Multiscale spatio-temporal statistics for global surface air temperatures

Finn Lindgren

UNIVERSITY OF
BATH

Bath, 1 May 2015

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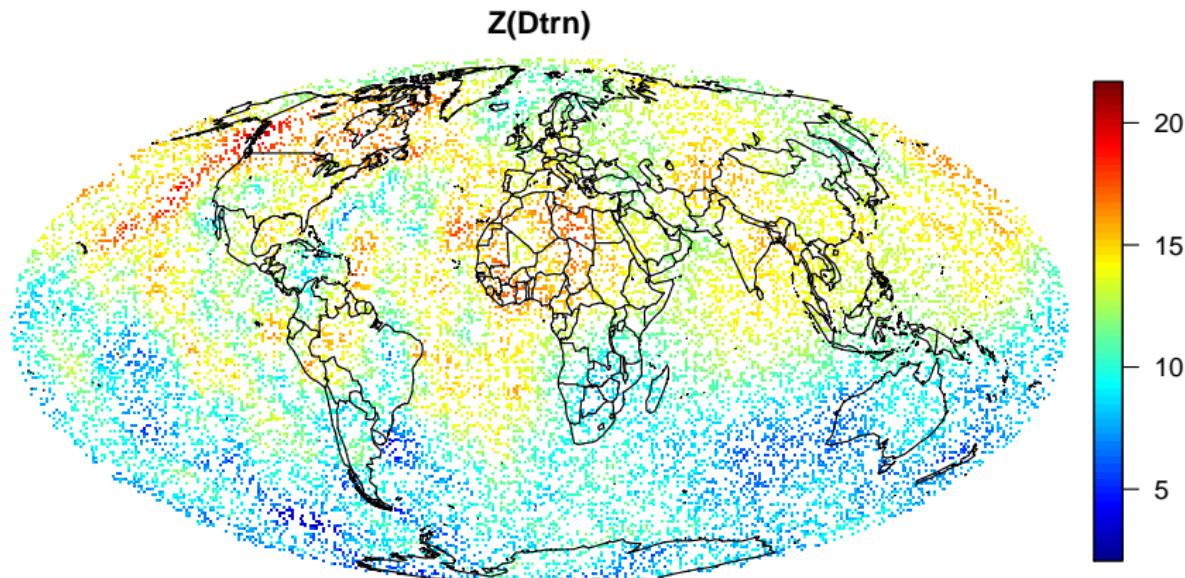
Illustration: Synthetic data mimicking satellite based CO_2 measurements.
Iregular data locations, uneven coverage, features on different scales.

Sparse spatial coverage of temperature measurements
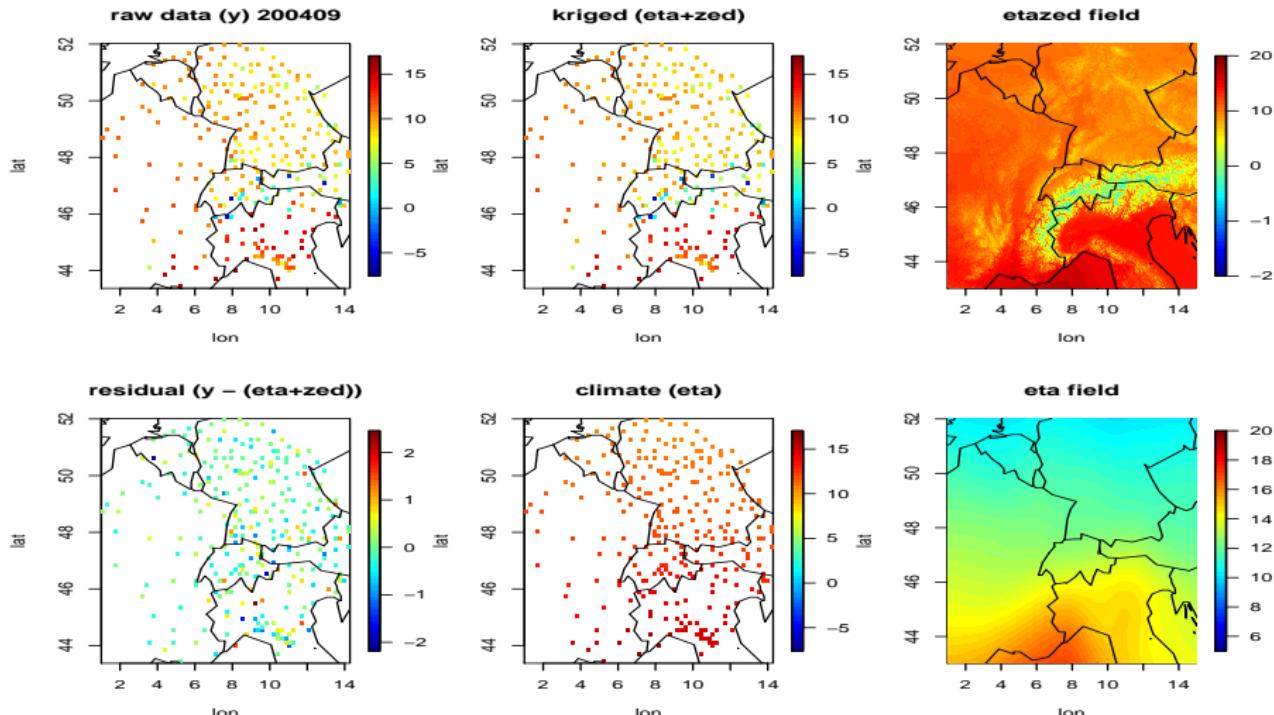
Regional observations: $\approx$ 20,000,000 from daily timeseries over 160 years

Note: This is a small *subset* of the full data!

EUSTACE (*EU Surface Temperatures for All Corners of Earth*)

*EUSTACE has received funding from the European Union's Horizon 2020
Programme for Research and Innovation, under Grant Agreement no 640171*

EUSTACE

EUSTACE will give publicly available daily estimates of surface air
temperature since 1850 across the globe for the first time by combining
surface and satellite data using novel statistical techniques.
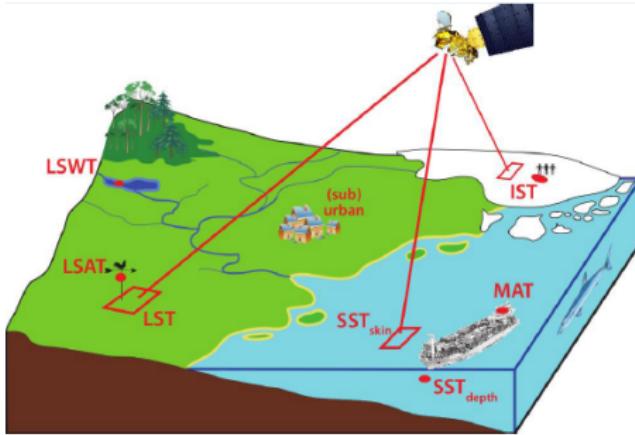
Nick A. Rayner, Renate Auchmann, Janette Bessembinder, Stefan Brönnimann, Yuri Brugnara, Laura Carrea, Darren Ghent, Elizabeth Good, Katie Herring, Jacob Høyer, John Kennedy, Albert Klein Tank, Finn Lindgren, Colin Morice, Chris Merchant, John Remedios, Ag Stephens and Rasmus Tonboe

contact: nick.rayner@metoffice.gov.uk

Royal Netherlands
Meteorological Institute
*Ministry of Infrastructure and the
Environment*

Science & Technology
Facilities Council

University of
Leicester

Met Office
Hadley Centre

University of
Reading

UNIVERSITY OF
BATH

UNIVERSITÄT
BERN
OESCHGER CENTRE
CLIMATE CHANGE RESEARCH

u^b

dmi

EUSTACE data and modelling workflow

- ▶ WP1: *Translate sensor data into local estimates of air temperatures*
  Relationships vary over land, ocean, ice, and lakes, and with season
  The estimation errors are spatially correlated
  Daily station data shows timeseries breaks in mean and distribution

- ▶ WP2: *Spatio-temporal blending of all data sources*
  Two approaches: "Advanced traditional" and "Ambitious"
  Point estimates of air temperature and uncertainty, and a random
  sample/ensemble from the posterior distribution

- ▶ WP3: *Validation of calibration models and data products*

Modelling challenges:

- ▶ Each data sources generates daily values of one (or more) of T_{max},
  T_{min}, T_{avg}, and T_{range} at irregular locations

- ▶ T_{max} and T_{min} are strongly non-Gaussian and dependent, but T_{avg} and
  T_{range} are less dependent.

- ▶ After compensating for the seasonal cycle, T_{avg} is close to Gaussian,
  but T_{range} is non-Gaussian.

Preliminary model framework

▶ Build a priori independent multi-component models for T_{avg} and T_{range}, in space ($\mathbf{s}$) and time (year t and time within year τ):

$$T_{\text{avg}}(\mathbf{s}, t, \tau) = \text{seasonal}_{\text{avg}}(\mathbf{s}, \tau) + \text{longterm}_{\text{avg}}(\mathbf{s}, t, \tau) + \text{shortterm}_{\text{avg}}(\mathbf{s}, t, \tau)$$

A transformation or copula model is needed for the range, non-linear $h(\cdot)$:

$$\log\left(\text{scaling}_{\text{range}}(\mathbf{s}, t, \tau)\right) = \text{seasonal}_{\text{range}}(\mathbf{s}, \tau) + \text{longterm}_{\text{range}}(\mathbf{s}, t, \tau)$$
$$T_{\text{range}}(\mathbf{s}, t, \tau) = \text{scaling}_{\text{range}}(\mathbf{s}, t, \tau) \ h\left(\text{shortterm}_{\text{range}}(\mathbf{s}, t, \tau)\right)$$

Each component is a Gaussian process with different spatial and temporal covariance properties.

▶ For observations of daily maxima and minima, define

$$T_{\text{min}} = T_{\text{avg}} - T_{\text{range}}/2$$
$$T_{\text{max}} = T_{\text{avg}} + T_{\text{range}}/2$$

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 $\mathbb{R}^d$

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

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

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

$\mathcal{W}(\cdot)$ 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
$\text{Cov}(\mathcal{W}(A), \mathcal{W}(B)) = |A \cap B|$, for all measurable $A, B \subseteq \mathbb{R}^d$. Not to be
confused with pointwise independent noise.)

Continuous and discrete Markov properties

Markov properties

S is a separating set for A and B: $u(A) \perp u(B) \mid u(S)$
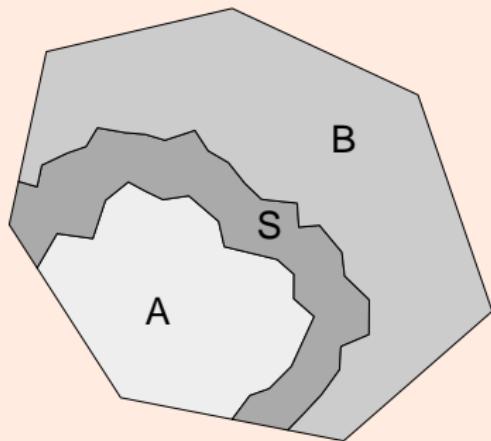
Solutions to
$$\left(\kappa^2 - \nabla \cdot \nabla\right)^{\alpha/2} u(\boldsymbol{s}) = \mathcal{W}(\boldsymbol{s})$$
are Markov when α is an integer.
(Rozanov, 1977)

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

B

S

A

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

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

Conditional distribution in a jointly Gaussian model

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

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

$\boldsymbol{Q}_{u|y} = \boldsymbol{Q}_u + \boldsymbol{A}^T \boldsymbol{Q}_{y|u} \boldsymbol{A}$ ($\sim$"Sparse iff ψ_k 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

Laplace approximations for non-Gaussian observations

Quadratic posterior log-likelihood approximation

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

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

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

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

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 latent Gaussian parts to some degree do scale to large non direct methods, but evaluating likelihoods becomes a very challenging problem, so EUSTACE will need to use alternative practical parameter estimation methods.

Climate and weather model (simplified)

▶ Climate process, simplified stochastic heat equation

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

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

▶ Weather anomaly, non-stationary spatial SPDE/GMRF
$\left(\kappa(\boldsymbol{s})^2 - \nabla \cdot \nabla\right) \left(\tau(\boldsymbol{s}) a(\boldsymbol{s}, t)\right) = \mathcal{W}_a(\boldsymbol{s}, t)$

▶ Temperature measurements from one or several sources
$y_i = a(\boldsymbol{s}_i, t_i) + z(\boldsymbol{s}_i, t_i) + \epsilon_i$, discretised into
$\boldsymbol{y} = \boldsymbol{A}(\boldsymbol{a} + (\boldsymbol{B} \otimes \boldsymbol{I})\boldsymbol{z}) + \boldsymbol{\epsilon}, \boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_{\epsilon}^{-1})$

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

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

Locally isotropic non-stationary precision construction

Finite element construction of basis weight precision

Non-stationary SPDE:
$$\left(\kappa(\boldsymbol{s})^2 - \nabla \cdot \nabla\right)\left(\tau(\boldsymbol{s})u(\boldsymbol{s})\right) = \mathcal{W}(\boldsymbol{s})$$

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
$$\boldsymbol{T} = \operatorname{diag}(\tau(\boldsymbol{s}_i)), \quad \boldsymbol{K} = \operatorname{diag}(\kappa(\boldsymbol{s}_i))$$
$$C_{ii} = \int \psi_i(\boldsymbol{s})\,d\boldsymbol{s}, \quad G_{ij} = \int \nabla\psi_i(\boldsymbol{s}) \cdot \nabla\psi_j(\boldsymbol{s})\,d\boldsymbol{s}$$
$$\boldsymbol{Q} = \boldsymbol{T}\left(\boldsymbol{K}^2\boldsymbol{C}\boldsymbol{K}^2 + \boldsymbol{K}^2\boldsymbol{G} + \boldsymbol{G}\boldsymbol{K}^2 + \boldsymbol{G}\boldsymbol{C}^{-1}\boldsymbol{G}\right)\boldsymbol{T}$$

For the temporally independent anomalies, we get $\boldsymbol{I} \otimes \boldsymbol{Q}_a$

GMRF precision for simplified stochastic heat equation

$$Q_z = M_2^{(t)} \otimes M_0^{(s)} + M_1^{(t)} \otimes M_1^{(s)} + M_0^{(t)} \otimes M_2^{(s)}$$

$$M_0^{(s)} = C + \gamma_{\mathcal{E}} \, G$$

$$M_1^{(s)} = G + \gamma_{\mathcal{E}} \, G \, C^{-1} \, G$$

$$M_2^{(s)} = G \, C^{-1} \, G + \gamma_{\mathcal{E}} \, G \, C^{-1} \, G \, C^{-1} \, G$$

Ignoring the degenerate aspect of the model, the precision structure can be used to formulate sampling as

$$Q_z z = \widetilde{L}_z w, \quad w \sim \mathcal{N}(0, I)$$

where $\widetilde{L}_z$ is a pseudo Cholesky factor,

$$\widetilde{L}_z = \left[\left[L_2^{(t)} \otimes L_C, \quad L_1^{(t)} \otimes L_G, \quad L_0^{(t)} \otimes G L_C^{-\top} \right], \right.$$
$$\left. \gamma_{\mathcal{E}}^{1/2} \left[L_2^{(t)} \otimes L_G, \quad L_1^{(t)} \otimes G L_C^{-\top}, \quad L_0^{(t)} \otimes G \, C^{-1} L_G \right] \right]$$

Posterior calculations

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

$$Q_{x|y} = \begin{bmatrix} I \otimes Q_a + A^\top Q_\epsilon A & -B \otimes Q_a \\ -B^\top \otimes Q_a & Q_z + B^\top B \otimes Q_a \end{bmatrix}$$

can be pseudo-Cholesky-factorised:

$$Q_{x|y} = \widetilde{L}_{x|y}\widetilde{L}_{x|y}^\top, \qquad \widetilde{L}_{x|y} = \begin{bmatrix} I \otimes L_a & 0 & A^\top L_\epsilon \\ -B \otimes L_a & \widetilde{L}_z & 0 \end{bmatrix}$$

Posterior expectation, samples, and marginal variances:

$$Q_{x|y}(\boldsymbol{\mu}_{x|y} - \boldsymbol{\mu}_x) = A^\top Q_\epsilon(\boldsymbol{y} - \boldsymbol{\mu}_x),$$

$$Q_{x|y}(\boldsymbol{x} - \boldsymbol{\mu}_{x|y}) = \widetilde{L}_{x|y}\boldsymbol{w}, \quad \boldsymbol{w} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}), \quad \text{or}$$

$$Q_{x|y}(\boldsymbol{x} - \boldsymbol{\mu}_x) = A^\top Q_\epsilon(\boldsymbol{y} - \boldsymbol{\mu}_x) + \widetilde{L}_{x|y}\boldsymbol{w}, \quad \boldsymbol{w} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}),$$

$$\text{Var}(x_i|\boldsymbol{y}) = \texttt{diag(inla.qinv(}Q_{x|y}\texttt{))} \quad \text{(requires Cholesky)}$$

Preconditioning for e.g. conjugate gradient solutions

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

Block Jacobi and Gauss-Seidel preconditioning

$$\text{Matrix split: } Q_{x|y} = L + D + L^\top$$

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

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

Remark: Block Gibbs sampling for a GMRF posterior

With $Q = Q_{x|y}$, $b = A^\top Q_\epsilon(y - \mu_x)$ and $\widetilde{x} = x - \mu_x$,

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

Multiscale Schur complement approximation

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

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

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

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

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

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

Triangulation mesh
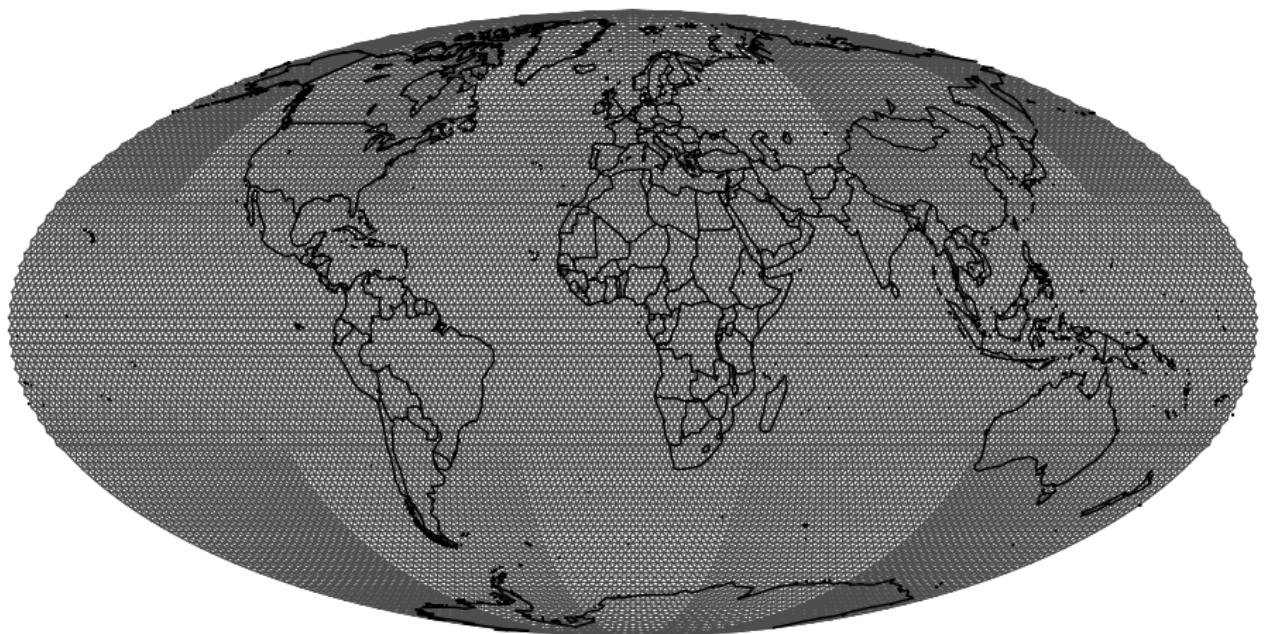
Finite element mesh (finer resolution needed!
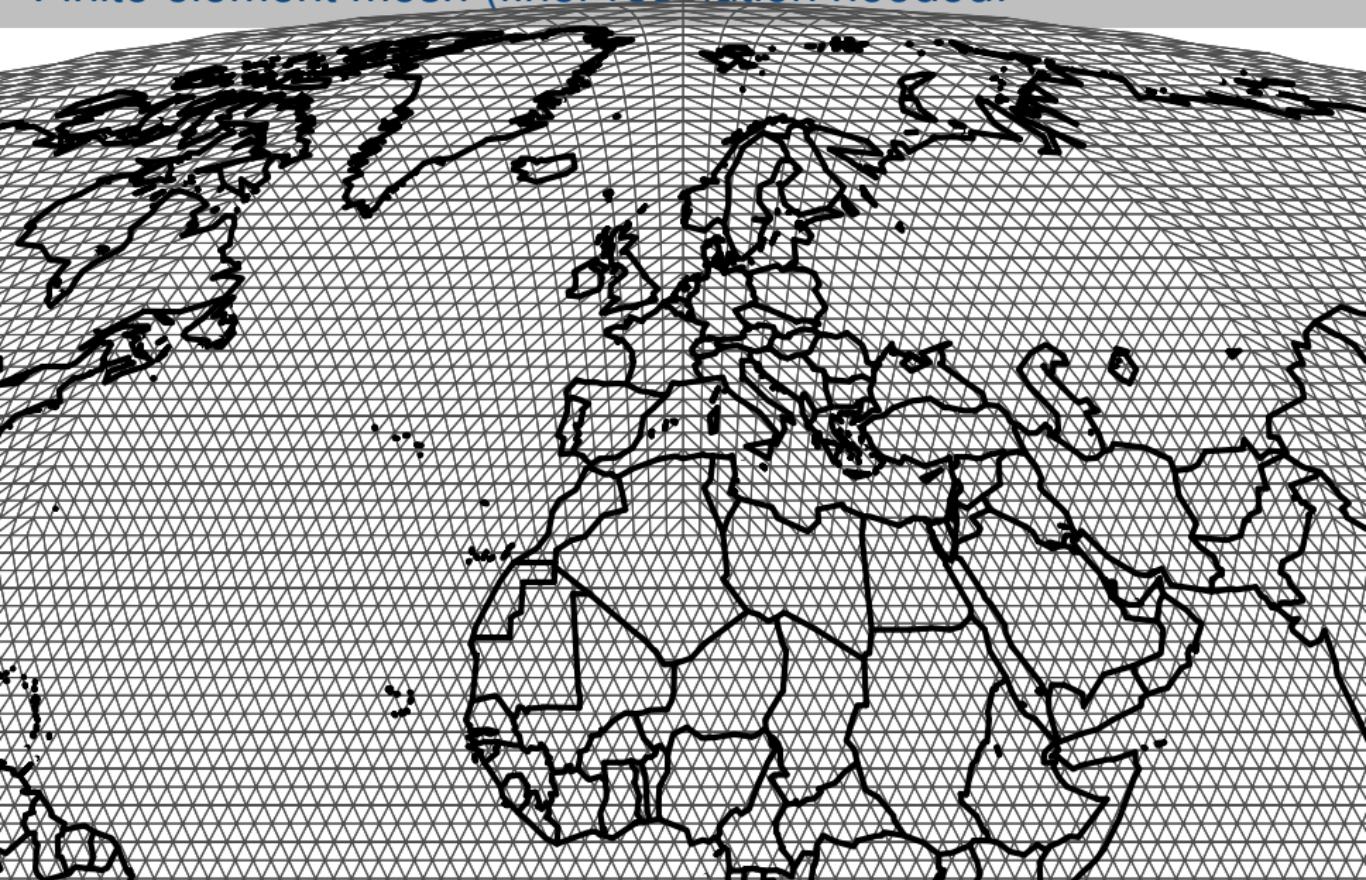
Domain decomposition

- ▶ 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

$\mathsf{E}(\boldsymbol{x}_i \mid \boldsymbol{y}) = \mathsf{E}\left(\mathsf{E}(\boldsymbol{x}_i \mid \boldsymbol{y}, \boldsymbol{x}_{\not\in\text{subblock}})\right)$

$\mathsf{Var}(\boldsymbol{x}_i \mid \boldsymbol{y}) = \mathsf{Var}\left(\boldsymbol{x}_i \mid \boldsymbol{y}, \boldsymbol{x}_{\not\in\text{subblock}}\right) + \mathsf{Var}\left(\mathsf{E}(\boldsymbol{x}_i \mid \boldsymbol{y}, \boldsymbol{x}_{\not\in\text{subblock}})\right)$

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'
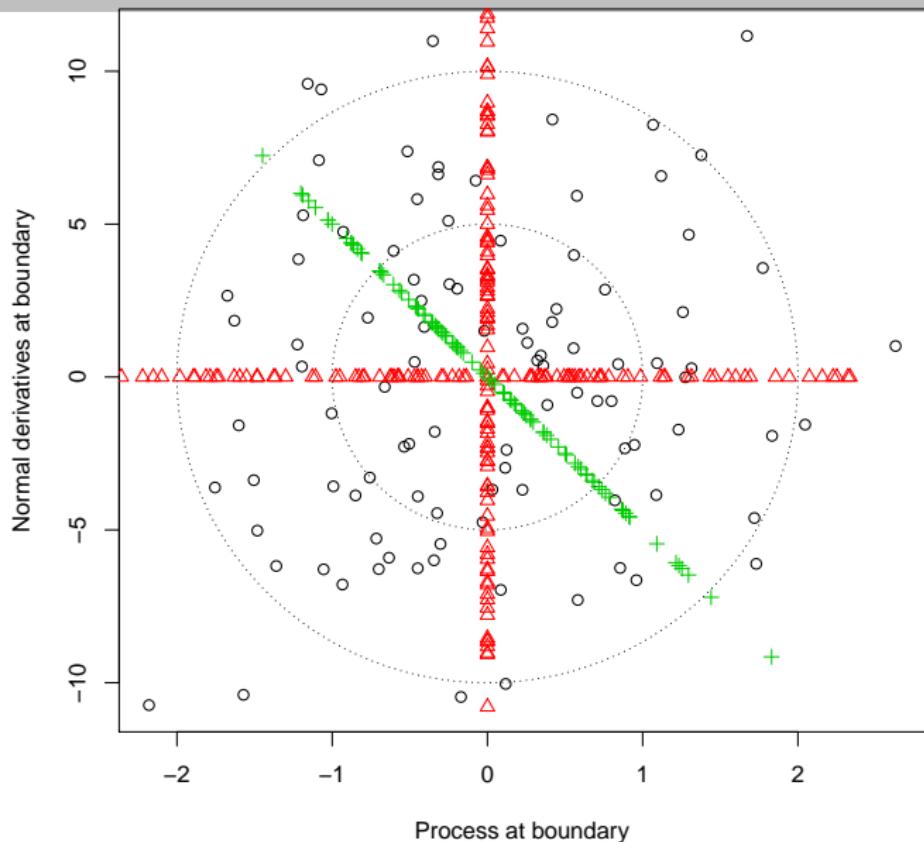
Stationary stochastic boundary adjustment

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:

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

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
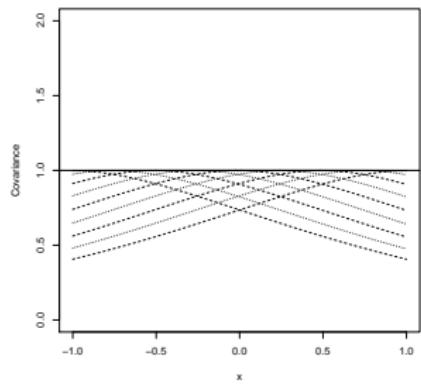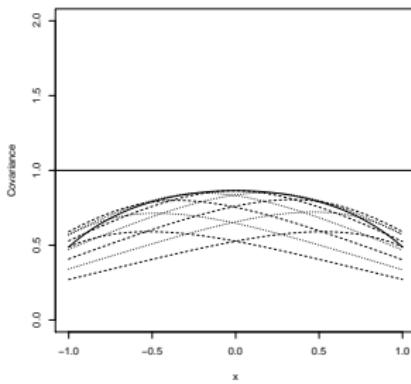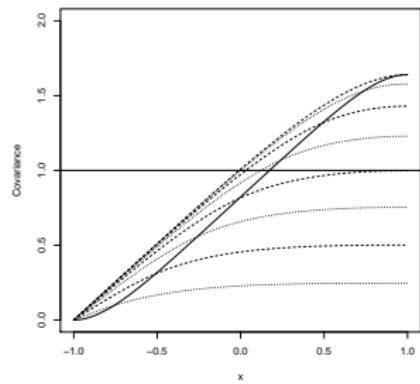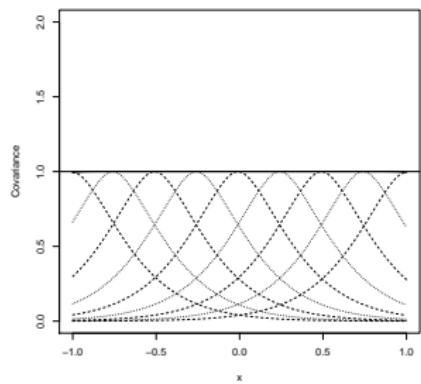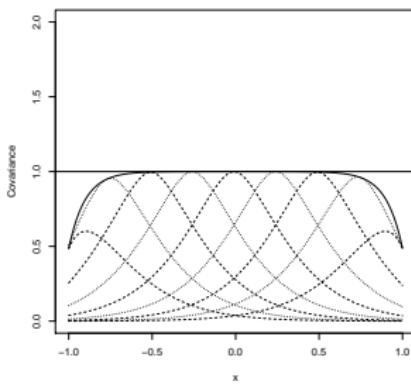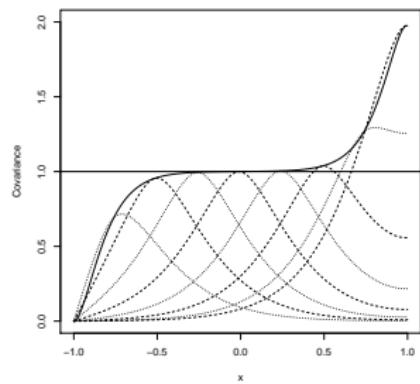
Square domain, stochastic boundary (variances)
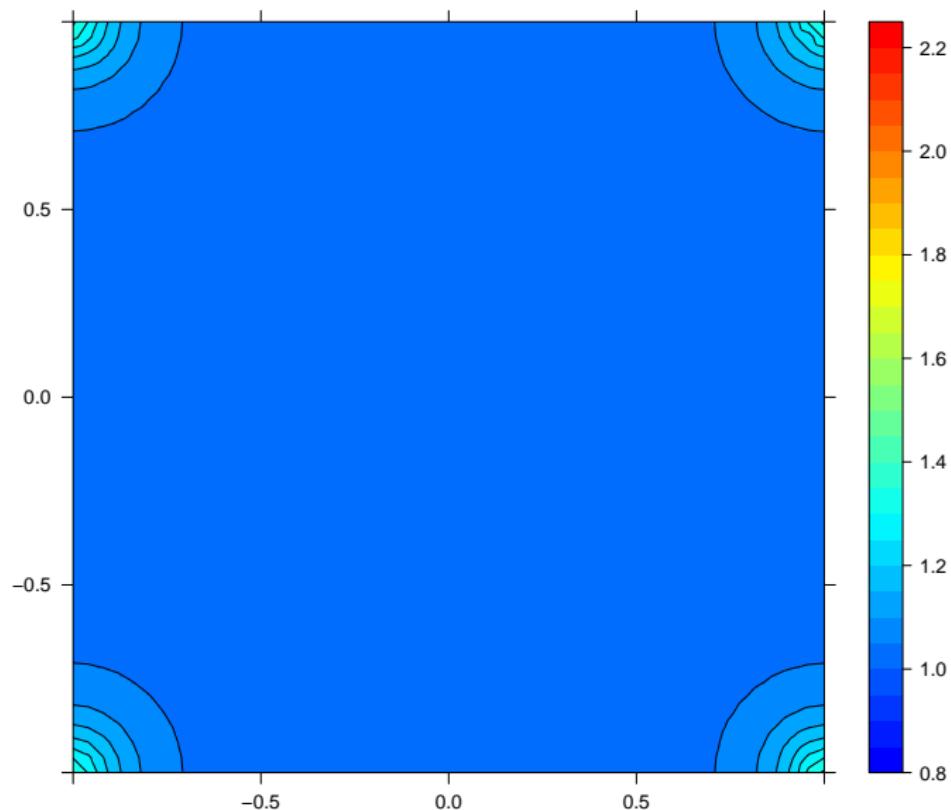
# Elliptical domain, stochastic boundary (variances)
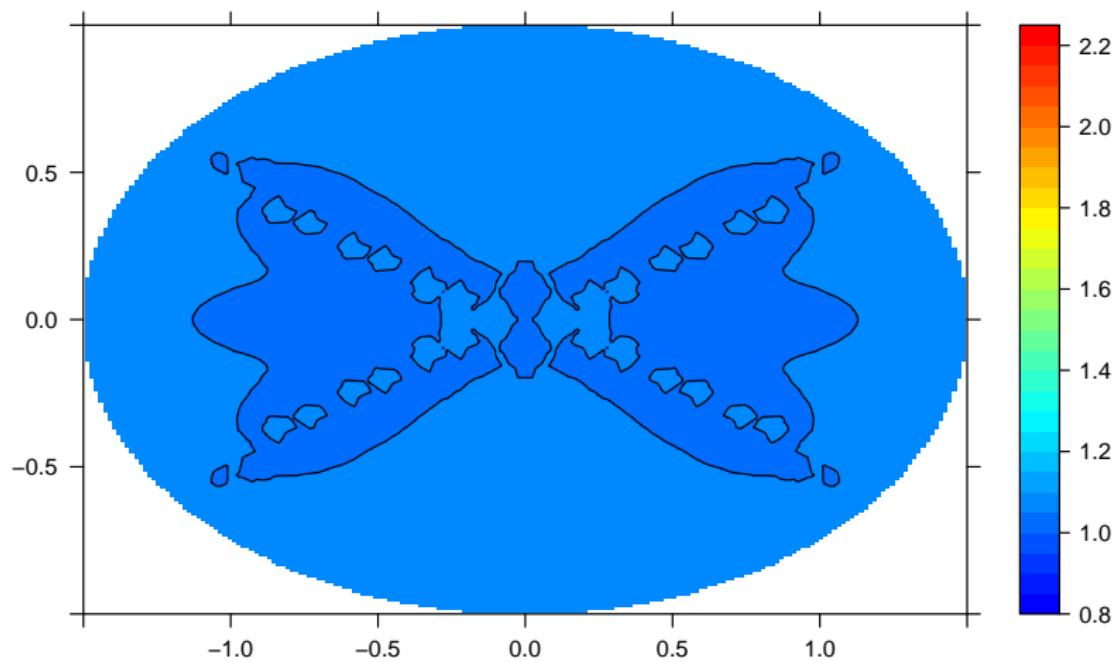
Non-stationary field
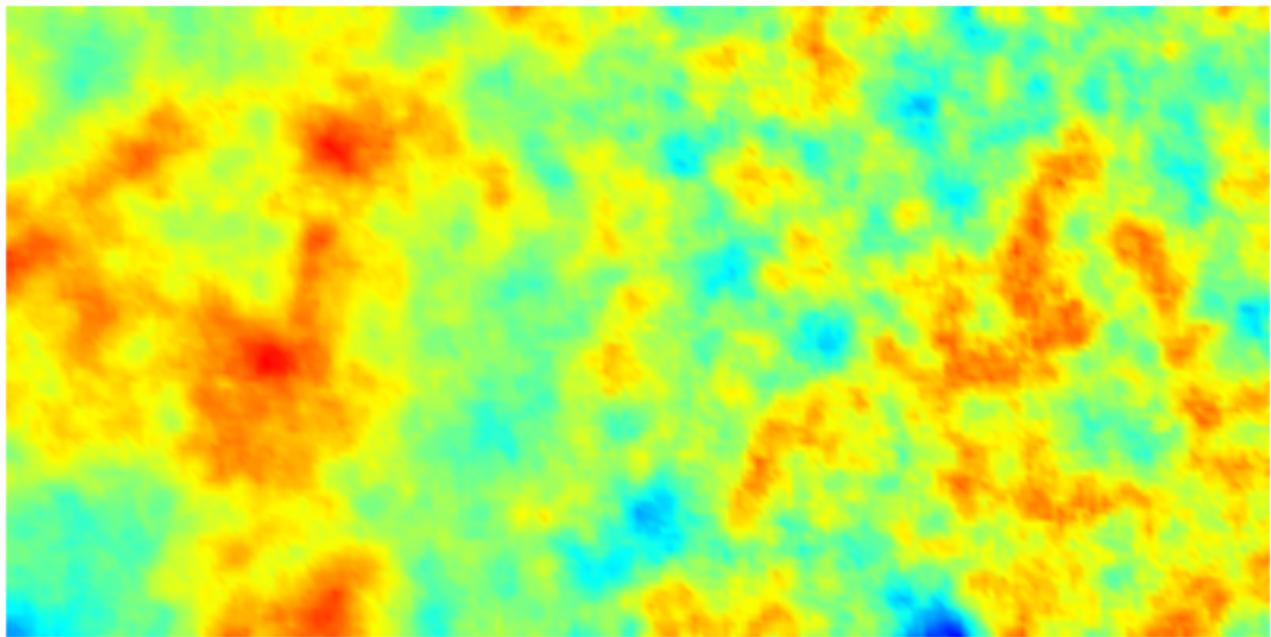
$$(\kappa(\boldsymbol{s})^2 - \nabla \cdot \nabla)u(\boldsymbol{s}) = \kappa(\boldsymbol{s})\mathcal{W}(\boldsymbol{s}), \quad \boldsymbol{s} \in \Omega$$

Anisotropic field on a globe via vector parameter field
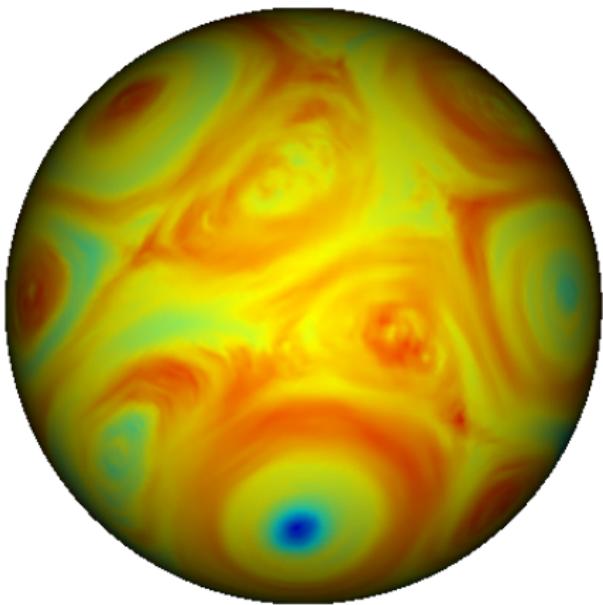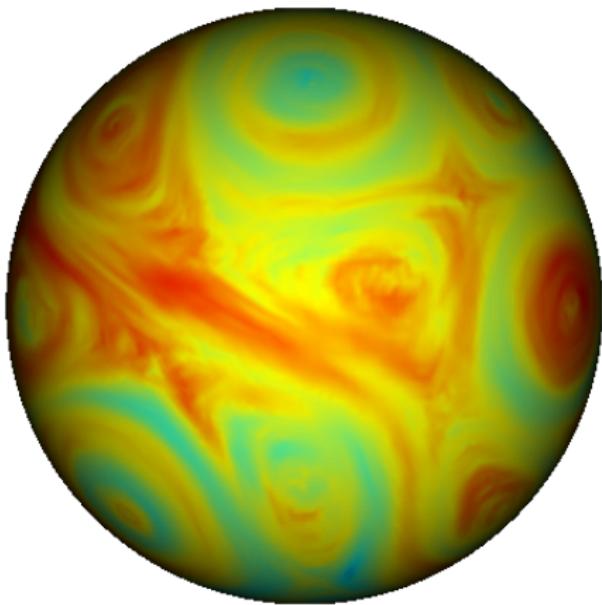
$$(\kappa(\boldsymbol{s})^2 - \nabla \cdot \boldsymbol{H}(\boldsymbol{s})\nabla)u(\boldsymbol{s}) = \kappa(\boldsymbol{s})\mathcal{W}(\boldsymbol{s}), \quad \boldsymbol{s} \in \Omega$$

Covariances for four reference points