Stochastic partial differential equations and numerical methods for large scale spatial statistics

Finn Lindgren

Big Data in Environmental Science
University of British Columbia, PIMS
Vancouver, 11 May 2015

“Big” data

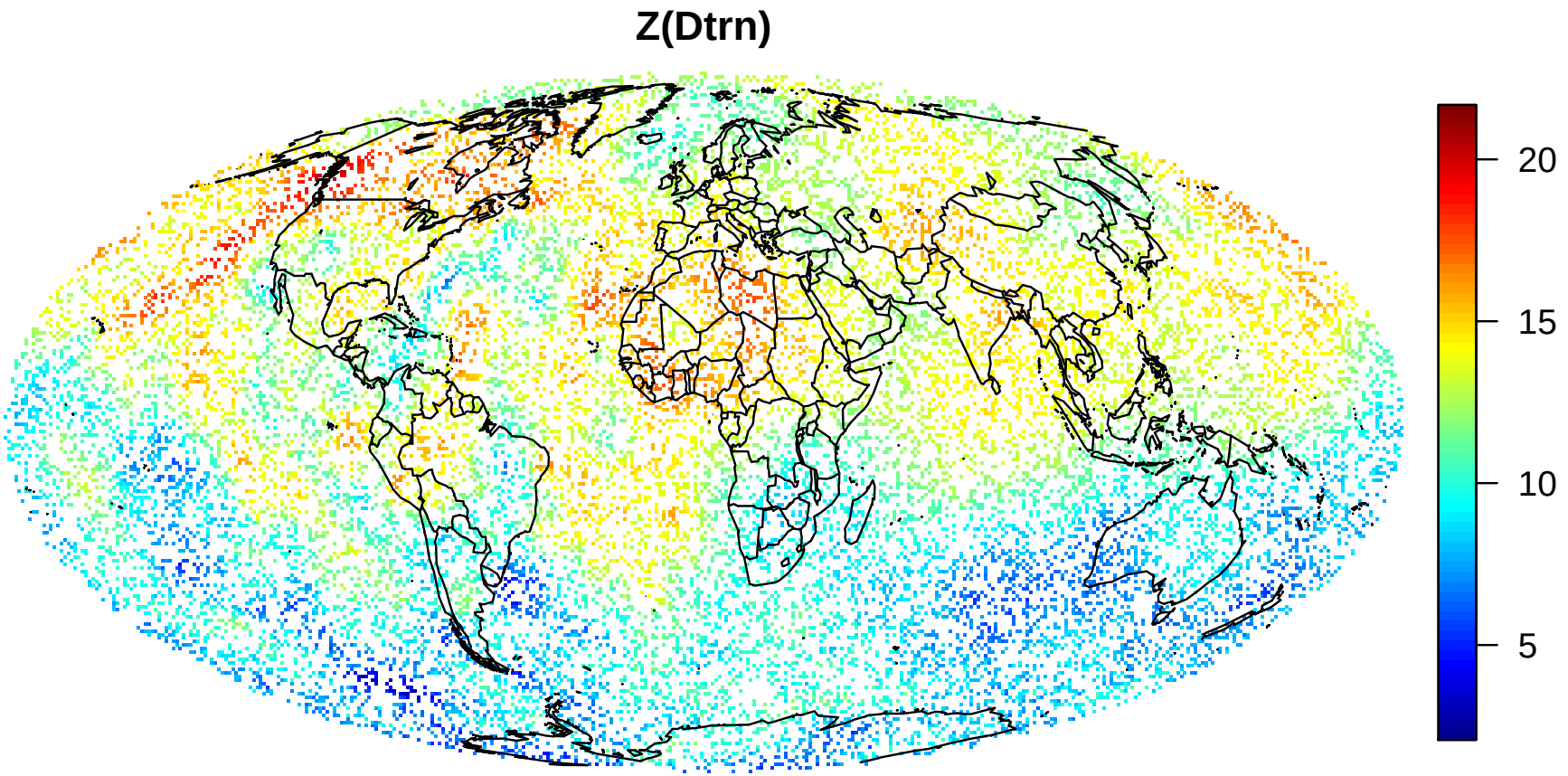
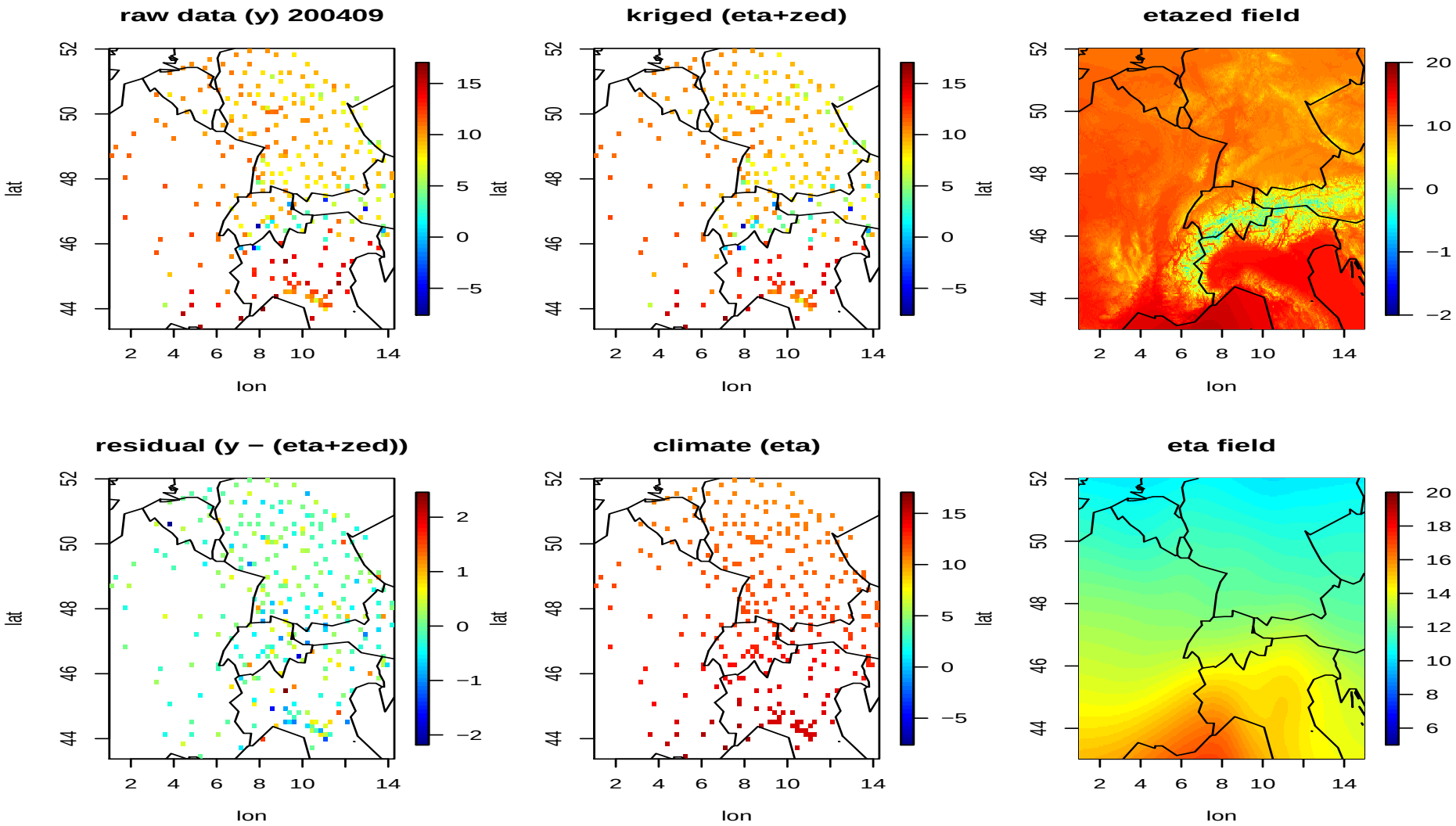


Illustration: Synthetic data mimicking satellite based CO_2 measurements. Iregular data locations, uneven coverage, features on different scales.

Sparse spatial coverage of temperature measurements



Regional observations: $\approx$ 20,000,000 from daily timeseries over 160 years

Note: This is a small *subset* of the full data! (More about this tomorrow!)

Spatio-temporal modelling framework

Spatial statistics framework

- Spatial domain Ω, or space-time domain $\Omega \times \mathbb{T}$, $\mathbb{T} \subset \mathbb{R}$.
- Random field $u(\boldsymbol{s})$, $\boldsymbol{s} \in \Omega$, or $u(\boldsymbol{s}, t)$, $(\boldsymbol{s}, t) \in \Omega \times \mathbb{T}$.
- Observations y_i. In the simplest setting, $y_i = u(\boldsymbol{s}_i) + \epsilon_i$, but more generally $y_i \sim$ GLMM, with $u(\cdot)$ as a structured random effect.
- Needed: models capturing stochastic dependence on multiple scales
- Partial solution: Basis function expansions, with large scale functions and covariates to capture static and slow structures, and small scale functions for more local variability

Two basic model and method components

- Stochastic models for $u(\cdot)$.
- Computationally efficient (i.e. avoid MCMC whenever possible) inference methods for the posterior distribution of $u(\cdot)$ given data $\boldsymbol{y}$.

Covariance functions and stochastic PDEs

The Matérn covariance family on $\mathbb{R}^d$

$$\mathsf{Cov}(u(\mathbf{0}), u(\boldsymbol{s})) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} (\kappa \|\boldsymbol{s}\|)^\nu K_\nu(\kappa \|\boldsymbol{s}\|)$$

Scale $\kappa > 0$, smoothness $\nu > 0$, variance $\sigma^2 > 0$

Whittle (1954, 1963): Matérn as SPDE solution

Matérn fields are the stationary solutions to the SPDE

$$\left(\kappa^2 - \nabla \cdot \nabla\right)^{\alpha/2} u(\boldsymbol{s}) = \mathcal{W}(\boldsymbol{s}), \quad \alpha = \nu + d/2$$

$\mathcal{W}(\cdot)$ white noise, $\nabla \cdot \nabla = \sum_{i=1}^d \frac{\partial^2}{\partial s_i^2}$, $\sigma^2 = \frac{\Gamma(\nu)}{\Gamma(\alpha)\kappa^{2\nu}(4\pi)^{d/2}}$

(Continuous domain white noise, $\mathsf{E}(\mathcal{W}(A)) = 0$, and $\mathsf{Cov}(\mathcal{W}(A), \mathcal{W}(B)) = |A \cap B|$, for all measurable $A, B \subseteq \mathbb{R}^d$. Not to be confused with pointwise independent noise.)

Basis function representations for Gaussian Matérn fields

Basis definitions

	Finite basis set ($k = 1, \ldots, n$)
Karhunen-Loève	$(\kappa^2 - \nabla \cdot \nabla)^{-\alpha} e_{\kappa,k}(\boldsymbol{s}) = \lambda_{\kappa,k} e_{\kappa,k}(\boldsymbol{s})$
Fourier	$-\nabla \cdot \nabla e_k(\boldsymbol{s}) = \lambda_k e_k(\boldsymbol{s})$
Convolution	$(\kappa^2 - \nabla \cdot \nabla)^{\alpha/2} g_\kappa(\boldsymbol{s}) = \delta(\boldsymbol{s})$
General/GMRF	$\psi_k(\boldsymbol{s})$

Field representations

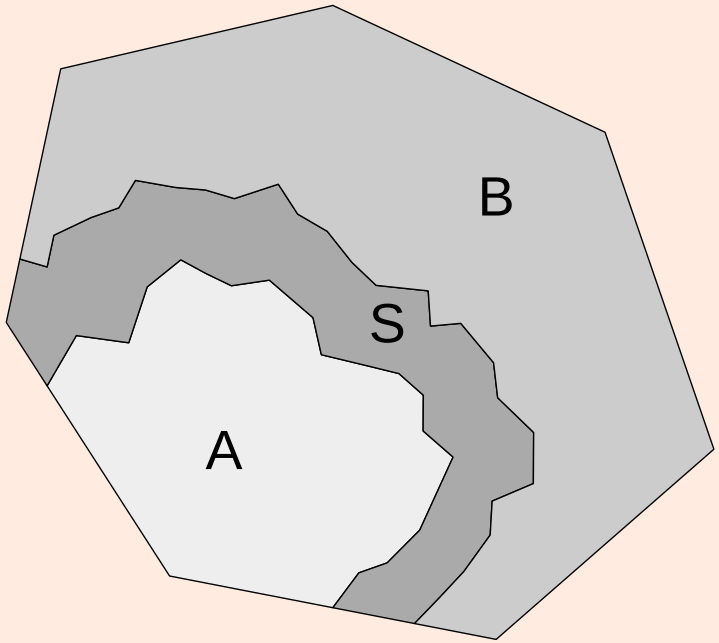
	Field $u(\boldsymbol{s})$	Weights
Karhunen-Loève	$\propto \sum_k e_{\kappa,k}(\boldsymbol{s}) z_k$	$z_k \sim \mathcal{N}(0, \lambda_{\kappa,k})$
Fourier	$\propto \sum_k e_k(\boldsymbol{s}) z_k$	$z_k \sim \mathcal{N}(0, (\kappa^2 + \lambda_k)^{-\alpha})$
Convolution	$\propto \sum_k g_\kappa(\boldsymbol{s} - \boldsymbol{s}_k) z_k$	$z_k \sim \mathcal{N}(0, \lvert\text{cell}_k\rvert)$
General/GMRF	$\propto \sum_k \psi_k(\boldsymbol{s}) u_k$	$\boldsymbol{u} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_\kappa^{-1})$

Note: Harmonic basis functions (as in the Fourier approach) give a diagonal $\boldsymbol{Q}_\kappa$, but lead to dense posterior precision matrices.

Continuous and discrete Markov properties

Markov properties

S is a separating set for A and B: $u(A) \perp u(B) \mid u(S)$



Solutions to

$$\left(\kappa^2 - \nabla \cdot \nabla\right)^{\alpha/2} u(\boldsymbol{s}) = \mathcal{W}(\boldsymbol{s})$$

are Markov when α is an integer.

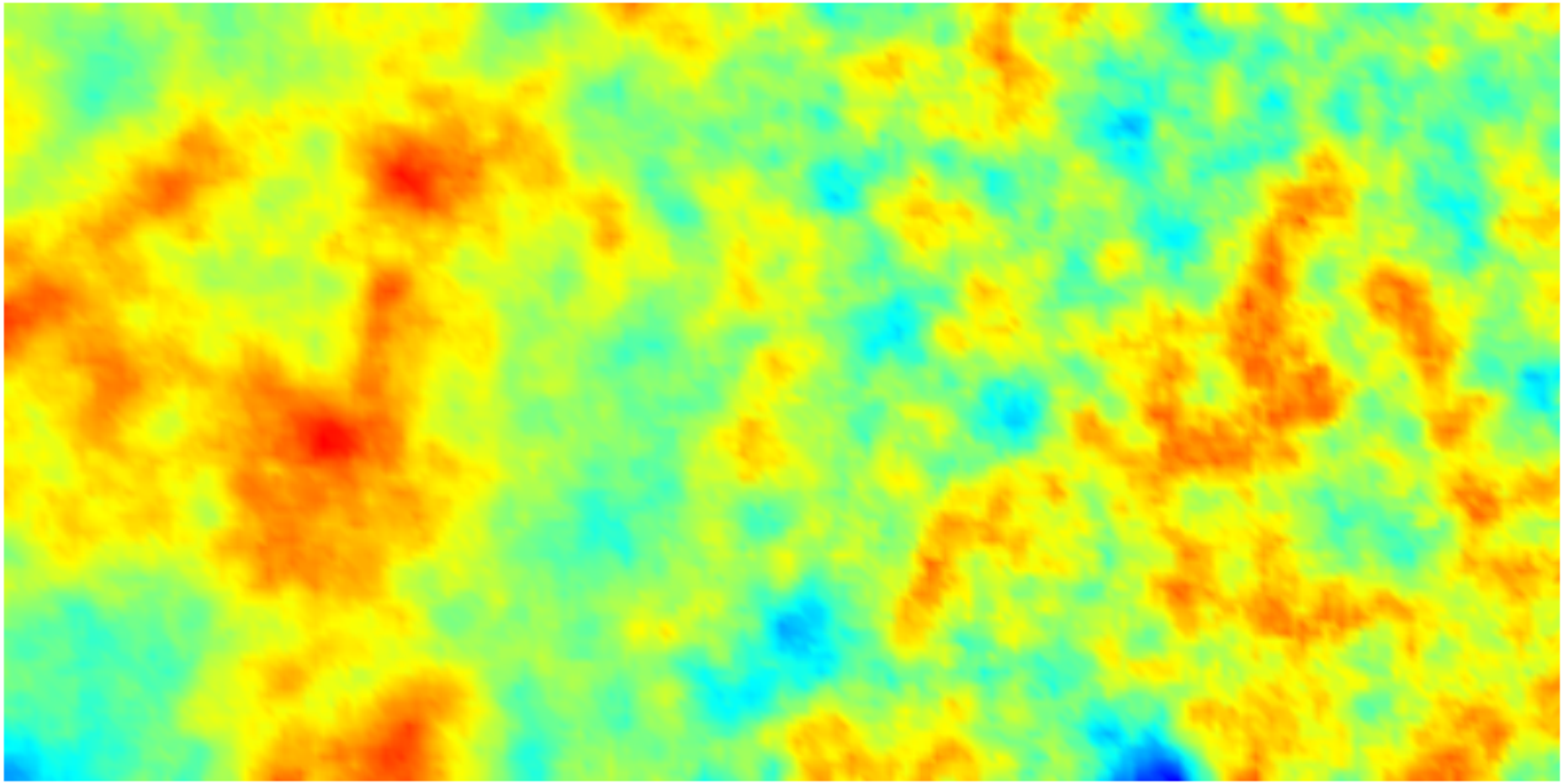
(Generally, when the reciprocal of the spectral density is a polynomial, Rozanov, 1977)

Discrete representations ($\boldsymbol{Q} = \boldsymbol{\Sigma}^{-1}$):

$$\boldsymbol{Q}_{AB} = \boldsymbol{0}$$
$$\boldsymbol{Q}_{A|S,B} = \boldsymbol{Q}_{AA}$$
$$\boldsymbol{\mu}_{A|S,B} = \boldsymbol{\mu}_A - \boldsymbol{Q}_{AA}^{-1}\boldsymbol{Q}_{AS}(\boldsymbol{u}_S - \boldsymbol{\mu}_S)$$

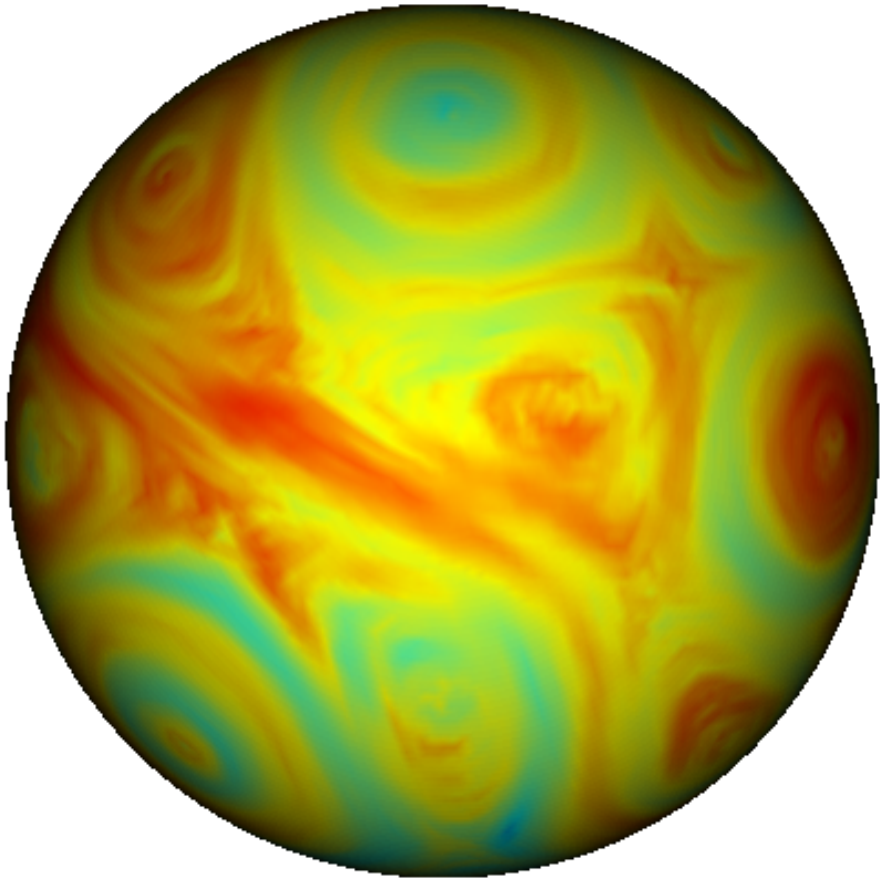
Fractional SPDEs can often be closely approximated by sums of a small number of Markov SPDEs, or even by a single (non-Matérn) Markov SPDE.

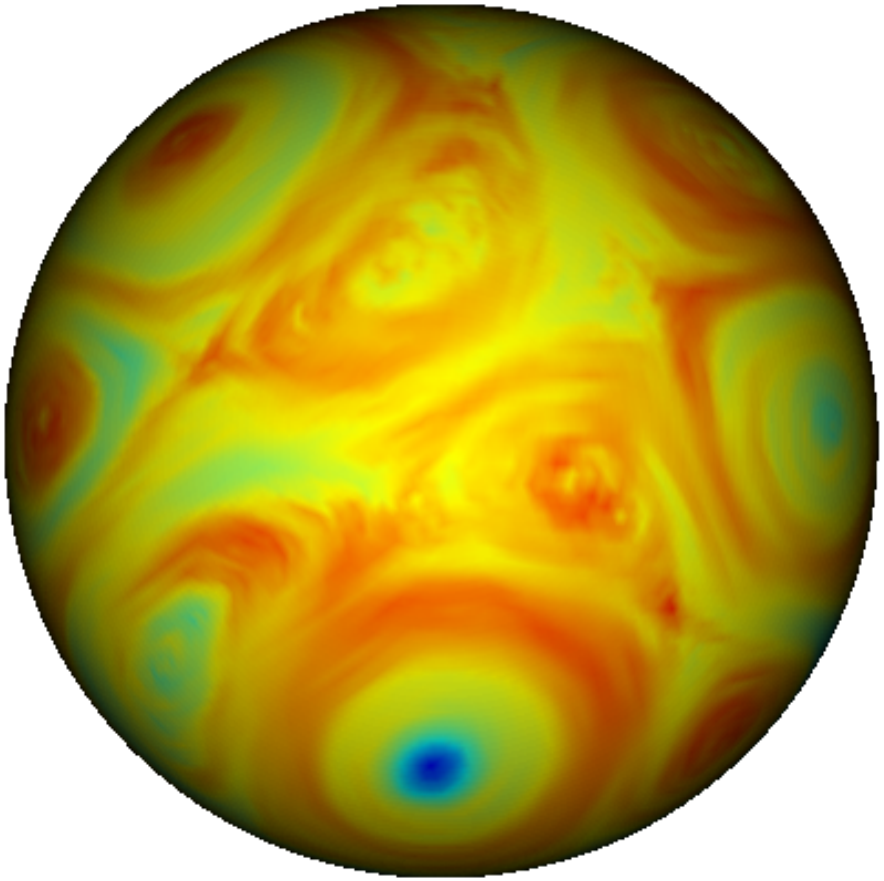
Beyond Matérn: Non-stationary field



$$(\kappa(\boldsymbol{s})^2 - \nabla \cdot \nabla) u(\boldsymbol{s}) = \kappa(\boldsymbol{s}) \mathcal{W}(\boldsymbol{s}), \quad \boldsymbol{s} \in \Omega$$

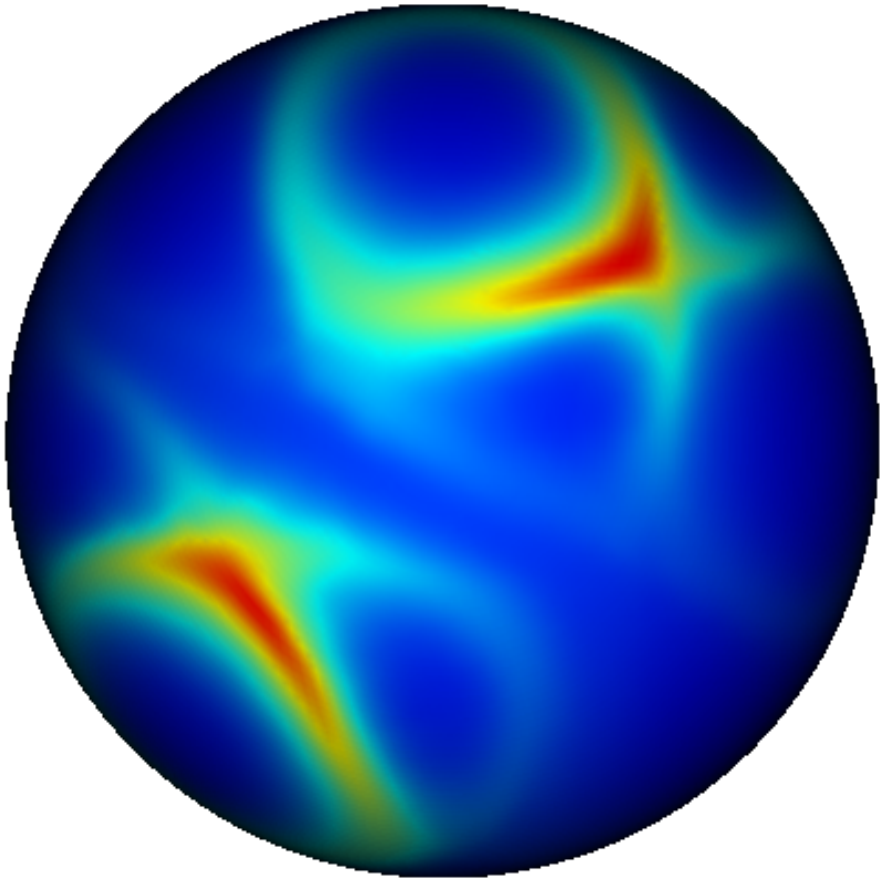
Beyond Matérn: Markov does *not* mean local dependence

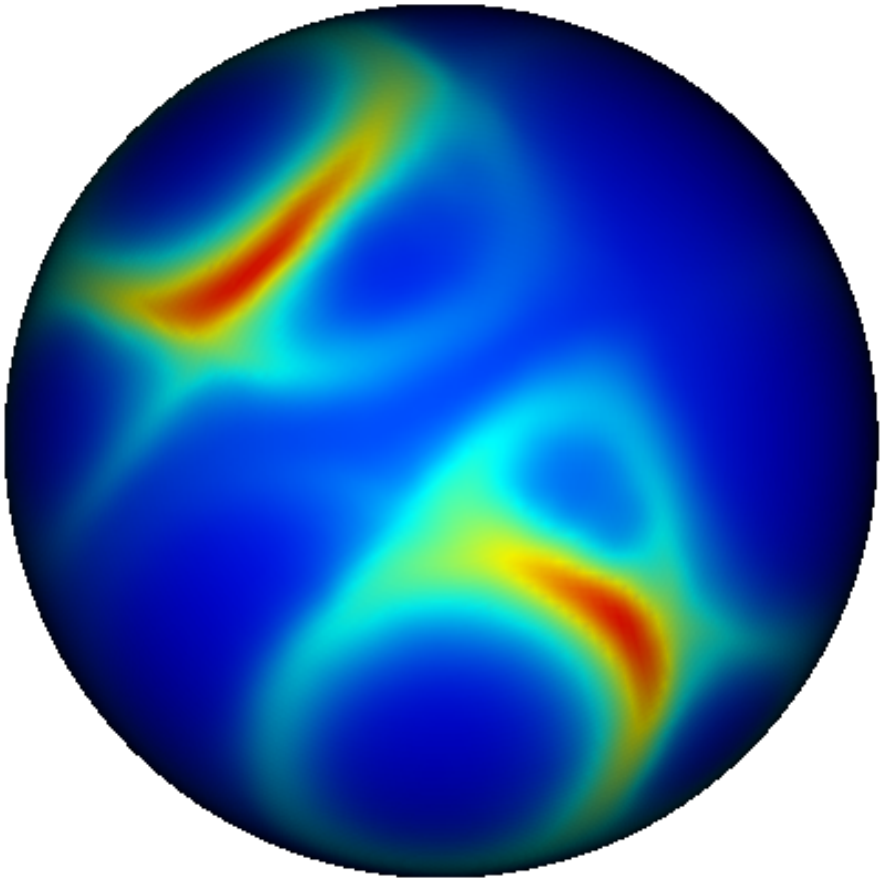




$$(\kappa(\boldsymbol{s})^2 - \nabla \cdot \boldsymbol{H}(\boldsymbol{s})\nabla)u(\boldsymbol{s}) = \kappa(\boldsymbol{s})\mathcal{W}(\boldsymbol{s}), \quad \boldsymbol{s} \in \Omega$$

Covariances for four reference points





Continuous domain Markov approximations

Continuous Markovian spatial models (Lindgren et al, 2011)

Local basis: $u(\boldsymbol{s}) = \sum_k \psi_k(\boldsymbol{s}) u_k$, (compact, piecewise linear)

Basis weights: $\boldsymbol{u} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}^{-1})$, sparse $\boldsymbol{Q}$ based on an SPDE

Special case: $(\kappa^2 - \nabla \cdot \nabla) u(\boldsymbol{s}) = \mathcal{W}(\boldsymbol{s}), \quad \boldsymbol{s} \in \Omega$

Precision: $\boldsymbol{Q} = \kappa^4 \boldsymbol{C} + 2\kappa^2 \boldsymbol{G} + \boldsymbol{G}_2$ $\quad (\kappa^4 + 2\kappa^2 |\boldsymbol{\omega}|^2 + |\boldsymbol{\omega}|^4)$

Conditional distribution in a jointly Gaussian model

$$\boldsymbol{u} \sim \mathcal{N}(\boldsymbol{\mu}_u, \boldsymbol{Q}_u^{-1}), \quad \boldsymbol{y}|\boldsymbol{u} \sim \mathcal{N}(\boldsymbol{A}\boldsymbol{u}, \boldsymbol{Q}_{y|u}^{-1}) \qquad (A_{ij} = \psi_j(\boldsymbol{s}_i))$$

$$\boldsymbol{u}|\boldsymbol{y} \sim \mathcal{N}(\boldsymbol{\mu}_{u|y}, \boldsymbol{Q}_{u|y}^{-1})$$

$$\boldsymbol{Q}_{u|y} = \boldsymbol{Q}_u + \boldsymbol{A}^T \boldsymbol{Q}_{y|u} \boldsymbol{A} \quad (\sim\text{"Sparse iff } \psi_k \text{ have compact support"})$$

$$\boldsymbol{\mu}_{u|y} = \boldsymbol{\mu}_u + \boldsymbol{Q}_{u|y}^{-1} \boldsymbol{A}^T \boldsymbol{Q}_{y|u} (\boldsymbol{y} - \boldsymbol{A}\boldsymbol{\mu}_u)$$

We've translated the spatial inference problem into sparse numerical linear algebra similar to finite element PDE solvers

The computational GMRF work-horse

Cholesky decomposition (Cholesky, 1924)

$$\boldsymbol{Q} = \boldsymbol{L}\boldsymbol{L}^\top, \quad \boldsymbol{L} \text{ lower triangular } (\sim \mathcal{O}(n^{(d+1)/2}) \text{ for } d = 1, 2, 3)$$

$$\boldsymbol{Q}^{-1}\boldsymbol{x} = \boldsymbol{L}^{-\top}\boldsymbol{L}^{-1}\boldsymbol{x}, \quad \text{via forward/backward substitution}$$

$$\log\det \boldsymbol{Q} = 2\log\det \boldsymbol{L} = 2\sum_i \log L_{ii}$$

André-Louis Cholesky (1875–1918)

"He invented, for the solution of the condition equations in the method of least squares, a very ingenious computational procedure which immediately proved extremely useful, and which most assuredly would have great benefits for all geodesists, if it were published some day." (Euology by Commandant Benoit, 1922)

Laplace approximations for non-Gaussian observations

Quadratic posterior log-likelihood approximation

$$p(\boldsymbol{u} \mid \boldsymbol{\theta}) \sim \mathcal{N}(\boldsymbol{\mu}_u, \boldsymbol{Q}_u^{-1}), \quad \boldsymbol{y} \mid \boldsymbol{u}, \boldsymbol{\theta} \sim p(\boldsymbol{y} \mid \boldsymbol{u})$$

$$p_G(\boldsymbol{u} \mid \boldsymbol{y}, \boldsymbol{\theta}) \sim \mathcal{N}(\widetilde{\boldsymbol{\mu}}, \widetilde{\boldsymbol{Q}}^{-1})$$

$$\mathbf{0} = \nabla_{\boldsymbol{u}} \left\{ \ln p(\boldsymbol{u} \mid \boldsymbol{\theta}) + \ln p(\boldsymbol{y} \mid \boldsymbol{u}) \right\}\big|_{\boldsymbol{u}=\widetilde{\boldsymbol{\mu}}}$$

$$\widetilde{\boldsymbol{Q}} = \boldsymbol{Q}_u - \nabla_{\boldsymbol{u}}^2 \ln p(\boldsymbol{y} \mid \boldsymbol{u})\big|_{\boldsymbol{u}=\widetilde{\boldsymbol{\mu}}}$$

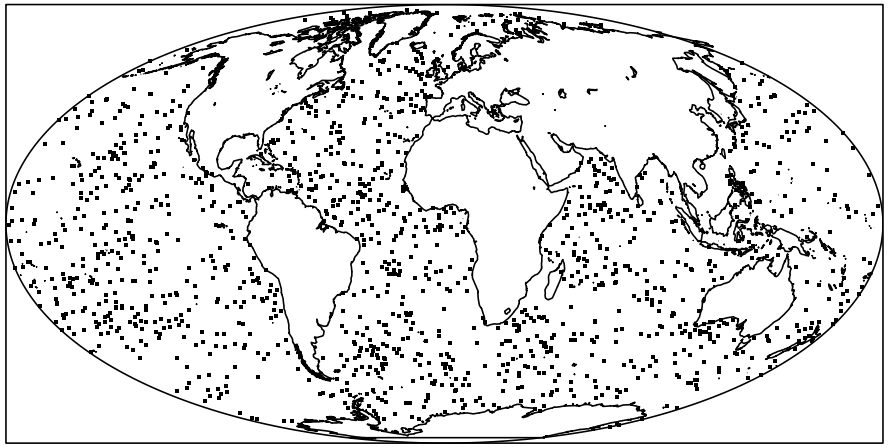
Direct Bayesian inference with INLA (`r-inla.org`)

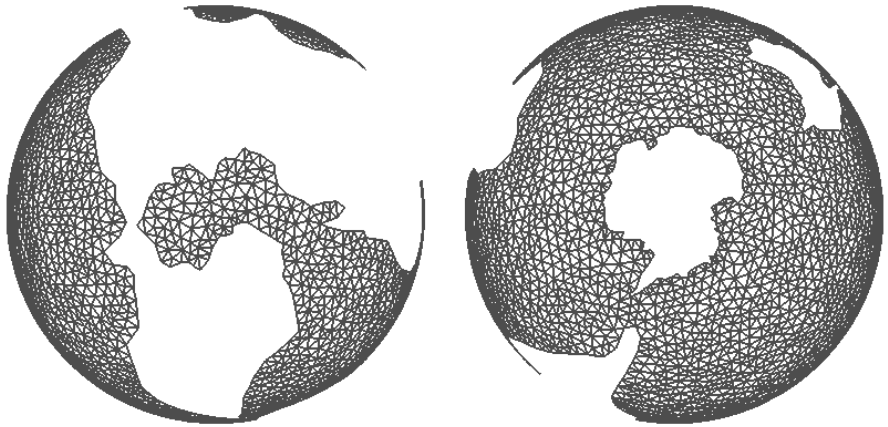
$$\widetilde{p}(\boldsymbol{\theta} \mid \boldsymbol{y}) \propto \left. \frac{p(\boldsymbol{\theta}) p(\boldsymbol{u} \mid \boldsymbol{\theta}) p(\boldsymbol{y} \mid \boldsymbol{u}, \boldsymbol{\theta})}{p_G(\boldsymbol{u} \mid \boldsymbol{y}, \boldsymbol{\theta})} \right|_{\boldsymbol{u}=\widetilde{\boldsymbol{\mu}}}$$

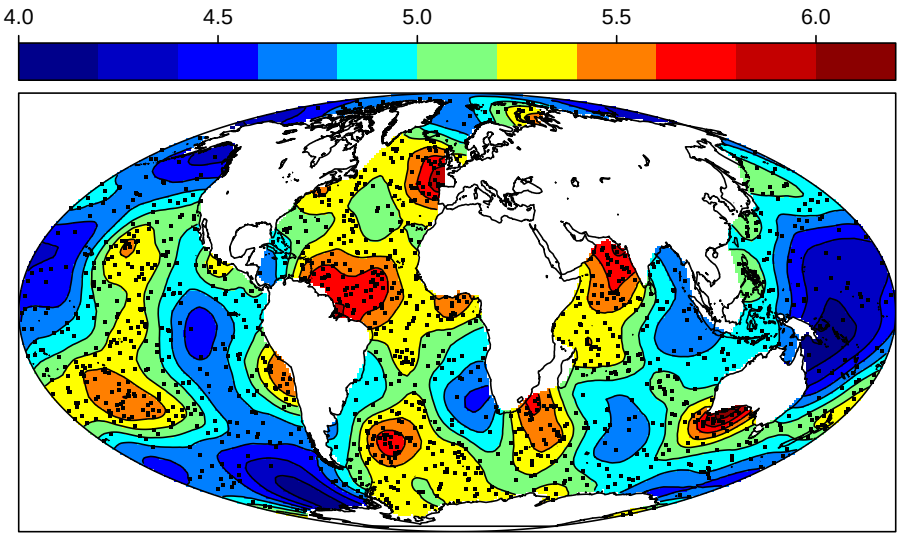
$$\widetilde{p}(\boldsymbol{u}_i \mid \boldsymbol{y}) \propto \int p_{GG}(\boldsymbol{u}_i \mid \boldsymbol{y}, \boldsymbol{\theta}) \widetilde{p}(\boldsymbol{\theta} \mid \boldsymbol{y}) \, d\boldsymbol{\theta}$$

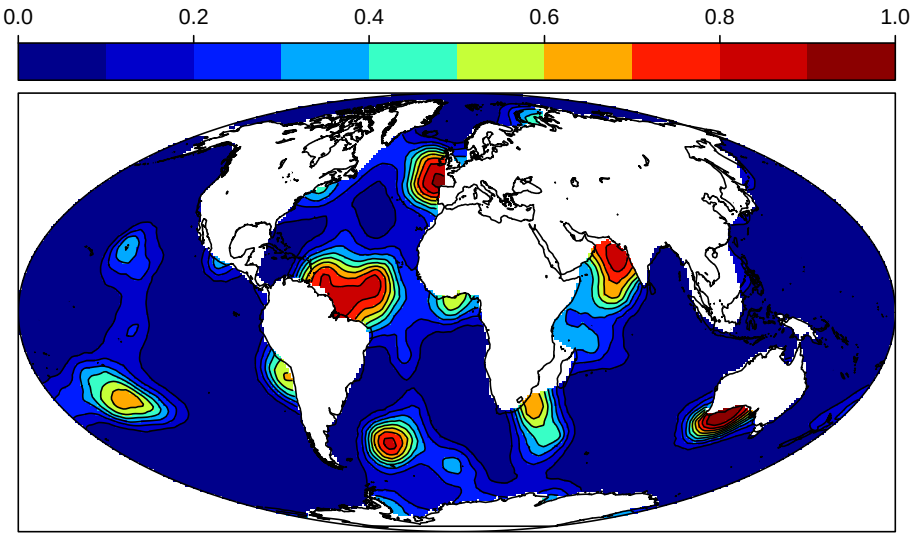
The main limiting factors for the INLA method are the number of latent variables and the number model parameters (and to a much lesser degree the number of observations). Here we'll concentrate on the latent variables.

SPDE based inference for point process data









A multiscale model example

- A temporally slow, simplified stochastic heat equation (non-separable)

$$\frac{\partial}{\partial t} z(\boldsymbol{s}, t) - \gamma_z \nabla \cdot \nabla z(\boldsymbol{s}, t) = \mathcal{E}(\boldsymbol{s}, t)$$

$$(1 - \gamma_{\mathcal{E}} \nabla \cdot \nabla) \mathcal{E}(\boldsymbol{s}, t) = \mathcal{W}_{\mathcal{E}}(\boldsymbol{s}, t)$$

- A temporally quick, spatially non-stationary SPDE/GMRF (separable)
  $\left(\frac{\partial}{\partial t} + \gamma_t\right) \left(\kappa(\boldsymbol{s})^2 - \nabla \cdot \nabla\right) \left(\tau(\boldsymbol{s}) a(\boldsymbol{s}, t)\right) = \mathcal{W}_a(\boldsymbol{s}, t)$
- Measurements
  $y_i = a(\boldsymbol{s}_i, t_i) + z(\boldsymbol{s}_i, t_i) + \epsilon_i$, discretised into
  $\boldsymbol{y} = \boldsymbol{A}(\boldsymbol{a} + (\boldsymbol{B} \otimes \boldsymbol{I})\boldsymbol{z}) + \boldsymbol{\epsilon}$, $\boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_\epsilon^{-1})$
  where $\boldsymbol{B}$ maps from long-term basis functions to short-term, and $\boldsymbol{A}$ maps from short-term basis functions to the observations.

The posterior precision can be formulated for $(a + z, z)|y$:

$$\boldsymbol{Q}_{(a+z,z)|y} = \begin{bmatrix} \boldsymbol{Q}_t \otimes \boldsymbol{Q}_a + \boldsymbol{A}^\top \boldsymbol{Q}_\epsilon \boldsymbol{A} & -\boldsymbol{Q}_t \boldsymbol{B} \otimes \boldsymbol{Q}_a \\ -\boldsymbol{B}^\top \boldsymbol{Q}_t \otimes \boldsymbol{Q}_a & \boldsymbol{Q}_z + \boldsymbol{B}^\top \boldsymbol{Q}_t \boldsymbol{B} \otimes \boldsymbol{Q}_a \end{bmatrix}$$

Locally isotropic non-stationary precision construction

Finite element construction of basis weight precision

Non-stationary SPDE:

$$(\kappa(\boldsymbol{s})^2 - \nabla \cdot \nabla)\,(\tau(\boldsymbol{s})u(\boldsymbol{s})) = \mathcal{W}(\boldsymbol{s})$$

The SPDE parameters are constructed via spatial covariates:

$$\log \tau(\boldsymbol{s}) = b_0^\tau(\boldsymbol{s}) + \sum_{j=1}^{p} b_j^\tau(\boldsymbol{s})\theta_j, \quad \log \kappa(\boldsymbol{s}) = b_0^\kappa(\boldsymbol{s}) + \sum_{j=1}^{p} b_j^\kappa(\boldsymbol{s})\theta_j$$

Finite element calculations give

$$\boldsymbol{T} = \operatorname{diag}(\tau(\boldsymbol{s}_i)), \quad \boldsymbol{K} = \operatorname{diag}(\kappa(\boldsymbol{s}_i))$$

$$C_{ii} = \int \psi_i(\boldsymbol{s})\, d\boldsymbol{s}, \quad G_{ij} = \int \nabla\psi_i(\boldsymbol{s}) \cdot \nabla\psi_j(\boldsymbol{s})\, d\boldsymbol{s}$$

$$\boldsymbol{Q} = \boldsymbol{T}\left(\boldsymbol{K}^2\boldsymbol{C}\boldsymbol{K}^2 + \boldsymbol{K}^2\boldsymbol{G} + \boldsymbol{G}\boldsymbol{K}^2 + \boldsymbol{G}\boldsymbol{C}^{-1}\boldsymbol{G}\right)\boldsymbol{T}$$

Combining this with an AR(1) discretisation of the temporal operator, we get $\boldsymbol{Q}_t \otimes \boldsymbol{Q}_a$. More on $\boldsymbol{Q}_z$ tomorrow.

Posterior calculations

Write $x = (a + z, z)$ for the full latent field.

$$\boldsymbol{Q}_{x|y} = \begin{bmatrix} \boldsymbol{Q}_t \otimes \boldsymbol{Q}_a + \boldsymbol{A}^\top \boldsymbol{Q}_\epsilon \boldsymbol{A} & -\boldsymbol{Q}_t \boldsymbol{B} \otimes \boldsymbol{Q}_a \\ -\boldsymbol{B}^\top \boldsymbol{Q}_t \otimes \boldsymbol{Q}_a & \boldsymbol{Q}_z + \boldsymbol{B}^\top \boldsymbol{Q}_t \boldsymbol{B} \otimes \boldsymbol{Q}_a \end{bmatrix}$$

can be pseudo-Cholesky-factorised:

$$\boldsymbol{Q}_{x|y} = \widetilde{\boldsymbol{L}}_{x|y} \widetilde{\boldsymbol{L}}_{x|y}^\top, \qquad \widetilde{\boldsymbol{L}}_{x|y} = \begin{bmatrix} \boldsymbol{L}_t \otimes \boldsymbol{L}_a & \boldsymbol{0} & \boldsymbol{A}^\top \boldsymbol{L}_\epsilon \\ -\boldsymbol{B}^\top \boldsymbol{L}_t \otimes \boldsymbol{L}_a & \widetilde{\boldsymbol{L}}_z & \boldsymbol{0} \end{bmatrix}$$

Posterior expectation, samples, and marginal variances (with $\widetilde{\boldsymbol{A}} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{0} \end{bmatrix}$):

$$\begin{aligned} \boldsymbol{Q}_{x|y}(\boldsymbol{\mu}_{x|y} - \boldsymbol{\mu}_x) &= \widetilde{\boldsymbol{A}}^\top \boldsymbol{Q}_\epsilon (\boldsymbol{y} - \widetilde{\boldsymbol{A}} \boldsymbol{\mu}_x), \\ \boldsymbol{Q}_{x|y}(\boldsymbol{x} - \boldsymbol{\mu}_{x|y}) &= \widetilde{\boldsymbol{L}}_{x|y} \boldsymbol{w}, \quad \boldsymbol{w} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}), \quad \text{or} \\ \boldsymbol{Q}_{x|y}(\boldsymbol{x} - \boldsymbol{\mu}_x) &= \widetilde{\boldsymbol{A}}^\top \boldsymbol{Q}_\epsilon (\boldsymbol{y} - \widetilde{\boldsymbol{A}} \boldsymbol{\mu}_x) + \widetilde{\boldsymbol{L}}_{x|y} \boldsymbol{w}, \quad \boldsymbol{w} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}), \\ \mathsf{Var}(x_i|\boldsymbol{y}) &= \texttt{diag(inla.qinv(}\boldsymbol{Q}_{x|y}\texttt{))} \quad \text{(requires Cholesky)} \end{aligned}$$

Preconditioning for e.g. conjugate gradient solutions

Solving $\boldsymbol{Q}\boldsymbol{x} = \boldsymbol{b}$ is equivalent to solving $\boldsymbol{M}^{-1}\boldsymbol{Q}\boldsymbol{x} = \boldsymbol{M}^{-1}\boldsymbol{b}$. Choosing $\boldsymbol{M}^{-1}$ as an approximate inverse to $\boldsymbol{Q}$ gives a less ill-conditioned system. Only the *action* of $\boldsymbol{M}^{-1}$ is needed, e.g. one or more fixed point iterations:

Block Jacobi and Gauss-Seidel preconditioning

Matrix split: $\boldsymbol{Q}_{x|y} = \boldsymbol{L} + \boldsymbol{D} + \boldsymbol{L}^\top$

Jacobi: $\boldsymbol{x}^{(k+1)} = \boldsymbol{D}^{-1}\left(-(\boldsymbol{L} + \boldsymbol{L}^\top)\boldsymbol{x}^{(k)} + \boldsymbol{b}\right)$

Gauss-Seidel: $\boldsymbol{x}^{(k+1)} = (\boldsymbol{L} + \boldsymbol{D})^{-1}\left(-\boldsymbol{L}^\top\boldsymbol{x}^{(k)} + \boldsymbol{b}\right)$

Remark: Block Gibbs sampling for a GMRF posterior

With $\boldsymbol{Q} = \boldsymbol{Q}_{x|y}$, $\boldsymbol{b} = \boldsymbol{A}^\top\boldsymbol{Q}_\epsilon(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{\mu}_x)$ and $\widetilde{\boldsymbol{x}} = \boldsymbol{x} - \boldsymbol{\mu}_x$,

$$\widetilde{\boldsymbol{x}}^{(k+1)} = (\boldsymbol{L} + \boldsymbol{D})^{-1}\left(-\boldsymbol{L}^\top\widetilde{\boldsymbol{x}}^{(k)} + \boldsymbol{b} + \widetilde{\boldsymbol{L}}_{\boldsymbol{D}}\boldsymbol{w}\right), \quad \boldsymbol{w} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$$

Gauss-Seidel and Gibbs are both inefficient on their own, but G-S leads to useful preconditioners. Convergence testing is much easier for linear solvers than for MCMC.

Multiscale Schur complement approximation

Solving $Q_{x|y}x = b$ can be formulated using two solves with the upper block $Q_t \otimes Q_a + A^\top Q_\epsilon A$, and one solve with the *Schur complement*

$$Q_z + B^\top Q_t B \otimes Q_a - B^\top Q_t \otimes Q_a \left(Q_t \otimes Q_a + A^\top Q_\epsilon A\right)^{-1} Q_t B \otimes Q_a$$

By mapping the fine scale model onto the coarse basis used for the coarse model, we get an *approximate* (and sparse) Schur solve via

$$\begin{bmatrix} \widetilde{Q}_B + \widetilde{B}^\top A^\top Q_\epsilon A \widetilde{B} & -\widetilde{Q}_B \\ -\widetilde{Q}_B & Q_z + \widetilde{Q}_B \end{bmatrix} \begin{bmatrix} \text{ignored} \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ \widetilde{b} \end{bmatrix}$$

where $\widetilde{B} = B \otimes I$, $\widetilde{Q}_B = B^\top Q_t B \otimes Q_a$, and the block matrix can be interpreted as the precision of a bivariate field on a common, coarse spatio-temporal scale.

Each block may be very large even on the coarse scale, and we still need a way to solve with the fine scale upper block!

Multigrid

Construct a sequence of increasingly detailed models, $\left(\boldsymbol{Q}^{(0)}, \boldsymbol{Q}^{(1)}, \ldots, \boldsymbol{Q}^{(L)}\right)$.

Basic idea:

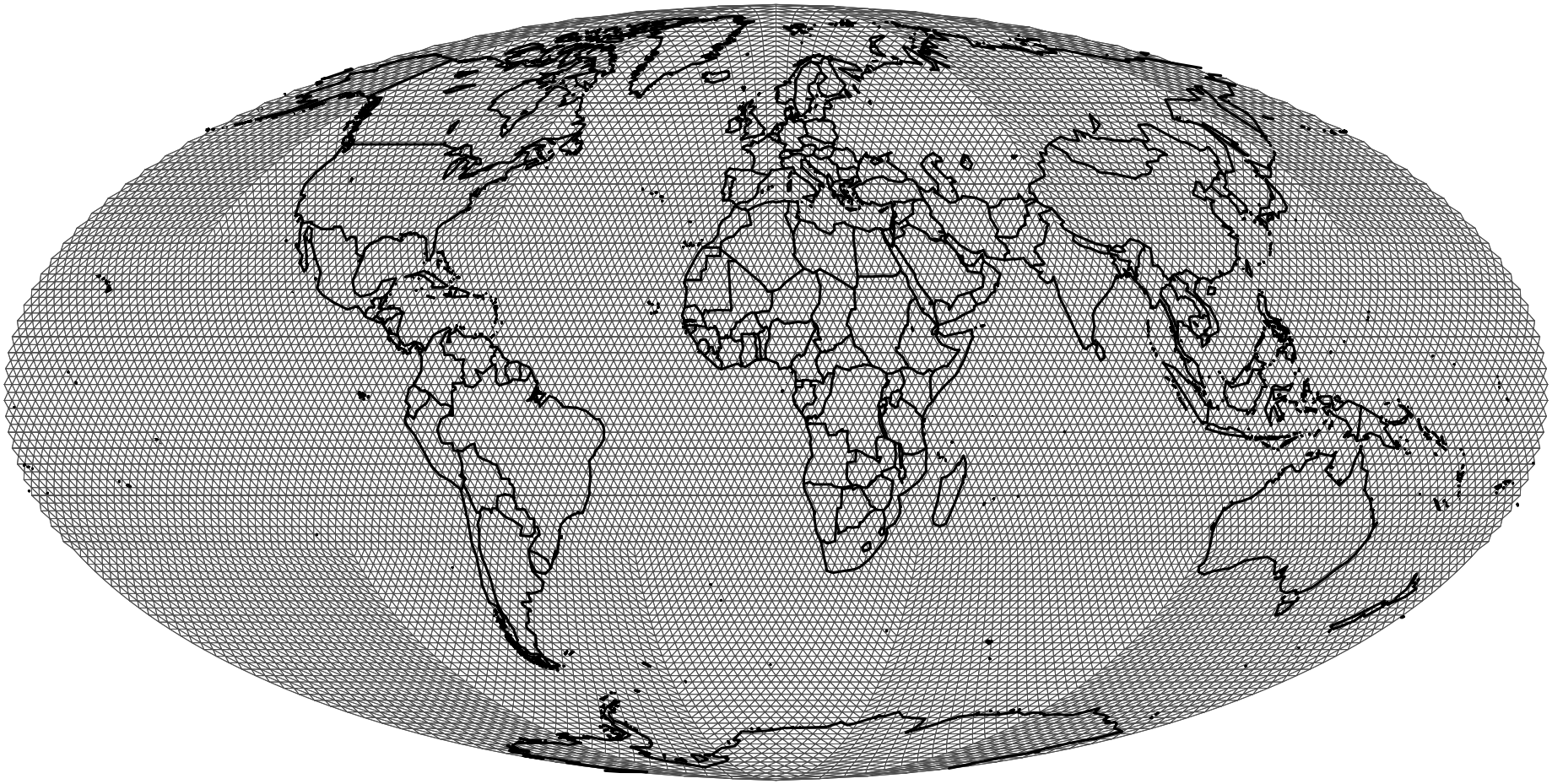
- On each level, a simple local fixed point iteration can eliminate small scale residual errors efficiently, but not large scale errors.
- Project the residual onto the next coarse level, where the large scale is now small, and then interpolate the result back onto the finer level.
- On the coarsest level, solve the exact problem.

Simple multigrid model traversal: $L = 4, 3, 2, 1, 0, 1, 2, 3, 4 = L$
Full multigrid: $L = 4, 3, 2, 1, 0, 1, 0, 1, 2, 1, 0, 1, 2, 3, 2, 1, 0, 1, 2, 3, 4 = L$

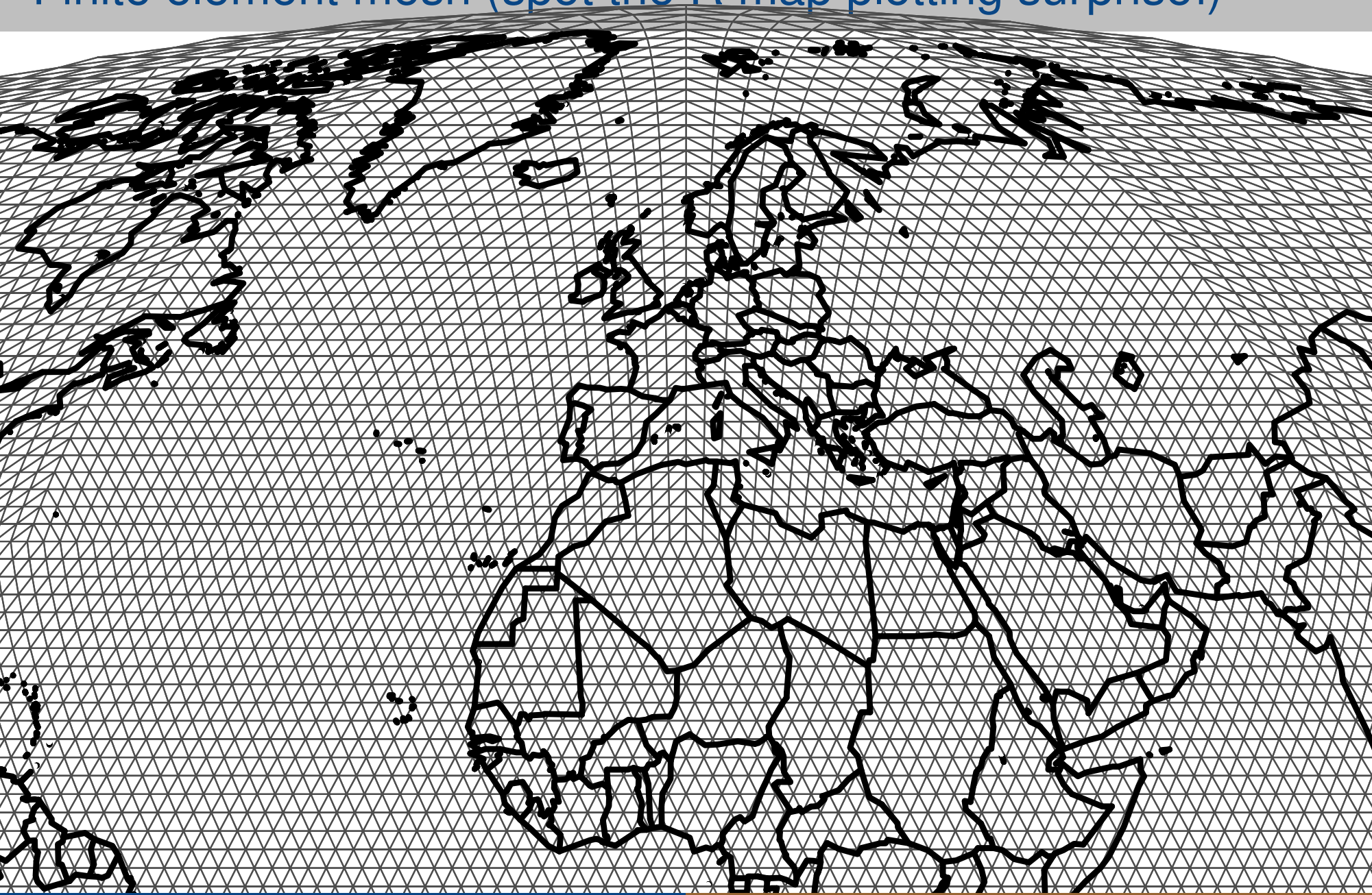
In theory, full multigrid can be $\mathcal{O}(n)$, but the constant may be very large, and toy model testing indicates that the high operator order is a problem.

Can be used as complete solver with small tolerance, or as preconditioner with large tolerance.

Finite element mesh



Finite element mesh (spot the R map plotting surprise!)



Domain decomposition

- Divide the domain into a collection of overlapping subdomain blocks
- Solve a local problem, e.g. the conditional solution, maintaining coherence by enforcing constraints on overlapping nodes.

Monte Carlo variance reduction for posterior variances

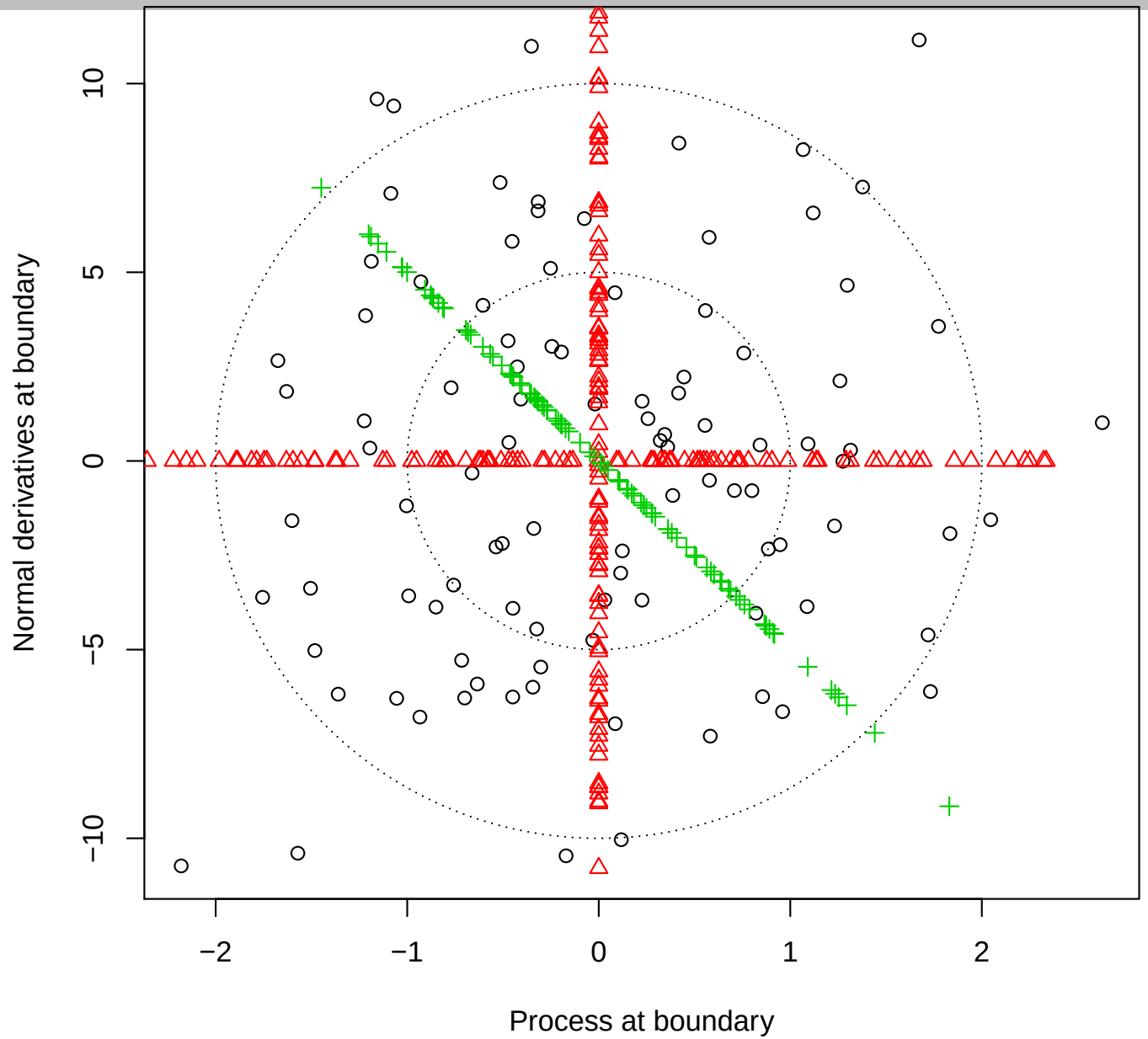
$$\mathsf{E}(\boldsymbol{x}_i \mid \boldsymbol{y}) = \mathsf{E}\left(\mathsf{E}(\boldsymbol{x}_i \mid \boldsymbol{y}, \boldsymbol{x}_{\notin \text{subblock}})\right)$$
$$\mathsf{Var}(\boldsymbol{x}_i \mid \boldsymbol{y}) = \mathsf{Var}\left(\boldsymbol{x}_i \mid \boldsymbol{y}, \boldsymbol{x}_{\notin \text{subblock}}\right) + \mathsf{Var}\left(\mathsf{E}(\boldsymbol{x}_i \mid \boldsymbol{y}, \boldsymbol{x}_{\notin \text{subblock}})\right)$$

Also works for linear combinations, with some complications

Subdomain boundary adjustment (new idea)

- Apply *stochastic boundary correction* for each subdomain
- Solve the full local problem, reusing the appropriate randomness for overlapping subdomains
- Blend the results for overlapping domains.
- Apply this as a preconditioner in an iterative solver

All deterministic boundary conditions are 'inappropriate'



Stationary stochastic boundary adjustment

Recall the Matérn generating SPDE

$$(\kappa^2 - \nabla \cdot \nabla)^{\alpha/2} u(\boldsymbol{s}) = \mathcal{W}(\boldsymbol{s})$$

RKHS inner product for Matérn precisions on $\mathbb{R}^d$:

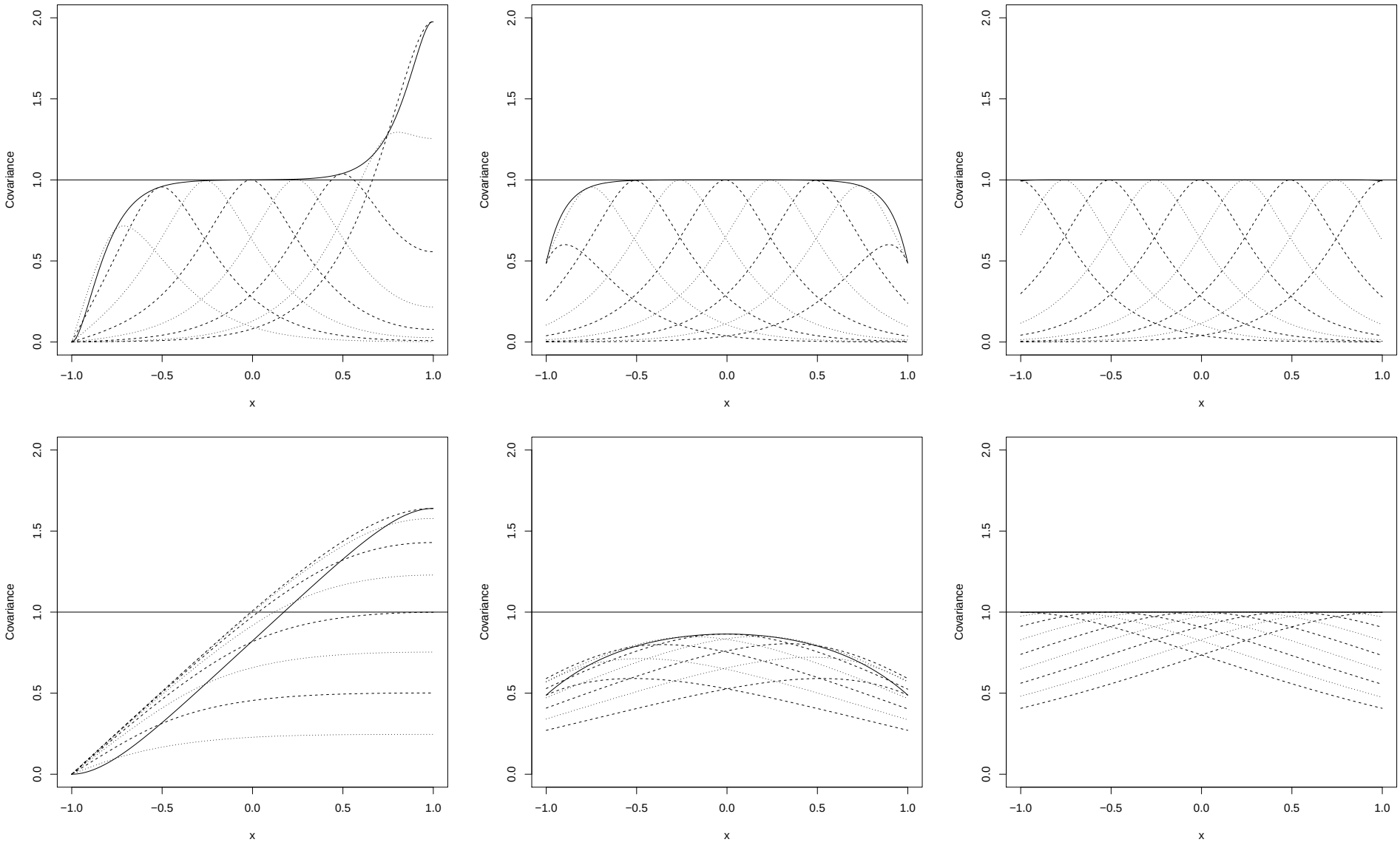
$$\langle f, g \rangle_{H(\Omega)} = \sum_{k=0}^{\alpha} \binom{\alpha}{k} \kappa^{2\alpha-2k} \left\langle \nabla^k f, \nabla^k g \right\rangle_{\Omega}$$

Boundary adjusted precision operator on a compact subdomain, where $\mathcal{P}$ projects onto the operator null-space:

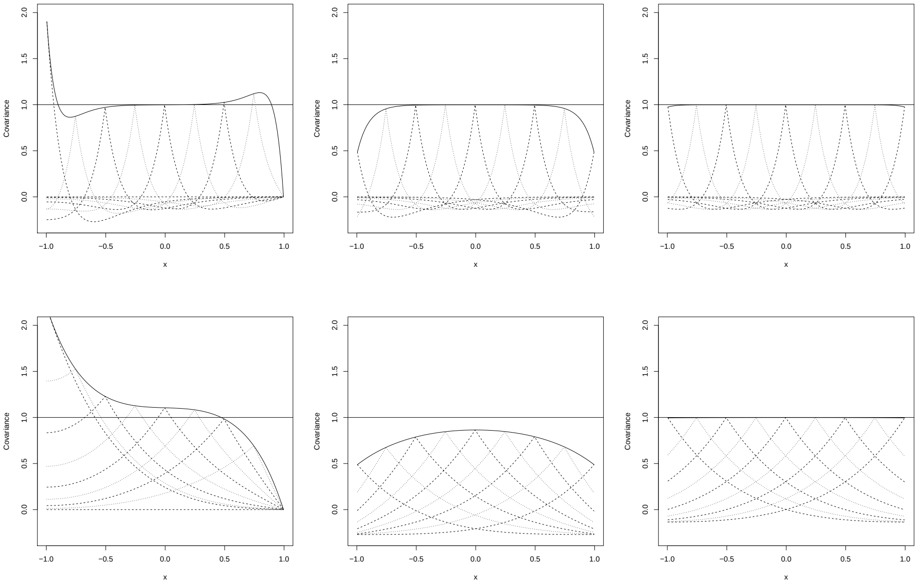
$$\begin{aligned} \mathcal{Q}_{\Omega}(f, g) &= \langle f, g \rangle_{H(\Omega)} - \langle \mathcal{P}f, \mathcal{P}g \rangle_{H(\Omega)} + \mathcal{Q}_{\mathcal{P};\partial\Omega}(\mathcal{P}f, \mathcal{P}g) \\ &= \langle f - \mathcal{P}f, g - \mathcal{P}g \rangle_{H(\Omega)} + \mathcal{Q}_{\mathcal{P};\partial\Omega}(\mathcal{P}f, \mathcal{P}g) \end{aligned}$$

Note that $\mathcal{Q}_{\mathcal{P};\partial\Omega}(\mathcal{P}f, \mathcal{P}g)$ may involve normal derivatives at the boundary.

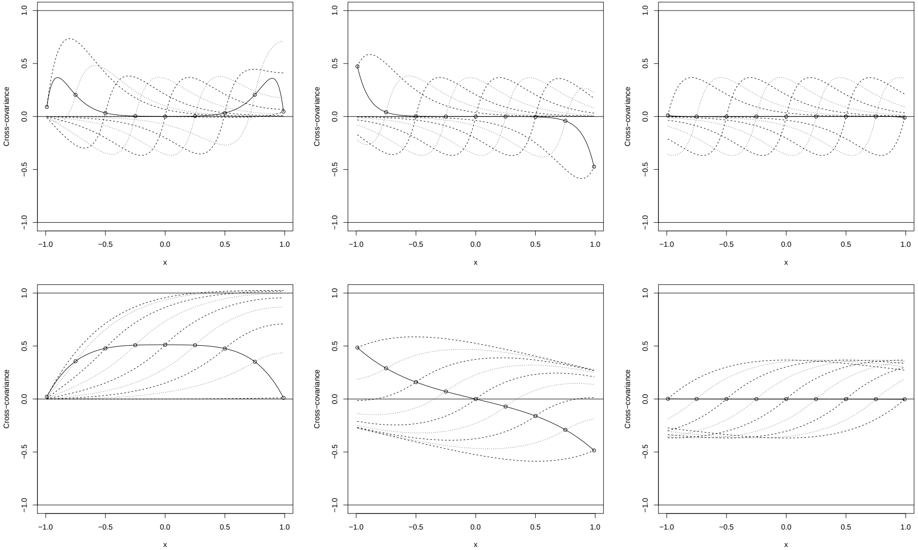
Covariances (D&N, Robin, Stoch) for $\kappa = 5$ and 1



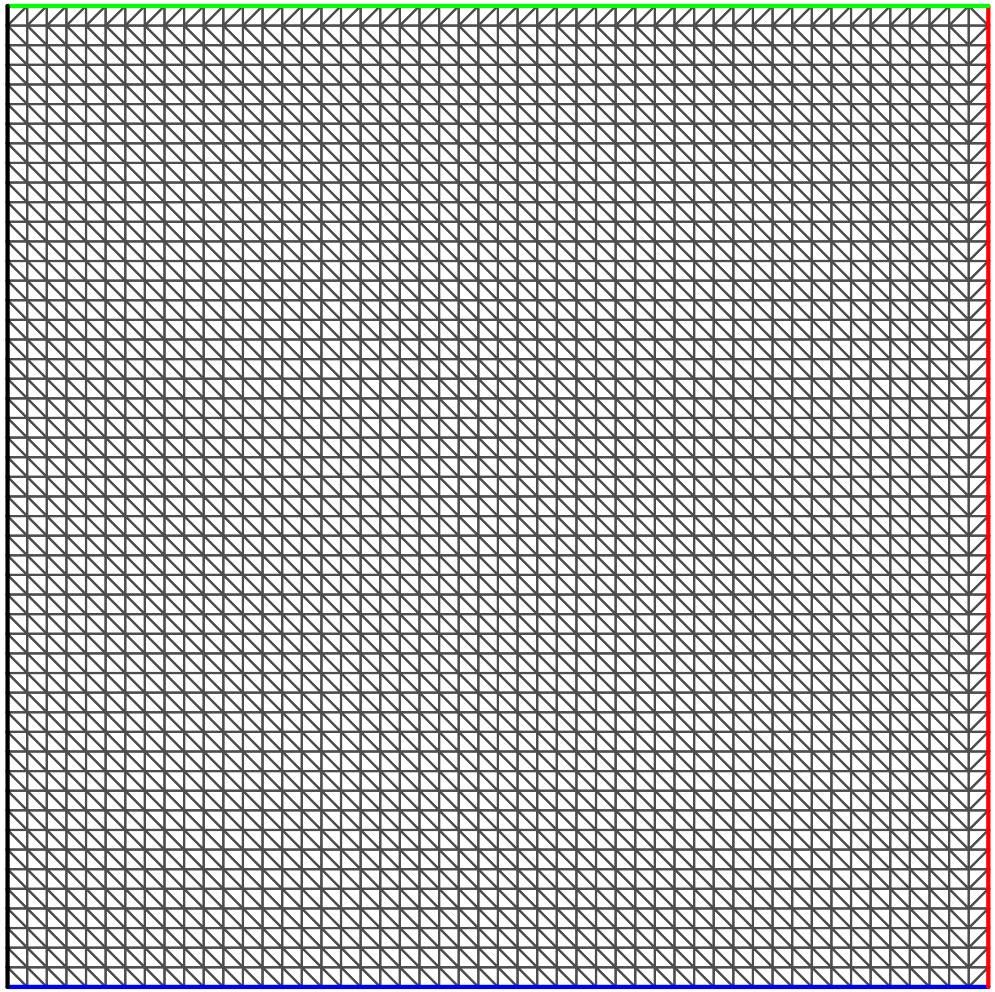
Derivative covariances (D&N, Robin, Stoch) for $\kappa = 5$ and 1



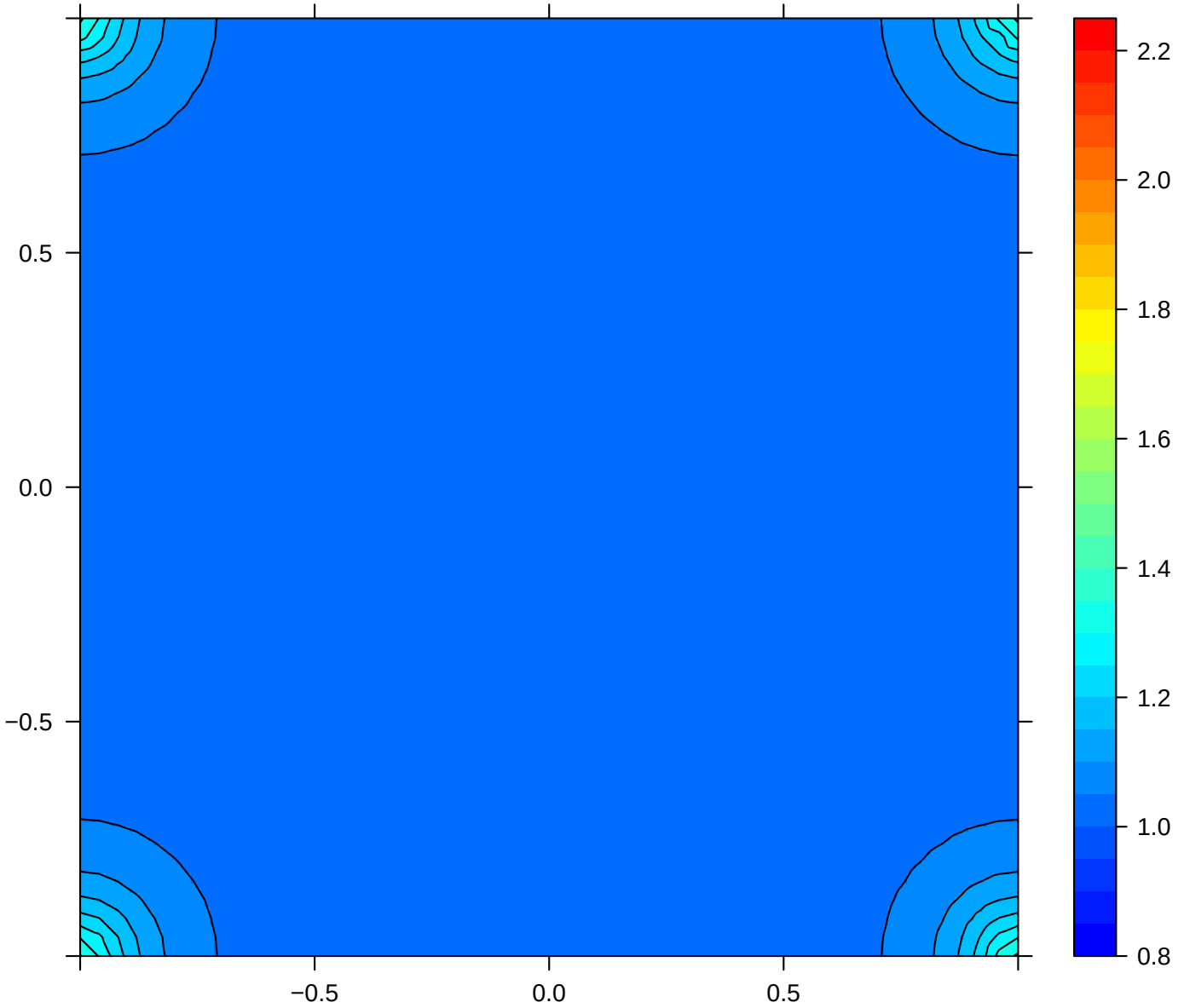
Process-derivative cross-covariances (D&N, Robin, Stoch)



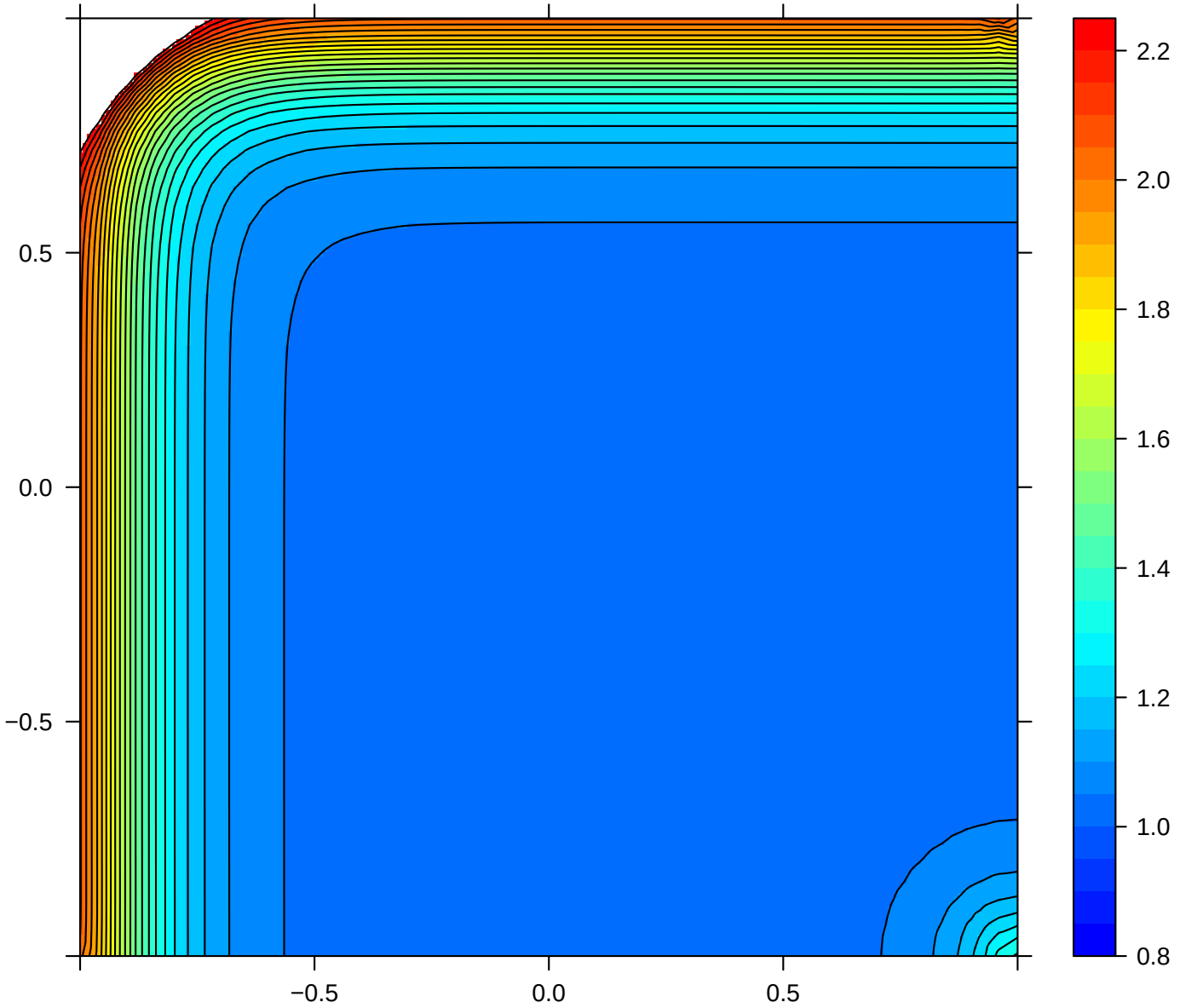
Square domain, basis triangulation



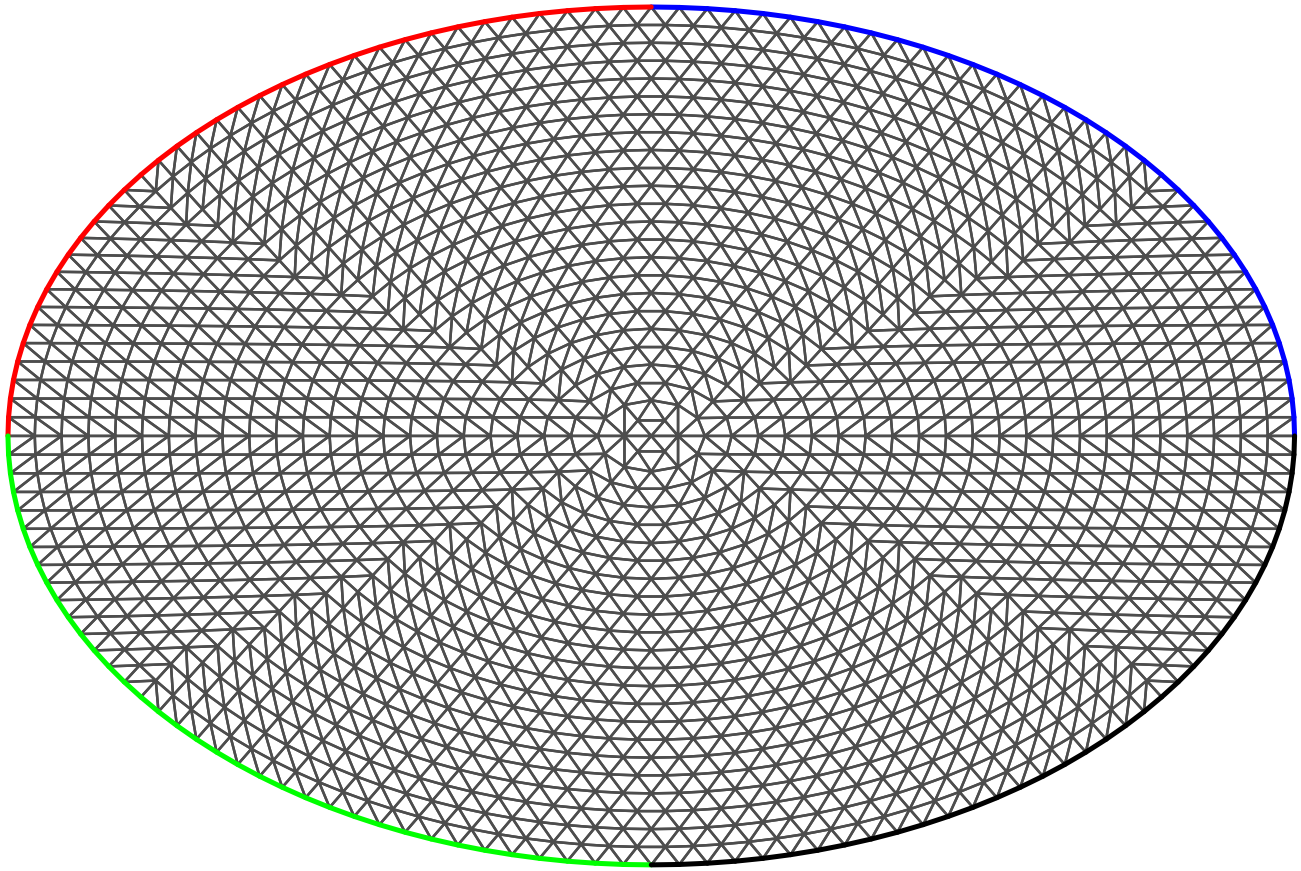
Square domain, stochastic boundary (variances)



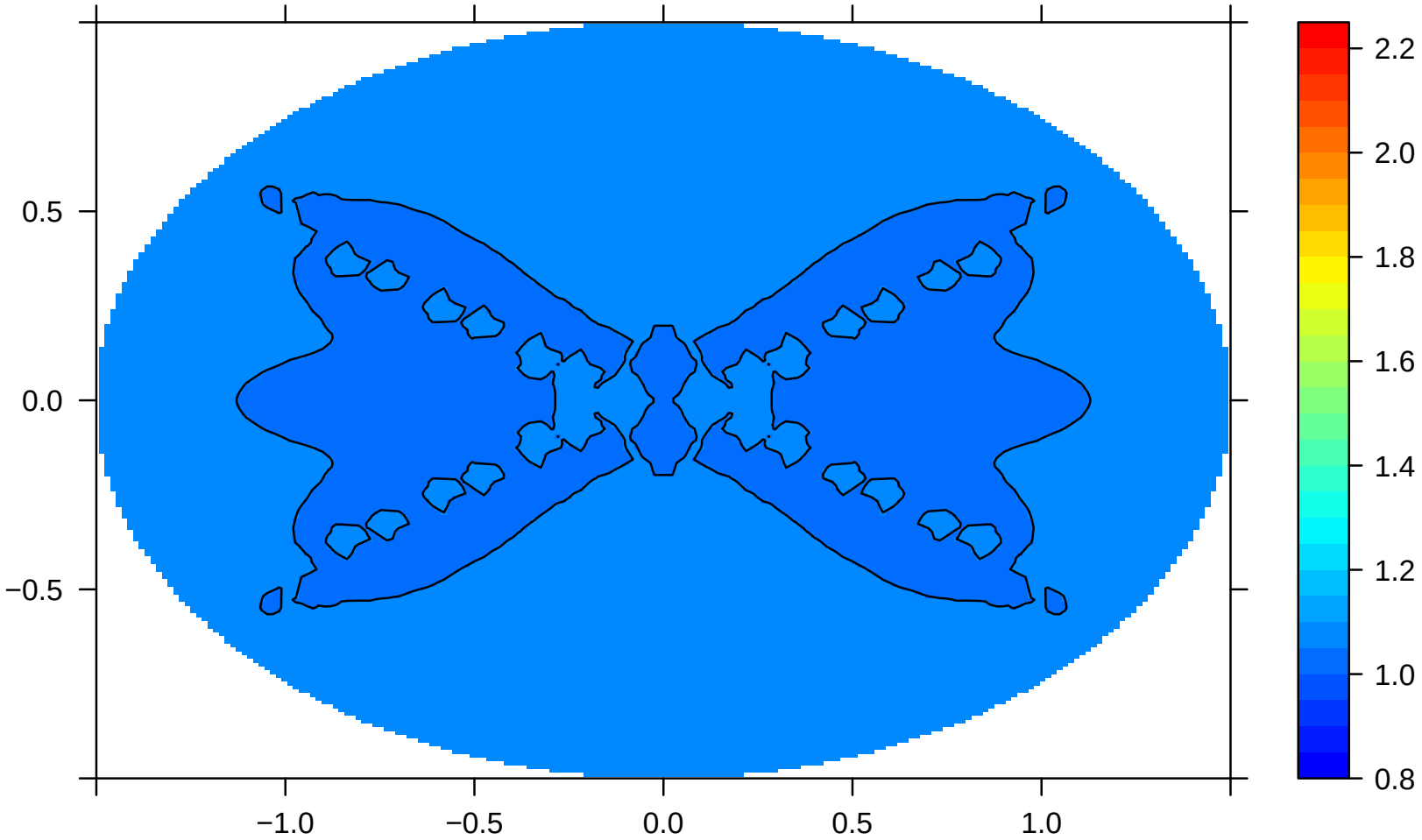
Square domain, mixed boundary (variances)



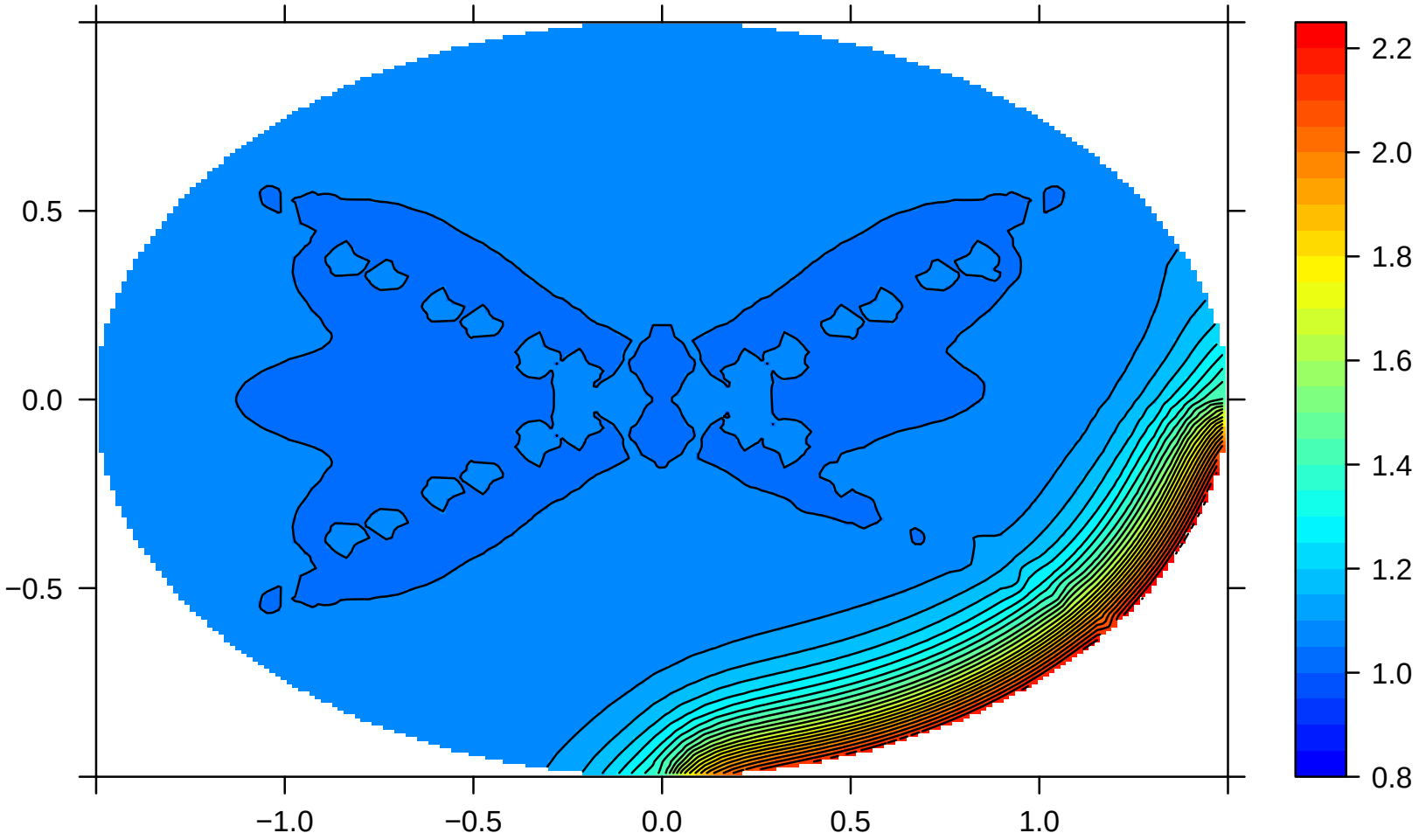
Elliptical domain, basis triangulation



Elliptical domain, stochastic boundary (variances)



Elliptical domain, mixed boundary (variances)



References

References (see also `r-inla.org`)

- F. Lindgren, 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), JRSSB, 73(4), 423–498.
- Preliminary draft book chapter on GMRF computation basics:
  `http://people.bath.ac.uk/fl353/tmp/gmrf.pdf`
- R. Ingebrigtsen, F. Lindgren, I. Steinsland (2013), *Spatial models with explanatory variables in the dependence structure*, Spatial Statistics, In Press (available online).
- G-A. Fuglstad, F. Lindgren, D. Simpson, H. Rue (2013),
  Exploring a new class of non-stationary spatial Gaussian random fields with varying local anisotropy, accepted for Statistica Sinica, arXiv:1304.6949
- D. Bolin, F. Lindgren (2013),
  Excursion and contour uncertainty regions for latent Gaussian models,
  JRSSB, in press. Accepted version at arXiv:1211.3946.
  CRAN package: `excursions`