Spatial Statistics with Image Analysis

Johan Lindström[1]

[1]Mathematical Statistics
Centre for Mathematical Sciences
Lund University

Lund
September 26, 2019

Outline

- Spatial Statistics with Image Analysis
  - Examples
  - Bayesian statistics
  - Hierarchical modelling
  - Estimation Procedures
- Spatial Statistics
  - Stochastic Fields
  - Gaussian Markov Random Fields
  - Image Reconstruction
  - Environmental Data
- Non-Gaussian Data
  - Examples
- Corrupted Pixels
- Learn more

Spatial Statistics

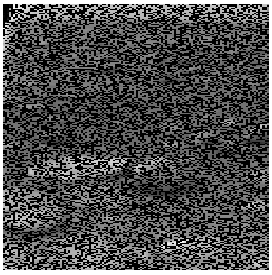
Many things vary in space and obsevations may depend on what happens at nearby locations. To modell and analys such data we need spatial dependence.

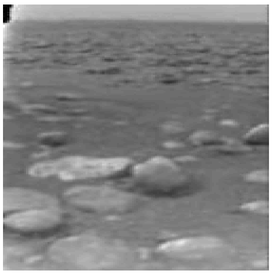
Spatial Interpolation

Given observations at some locations (pixels),
$y(\mathbf{u}_i),\ i = 1 \ldots n$
we want to make statements about the value at unobserved location(s), $x(\mathbf{u}_0)$.

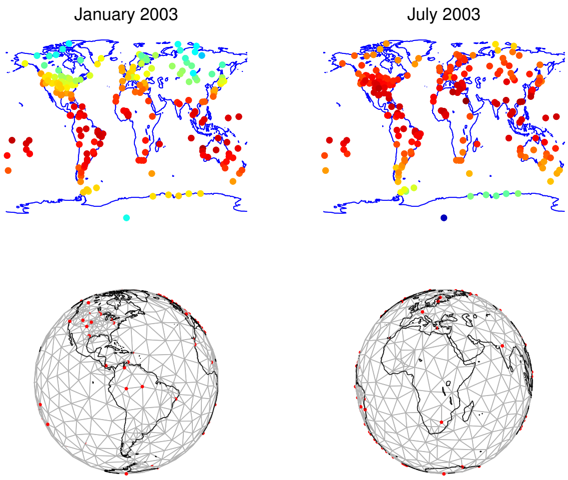
Stationary Stochastic Processes (FMSF10) in 2+ dimensions!

Examples: Image Reconstruction

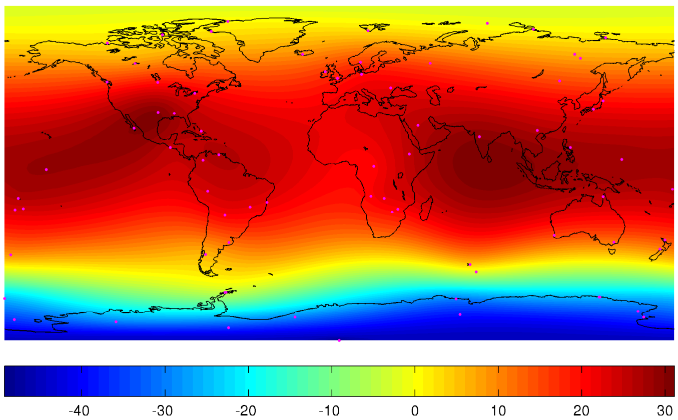


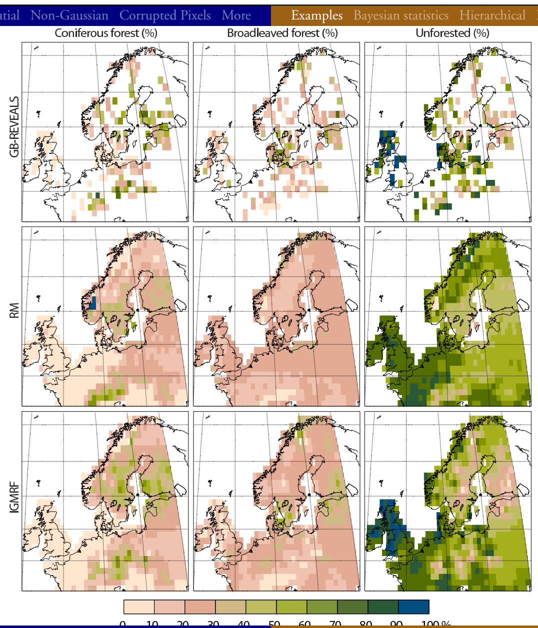


Global Temperature — Data



Global Temperature — Reconstruction





Bayesian modelling

A Bayesian model consists of

- A **prior**, "a priori", model for reality, $\boldsymbol{x}$, given by the probability density $\pi(\boldsymbol{x})$.
- A conditional **model for data**, $\boldsymbol{y}$, given reality, with density $p(\boldsymbol{y}|\boldsymbol{x})$.

The **prior** can be expanded into several **layers** creating a **Bayesian hierarchical model**.

Bayes' Formula

How should the **prior** and **data model** be combined to make statements about the reality $\boldsymbol{x}$, given observations of $\boldsymbol{y}$?

Bayes' Formula

$$p(\boldsymbol{x}|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|\boldsymbol{x})\pi(\boldsymbol{x})}{p(\boldsymbol{y})} = \frac{p(\boldsymbol{y}|\boldsymbol{x})\pi(\boldsymbol{x})}{\int_{\boldsymbol{x}'\in\Omega} p(\boldsymbol{y}|\boldsymbol{x}')\pi(\boldsymbol{x}')\,\mathrm{d}\boldsymbol{x}'}$$

$p(\boldsymbol{x}|\boldsymbol{y})$ is called the **posterior**, or "a posteriori", distribution.

Often, only the proportionality relation

$$p(\boldsymbol{x}|\boldsymbol{y}) \propto p(\boldsymbol{x},\boldsymbol{y}) = p(\boldsymbol{y}|\boldsymbol{x})\pi(\boldsymbol{x})$$

is needed, when seen as a function of x.

Hierarchical Models

- We often have some **prior knowledge** of the reality.
- Given knowledge of the true reality, what can we say about images and other data?
- Construct a model for observations given that we know the truth.
- Given data, what can we say about the unknown reality?

This is an **inverse problem**.

Bayesian hierarchical modelling (BHM)

A hierarchical model is constructed by systematically considering components/features of the data, and how/why these features arise.

Bayesian hierarchical modelling

A Bayesian hierarchical model typically consists of (at least)

Data model, $p(\boldsymbol{y}|\boldsymbol{x})$: Describing how **observations** arise assuming **known latent variables** $\boldsymbol{x}$.

Latent model, $p(\boldsymbol{x}|\vartheta)$: Describing how the **latent variables** (reality) behaves, assuming **known parameters**.

Parameters, $\pi(\vartheta)$: Describing our, sometimes vauge, **prior knowledge** of the parameters.

Estimation Procedures

Maximum A Posteriori (MAP): Maximise the posterior distribution $p(\boldsymbol{x}|\boldsymbol{y})$ with respect to $\boldsymbol{x}$.

- Standard optimisation methods
- Specialised procedures, using the model structure

Simulation: Simulate samples from the posterior distribution $p(\boldsymbol{x}|\boldsymbol{y})$.

- Markov chain Monte Carlo (MCMC)
- Gibbs sampling

Image Reconstruction

Spatial Interpolation

Given observations at some locations (pixels), $y(\boldsymbol{u}_i),\ i = 1 \dots n$
we want to make statements about the value at unobserved location(s), $x(\boldsymbol{u}_0)$.

The typical model consists of a **latent Gaussian field**

$$\boldsymbol{x} \in \mathsf{N}(\mu, \Sigma),$$

observed at locations $\boldsymbol{u}_i$, $i = 1, \dots, n$, with additive **Gaussian noise** (**nugget** or small scale variability)

$$y_i = x(\boldsymbol{u}_i) + \varepsilon_i \qquad \varepsilon_i \in \mathsf{N}\left(0, \sigma_\varepsilon^2\right).$$

Stochastic Fields

To perform the reconstruction (interpolation) we need a model for the **spatial dependence** between locations (pixels).

1. Assume a **latent Gaussian field**
$$\boldsymbol{x} \in \mathsf{N}(\mu, \Sigma).$$
2. Assume a regresion model for $\mu = \boldsymbol{B}\beta$.
3. Assume a parametric (**stationary**) model for the dependence (covariance)
$$\Sigma_{i,j} = \mathsf{C}(x(\boldsymbol{u}_i), x(\boldsymbol{u}_j)) = r(\boldsymbol{u}_i, \boldsymbol{u}_j; \vartheta) = r(\|\boldsymbol{u}_i - \boldsymbol{u}_j\|; \vartheta).$$
$r(\boldsymbol{u}_i, \boldsymbol{u}_j; \vartheta)$ is called the **covariance function**.

A local model

Instead of specifying the covariance function we could consider the **local behaviour** of pixels. A popular model is the **conditional autoregressive, CAR(1)** model.

$$x_{ij} = \frac{1}{4 + \kappa^2}\left(x_{i-1,j} + x_{i+1,j} + x_{i,j-1} + x_{i,j+1}\right) + \varepsilon,$$
$$\varepsilon \in \mathsf{N}\left(0, \frac{1}{\tau^2}\right).$$

This corresponds to a model for $\boldsymbol{x}$ where

$$\boldsymbol{x} \in \mathsf{N}\left(0, \boldsymbol{Q}^{-1}\right),$$

where $\boldsymbol{Q}$ is called the precision matrix

Matérn covariances

The Matérn covariance family

The covariance between two points at distance $\|\boldsymbol{h}\|$ is

$$r_M(\boldsymbol{h}) = \frac{\sigma^2}{\Gamma(\nu)\,2^{\nu-1}}(\varkappa\|\boldsymbol{h}\|)^{\nu} K_{\nu}(\varkappa\|\boldsymbol{h}\|)$$

Fields with Matérn covariances are solutions to a **Stochastic Partial Differential Equation (SPDE)** (Whittle, 1954),

$$\left(\varkappa^2 - \Delta\right)^{\alpha/2} x(\boldsymbol{u}) = \mathcal{W}(\boldsymbol{u}).$$

Lattice on $\mathbb{R}^2$

Order $\alpha = 1$ ($\nu = 0$):

$$\varkappa^2 \underbrace{\begin{bmatrix} & & \\ & 1 & \\ & & \end{bmatrix}}_{(\boldsymbol{C})} + \underbrace{\begin{bmatrix} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{bmatrix}}_{\approx -\Delta\ (\boldsymbol{G})}$$

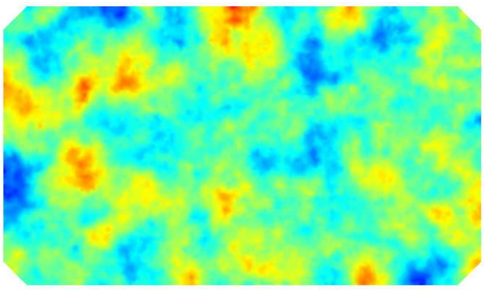
Order $\alpha = 2$ ($\nu = 1$):

$$\varkappa^4 \underbrace{\begin{bmatrix} & & \\ & 1 & \\ & & \end{bmatrix}}_{(\boldsymbol{C})} + 2\varkappa^2 \underbrace{\begin{bmatrix} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{bmatrix}}_{\approx -\Delta\ (\boldsymbol{G})} + \underbrace{\begin{bmatrix} & & 1 & & \\ & 2 & -8 & 2 & \\ 1 & -8 & 20 & -8 & 1 \\ & 2 & -8 & 2 & \\ & & 1 & & \end{bmatrix}}_{\approx \Delta^2\ (\boldsymbol{G}_2 = \boldsymbol{G}\boldsymbol{C}^{-1}\boldsymbol{G})}$$

Spatial models for data

GMRF representations of SPDEs can be constructed for oscillating, anisotropic, non-stationary, non-separable spatio-temporal, and multivariate fields on manifolds.

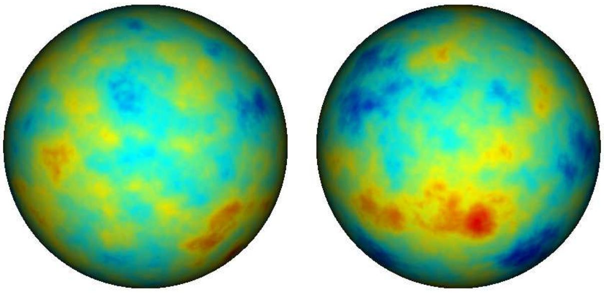
$$(\varkappa^2 - \Delta)(\tau x(\boldsymbol{u})) = \mathcal{W}(\boldsymbol{u}), \quad \boldsymbol{u} \in \mathbb{R}^d$$



Spatial models for data

GMRF representations of SPDEs can be constructed for oscillating, anisotropic, non-stationary, non-separable spatio-temporal, and multivariate fields on manifolds.

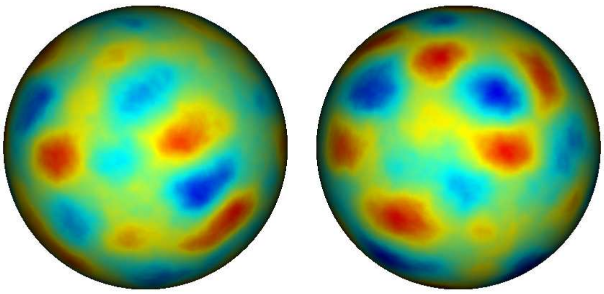
$$(\varkappa^2 - \Delta)(\tau \boldsymbol{x}(\boldsymbol{u})) = \mathcal{W}(\boldsymbol{u}), \quad \boldsymbol{u} \in \Omega$$



Spatial models for data

GMRF representations of SPDEs can be constructed for oscillating, anisotropic, non-stationary, non-separable spatio-temporal, and multivariate fields on manifolds.

$$(\varkappa^2 e^{i\pi\vartheta} - \Delta)(\tau \boldsymbol{x}(\boldsymbol{u})) = \mathcal{W}(\boldsymbol{u}), \quad \boldsymbol{u} \in \Omega$$



Spatial models for data

GMRF representations of SPDEs can be constructed for oscillating, anisotropic, non-stationary, non-separable spatio-temporal, and multivariate fields on manifolds.

$$(\varkappa_{\boldsymbol{u}}^2 + \nabla \cdot \boldsymbol{m}_{\boldsymbol{u}} - \nabla \cdot \boldsymbol{M}_{\boldsymbol{u}} \nabla)(\tau_{\boldsymbol{u}} x(\boldsymbol{u})) = \mathcal{W}(\boldsymbol{u}), \quad \boldsymbol{u} \in \Omega$$

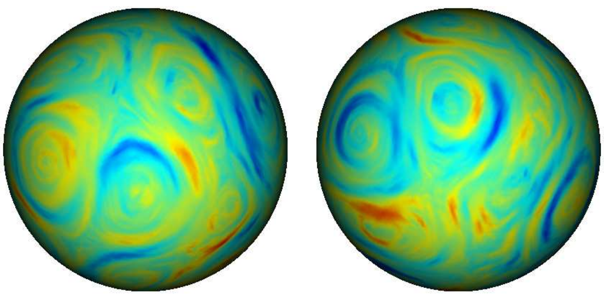


Image Reconstruction II

Model with observations, $\boldsymbol{y}$, and latent field, $\boldsymbol{x}$,

$$\boldsymbol{y}|\boldsymbol{x} \in \mathsf{N}\left(\boldsymbol{A}\boldsymbol{x}, \sigma^2\boldsymbol{I}\right) \qquad \boldsymbol{x} \in \mathsf{N}\left(\mu, \boldsymbol{Q}^{-1}\right).$$

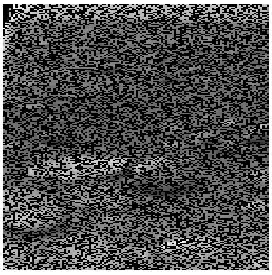
and $\boldsymbol{Q} = \kappa^2\boldsymbol{C} + \boldsymbol{G}$ or $\boldsymbol{Q} = \kappa^4\boldsymbol{C} + 2\kappa^2\boldsymbol{G} + \boldsymbol{G}\boldsymbol{C}^{-1}\boldsymbol{G}$.

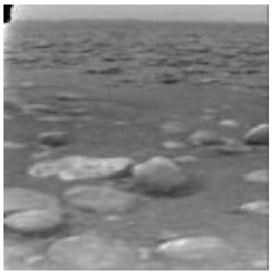
Interpolation using a GMRF

$$\mathsf{E}\left(\boldsymbol{x}|\boldsymbol{y}\right) = \mu + \frac{1}{\sigma^2}\boldsymbol{Q}_{x|y}^{-1}\boldsymbol{A}^\top\left(\boldsymbol{y} - \boldsymbol{A}\mu\right)$$

$$\mathsf{V}\left(\boldsymbol{x}|\boldsymbol{y}\right) = \boldsymbol{Q}_{x|y}^{-1} = \left(\boldsymbol{Q} + \frac{1}{\sigma^2}\boldsymbol{A}^\top\boldsymbol{A}\right)^{-1}$$

Image Reconstruction





Satellite Data — Vegetation

January 1999

July 1999

Satellite Data — Models

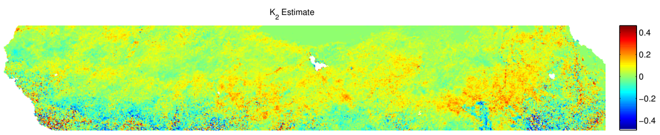
Independent pixels:
For each pixel, do a standard linear regression.

$$x(t) = k \cdot t + m + \varepsilon_t \qquad \varepsilon_t \in \mathsf{N}\left(0, \sigma^2\right)$$

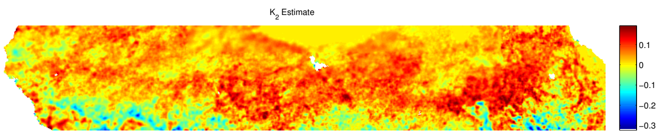
Dependent pixels:
But neighbouring pixels probably behave similarly. Account for dependence in regression coefficients.

$$x(s,t) = k(s) \cdot t + m(s) + \varepsilon_{s,t} \qquad \varepsilon_{s,t} \in \mathsf{N}\left(0, \sigma^2\right)$$
$$k(s) \in \mathsf{N}\left(0, \boldsymbol{Q}_k^{-1}\right) \qquad m(s) \in \mathsf{N}\left(0, \boldsymbol{Q}_m^{-1}\right)$$

Satellite Data — Trend in Vegetation David Bolin



Independent estimates



Correlated estimates

Non-Gaussian Data

Bayesian hierarchical modelling

A Bayesian hierarchical model typically consists of (at least)

Data model, $p(\boldsymbol{y}|\boldsymbol{x})$: Describing how **observations** arise assuming **known latent variables $\boldsymbol{x}$**.

Latent model, $p(\boldsymbol{x}|\vartheta)$: $\boldsymbol{x} \in \mathsf{N}\left(\mu, \boldsymbol{Q}^{-1}\right)$.

Parameters, $\pi(\vartheta)$

So far we have assumed Gaussian observations

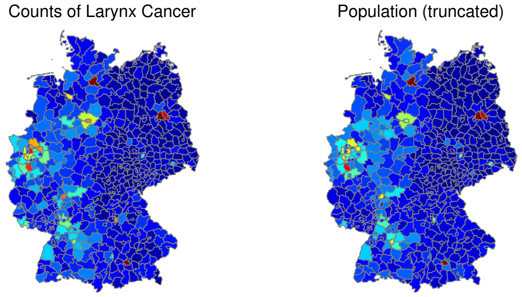
$$\boldsymbol{y}|\boldsymbol{x} \in \mathsf{N}\left(\boldsymbol{A}\boldsymbol{x}, \sigma^2\boldsymbol{I}\right)$$

However we could (almost) as easily handle

$$\boldsymbol{y}_i|\boldsymbol{x} \in \mathsf{F}\left(g(\boldsymbol{A}\boldsymbol{x});\ \vartheta\right)$$

Larynx Cancer — Count data

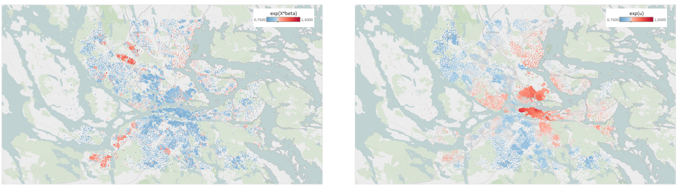
Given counts of larynx cancer cases, y_i, and population in each region, E_i, we want to estimate the risk of cancer.



$$y_i|x_i \in \mathrm{Po}\left(E_i \exp(x_i)\right)$$

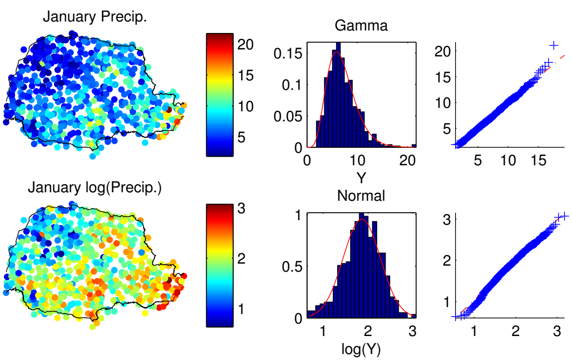
Insurance Claims — Count data Oscar Tufvesson

Given the number of insurance claims, y_i, we want to estimate the risk of an accident.



$$y_i|\eta_i \in \mathrm{Po}\left(E_i \exp(\eta_i)\right)$$
$$\boldsymbol{\eta} = \boldsymbol{B}\beta + \boldsymbol{x}$$

Parana Rainfall — Positive data

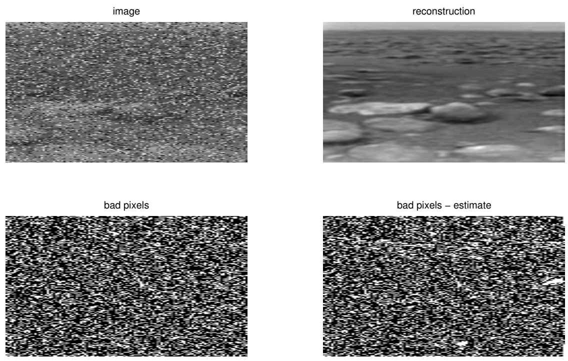


$$y_i|x_i \in \Gamma\left(b, \frac{e^{x_i}}{b}\right)$$

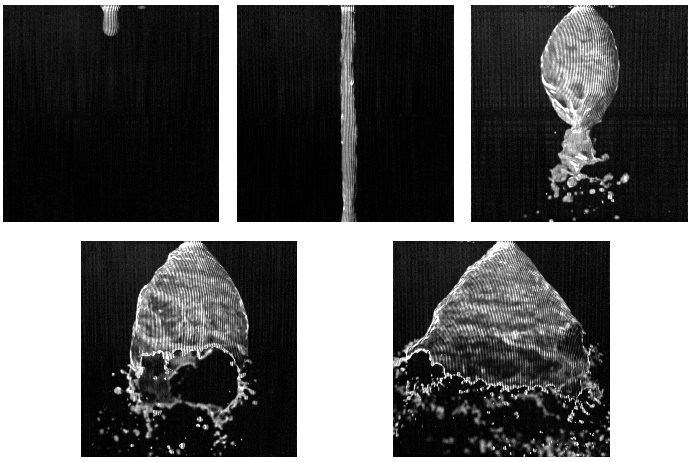
Image Reconstruction — Corrupted Pixels

- Typically we don't know which pixels are bad.
- A better model is then
  - Assume an underlying image, $\boldsymbol{x}$.
  - Assume an indicator image for **bad** pixels, $\boldsymbol{z}$.
  - Given the indicator we either observe the correct pixel value from $\boldsymbol{x}$ or noise.
- Use Bayes' formula to compute the distribution for the unknown image (and indicator) given observations and parameters.

Image Reconstruction — Corrupted pixels



Potential Thesis Project — Images from Combustion Physics



Learn more!

What?
Spatial statistics with image analysis, FMSN20

When?
HT2-2019, October–December

Where?
Information and Matlab files will be available at
`www.maths.lth.se/matstat/kurser/fmsn20masm25/`
(currently the 2018 webpage, updated soon)

Who?
Lecturer: Johan Lindström
MH:319