle (1954), Besag (1974), Besag (1981), Besag and Kooperberg (1995) and Besag and Mondal (2005).

## $\mathbf{Q}_\alpha(\tau, \kappa)$: grid and piecewise linear basis

- $\alpha = 1$: $\tau \mathbf{K}_1(\kappa) = \tau(\kappa^2 \mathbf{C} + \mathbf{G})$
  - d=1, $u_1, u_2, ..., u_n$, two neighbours

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

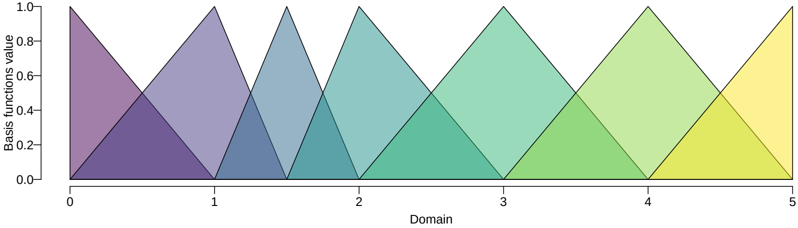
## $\mathbf{Q}_\alpha(\tau,\kappa)$: grid and piecewise linear basis

- $\alpha = 1$: $\tau\mathbf{K}_1(\kappa) = \tau(\kappa^2\mathbf{C} + \mathbf{G})$
  - d=1, $u_1, u_2, ..., u_n$, two neighbours

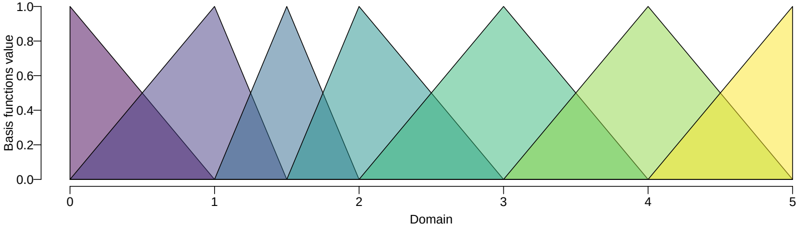
$$\tau^2 \begin{bmatrix} 1+\kappa^2 & -1 & & & \\ -1 & 2+\kappa^2 & -1 & & \\ & & \ddots & & \\ & & -1 & 2+\kappa^2 & -1 \\ & & & -1 & 1+\kappa^2 \end{bmatrix}$$

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

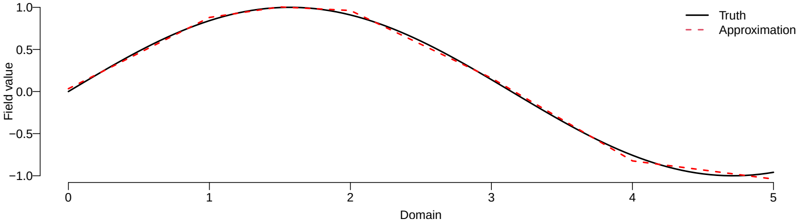
# Piecewise linear basis, Finite Element Method (FEM): 1d



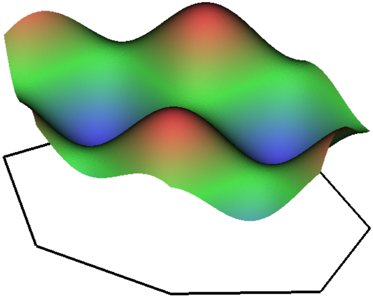
# Piecewise linear basis, Finite Element Method (FEM): 1d

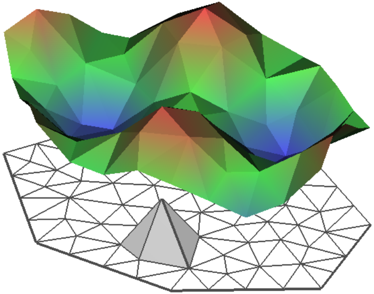



- $u(\mathbf{s}) \approx \sum_{k=1}^{m} \psi_k(\mathbf{s}) u_k = \mathbf{A}(\mathbf{s}, \mathbf{s}_0) u(\mathbf{s}_0)$,
  - $\psi_k$: basis functions evaluated at data locations $\mathbf{s}$
  - $u_k$: the process at the discretization points $\mathbf{s}_0$

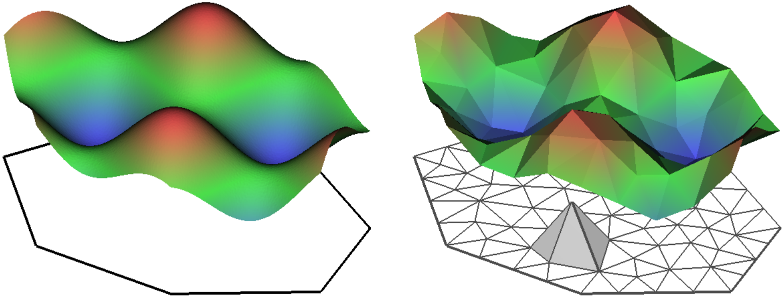


# Piecewise linear basis, Finite Element Method: 2d





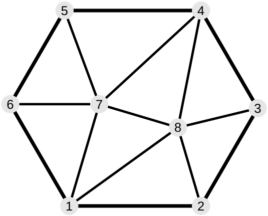
# Piecewise linear basis, Finite Element Method: 2d



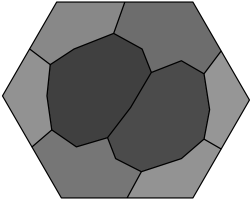
- $u(\mathbf{s}) \approx \sum_{k=1}^{m} \psi_k(\mathbf{s}) u_k = \mathbf{A}(\mathbf{s}, \mathbf{s}_0) u(\mathbf{s}_0)$,
  - $\psi_k$: basis functions evaluated at data locations $\mathbf{s}$
  - $u_k$: the process at the discretization points $\mathbf{s}_0$

# Piecewise linear basis, FEM matrices

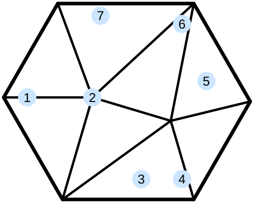
**Mesh nodes**




**Dual mesh**

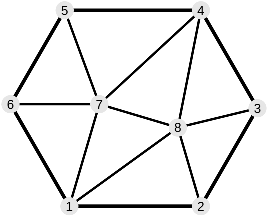


**Data locations**

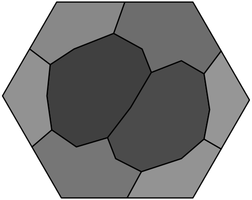


# Piecewise linear basis, FEM matrices

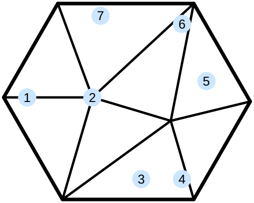
**Mesh nodes**



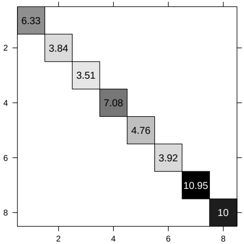

**Dual mesh**



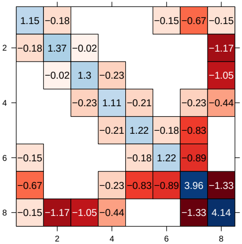
**Data locations**



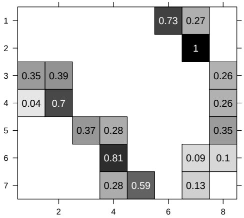

**C**




**G**




**A**



# The log Gaussian Cox process model

# Point pattern and intensity

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

# Point pattern and intensity

- 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 $\mathcal{R} \subset \mathcal{D}$: $y_{\mathcal{R}} \sim \text{Poisson}(n_{\mathcal{R}})$
  - $n_{\mathcal{R}} = \int_{\mathcal{R}} \lambda(\mathbf{l}) \partial \mathbf{l}$

# Point pattern and intensity

- 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

# Point pattern and intensity

- 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) .
  - $u(.|\theta)$, $\theta$ are GP parameters

# LGCP inference

- The log-likelihood function:

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

# 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

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

  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

# References


Besag, J. 1974. "Spatial Interaction and the Statistical Analysis of Lattice Systems." *JRSS-B* 36 (2): 192–236.

———. 1981. "On a System of Two-Dimensional Recurrence Equations." *JRSS-B* 43 (3): 302–9.

Besag, J., and C. Kooperberg. 1995. "On Conditional and Intrinsic Autoregression." *Biometrika* 82 (4): 733–46.

Besag, J., and D. Mondal. 2005. "First-Order Intrinsic Autoregressions and the de Wijs Process." *Biometrika* 92 (4): 909–20.

Lindgren, F., H. Rue, and J. Lindström. 2011. "An Explicit Link Between Gaussian Fields and Gaussian Markov Random Fields: The Stochastic Partial Differential Equation Approach (with Discussion)." *JRSS-B* 73 (4): 423–98.

Matérn, B. 1960. "Spatial Variation-Stochastic Models and Their Application to Some Problems in Forest Surveys and Other Sampling Investigations. Meddelanden Fran Statens Skogsforskningsintitut, Almaenna Foerlaget, Stockholm. (1986), 49 (5)." Springer, Berlin.

Møller, J., A. R. Syversveen, and R. P. Waagepetersen. 1998. "Log Gaussian Cox Processes." *SJS* 25: 451–82.

Simpson, D. P., J. B. Illian, F. Lindren, S. H Sørbye, and H. Rue. 2016. "Going Off Grid: Computationally Efficient Inference for Log-Gaussian Cox Processes." *Biometrika* 103 (1): 49–70.

Simpson, D., F. Lindgren, and H. Rue. 2011. "In Order to Make Spatial Statistics Computationally Feasible, We Need to Forget about the Covariance Function." *Environmetrics*, no. 23: 65–74. https://doi.org/https://doi.org/10.1002/env.1137.

———. 2012. "Think Continuous: Markovian Gaussian Models in Spatial Statistics." *Spatial Statistics* 1: 16–29. https://doi.org/http://dx.doi.org/10.1016/j.spasta.2012.02.003.

Whittle, P. 1954. "On Stationary Processes in the Plane." *Biometrika* 41 (3/4): 434–49.

———. 1963. "Stochastic-Processes in Several Dimensions." *Bulletin of the International Statistical Institute* 40 (2): 974–94.