Bayesian Modelling for Spatial Data

“Hierarchical Modeling and Analysis for Spatial Data”

Sudipto Banerjee, Bradley P. Carlin and Alan E. Gelfand

Eva Sabiote and Katharina Krämer

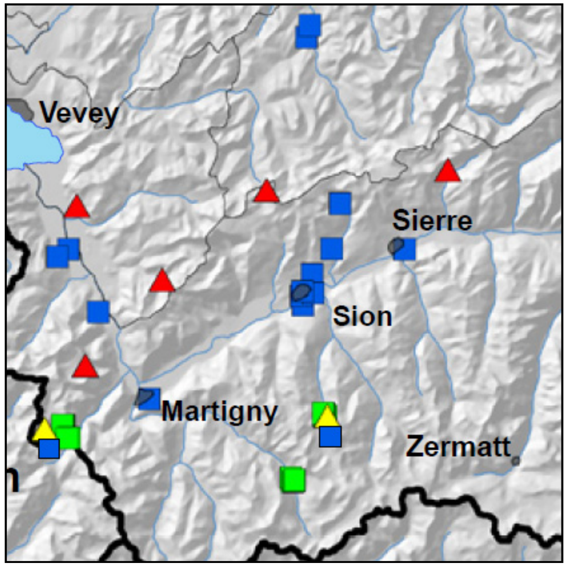
ETH Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zürich

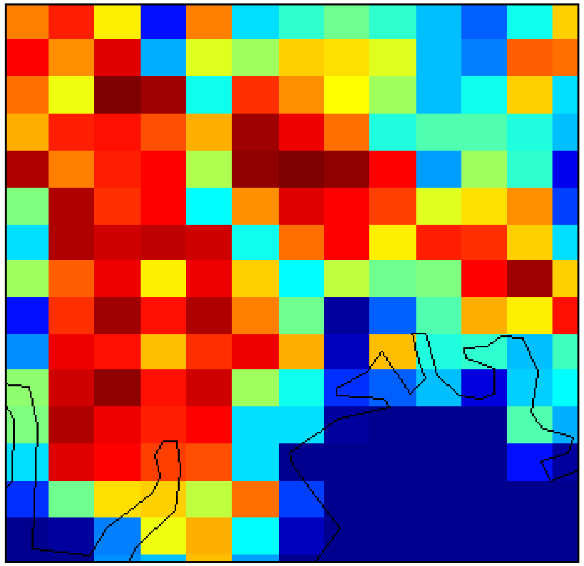
ibk Institute of Structural Engineering
Group Risk and Safety

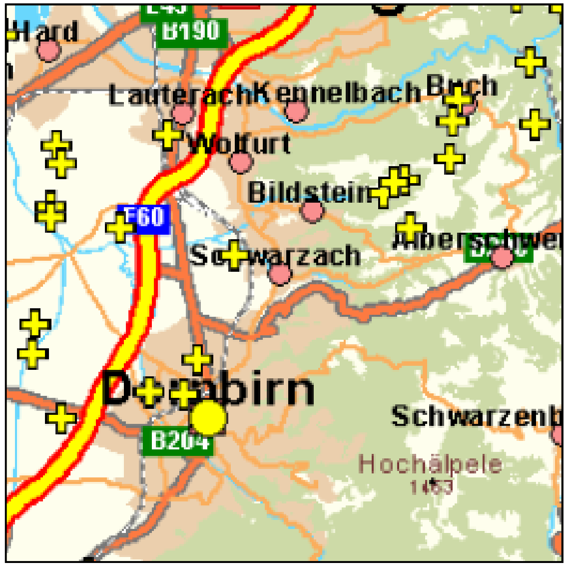
Introduction

- Types of spatial data, Problem setting
- Important properties of random fields
- Introduction to classical approach
- Basic model for homogenous random field
- Bayesian modelling: Inference and prediction
- Non-hierarchical vs. hierarchical approach
- Generalizations

3 Types of spatial data sets:







Point-referenced data

Known locations

$$Y(\mathbf{s}), \quad \mathbf{s} \in \Re^r$$

Aereal data

Aeral units ("blocks")

$$Y(B_i), \quad B_i \in D$$

Point pattern data

Random locations

$$Y(\mathbf{s}) = 1, \quad \mathbf{s} \in D$$

Objectives:

- Modelling the spatial variation of a random Variable $Y(\mathbf{s})$ at locations $\mathbf{s} \in D$.
- Description of trends and spatial correlation.
- Prediction for unobserved times / locations.

Examples:

- Soil properties on a construction site.
- Chloride concentration on a concrete surface.

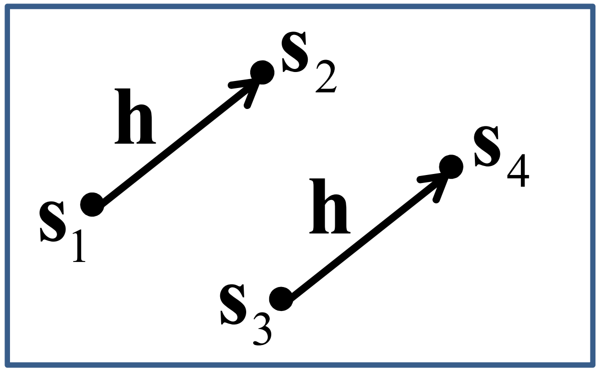
Stationarity

- Strict Stationarity:

$$dist(Y(\mathbf{s}_1),\ldots,Y(\mathbf{s}_n)) = dist(Y(\mathbf{s}_1+\mathbf{h}),\ldots,Y(\mathbf{s}_n+\mathbf{h})) \quad \forall\,\mathbf{h}\in\Re^r$$

- Weak Stationarity:

$$\mu(\mathbf{s}) = const. \qquad Cov(Y(\mathbf{s}),Y(\mathbf{s}+\mathbf{h})) = C(\mathbf{h})$$



Homogeneity

- Intrinsic Stationarity

$$E[Y(\mathbf{s}+\mathbf{h})-Y(\mathbf{s})]=0$$

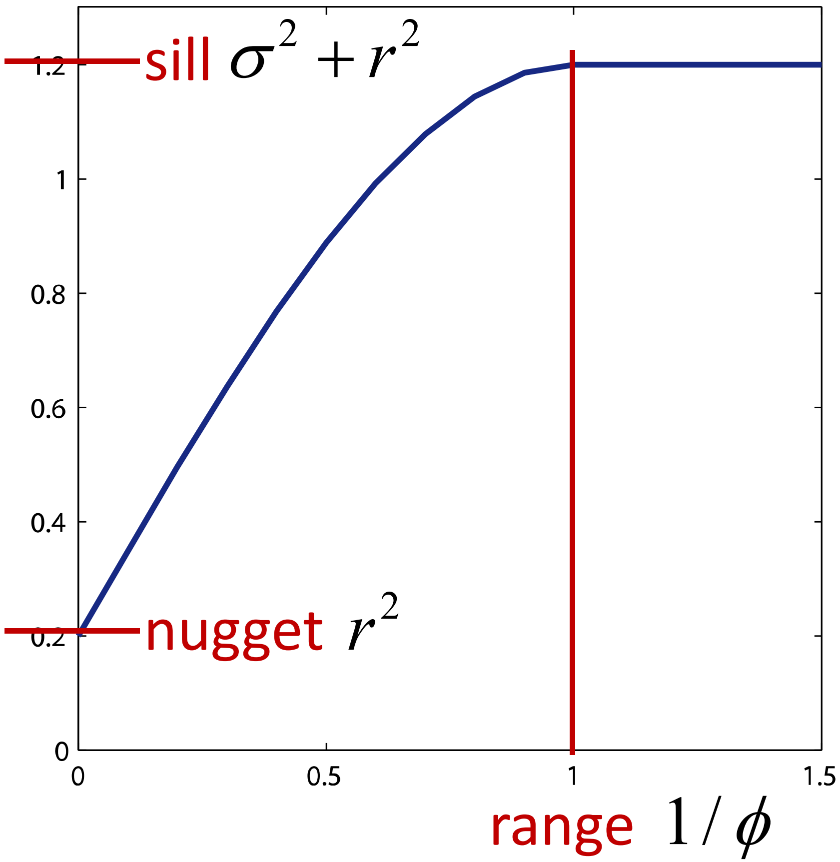
$$Var[Y(\mathbf{s}+\mathbf{h})-Y(\mathbf{s})]=2\gamma(\mathbf{h})$$ Variogram

- Isotropy

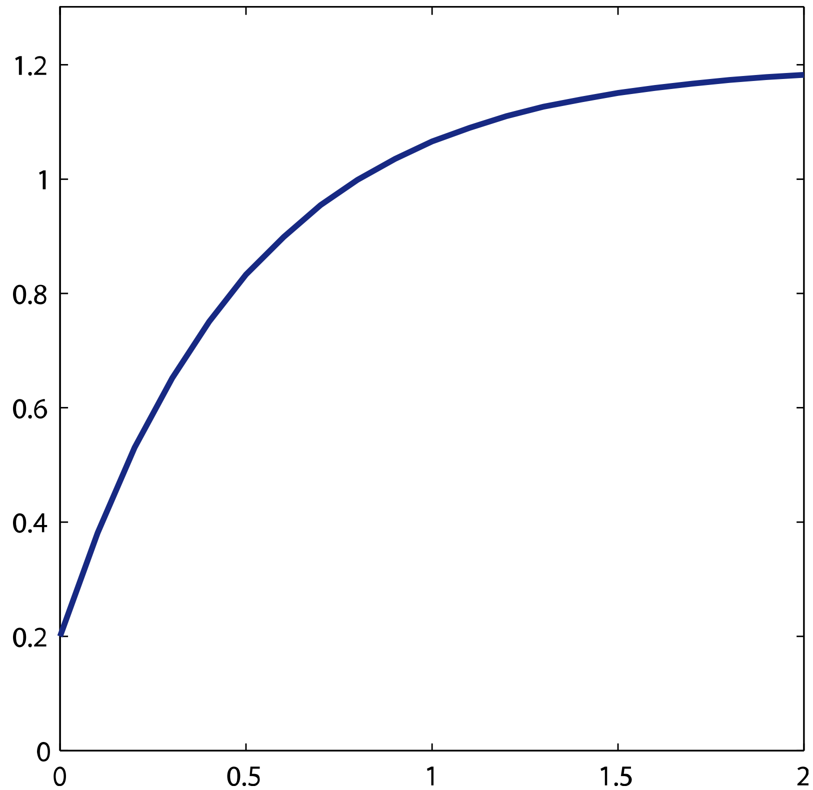
$$\gamma(\mathbf{h})=\gamma(\|\mathbf{h}\|)$$

⇨ Simple parametric variograms can be used to model spatial correlation structure

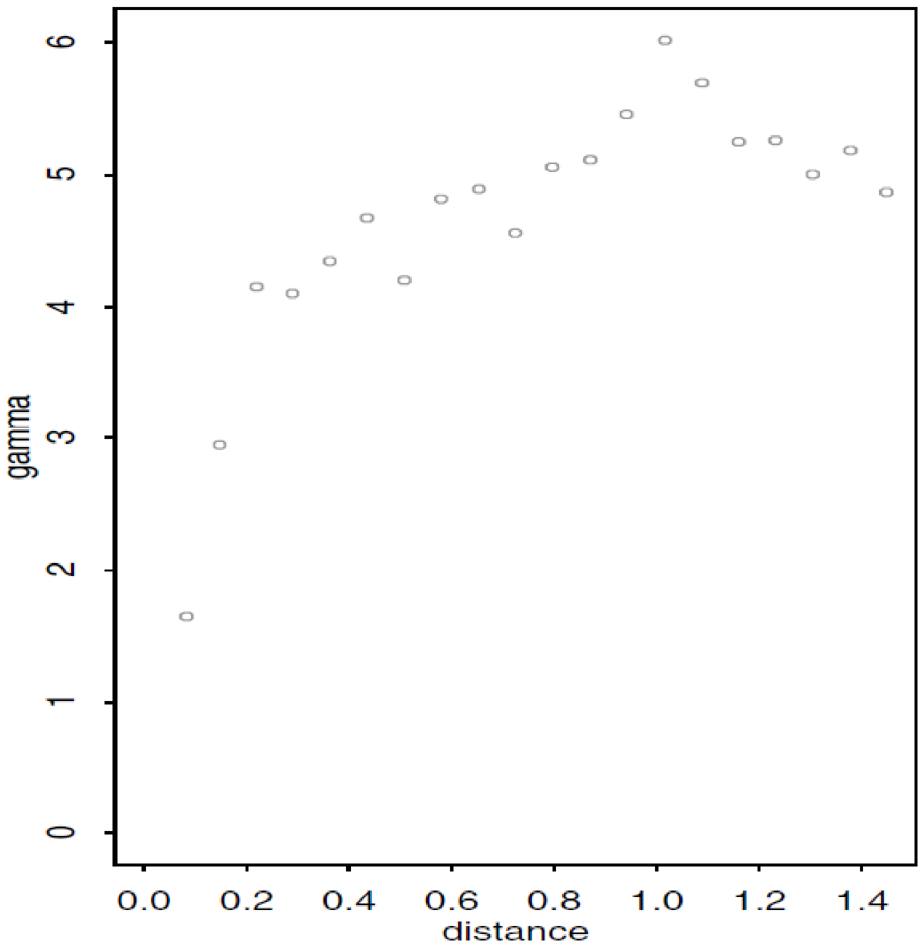
Spherical semivariogram



Exponential semivariogram



Empirical semivariograms



Empirical semivariogram:

$$\hat{\gamma}(t) = \frac{1}{2N(t)} \sum_{(\mathbf{s}_i, \mathbf{s}_j \notin N(t))} [Y(\mathbf{s}_i) - Y(\mathbf{s}_j)]^2$$

- Similar to histogram construction.
- Sensitive to outliers.
- $N(t)$ varies across bins.

Objective:

- Spatial prediction given data:

  Predict $Y(\mathbf{s}_0)$ given $\mathbf{Y} = (Y(\mathbf{s}_1), \ldots, Y(\mathbf{s}_1))$

Approach:

- Classical Kriging: Minimum mean-squared error estimate.
- Bayesian Kriging: Specify prior distribution for model parameters, Bayesian updating .

Example of a simple spatial process

Basic Model:

$$Y(\mathbf{s}) = \mu(\mathbf{s}) + \underbrace{\omega(\mathbf{s}) + \varepsilon(\mathbf{s})}$$

Spatial & nonspatial variation

Assumptions:

- Gaussian spatial process: $\mathbf{Y} \sim MVN$
- Homogeneity (stationarity and isotropy).
- No covariates.

Example of a simple spatial process

Basic Model:

$$Y(\mathbf{s}) = \mu(\mathbf{s}) + \omega(\mathbf{s}) + \varepsilon(\mathbf{s})$$

$$\mu(\mathbf{s}) = \mathbf{x}^T(\mathbf{s})\boldsymbol{\beta} \equiv \mu \quad const.$$

$$\omega(\mathbf{s}) = f(\sigma^2, \phi)$$ Partial sill, range

$$\varepsilon(\mathbf{s}) = f(r^2)$$ Nugget

$$\mathbf{Y}\,|\,\boldsymbol{\theta} \sim MVN(\mu, \underbrace{\sigma^2 H(\phi)} + \underbrace{r^2 I})$$

Spatial & nonspatial variation

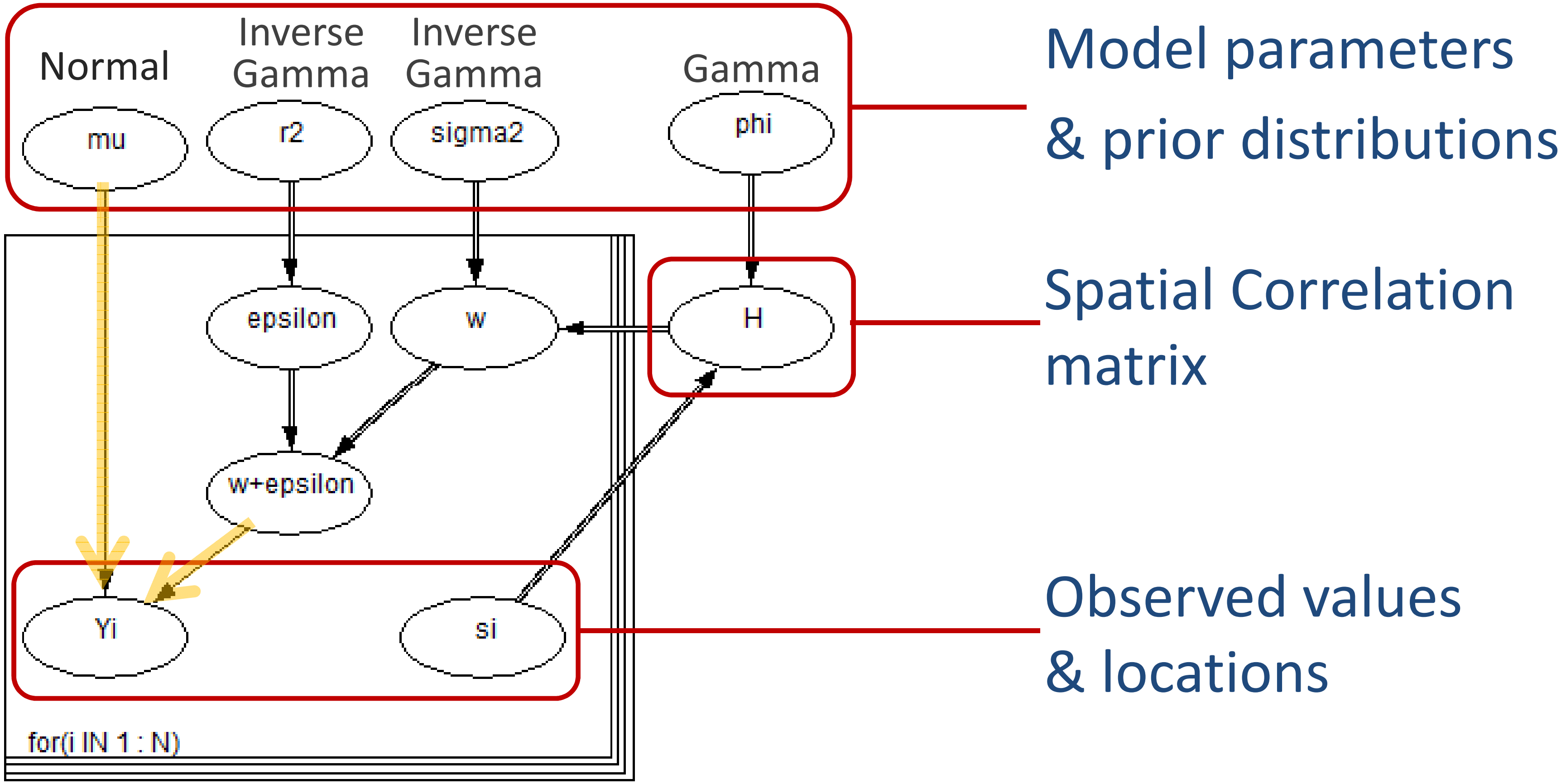
with correlation matrix

$$H_{ij} = \rho(\mathbf{s}_i - \mathbf{s}_j; \phi) = \exp(-\phi \|\mathbf{s}_i - \mathbf{s}_j\|)$$

(here: exponential model)

and prior distributions for μ , σ^2, ϕ and r^2

$$\mathbf{Y}|\boldsymbol{\theta} \sim MVN(\mu, \sigma^2 H(\phi) + r^2 I)$$



Bayesian modelling - Inference

likelihood prior

posterior
$$\pi_1(\boldsymbol{\theta}|\mathbf{y}) = \frac{f(\mathbf{y}|\boldsymbol{\theta})\pi_0(\boldsymbol{\theta})}{\int f(\mathbf{y}|\boldsymbol{\theta})\pi_0(\boldsymbol{\theta})\,d\boldsymbol{\theta}}$$

$$\begin{aligned}
\pi_1(\mu,\sigma^2,\phi,r^2|\mathbf{y}) & \\
&\propto f(\mathbf{y}|\mu,\sigma^2,\phi,r^2)\pi_0(\mu,\sigma^2,\phi,r^2) \\
&= f(\mathbf{y}|\mu,\sigma^2,\phi,r^2)\pi_0(\mu)\pi_0(\sigma^2)\pi_0(\phi)\pi_0(r^2)
\end{aligned}$$

With the posterior for $\boldsymbol{\theta}$, predictions $Y(\mathbf{s}_0)$ at new locations $\mathbf{s}_0$ are straight-forward:

$$f(y_0|\mathbf{y}) = \int f(y_0, \boldsymbol{\theta}|\mathbf{y})d\boldsymbol{\theta} = \int f(y_0|\mathbf{y}, \boldsymbol{\theta})\pi_1(\boldsymbol{\theta}|\mathbf{y})d\boldsymbol{\theta}$$

The spatial correlation matrix needs to be extended to include the new locations

$$\mathbf{Y}, \mathbf{Y_0}|\boldsymbol{\theta} \sim MVN(\mu, \sigma^2 H(\phi) + r^2 I)$$

$$\mathbf{Y} \big| \boldsymbol{\theta}, \mathbf{W} \sim MVN(\mu + \mathbf{W}, r^2 I)$$

with

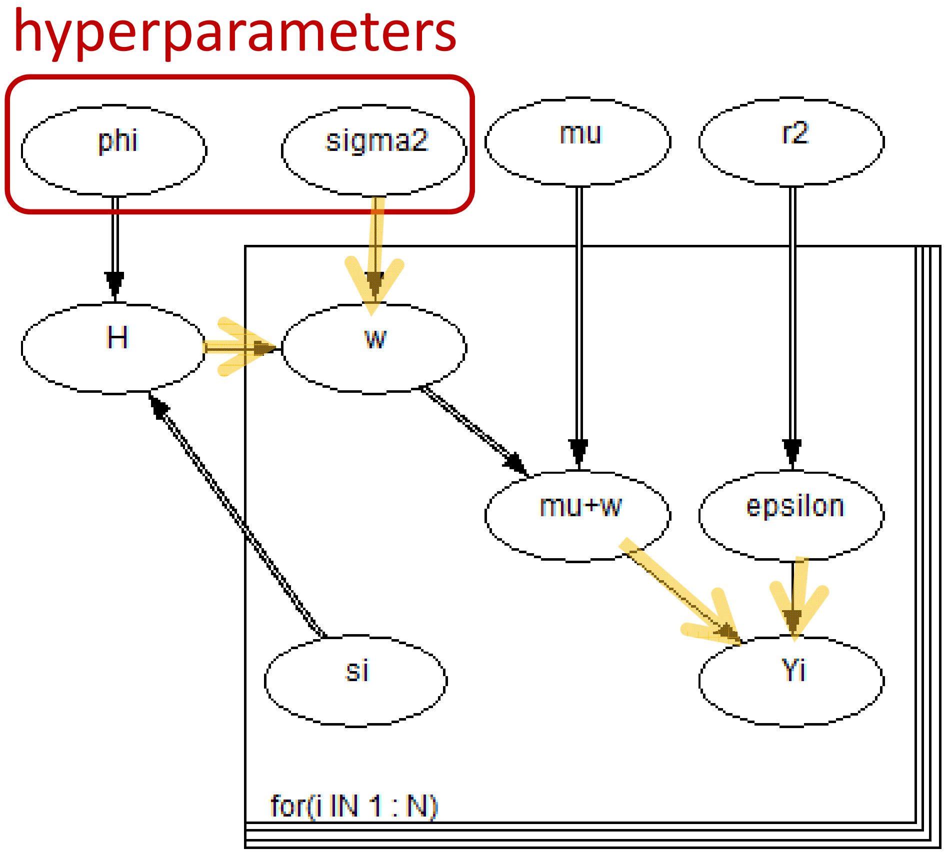
$$\mathbf{W} \big| \sigma^2, \phi \sim MVN(\mathbf{0}, \sigma^2 H(\phi))$$

with prior distributions for μ and r^2

and for the hyperparameters σ^2 and ϕ

$$\mathbf{Y} \mid \boldsymbol{\theta}, \mathbf{W} \sim MVN(\mu + \mathbf{W}, r^2 I)$$

$$\mathbf{W} \mid \sigma^2, \phi \sim$$
$$MVN(\mathbf{0}, \sigma^2 H(\phi))$$



Spatial dependence is modelled on the second hierarchical level

Hierarchical bayesian modelling

Hierarchical modelling:

Specification of priors in two stages

hyperparameters

$$\pi_1(\mu, r^2, \sigma^2, \phi \mid \mathbf{y})$$

$$\propto \int \underbrace{f(\mathbf{y} \mid \mu, r^2, \mathbf{W})}\, \underbrace{\pi_0(\mu, r^2)}_{\text{"normal" prior}}\, \underbrace{f(\mathbf{W} \mid \sigma^2, \phi)}_{\text{first-stage prior}}\, \underbrace{\pi_0(\sigma^2, \phi)}_{\text{hyper-prior}}\, d\mathbf{W}$$

Modelling in WinBUGS

- Model can be directly implemented using stochastic and logical links.
- Faster: GeoBUGS function “spatial”: Bayesian Gaussian kriging model

Try yourself! Code will be provided by email

WinBUGS:

http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml

http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/winbugsthemovie.html

What is the difference?

- **Mathematically:** No difference.

- **Computationally:** Hierarchical model is more demanding.

- **Conceptually:** Direct (first-stage) modelling of spatial dependence in the error term versus spatial modelling in the mean of the process

Generalizations

- Include covariates in the mean term (no constant mean).
- Model anisotrophy with the aid of more complex Correlation functions.
- Transform linear predictor / data for more general use.

- Generalized linear spatial process modelling.

Generalized Linear Models

Extend linear modelling to non-normal setting.

Idea:

- The outcomes Y_i are conditionally independent given the distribution parameters μ_i and ψ.
- $E(Y_i \mid \mathbf{X}_i) = \mu_i$ is dependent on a set of explanatory variables, $\mathbf{X}_i$.

$$p(\mathbf{y} \mid \mathbf{X}, \boldsymbol{\beta}, \boldsymbol{\psi}) = \prod_{i=1}^{n} p(y_i \mid (\mathbf{X}\boldsymbol{\beta})_i, \psi)$$

Specification in 3 stages:

- Probability distribution for the outcome variable Y with mean $E[y|X]=\mu$.
- Link function: $g(\cdot)$; $\mu = g^{-1}(\eta) = g^{-1}(X\beta)$
- Linear predictor: $\eta = X\beta$

$$p(\mathbf{y}|\mathbf{X},\boldsymbol{\beta},\boldsymbol{\psi}) = \prod_{i=1}^{n} p(y_i|\mu_i,\psi)$$

$$= \prod_{i=1}^{n} p(y_i|g^{-1}((\mathbf{X}\boldsymbol{\beta})_i),\psi)$$

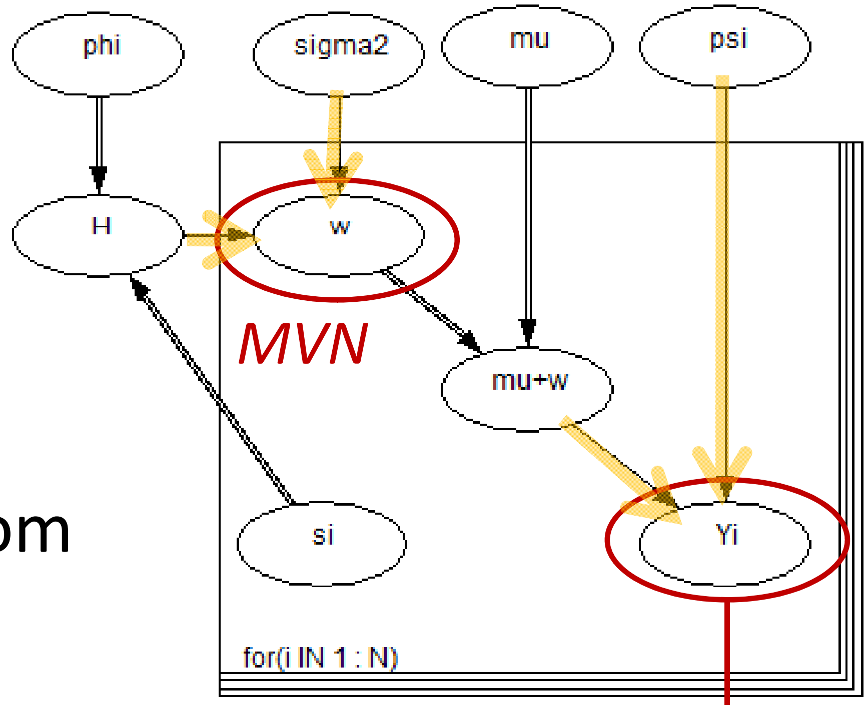
Generalized linear spatial process modelling

Model spatial dependence in the linear predictor:

$$\eta(\mathbf{s}_i) = \mathbf{x}^T(\mathbf{s})\boldsymbol{\beta} + w(\mathbf{s}_i)$$

$$\mathbf{W} \sim MVN(0, \sigma^2 H(\phi))$$

(here for homogenous random field without covariates, i.e. $\mathbf{x}^T(\mathbf{s})\boldsymbol{\beta} \equiv \mu$)



Any probability model in the exponential family

Binary data: $Y(\mathbf{s}_i) \in \{0;1\}$

$$\underbrace{P(Y(\mathbf{s}_i)=1 \mid p(\mathbf{s}_i)) = p(\mathbf{s}_i)}_{\text{probability model}} = \underbrace{\operatorname{logit}^{-1}}_{\text{link}}\underbrace{(\mathbf{x}^T(\mathbf{s})\boldsymbol{\beta} + w(\mathbf{s}_i))}_{\text{linear predictor}}$$

$$\mathbf{W} \sim MVN(0, \sigma^2 H(\phi))$$

$$\underbrace{\operatorname{logit}^{-1}(u) = \frac{e^u}{1+e^u}}_{\text{transform continuous variable to (0;1)}}$$

⇨Spatial dependence is modelled in the mean of the process.

Summary

- Random fields as a means to model spatial variation and correlation.
- Modelling with a Bayesian approach: Direct vs. hierarchical.
- Hierarchical modelling not necessary with a Gaussian spatial process.
- Hierarchical approach facilitates further generalizations

THANKS!