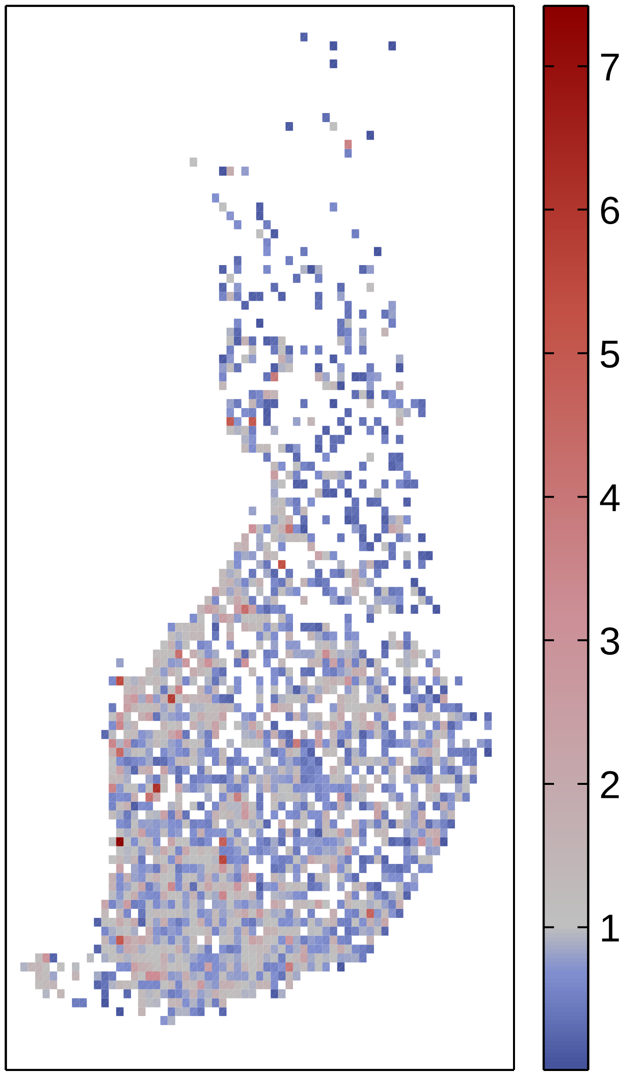
Sparse Log Gaussian Processes via MCMC for Spatial Epidemiology

Aki Vehtari and Jarno Vanhatalo*

HELSINKI UNIVERSITY OF TECHNOLOGY
Laboratory of Computational Engineering

Spatial epidemiology

- Mortality data from Statistics Finland
- Large scale:
  - over million deaths with various reasons
  - 30 years (one month accuracy)
  - in quarter million locations (250m accuracy)
- Complex
  - spatial effects
  - temporal effects
  - covariates



Example of raw data

The model

- The mortality is modeled as a Poisson process with mean $E\mu$

$$\mathbf{Y} \sim \mathrm{Poisson}(\mathbf{E}\mu),$$

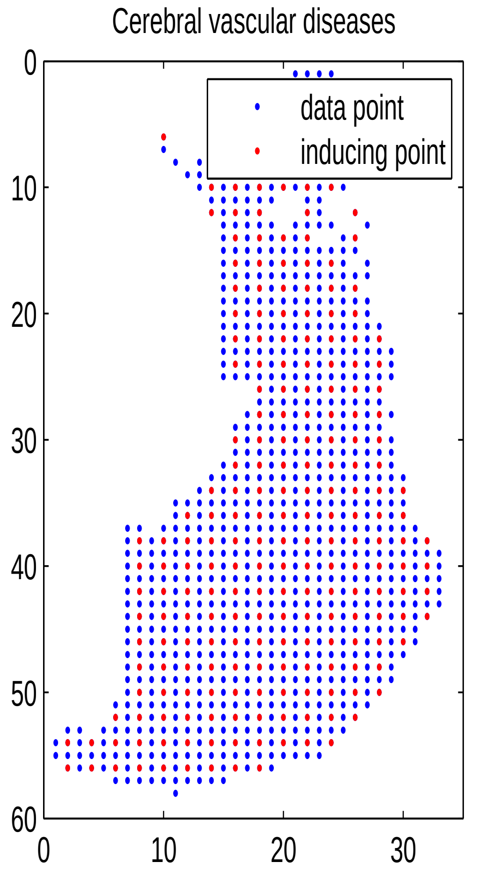
where E is the standardised expected number of deaths

- $\log(\mu)$ is given a GP prior with zero mean

$$\log(\mu) = \mathbf{f}(\mathbf{x}_i, \mathbf{x}_j) \sim \mathcal{GP}(\mathbf{0}, \mathbf{k}(\mathbf{x}_i, \mathbf{x}_j))$$

FITC

- A fully independent training conditional (FITC) (Snelson & Ghahramani, 2006; Quinonero-Candela & Rasmussen, 2005) sparse approximation is used to speed up GP computations



Approximate conditional posterior of latent values

- Following Christensen et al (2006), approximate posterior precision is obtained as

$$\Sigma^{-1} = \mathbf{K}^{-1} + \Sigma_{\mathrm{l}}^{-1},$$

where

$$\Sigma_{\mathrm{l}}^{-1} \approx -\frac{\partial^2 \log(\mathrm{Poisson}(E\lambda))}{\partial f^2} = E\mu,$$

where μ is approximated with its prior mean 1

Transformation when K is reduced rank

- Using matrix inversion lemma and eigen decomposition following equations for transformation are obtained

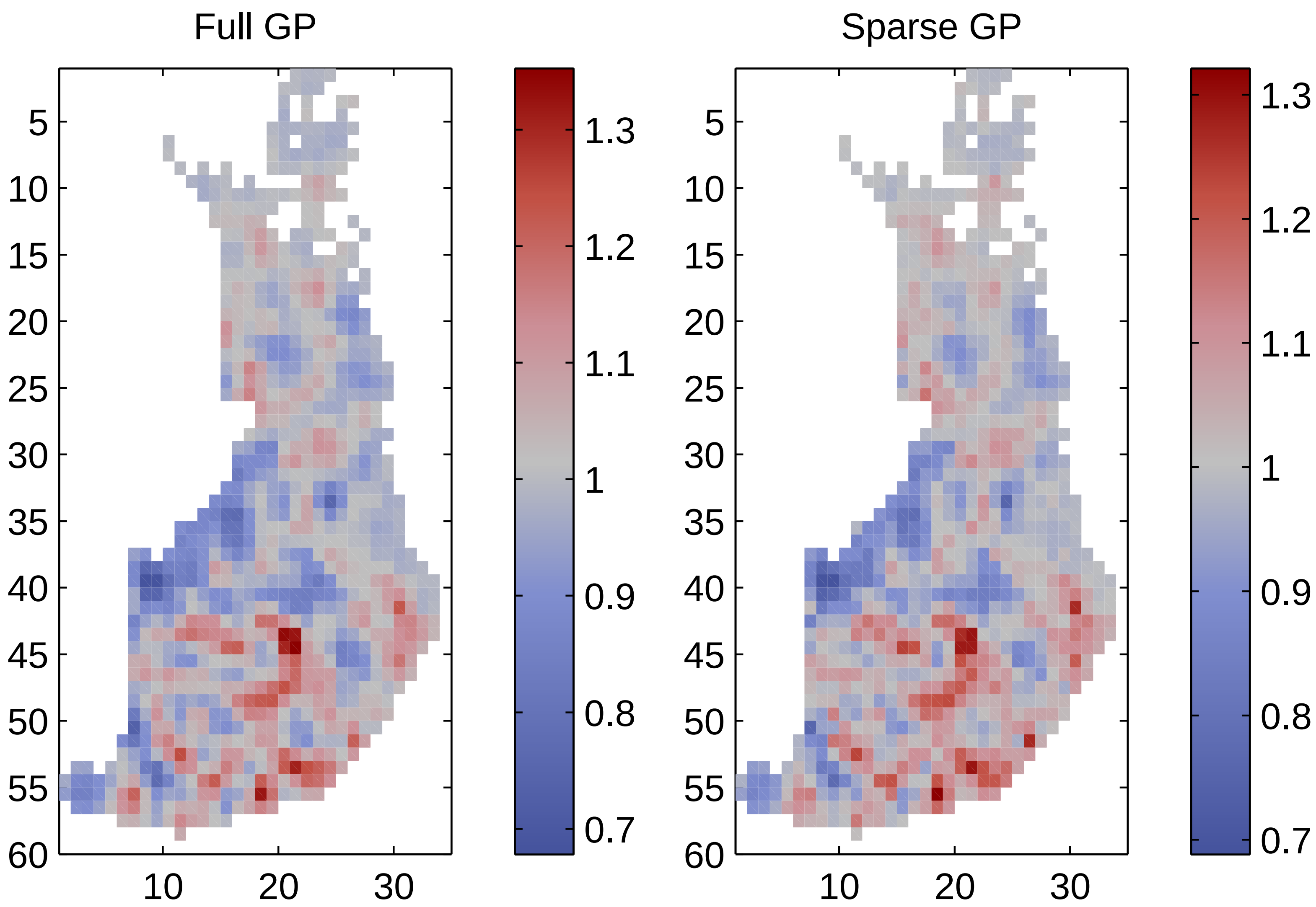
$$\mathbf{USU}^{\mathrm{T}} = \hat{\Lambda}^{1/2}\Lambda^{-1}\mathbf{K}_{\mathrm{f,u}}\left(\mathbf{K}_{\mathrm{u,u}} + \mathbf{K}_{\mathrm{u,f}}\Lambda^{-1}\mathbf{K}_{\mathrm{f,u}}\right)^{-1}\mathbf{K}_{\mathrm{u,f}}\Lambda^{-1}\hat{\Lambda}^{1/2}*$$

$$\mathbf{f} = \hat{\Lambda}^{1/2}(\tilde{\mathbf{f}} + \mathbf{UD}^{-1}\mathbf{U}^{\mathrm{T}}\tilde{\mathbf{f}} - \mathbf{UU}^{\mathrm{T}}\tilde{\mathbf{f}})$$

$$\tilde{\mathbf{f}} = \hat{\Lambda}^{-1/2}\mathbf{f} + \mathbf{UDU}^{\mathrm{T}}\hat{\Lambda}^{-1/2}\mathbf{f} - \mathbf{UU}^{\mathrm{T}}\hat{\Lambda}^{-1/2}\mathbf{f},$$

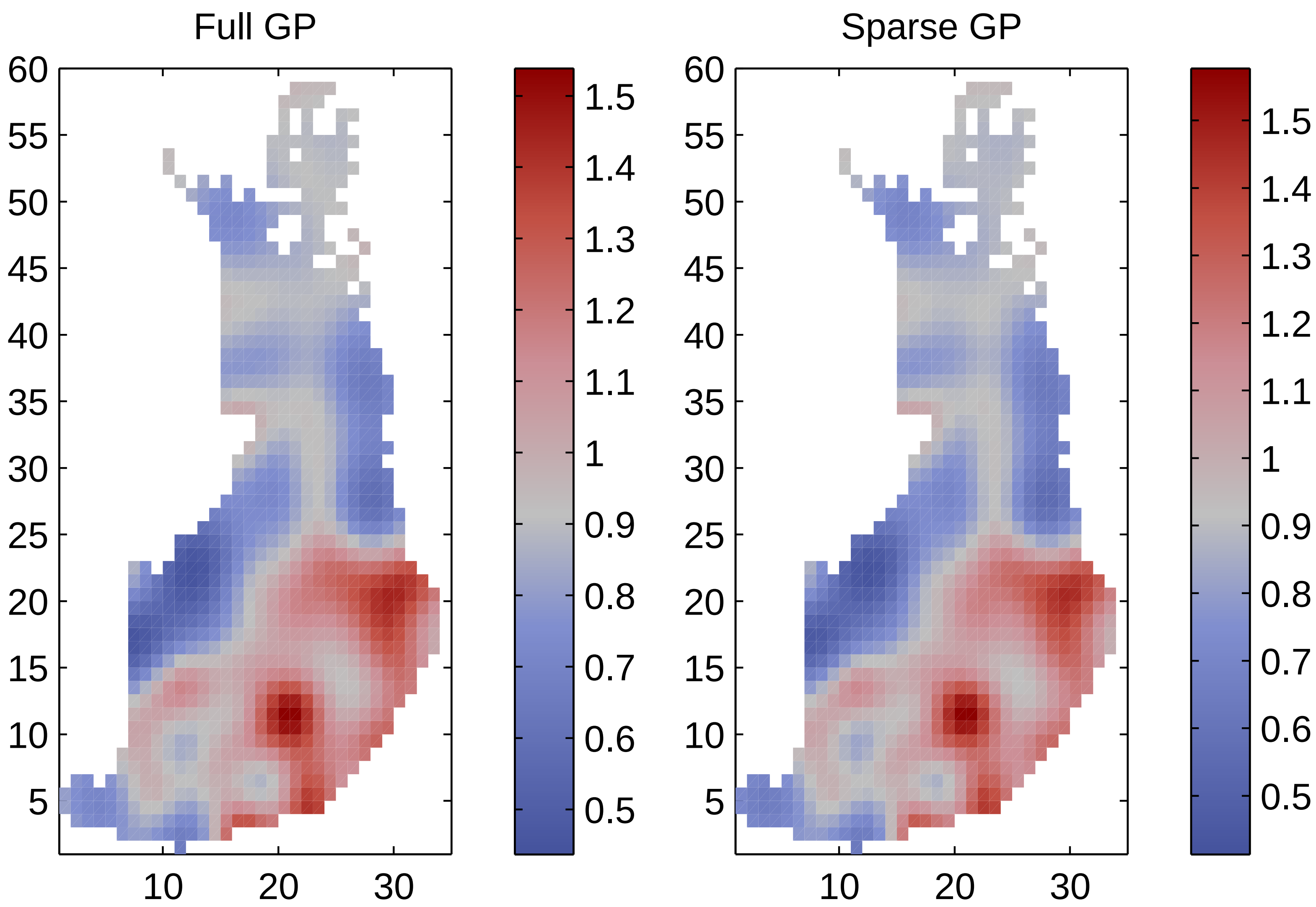
where $\mathbf{U}$ and $\mathbf{S}$ are matrices of eigenvectors and eigenvalues of the right hand side of the * respectively. $\mathbf{D}_{ii} = \sqrt{1 - \mathbf{S}_{ii}}$ and $\hat{\Lambda} = \left(\Sigma_l^{-1} + \Lambda^{-1}\right)^{-1}$.

Example 1



Full GP
Sparse GP
5
10
15
20
25
30
35
40
45
50
55
60
10
20
30
1.3
1.2
1.1
1
0.9
0.8
0.7

Example 2



Full GP
Sparse GP
60
55
50
45
40
35
30
25
20
15
10
5
10
20
30
1.5
1.4
1.3
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5