Stochastic adventures in space and time

Finn Lindgren (finn.lindgren@ed.ac.uk)

with Elias Krainski, David Bolin, Haakon Bakka, Haavard Rue
Rafael Arce Guillen, Stefanie Muff, Ulrike Schlägel

THE UNIVERSITY *of* EDINBURGH

Edinburgh, 2022-11-14

The eternal quest for spatial dependence models

- Gaussian random field: $u(\mathbf{s})$, $\mathbf{s} \in \mathcal{D}$ (subset of $\mathbb{R}^d$ or a manifold)
- Moment characterisation:
  - Expectation $\mu(\mathbf{s}) = \mathbb{E}[u(\mathbf{s})]$
  - Covariance $\mathcal{R}(\mathbf{s}, \mathbf{s}') = \mathbb{C}[u(\mathbf{s}), u(\mathbf{s}')]$, symmetric positive definite function.
- Precision operator; inverse covariance: $\mathcal{Q} = \mathcal{R}^{-1}$
  In practice, easier conditions for valid models
- Reproducing Kernel Hilbert Space (RKHS) $H_\mathcal{Q}$: Inner product

$$\langle f, g \rangle_{H_\mathcal{Q}} = \langle f, \mathcal{Q} g \rangle_\mathcal{D}$$

  and squared norm $\|f\|^2 = \langle f, f \rangle_{H_\mathcal{Q}}$
- $m(\cdot) = \mathbb{E}(u(\cdot) - \mu(\cdot) | \{u(\mathbf{s}_k)\}) \in H_\mathcal{Q}$ but $u(\cdot) - \mu(\cdot) \notin H_\mathcal{Q}$; the process is less smooth!
- Spatial and spatio-temporal stochastic PDEs generate random field models:

$$\mathcal{L} u(\mathbf{s}) = \dot{\mathcal{W}}(\mathbf{s})$$

$$\mathcal{Q}_u = \mathcal{L}^* \mathcal{L}$$

$$\langle f, g \rangle_{H_\mathcal{Q}} = \langle \mathcal{L} f, \mathcal{L} g \rangle_\mathcal{D}$$

  Can work directly with the precision or inner product; no need to know the covariance!

The eternal quest for spatial dependence models

- Gaussian random field: $u(\mathbf{s})$, $\mathbf{s} \in \mathcal{D}$ (subset of $\mathbb{R}^d$ or a manifold)
- Moment characterisation:
  - Expectation $\mu(\mathbf{s}) = \mathbb{E}[u(\mathbf{s})]$
  - Covariance $\mathcal{R}(\mathbf{s}, \mathbf{s}') = \mathbb{C}[u(\mathbf{s}), u(\mathbf{s}')]$, symmetric positive definite function.
- Precision operator; inverse covariance: $\mathcal{Q} = \mathcal{R}^{-1}$

  In practice, easier conditions for valid models
- Reproducing Kernel Hilbert Space (RKHS) $H_\mathcal{Q}$: Inner product

$$\langle f, g \rangle_{H_\mathcal{Q}} = \langle f, \mathcal{Q}g \rangle_\mathcal{D}$$

  and squared norm $\|f\|^2 = \langle f, f \rangle_{H_\mathcal{Q}}$
- $m(\cdot) = \mathbb{E}(u(\cdot) - \mu(\cdot) | \{u(\mathbf{s}_k)\}) \in H_\mathcal{Q}$ but $u(\cdot) - \mu(\cdot) \notin H_\mathcal{Q}$; the process is less smooth!
- Spatial and spatio-temporal stochastic PDEs generate random field models:

$$\mathcal{L}u(\mathbf{s}) = \dot{\mathcal{W}}(\mathbf{s})$$

$$\mathcal{Q}_u = \mathcal{L}^*\mathcal{L}$$

$$\langle f, g \rangle_{H_\mathcal{Q}} = \langle \mathcal{L}f, \mathcal{L}g \rangle_\mathcal{D}$$

  Can work directly with the precision or inner product; no need to know the covariance!

The eternal quest for spatial dependence models

- Gaussian random field: $u(\mathbf{s})$, $\mathbf{s} \in \mathcal{D}$ (subset of $\mathbb{R}^d$ or a manifold)
- Moment characterisation:
  - Expectation $\mu(\mathbf{s}) = \mathbb{E}[u(\mathbf{s})]$
  - Covariance $\mathcal{R}(\mathbf{s}, \mathbf{s}') = \mathbb{C}[u(\mathbf{s}), u(\mathbf{s}')]$, symmetric positive definite function.
- Precision operator; inverse covariance: $\mathcal{Q} = \mathcal{R}^{-1}$

  In practice, easier conditions for valid models
- Reproducing Kernel Hilbert Space (RKHS) $H_{\mathcal{Q}}$: Inner product

$$\langle f, g \rangle_{H_{\mathcal{Q}}} = \langle f, \mathcal{Q} g \rangle_{\mathcal{D}}$$

  and squared norm $\|f\|^2 = \langle f, f \rangle_{H_{\mathcal{Q}}}$
- $m(\cdot) = \mathbb{E}(u(\cdot) - \mu(\cdot) | \{u(\mathbf{s}_k)\}) \in H_{\mathcal{Q}}$ but $u(\cdot) - \mu(\cdot) \notin H_{\mathcal{Q}}$; the process is less smooth!
- Spatial and spatio-temporal stochastic PDEs generate random field models:

$$\mathcal{L} u(\mathbf{s}) = \dot{\mathcal{W}}(\mathbf{s})$$

$$\mathcal{Q}_u = \mathcal{L}^* \mathcal{L}$$

$$\langle f, g \rangle_{H_{\mathcal{Q}}} = \langle \mathcal{L} f, \mathcal{L} g \rangle_{\mathcal{D}}$$

  Can work directly with the precision or inner product; no need to know the covariance!

The eternal quest for spatial dependence models

- Gaussian random field: $u(\mathbf{s})$, $\mathbf{s} \in \mathcal{D}$ (subset of $\mathbb{R}^d$ or a manifold)
- Moment characterisation:
  - Expectation $\mu(\mathbf{s}) = \mathbb{E}[u(\mathbf{s})]$
  - Covariance $\mathcal{R}(\mathbf{s}, \mathbf{s}') = \mathbb{C}[u(\mathbf{s}), u(\mathbf{s}')]$, symmetric positive definite function.
- Precision operator; inverse covariance: $\mathcal{Q} = \mathcal{R}^{-1}$
  In practice, easier conditions for valid models
- Reproducing Kernel Hilbert Space (RKHS) $H_{\mathcal{Q}}$: Inner product

$$\langle f, g \rangle_{H_{\mathcal{Q}}} = \langle f, \mathcal{Q} g \rangle_{\mathcal{D}}$$

  and squared norm $\|f\|^2 = \langle f, f \rangle_{H_{\mathcal{Q}}}$
- $m(\cdot) = \mathbb{E}(u(\cdot) - \mu(\cdot) | \{u(\mathbf{s}_k)\}) \in H_{\mathcal{Q}}$ but $u(\cdot) - \mu(\cdot) \notin H_{\mathcal{Q}}$; the process is less smooth!
- Spatial and spatio-temporal stochastic PDEs generate random field models:

$$\mathcal{L} u(\mathbf{s}) = \dot{\mathcal{W}}(\mathbf{s})$$

$$\mathcal{Q}_u = \mathcal{L}^* \mathcal{L}$$

$$\langle f, g \rangle_{H_{\mathcal{Q}}} = \langle \mathcal{L} f, \mathcal{L} g \rangle_{\mathcal{D}}$$

  Can work directly with the precision or inner product; no need to know the covariance!

The eternal quest for spatial dependence models

- Gaussian random field: $u(\mathbf{s})$, $\mathbf{s} \in \mathcal{D}$ (subset of $\mathbb{R}^d$ or a manifold)
- Moment characterisation:
  - Expectation $\mu(\mathbf{s}) = \mathbb{E}[u(\mathbf{s})]$
  - Covariance $\mathcal{R}(\mathbf{s}, \mathbf{s}') = \mathbb{C}[u(\mathbf{s}), u(\mathbf{s}')]$, symmetric positive definite function.
- Precision operator; inverse covariance: $\mathcal{Q} = \mathcal{R}^{-1}$

  In practice, easier conditions for valid models
- Reproducing Kernel Hilbert Space (RKHS) $H_\mathcal{Q}$: Inner product

$$\langle f, g \rangle_{H_\mathcal{Q}} = \langle f, \mathcal{Q} g \rangle_\mathcal{D}$$

  and squared norm $\|f\|^2 = \langle f, f \rangle_{H_\mathcal{Q}}$
- $m(\cdot) = \mathbb{E}(u(\cdot) - \mu(\cdot) | \{u(\mathbf{s}_k)\}) \in H_\mathcal{Q}$ but $u(\cdot) - \mu(\cdot) \notin H_\mathcal{Q}$; the process is less smooth!
- Spatial and spatio-temporal stochastic PDEs generate random field models:

$$\mathcal{L} u(\mathbf{s}) = \dot{\mathcal{W}}(\mathbf{s})$$

$$\mathcal{Q}_u = \mathcal{L}^* \mathcal{L}$$

$$\langle f, g \rangle_{H_\mathcal{Q}} = \langle \mathcal{L} f, \mathcal{L} g \rangle_\mathcal{D}$$

  Can work directly with the precision or inner product; no need to know the covariance!

Direct Bayesian inference: The inner core of the Integrated Nested Laplace Approximation method

- Latent Gaussian model structure

$$\boldsymbol{\theta} \sim p(\boldsymbol{\theta}) \quad \text{(precision parameters)} \qquad \eta(\mathbf{s}) = \sum_{k=1}^{n} \psi(\mathbf{s}) u_k \quad \text{(predictor)}$$

$$\boldsymbol{u}|\boldsymbol{\theta} \sim \mathsf{N}[\boldsymbol{\mu}_u, \boldsymbol{Q}_u^{-1}] \quad \text{(latent field)} \qquad \boldsymbol{y}|\boldsymbol{\theta}, \boldsymbol{u} \sim p(\boldsymbol{y}|\boldsymbol{\theta}, \eta) \quad \text{(observations)}$$

- Conditional log-posterior mode ($\boldsymbol{\mu}_{u|y}$) and Hessian ($\boldsymbol{Q}_{u|y}$), for each $\boldsymbol{\theta}$, by iteration:

$$\boldsymbol{g}_y^* = -\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{u}} \log p(\boldsymbol{y}|\boldsymbol{\theta}, \eta)\Big|_{\boldsymbol{u}=\boldsymbol{u}^*}$$

$$\boldsymbol{H}_y^* = -\frac{\mathrm{d}^2}{\mathrm{d}\boldsymbol{u}\mathrm{d}\boldsymbol{u}^\top} \log p(\boldsymbol{y}|\boldsymbol{\theta}, \eta)\Big|_{\boldsymbol{u}=\boldsymbol{u}^*}$$

$$\boldsymbol{Q}_{u|y} = \boldsymbol{Q}_u + \boldsymbol{H}_y^*$$

$$\boldsymbol{Q}_{u|y}(\boldsymbol{\mu}_{u|y} - \boldsymbol{\mu}_u) = \boldsymbol{H}_u^*(\boldsymbol{u}^* - \boldsymbol{\mu}_u) - \boldsymbol{g}_y^*$$

Direct Bayesian inference: The inner core of the Integrated Nested Laplace Approximation method

- Latent Gaussian model structure

$$\boldsymbol{\theta} \sim p(\boldsymbol{\theta}) \quad \text{(precision parameters)} \qquad \eta(\mathbf{s}) = \sum_{k=1}^{n} \psi(\mathbf{s}) u_k \quad \text{(predictor)}$$

$$\boldsymbol{u}|\boldsymbol{\theta} \sim \mathrm{N}[\boldsymbol{\mu}_u, \boldsymbol{Q}_u^{-1}] \quad \text{(latent field)} \qquad \boldsymbol{y}|\boldsymbol{\theta},\boldsymbol{u} \sim p(\boldsymbol{y}|\boldsymbol{\theta},\eta) \quad \text{(observations)}$$

- Conditional log-posterior mode ($\boldsymbol{\mu}_{u|y}$) and Hessian ($\boldsymbol{Q}_{u|y}$), for each $\boldsymbol{\theta}$, by iteration:

$$\boldsymbol{g}_y^* = -\frac{\mathrm{d}}{\mathrm{d}\boldsymbol{u}} \log p(\boldsymbol{y}|\boldsymbol{\theta},\eta)\Big|_{\boldsymbol{u}=\boldsymbol{u}^*}$$

$$\boldsymbol{H}_y^* = -\frac{\mathrm{d}^2}{\mathrm{d}\boldsymbol{u}\mathrm{d}\boldsymbol{u}^\top} \log p(\boldsymbol{y}|\boldsymbol{\theta},\eta)\Big|_{\boldsymbol{u}=\boldsymbol{u}^*}$$

$$\boldsymbol{Q}_{u|y} = \boldsymbol{Q}_u + \boldsymbol{H}_y^*$$

$$\boldsymbol{Q}_{u|y}(\boldsymbol{\mu}_{u|y} - \boldsymbol{\mu}_u) = \boldsymbol{H}_u^*(\boldsymbol{u}^* - \boldsymbol{\mu}_u) - \boldsymbol{g}_y^*$$

Spatio-temporal separability for functions, covariances, and precisions

- Functional separability for $\mathbf{s} \in \mathcal{D}$ and $t \in \mathcal{T}$
  - Addition: $w(\mathbf{s}, t) = u(\mathbf{s}) + v(t)$
  - Multiplication $w(\mathbf{s}, t) = u(\mathbf{s})v(t)$ (degrees of freedom $|\mathcal{D}| + |\mathcal{T}|$)
- Covariance separability
  - Addition: $\mathcal{R}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \mathcal{R}_u(\mathbf{s}, \mathbf{s}') + \mathcal{R}_v(t, t')$
  - Multiplication $\mathcal{R}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \mathcal{R}_u(\mathbf{s}, \mathbf{s}')\mathcal{R}_v(t, t')$ (degrees of freedom $|\mathcal{D}| \cdot |\mathcal{T}|$)
  - Simple to construct, but with some unrealistic properties
- "Precision separability"
  - Addition: $\mathcal{Q}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \mathcal{Q}_u(\mathbf{s}, \mathbf{s}') + \mathcal{Q}_v(t, t')$
    but $\mathcal{R}_w = \mathcal{Q}_w^{-1} = (\mathcal{Q}_u + \mathcal{Q}_v)^{-1}$ (not covariance separable!)
  - Multiplication $\mathcal{Q}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \mathcal{Q}_u(\mathbf{s}, \mathbf{s}')\mathcal{Q}_v(t, t')$
    and $\mathcal{R}_w = \mathcal{Q}_w^{-1} = \mathcal{Q}_u^{-1}\mathcal{Q}_v^{-1}$ (covariance separable!)

Idea: A simple more general structure, combining multiplication and addition of precisions:

$$\mathcal{Q}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \sum_k \mathcal{Q}_{u_k}(\mathbf{s}, \mathbf{s}')\mathcal{Q}_{v_k}(t, t')$$

Question: Are there interpretable process constructions that lead to this structure?

Spatio-temporal separability for functions, covariances, and precisions

- Functional separability for $\mathbf{s} \in \mathcal{D}$ and $t \in \mathcal{T}$
  - Addition: $w(\mathbf{s}, t) = u(\mathbf{s}) + v(t)$
  - Multiplication $w(\mathbf{s}, t) = u(\mathbf{s})v(t)$ (degrees of freedom $|\mathcal{D}| + |\mathcal{T}|$)
- Covariance separability
  - Addition: $\mathcal{R}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \mathcal{R}_u(\mathbf{s}, \mathbf{s}') + \mathcal{R}_v(t, t')$
  - Multiplication $\mathcal{R}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \mathcal{R}_u(\mathbf{s}, \mathbf{s}')\mathcal{R}_v(t, t')$ (degrees of freedom $|\mathcal{D}| \cdot |\mathcal{T}|$)
  - Simple to construct, but with some unrealistic properties
- "Precision separability"
  - Addition: $\mathcal{Q}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \mathcal{Q}_u(\mathbf{s}, \mathbf{s}') + \mathcal{Q}_v(t, t')$
    but $\mathcal{R}_w = \mathcal{Q}_w^{-1} = (\mathcal{Q}_u + \mathcal{Q}_v)^{-1}$ (not covariance separable!)
  - Multiplication $\mathcal{Q}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \mathcal{Q}_u(\mathbf{s}, \mathbf{s}')\mathcal{Q}_v(t, t')$
    and $\mathcal{R}_w = \mathcal{Q}_w^{-1} = \mathcal{Q}_u^{-1}\mathcal{Q}_v^{-1}$ (covariance separable!)

Idea: A simple more general structure, combining multiplication and addition of precisions:

$$\mathcal{Q}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \sum_k \mathcal{Q}_{u_k}(\mathbf{s}, \mathbf{s}')\mathcal{Q}_{v_k}(t, t')$$

Question: Are there interpretable process constructions that lead to this structure?

Spatio-temporal separability for functions, covariances, and precisions

- Functional separability for $\mathbf{s} \in \mathcal{D}$ and $t \in \mathcal{T}$
  - Addition: $w(\mathbf{s}, t) = u(\mathbf{s}) + v(t)$
  - Multiplication $w(\mathbf{s}, t) = u(\mathbf{s})v(t)$ (degrees of freedom $|\mathcal{D}| + |\mathcal{T}|$)
- Covariance separability
  - Addition: $\mathcal{R}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \mathcal{R}_u(\mathbf{s}, \mathbf{s}') + \mathcal{R}_v(t, t')$
  - Multiplication $\mathcal{R}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \mathcal{R}_u(\mathbf{s}, \mathbf{s}')\mathcal{R}_v(t, t')$ (degrees of freedom $|\mathcal{D}| \cdot |\mathcal{T}|$)
  - Simple to construct, but with some unrealistic properties
- "Precision separability"
  - Addition: $\mathcal{Q}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \mathcal{Q}_u(\mathbf{s}, \mathbf{s}') + \mathcal{Q}_v(t, t')$
    but $\mathcal{R}_w = \mathcal{Q}_w^{-1} = (\mathcal{Q}_u + \mathcal{Q}_v)^{-1}$ (not covariance separable!)
  - Multiplication $\mathcal{Q}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \mathcal{Q}_u(\mathbf{s}, \mathbf{s}')\mathcal{Q}_v(t, t')$
    and $\mathcal{R}_w = \mathcal{Q}_w^{-1} = \mathcal{Q}_u^{-1}\mathcal{Q}_v^{-1}$ (covariance separable!)

Idea: A simple more general structure, combining multiplication and addition of precisions:

$$\mathcal{Q}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \sum_k \mathcal{Q}_{u_k}(\mathbf{s}, \mathbf{s}')\mathcal{Q}_{v_k}(t, t')$$

Question: Are there interpretable process constructions that lead to this structure?

Spatio-temporal separability for functions, covariances, and precisions

- Functional separability for $\mathbf{s} \in \mathcal{D}$ and $t \in \mathcal{T}$
  - Addition: $w(\mathbf{s}, t) = u(\mathbf{s}) + v(t)$
  - Multiplication $w(\mathbf{s}, t) = u(\mathbf{s})v(t)$ (degrees of freedom $|\mathcal{D}| + |\mathcal{T}|$)
- Covariance separability
  - Addition: $\mathcal{R}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \mathcal{R}_u(\mathbf{s}, \mathbf{s}') + \mathcal{R}_v(t, t')$
  - Multiplication $\mathcal{R}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \mathcal{R}_u(\mathbf{s}, \mathbf{s}')\mathcal{R}_v(t, t')$ (degrees of freedom $|\mathcal{D}| \cdot |\mathcal{T}|$)
  - Simple to construct, but with some unrealistic properties
- "Precision separability"
  - Addition: $\mathcal{Q}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \mathcal{Q}_u(\mathbf{s}, \mathbf{s}') + \mathcal{Q}_v(t, t')$
    but $\mathcal{R}_w = \mathcal{Q}_w^{-1} = (\mathcal{Q}_u + \mathcal{Q}_v)^{-1}$ (not covariance separable!)
  - Multiplication $\mathcal{Q}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \mathcal{Q}_u(\mathbf{s}, \mathbf{s}')\mathcal{Q}_v(t, t')$
    and $\mathcal{R}_w = \mathcal{Q}_w^{-1} = \mathcal{Q}_u^{-1}\mathcal{Q}_v^{-1}$ (covariance separable!)

Idea: A simple more general structure, combining multiplication and addition of precisions:

$$\mathcal{Q}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \sum_k \mathcal{Q}_{u_k}(\mathbf{s}, \mathbf{s}')\mathcal{Q}_{v_k}(t, t')$$

Question: Are there interpretable process constructions that lead to this structure?

Spatio-temporal separability for functions, covariances, and precisions

- Functional separability for $\mathbf{s} \in \mathcal{D}$ and $t \in \mathcal{T}$
  - Addition: $w(\mathbf{s}, t) = u(\mathbf{s}) + v(t)$
  - Multiplication $w(\mathbf{s}, t) = u(\mathbf{s})v(t)$ (degrees of freedom $|\mathcal{D}| + |\mathcal{T}|$)
- Covariance separability
  - Addition: $\mathcal{R}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \mathcal{R}_u(\mathbf{s}, \mathbf{s}') + \mathcal{R}_v(t, t')$
  - Multiplication $\mathcal{R}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \mathcal{R}_u(\mathbf{s}, \mathbf{s}')\mathcal{R}_v(t, t')$ (degrees of freedom $|\mathcal{D}| \cdot |\mathcal{T}|$)
  - Simple to construct, but with some unrealistic properties
- "Precision separability"
  - Addition: $\mathcal{Q}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \mathcal{Q}_u(\mathbf{s}, \mathbf{s}') + \mathcal{Q}_v(t, t')$
    but $\mathcal{R}_w = \mathcal{Q}_w^{-1} = (\mathcal{Q}_u + \mathcal{Q}_v)^{-1}$ (not covariance separable!)
  - Multiplication $\mathcal{Q}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \mathcal{Q}_u(\mathbf{s}, \mathbf{s}')\mathcal{Q}_v(t, t')$
    and $\mathcal{R}_w = \mathcal{Q}_w^{-1} = \mathcal{Q}_u^{-1}\mathcal{Q}_v^{-1}$ (covariance separable!)

Idea: A simple more general structure, combining multiplication and addition of precisions:

$$\mathcal{Q}_w[(\mathbf{s}, t), (\mathbf{s}', t')] = \sum_k \mathcal{Q}_{u_k}(\mathbf{s}, \mathbf{s}')\mathcal{Q}_{v_k}(t, t')$$

Question: Are there interpretable process constructions that lead to this structure?

From temporal random walks to spatio-temporal diffusion

- Spatial Whittle-Matérn models with $\mathcal{L}_s = \gamma_s^2 - \Delta$:

$$\mathcal{L}_s^{\alpha_s/2} u(\mathbf{s}) = \dot{\mathcal{W}}(\mathbf{s}) \quad \text{(spatial white noise)}$$

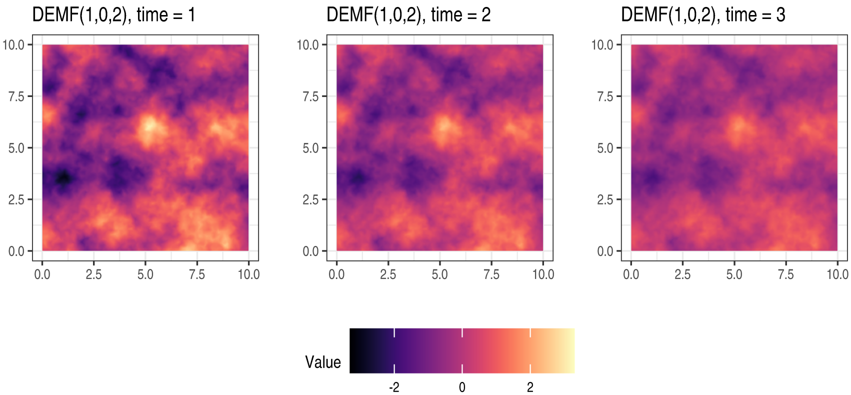
  Precision $\mathcal{Q} = \mathcal{L}_s^{\alpha_s}$, Matérn covariance on $\mathbb{R}^d$.

- Separable space-time model (separable vector Ornstein-Uhlenbeck/AR(1) process):

$$\left(\frac{\partial}{\partial t} + \kappa\right) \mathcal{L}_s^{\alpha_s/2} u(\mathbf{s}, t) = \dot{\mathcal{W}}(\mathbf{s}, t) \quad \text{(spatio-temporal white noise)}$$

  Precision $\mathcal{Q} = \left(\kappa^2 - \frac{\partial^2}{\partial t^2}\right) \mathcal{L}_s^{\alpha_s}$, covariance is a product of a temporal Matérn kernel and the spatial covariance.

Prediction



Conditional expectations into the future decay pointwise towards zero; no spatial dynamics.

Diffusion extension of Matérn fields (DEMF)

- Non-separable space-time DEMF$(\alpha_t, \alpha_s, \alpha_e)$ model for $(\mathbf{s}, t) \in \mathcal{D} \times \mathcal{T}$:

$$\gamma_e \mathcal{L}_s^{\alpha_e/2} \left(\gamma_t \frac{\partial}{\partial t} + \mathcal{L}_s^{\alpha_s/2} \right)^{\alpha_t} u(\mathbf{s}, t) = \gamma_e \mathcal{L}_s^{\alpha_e/2} \left(-\gamma_t^2 \frac{\partial^2}{\partial t^2} + \mathcal{L}_s^{\alpha_s} \right)^{\alpha_t/2} u(\mathbf{s}, t) = \dot{\mathcal{W}}(\mathbf{s}, t),$$

  where $\gamma_e, \gamma_t > 0$, and $\alpha_t > 0$, $\alpha_s, \alpha_e \geq 0$.
- In the stationary case, the resulting field has Matérn covariance for every time point
- The spatial smoothness is $\nu_s = \alpha_s(\alpha_t - 1/2) + \alpha_e - d/2$
- The temporal smoothness is $\nu_t = \min[\alpha_t - 1/2, \nu_s/\alpha_s]$.
- Non-separability parameter: $\beta_s = 1 - \frac{\alpha_e}{\nu_s + d/2} \in [0, 1]$
- Tensor product basis discretisation for integer α_t gives precision matrix structure

$$\boldsymbol{Q} = \gamma_e^2 \sum_{k=0}^{2\alpha_t} \gamma_t^k \boldsymbol{J}_{\alpha_t, k/2} \otimes \boldsymbol{K}_{\alpha_s(\alpha_t - k/2) + \alpha_e}$$

  where $\boldsymbol{J}_{\cdot,\cdot}$ are purely temporal and $\boldsymbol{K}_{\cdot}$ are purely spatial.
  This is what we were looking for!

Diffusion extension of Matérn fields (DEMF)

- Non-separable space-time DEMF$(\alpha_t, \alpha_s, \alpha_e)$ model for $(\mathbf{s}, t) \in \mathcal{D} \times \mathcal{T}$:

$$\gamma_e \mathcal{L}_s^{\alpha_e/2} \left(\gamma_t \frac{\partial}{\partial t} + \mathcal{L}_s^{\alpha_s/2} \right)^{\alpha_t} u(\mathbf{s}, t) = \gamma_e \mathcal{L}_s^{\alpha_e/2} \left(-\gamma_t^2 \frac{\partial^2}{\partial t^2} + \mathcal{L}_s^{\alpha_s} \right)^{\alpha_t/2} u(\mathbf{s}, t) = \dot{\mathcal{W}}(\mathbf{s}, t),$$

  where $\gamma_e, \gamma_t > 0$, and $\alpha_t > 0$, $\alpha_s, \alpha_e \geq 0$.
- In the stationary case, the resulting field has Matérn covariance for every time point
- The spatial smoothness is $\nu_s = \alpha_s(\alpha_t - 1/2) + \alpha_e - d/2$
- The temporal smoothness is $\nu_t = \min[\alpha_t - 1/2, \nu_s/\alpha_s]$.
- Non-separability parameter: $\beta_s = 1 - \frac{\alpha_e}{\nu_s + d/2} \in [0, 1]$
- Tensor product basis discretisation for integer α_t gives precision matrix structure

$$\boldsymbol{Q} = \gamma_e^2 \sum_{k=0}^{2\alpha_t} \gamma_t^k \boldsymbol{J}_{\alpha_t, k/2} \otimes \boldsymbol{K}_{\alpha_s(\alpha_t - k/2) + \alpha_e}$$

  where $\boldsymbol{J}_{\cdot,\cdot}$ are purely temporal and $\boldsymbol{K}_{\cdot}$ are purely spatial.
  This is what we were looking for!

Diffusion extension of Matérn fields (DEMF)

- Non-separable space-time DEMF$(\alpha_t, \alpha_s, \alpha_e)$ model for $(\mathbf{s}, t) \in \mathcal{D} \times \mathcal{T}$:

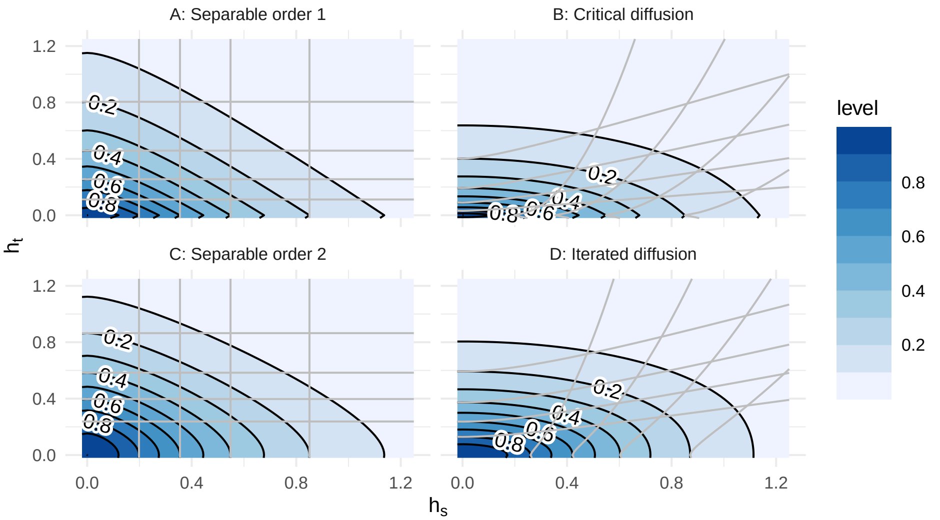
$$\gamma_e \mathcal{L}_s^{\alpha_e/2} \left(\gamma_t \frac{\partial}{\partial t} + \mathcal{L}_s^{\alpha_s/2} \right)^{\alpha_t} u(\mathbf{s}, t) = \gamma_e \mathcal{L}_s^{\alpha_e/2} \left(-\gamma_t^2 \frac{\partial^2}{\partial t^2} + \mathcal{L}_s^{\alpha_s} \right)^{\alpha_t/2} u(\mathbf{s}, t) = \dot{\mathcal{W}}(\mathbf{s}, t),$$

  where $\gamma_e, \gamma_t > 0$, and $\alpha_t > 0$, $\alpha_s, \alpha_e \geq 0$.
- In the stationary case, the resulting field has Matérn covariance for every time point
- The spatial smoothness is $\nu_s = \alpha_s(\alpha_t - 1/2) + \alpha_e - d/2$
- The temporal smoothness is $\nu_t = \min[\alpha_t - 1/2, \nu_s/\alpha_s]$.
- Non-separability parameter: $\beta_s = 1 - \frac{\alpha_e}{\nu_s + d/2} \in [0, 1]$
- Tensor product basis discretisation for integer α_t gives precision matrix structure

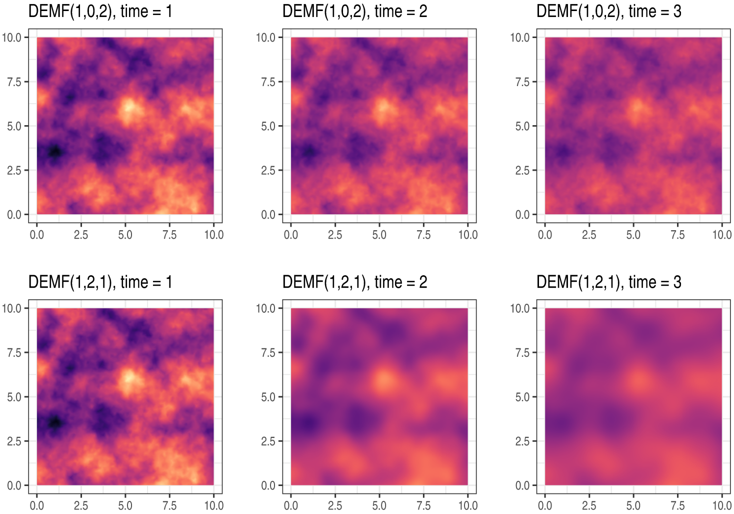
$$\boldsymbol{Q} = \gamma_e^2 \sum_{k=0}^{2\alpha_t} \gamma_t^k \boldsymbol{J}_{\alpha_t, k/2} \otimes \boldsymbol{K}_{\alpha_s(\alpha_t - k/2) + \alpha_e}$$

  where $\boldsymbol{J}_{\cdot,\cdot}$ are purely temporal and $\boldsymbol{K}_{\cdot}$ are purely spatial.
  This is what we were looking for!

Non-separable covariances, from spectral inversion



Prediction


DEMF(1,0,2), time = 1
DEMF(1,0,2), time = 2
DEMF(1,0,2), time = 3
DEMF(1,2,1), time = 1
DEMF(1,2,1), time = 2
DEMF(1,2,1), time = 3
0.0
2.5
5.0
7.5
10.0

Step selection analysis with telemetry data

Goal: Understand sequential movement decisions

- The general movement capacity of an animal. Expressed by a movement kernel:

$$K(\mathbf{y}_t|\mathbf{y}_{t-1},\mathbf{y}_{t-2},\boldsymbol{\theta}) = K_{\text{length}}(\mathbf{y}_t|\mathbf{y}_{t-1},\boldsymbol{\theta})K_{\text{angle}}(\mathbf{y}_t|\mathbf{y}_{t-1},\mathbf{y}_{t-2},\boldsymbol{\theta}), \quad \mathbf{y}_\cdot \in \mathcal{D} \subset \mathbb{R}^2$$

- Selection behaviour of the animal. Modelled by a resource selection function (RSF):

$$\xi(\mathbf{s}) = \exp[\eta(\mathbf{s})] = \exp[\beta_1 X_1(\mathbf{s}) + \ldots + \beta_p X_p(\mathbf{s}) + u(\mathbf{s})], \quad \mathbf{s} \in \mathcal{D}$$

  Spatially (or spatio-temporally) varying covariates $X_\cdot$ and a residual random field $u(\mathbf{s})$.

- Combined normalised conditional observation density function:

$$f_{t|<t}(\mathbf{y}_t|\boldsymbol{\theta},\eta) = \frac{K(\mathbf{y}_t|\mathbf{y}_{<t},\boldsymbol{\theta})\exp[\eta(\mathbf{y}_t)]}{\int_{\mathcal{D}} K(\mathbf{s}|\mathbf{y}_{<t},\boldsymbol{\theta})\exp[\eta(\mathbf{s})]\,\mathrm{d}\mathbf{s}}$$

Step selection analysis with telemetry data

Goal: Understand sequential movement decisions

- The general movement capacity of an animal. Expressed by a movement kernel:

$$K(\mathbf{y}_t|\mathbf{y}_{t-1},\mathbf{y}_{t-2},\boldsymbol{\theta}) = K_{\text{length}}(\mathbf{y}_t|\mathbf{y}_{t-1},\boldsymbol{\theta})K_{\text{angle}}(\mathbf{y}_t|\mathbf{y}_{t-1},\mathbf{y}_{t-2},\boldsymbol{\theta}), \quad \mathbf{y}_\cdot \in \mathcal{D} \subset \mathbb{R}^2$$

- Selection behaviour of the animal. Modelled by a resource selection function (RSF):

$$\xi(\mathbf{s}) = \exp[\eta(\mathbf{s})] = \exp[\beta_1 X_1(\mathbf{s}) + \ldots + \beta_p X_p(\mathbf{s}) + u(\mathbf{s})], \quad \mathbf{s} \in \mathcal{D}$$

  Spatially (or spatio-temporally) varying covariates $X_\cdot$ and a residual random field $u(\mathbf{s})$.

- Combined normalised conditional observation density function:

$$f_{t|<t}(\mathbf{y}_t|\boldsymbol{\theta},\eta) = \frac{K(\mathbf{y}_t|\mathbf{y}_{<t},\boldsymbol{\theta})\exp[\eta(\mathbf{y}_t)]}{\int_{\mathcal{D}} K(\mathbf{s}|\mathbf{y}_{<t},\boldsymbol{\theta})\exp[\eta(\mathbf{s})]\,\mathrm{d}\mathbf{s}}$$

Step selection analysis with telemetry data

Goal: Understand sequential movement decisions

- The general movement capacity of an animal. Expressed by a movement kernel:

$$K(\mathbf{y}_t|\mathbf{y}_{t-1},\mathbf{y}_{t-2},\boldsymbol{\theta}) = K_{\text{length}}(\mathbf{y}_t|\mathbf{y}_{t-1},\boldsymbol{\theta})K_{\text{angle}}(\mathbf{y}_t|\mathbf{y}_{t-1},\mathbf{y}_{t-2},\boldsymbol{\theta}), \quad \mathbf{y}_\cdot \in \mathcal{D} \subset \mathbb{R}^2$$

- Selection behaviour of the animal. Modelled by a resource selection function (RSF):

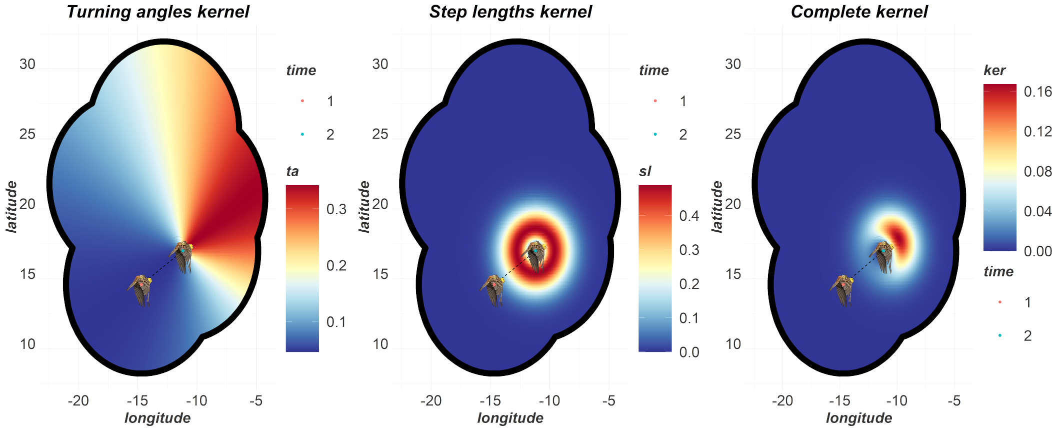
$$\xi(\mathbf{s}) = \exp[\eta(\mathbf{s})] = \exp[\beta_1 X_1(\mathbf{s}) + \ldots + \beta_p X_p(\mathbf{s}) + u(\mathbf{s})], \quad \mathbf{s} \in \mathcal{D}$$

  Spatially (or spatio-temporally) varying covariates $X_\cdot$ and a residual random field $u(\mathbf{s})$.
- Combined normalised conditional observation density function:

$$f_{t|<t}(\mathbf{y}_t|\boldsymbol{\theta},\eta) = \frac{K(\mathbf{y}_t|\mathbf{y}_{<t},\boldsymbol{\theta})\exp[\eta(\mathbf{y}_t)]}{\int_{\mathcal{D}} K(\mathbf{s}|\mathbf{y}_{<t},\boldsymbol{\theta})\exp[\eta(\mathbf{s})]\,\mathrm{d}\mathbf{s}}$$

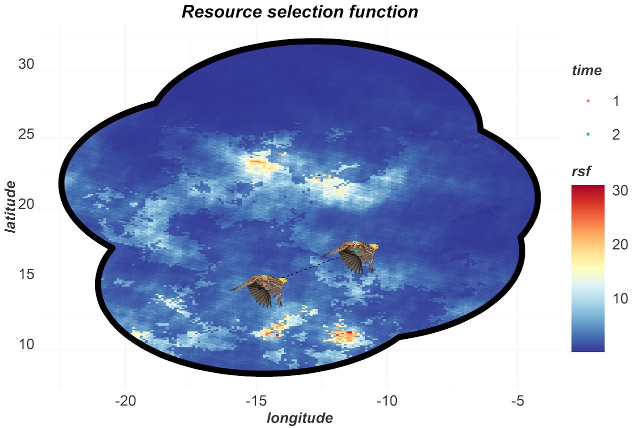
Movement kernel

Movement capacity of an animal:



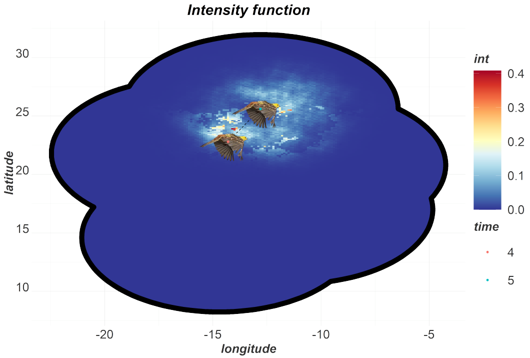
Resource selection function

Spatial features in the study area:

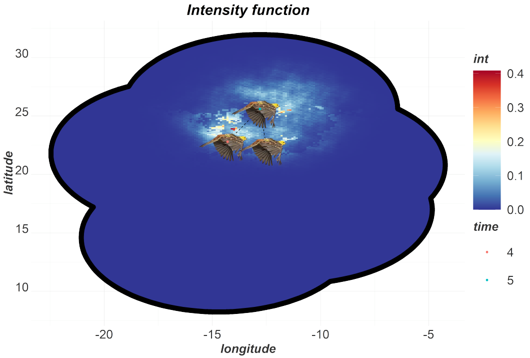


Combined effect

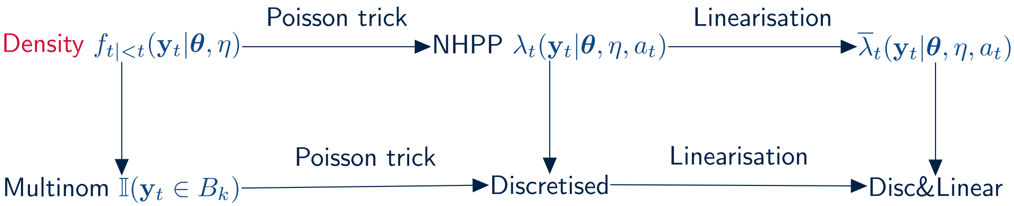
Intensity function:



Movement decision!:



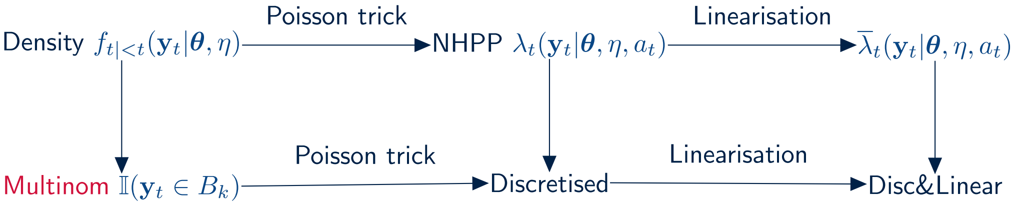
From movement kernel to discretised point process likelihood



$$f_{t|<t}(\mathbf{y}_t|\boldsymbol{\theta},\eta) = \frac{K(\mathbf{y}_t|\mathbf{y}_{<t},\boldsymbol{\theta})\exp[\eta(\mathbf{y}_t)]}{\int_{\mathcal{D}} K(\mathbf{s}|\mathbf{y}_{<t},\boldsymbol{\theta})\exp[\eta(\mathbf{s})]\,\mathrm{d}\mathbf{s}}$$

Problem: Inconvenient normalisation integral.

From movement kernel to discretised point process likelihood



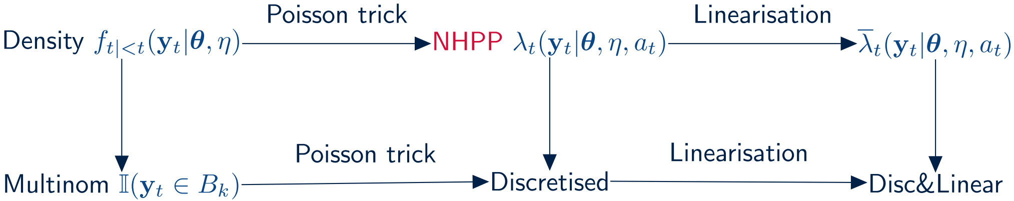
$$
f_{t|<t}(\mathbf{y}_t|\boldsymbol{\theta},\eta) = \frac{K(\mathbf{y}_t|\mathbf{y}_{<t},\boldsymbol{\theta})\exp[\eta(\mathbf{y}_t)]}{\int_{\mathcal{D}} K(\mathbf{s}|\mathbf{y}_{<t},\boldsymbol{\theta})\exp[\eta(\mathbf{s})]\,\mathrm{ds}}
$$

Previous approach: Subdivide space into disjoint sets B_k, with $\mathcal{D}=\cup_{k=1}^N B_k$.

$$
\mathbf{z}_t = [\mathbb{I}(\mathbf{y}_t\in B_1),\dots,\mathbb{I}(\mathbf{y}_t\in B_N)] \sim \text{Multinomial}\,(1,\{p_k,k=1,\dots,N\})
$$

$$
p_k = \mathbb{P}(\mathbf{y}_t\in B_k|\mathbf{y}_{<t},\boldsymbol{\theta},\eta) = \int_{B_k} f_{t|<t}(\mathbf{s}|\boldsymbol{\theta},\eta)\,\mathrm{ds}
$$

From movement kernel to discretised point process likelihood

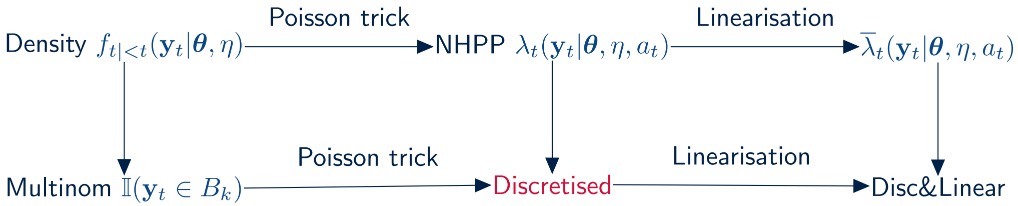


$$\lambda_t(\mathbf{y}_t|\boldsymbol{\theta},\eta,a_t)=K(\mathbf{y}_t|\mathbf{y}_{<t},\boldsymbol{\theta})\exp[\eta(\mathbf{y}_t)+a_t],\quad a_t\sim\text{Unif}(\mathbb{R})$$

$$l(\{\mathbf{y}_t\}|\boldsymbol{\theta},\eta,\{a_t\})=-\sum_t\int_{\mathcal{D}}\lambda_t(\mathbf{s}|\boldsymbol{\theta},\eta,a_t)\,\mathrm{d}\mathbf{s}+\sum_t\log\lambda_t(\mathbf{y}_t|\boldsymbol{\theta},\eta,a_t)$$

Non-homogeneous Poisson point process with a single point observation for each t.
a_t replaces the explicit density normalisation by *estimating* it.
The posterior distribution for $\boldsymbol{\theta}$, $\beta_\cdot$, and $u(\cdot)$ is unchanged!

From movement kernel to discretised point process likelihood

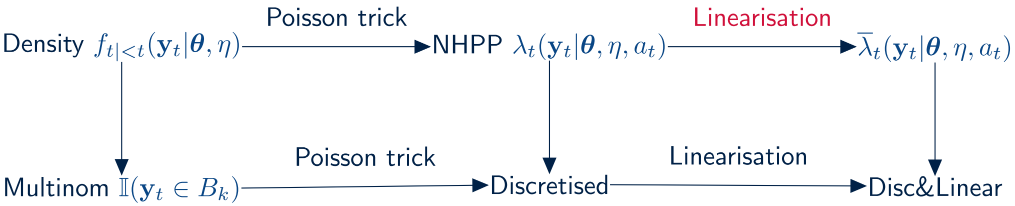


$$\lambda_t(\mathbf{y}_t|\boldsymbol{\theta},\eta,a_t) = K(\mathbf{y}_t|\mathbf{y}_{<t},\boldsymbol{\theta})\exp[\eta(\mathbf{y}_t)+a_t], \quad a_t\sim\text{Unif}(\mathbb{R})$$

$$l(\{\mathbf{y}_t\}|\boldsymbol{\theta},\eta,\{a_t\}) \approx -\sum_t\sum_k \lambda_t(\mathbf{s}_k|\boldsymbol{\theta},\eta,a_t)w_k + \sum_t \log\lambda_t(\mathbf{y}_t|\boldsymbol{\theta},\eta,a_t)$$

Integration points and weigths $(\mathbf{s}_k, w_k)$, adapted to the spatial model resolution.

From movement kernel to discretised point process likelihood

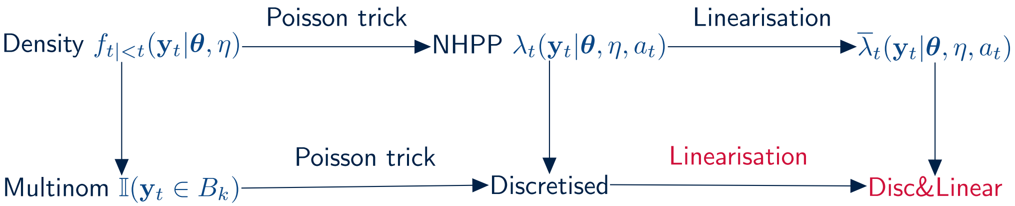


$$\log \overline{\lambda}(\mathbf{y}_t|\boldsymbol{\theta},\eta,a_t) = \log K(\mathbf{y}_t|\mathbf{y}_{<t},\boldsymbol{\theta}_0) + \frac{\mathrm{d}\log K(\mathbf{y}_t|\mathbf{y}_{<t},\boldsymbol{\theta})}{\mathrm{d}\boldsymbol{\theta}}(\boldsymbol{\theta}-\boldsymbol{\theta}_0) + \eta(\mathbf{y}_t) + a_t$$

$$l(\{\mathbf{y}_t\}|\boldsymbol{\theta},\eta,\{a_t\}) = -\sum_t \int_{\mathcal{D}} \overline{\lambda}_t(\mathbf{s}|\boldsymbol{\theta},\eta,a_t)\,\mathrm{d}\mathbf{s} + \sum_t \log \overline{\lambda}_t(\mathbf{y}_t|\boldsymbol{\theta},\eta,a_t)$$

(Iterative) linearisation to a log-linear point process intensity allows more general movement kernel parameterisation.
(Preliminary theory: posterior approximation related to Fischer scoring)

From movement kernel to discretised point process likelihood



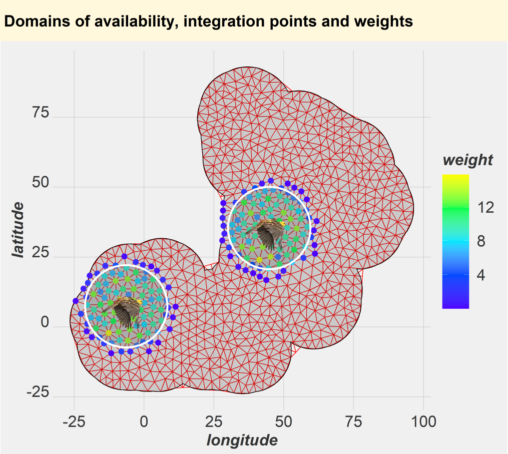
$$\log\overline{\lambda}(\mathbf{y}_t|\boldsymbol{\theta},\eta,a_t)=\log K(\mathbf{y}_t|\mathbf{y}_{<t},\boldsymbol{\theta}_0)+\frac{\mathrm{d}\log K(\mathbf{y}_t|\mathbf{y}_{<t},\boldsymbol{\theta})}{\mathrm{d}\boldsymbol{\theta}}(\boldsymbol{\theta}-\boldsymbol{\theta}_0)+\eta(\mathbf{y}_t)+a_t$$

$$l(\{\mathbf{y}_t\}|\boldsymbol{\theta},\eta,\{a_t\})\approx-\sum_t\sum_k\overline{\lambda}_t(\mathbf{s}_k|\boldsymbol{\theta},\eta,a_t)w_k+\sum_t\log\overline{\lambda}_t(\mathbf{y}_t|\boldsymbol{\theta},\eta,a_t)$$

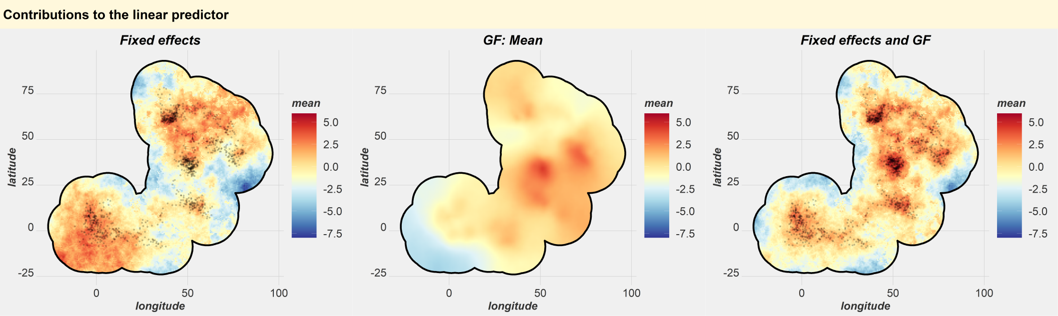
This is *almost* a log-linear Poisson count log-likelihood;
In $-E\lambda+y\log(E\lambda)$, R-INLA allows us to specify the two terms separately, without pairing them up with equal E and λ values.

Mesh, integration points and weights

- Restricted domain of availability at each time point: Disk with radius (at least) equal to the maximum observed step length
- Integration points: At mesh nodes to ensure stability
- Deterministic integration: Previous Monte Carlo strategies are inefficient and unstable

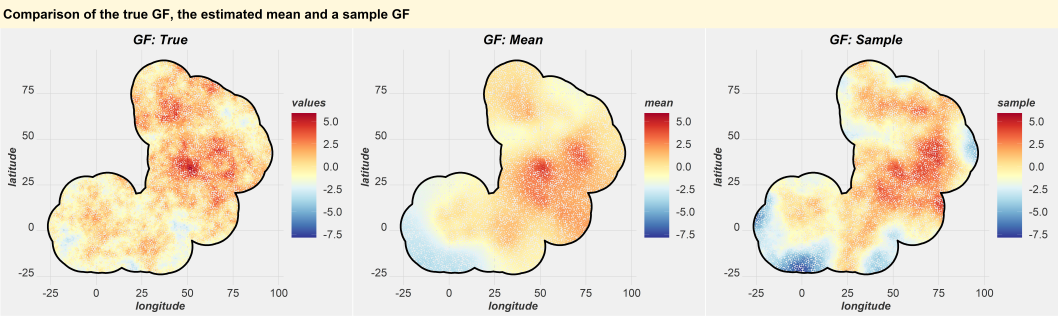


Estimated log-intensity function



The Gaussian random field (GF) contribution improves the estimated animal abundance

Estimated Gaussian random field (GF)



Posterior samples can be used to quantify uncertainty of the fields and linear/nonlinear functionals of the fields.

Note: Recall that conditional means are fundamentally smoother than conditional realisations!

Summary

- Non-separability and non-stationarity are distinct concepts; both needed
- (Relatively) simple stochastic PDEs provide useful building blocks
- Computational methods need to handle hierarchical structures, not just additive noise.
- The Poisson trick & iterative linearisation allows `inlabru` to estimate new model classes
- The SPDE approach for Gaussian and non-Gaussian fields: 10 years and still running (Lindgren et al, 2020, Spatial Statistics)
  https://arxiv.org/abs/2111.01084
  https://doi.org/10.1016/j.spasta.2022.100599
- A diffusion-based spatio-temporal extension of Gaussian Matérn fields (Lindgren et al)
  https://arxiv.org/abs/2006.04917
- `inlabru` documentation and examples:
  https://inlabru-org.github.io/inlabru/
- Related work: Stefanie Muff, Johannes Signer, John Fieberg (2020, Journal of Animal Ecology)
  https://doi.org/10.1111/1365-2656.13087

Summary

- Non-separability and non-stationarity are distinct concepts; both needed
- (Relatively) simple stochastic PDEs provide useful building blocks
- Computational methods need to handle hierarchical structures, not just additive noise.
- The Poisson trick & iterative linearisation allows `inlabru` to estimate new model classes
- The SPDE approach for Gaussian and non-Gaussian fields: 10 years and still running (Lindgren et al, 2020, Spatial Statistics)
  https://arxiv.org/abs/2111.01084
  https://doi.org/10.1016/j.spasta.2022.100599
- A diffusion-based spatio-temporal extension of Gaussian Matérn fields (Lindgren et al)
  https://arxiv.org/abs/2006.04917
- `inlabru` documentation and examples:
  https://inlabru-org.github.io/inlabru/
- Related work: Stefanie Muff, Johannes Signer, John Fieberg (2020, Journal of Animal Ecology)
  https://doi.org/10.1111/1365-2656.13087