

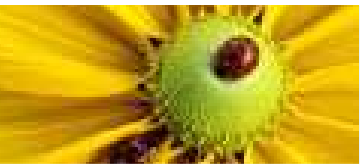
## REML Estimation and Linear Mixed Models

### 4. Geostatistics and linear mixed models for spatial data

Sue Welham

Rothamsted Research  
Harpenden UK AL5 2JQ

December 1, 2008



# Introduction

## Introduction

Geostatistics review

Spatial mixed models

Field experiments

Observational data

References

Exercise

We will start by reviewing the principles of geo-statistics, and then show how these can be implemented within a linear mixed model framework with REML estimation.

We will then look at two types of spatial analysis

- spatial analysis of designed field experiments
  - ◆ usually on a regular grid
  - ◆ aim of spatial analysis is to get better SEDs for treatments
  
- analysis of observational spatial data *i.e.* data observed at several spatial locations
  - ◆ may be regular or irregular grid
  - ◆ one aim of analysis is to identify patterns of spatial covariation between sample points



# Geostatistics

Introduction

Geostatistics review

**Introduction**

Definitions

Isotropy

Covariance functions

The variogram

Empirical variogram

Estimation

Nugget effect

Kriging

Spatial mixed models

Field experiments

Observational data

References

Exercise

Geostatistics is concerned with the analysis of data observed at several spatial locations, particularly with respect to prediction at unobserved locations (kriging)

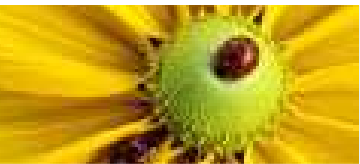
- historical motivation for development comes from mining
  - ◆ seek to predict the most profitable areas to mine based on data from soil surveys
  - ◆ better predictions obtained if we recognise and use spatial correlation
  - ◆ terminology reflects origins

Main aim may be to successfully account for spatial correlation rather than prediction: in either case, good estimation of the spatial correlation function is key.

Traditionally, geostatistical data considered as observations of a spatial random process with constant mean, but no other covariates

- initially, we work within this simple structure, then extend to the more general case.

Acknowledgements: much of the material in this review comes from Haskard (2007) (<http://digital.library.adelaide.edu.au/dspace/handle/2440/47972> online address), and Cressie (1993).



# Definitions

Introduction

Geostatistics review

Introduction

Definitions

Isotropy

Covariance functions

The variogram

Empirical variogram

Estimation

Nugget effect

Kriging

Spatial mixed models

Field experiments

Observational data

References

Exercise

We define  $y(\mathbf{s})$  to be the observation of a random variable at location  $\mathbf{s}$ , where

- $\mathbf{s}$  may be multi-dimensional - usually denotes location in one or two, occasionally three dimensions

A spatial sample of  $n$  points is written as  $\mathbf{y} = (y_1 \dots y_n)' = (y(\mathbf{s}_1) \dots y(\mathbf{s}_n))'$  where  $\mathbf{s}_i$  is the location of sample  $i$ .

Cressie (1993) suggests a decomposition of spatial data as

$$y(\mathbf{s}) = \mu(\mathbf{s}) + e(\mathbf{s})$$

- $\mu(\mathbf{s})$  is deterministic mean structure or trend, sometimes called large-scale non-stochastic variation
- $e(\mathbf{s})$  zero-mean stochastic component, sometimes called small-scale stochastic variation

or

$$y(\mathbf{s}) = \mu(\mathbf{s}) + v(\mathbf{s}) + \eta(\mathbf{s})$$

- $v(\mathbf{s})$  correlated component of stochastic variation
- $\eta(\mathbf{s})$  uncorrelated component of stochastic variation



## Definitions (2)

Introduction

Geostatistics review

Introduction

Definitions

Isotropy

Covariance functions

The variogram

Empirical variogram

Estimation

Nugget effect

Kriging

Spatial mixed models

Field experiments

Observational data

References

Exercise

Initially we assume  $\mu(\mathbf{s}) = \mu$ , ie. constant mean.

Note the potential lack of identifiability between large-scale and small-scale trend

- many trends could be modelled realistically within either context

Cressie (1993) advises against putting too much emphasis into large-scale trend (fixed terms)

- can lead to over-fitting and spurious predictions

In traditional applications, only one realisation of the spatial process is observed. Some assumptions are required to make inferences.

Usual assumption: stationarity (second order or weak stationarity)

- the mean of the process  $\mu(\mathbf{s})$  does not depend on the location  $\mathbf{s}$
- the spatial covariance function depends only on the spatial separation of the points, ie.  $\text{cov}(y(\mathbf{s}_i), y(\mathbf{s}_j)) = f(\mathbf{s}_i - \mathbf{s}_j)$

This implies that the mean is constant, and that the covariance is constant for a pair of points with a set spatial displacement (may depend on direction as well as distance).

## Definitions (3)

Other forms of stationarity can also be defined:

- Strong stationarity

- ◆ all moments of the joint distribution of  $y(\mathbf{s}_i)$  and  $y(\mathbf{s}_j)$  depend only on the spatial separation  $\mathbf{s}_i - \mathbf{s}_j$

- Intrinsic stationarity

- ◆ defined in terms of the differences  $y(\mathbf{s}_i) - y(\mathbf{s}_j)$  such that
- ◆ their mean is zero
- ◆ their covariance function depends only on their spatial separation.

With intrinsic stationarity, the mean must be constant, but the variance of the process may be non-constant. Intrinsic stationarity plus constant variance  $\Rightarrow$  second-order stationarity.

We will use the term 'stationarity' to mean 'second-order stationarity'.

Introduction

Geostatistics review

Introduction

Definitions

Isotropy

Covariance functions

The variogram

Empirical variogram

Estimation

Nugget effect

Kriging

Spatial mixed models

Field experiments

Observational data

References

Exercise



# Isotropy

Introduction

Geostatistics review

Introduction

Definitions

Isotropy

Covariance functions

The variogram

Empirical variogram

Estimation

Nugget effect

Kriging

Spatial mixed models

Field experiments

Observational data

References

Exercise

A stationary process is isotropic if

$$\text{cov}(y(\mathbf{s}_i), y(\mathbf{s}_j)) = f(h)$$

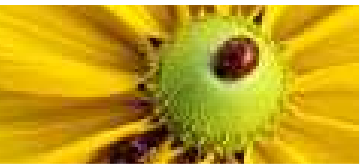
where  $h = \|\mathbf{s}_i - \mathbf{s}_j\|$ , ie. the covariance function depends on the distance between their points and not the direction of the displacement.

A spatial process that is not isotropic is called anisotropic.

Clearly a process can only be anisotropic in  $> 1$  dimension.

Examples of processes where anisotropy may occur:

- soil properties due to underlying geological structures
- impact of historical cultivation practices with single predominant direction
- species distributions where latitude (determining temperature and daylength) is more important than longitude



# Spatial covariance function

Introduction

Geostatistics review

Introduction

Definitions

Isotropy

Covariance functions

The variogram

Empirical variogram

Estimation

Nugget effect

Kriging

Spatial mixed models

Field experiments

Observational data

References

Exercise

The spatial covariance function is defined as

$$C(\mathbf{s}_i, \mathbf{s}_j) = \text{cov}(y(\mathbf{s}_i), y(\mathbf{s}_j)) = E[(y(\mathbf{s}_i) - \mu(\mathbf{s}_i))(y(\mathbf{s}_j) - \mu(\mathbf{s}_j))]$$

For model-based geo-statistical analysis, a parametric form of model is proposed for the covariance function.

For a (second-order) stationary process

$$\mu(\mathbf{s}_i) = \mu$$

$$C(\mathbf{s}_i, \mathbf{s}_j) = C(\mathbf{h})$$

$$\text{var}(y(\mathbf{s}_i)) = C(0)$$

$$C(\mathbf{s}_i, \mathbf{s}_j) = C(\mathbf{s}_j, \mathbf{s}_i); C(\mathbf{h}) = C(-\mathbf{h})$$

Note: here  $\mathbf{h}$  is a directional vector.

The spatial covariance function is

- typically continuous - decreasing as spatial displacement increases
- positive definite, such that  $\text{var}(\mathbf{a}'\mathbf{y}) > 0$  for  $\mathbf{a} \neq \mathbf{0}$
- usually non-negative - with the exception of periodic processes





## Spatial covariance function (2)

Introduction

Geostatistics review

Introduction

Definitions

Isotropy

Covariance functions

The variogram

Empirical variogram

Estimation

Nugget effect

Kriging

Spatial mixed models

Field experiments

Observational data

References

Exercise

For an isotropic process,  $C(\mathbf{h}) = C(h)$ .

This gives a univariate covariance function, and hence correlation function.

Common correlation functions,  $\rho(h)$  such that  $C(h) = \sigma^2 \rho(h)$ :

■ exponential (power):  $\rho(h) = \exp(-|h|/\phi)$

■ gaussian:  $\rho(h) = \exp(-(|h|/\phi)^2)$

■ spherical:  $\rho(h) = \begin{cases} 1 - \frac{3|h|}{2\phi} + \frac{1}{2}\left(\frac{|h|}{\phi}\right)^3 & 0 \leq |h| < \phi \\ 0 & |h| \geq \phi \end{cases}$

■ bounded linear:  $\rho(h) = \begin{cases} 1 - \frac{|h|}{\phi} & 0 \leq |h| < \phi \\ 0 & |h| \geq \phi \end{cases}$

Note that some univariate models do not generalize to a valid (positive-definite) covariance function in higher dimensions, eg. bounded linear model is valid only for 1 dimension.

## Spatial covariance function (3)

Introduction

Geostatistics review

Introduction

Definitions

Isotropy

Covariance functions

The variogram

Empirical variogram

Estimation

Nugget effect

Kriging

Spatial mixed models

Field experiments

Observational data

References

Exercise

Note that:

- the exponential model is a generalization of the auto-regressive (AR) process to non-integer (and multi-dimensional) lags, since  $\rho(h) = \theta^{|h|}$  where  $\theta = e^{-1/\phi}$
- the bounded linear and spherical correlation functions have a finite range and are zero for  $|h| > \phi$
- the exponential and gaussian models are zero only in the limit as  $|h| \rightarrow \infty$ , but are effectively zero (close to zero) at finite distances

We will also briefly consider the Matern correlation function, defined by

$$\rho(h; \zeta, \nu) = \frac{1}{2^{\nu-1}\Gamma(\nu)} \left(\frac{h}{\zeta}\right)^{\nu} K_{\nu} \left(\frac{h}{\zeta}\right)$$

where  $\Gamma()$  is the Gamma function and  $K_{\nu}()$  is the modified Bessel function of the third order.

This Bessel function takes a simple form for  $\nu = m + \frac{1}{2}$  with  $m \in \mathbb{Z}^+$  (see Haskard, 2007, for details).



# Matern correlation function

Introduction

Geostatistics review

Introduction

Definitions

Isotropy

Covariance functions

The variogram

Empirical variogram

Estimation

Nugget effect

Kriging

Spatial mixed models

Field experiments

Observational data

References

Exercise

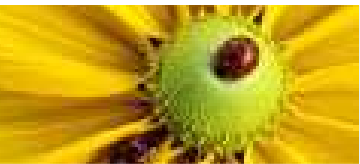
Special cases of the Matern correlation function occur for

- $\nu = \frac{1}{2}$ :  $\rho(h; \zeta, \nu = \frac{1}{2}) = e^{-h/\zeta}$  ie. exponential correlation function
- $\nu = \frac{3}{2}$ :  $\rho(h; \zeta, \nu = \frac{3}{2}) = e^{-h/\zeta} \left( \frac{h}{\zeta} + 1 \right)$
- $\nu = \frac{5}{2}$ :  $\rho(h; \zeta, \nu = \frac{5}{2}) = e^{-h/\zeta} \left\{ \frac{1}{3} \left( \frac{h}{\zeta} \right)^2 + \frac{h}{\zeta} + 1 \right\}$
- if  $\nu \rightarrow \infty$  with  $\zeta \rightarrow 0$  such that  $\beta = 2\nu^{1/2}\zeta$  remains constant, then  $\rho(h; \beta) = \exp(-(h/\beta)^2)$ , ie. gaussian correlation function

The Matern correlation function is therefore a generalization of several common correlation functions. Its parameters can be interpreted as

- $\nu$ , the smoothness of the spatial process (smoothness increases with  $\nu$ )
- $\zeta$ , the range of the correlation function

Note that the physical distance at which correlation decays to a given value depends upon both  $\nu$  and  $\zeta$ .



# Separability

Introduction

Geostatistics review

Introduction

Definitions

Isotropy

Covariance functions

The variogram

Empirical variogram

Estimation

Nugget effect

Kriging

Spatial mixed models

Field experiments

Observational data

References

Exercise

If we can write

$$\mathbf{h} = \begin{pmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \end{pmatrix}$$

then the spatial correlation function is called separable if it can be written as

$$\rho(\mathbf{h}) = \rho_1(\mathbf{h}_1)\rho_2(\mathbf{h}_2).$$

So the covariance function is a simple product of the composite correlation functions.

This corresponds to the case of two independent processes in different directions.



# The variogram

Introduction

Geostatistics review

Introduction

Definitions

Isotropy

Covariance functions

The variogram

Empirical variogram

Estimation

Nugget effect

Kriging

Spatial mixed models

Field experiments

Observational data

References

Exercise

The variogram only exists for a spatial process with intrinsic stationarity, and is defined as

$$\gamma(\mathbf{h}) = \frac{1}{2} \text{var} (y(\mathbf{s} + \mathbf{h}) - y(\mathbf{s})).$$

Because of the multiplier, this function was historically called the semi-variogram.

Note that this function  $\gamma()$  is not the variance ratios used elsewhere.

With intrinsic stationarity

$$\gamma(\mathbf{h}) = \frac{1}{2} V(\mathbf{s} + \mathbf{h}) + \frac{1}{2} V(\mathbf{s}) - C(\mathbf{h})$$

where  $V(\mathbf{s}_i)$  is the variance of the process at position  $\mathbf{s}_i$ .

With (second-order) stationarity

$$\begin{aligned} \gamma(\mathbf{h}) &= \frac{1}{2} C(0) + \frac{1}{2} C(0) - C(\mathbf{h}) \\ &= C(0) - C(\mathbf{h}) \\ &= C(0) \left( 1 - \frac{C(\mathbf{h})}{C(0)} \right) \\ &= \sigma^2 (1 - \rho(\mathbf{h})) \end{aligned}$$

# Features of variograms

Introduction

Geostatistics review

Introduction

Definitions

Isotropy

Covariance functions

The variogram

Empirical variogram

Estimation

Nugget effect

Kriging

Spatial mixed models

Field experiments

Observational data

References

Exercise

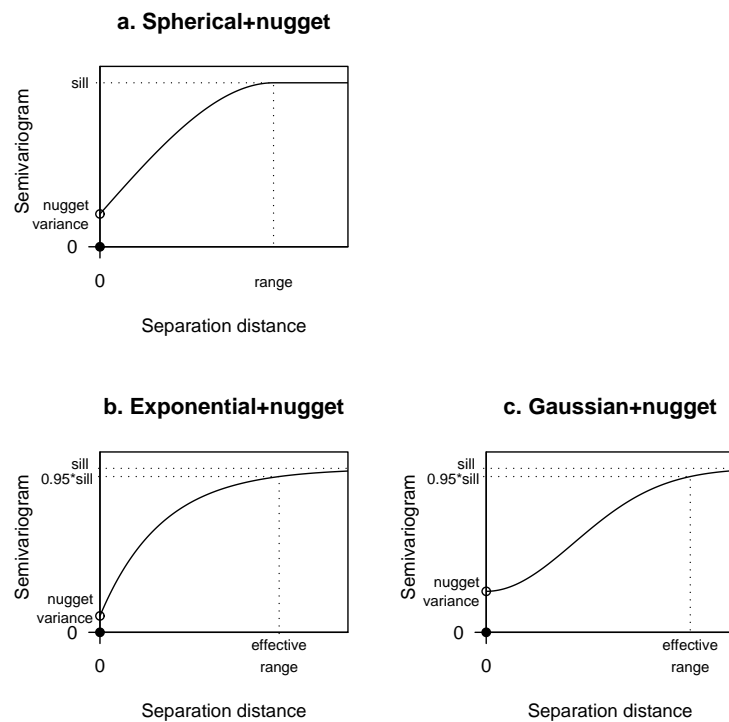


Figure 1: Fig 2.2 from Haskard (2007)

- sill = value at which variogram levels off
- range = lag at which the variogram reaches the sill
- effective range = 95% of the sill
- nugget variance = value of variogram at distance 0



## Features of variograms (2)

Introduction

Geostatistics review

Introduction

Definitions

Isotropy

Covariance functions

The variogram

Empirical variogram

Estimation

Nugget effect

Kriging

Spatial mixed models

Field experiments

Observational data

References

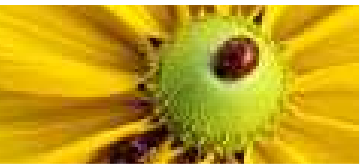
Exercise

### Anisotropy:

- implies different variograms in different directions
- usually assume common variances (hence sill) in different directions
- usually assume different ranges may apply in different directions
- can look at variogram in two directions directly or calculate several one-dimensional variograms in different directions

### Nugget effect:

- spatial separation  $\downarrow 0$  but variogram does not decrease to zero
- but  $\gamma(0) = 0$  by definition
- we will return to this



# Features of variograms (3)

- Introduction
- Geostatistics review
- Introduction
- Definitions
- Isotropy
- Covariance functions
- The variogram**
- Empirical variogram
- Estimation
- Nugget effect
- Kriging
- Spatial mixed models
- Field experiments
- Observational data
- References
- Exercise

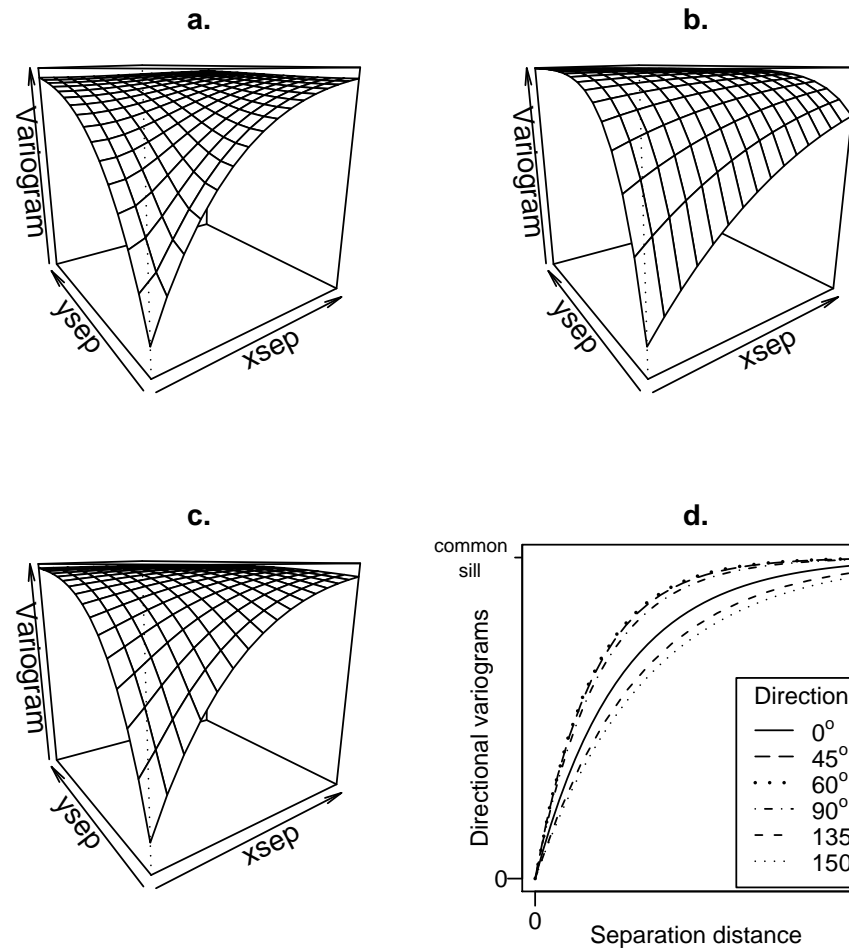


Figure 2: Fig 2.3 from Haskard (2007)





# Empirical variogram

Introduction

Geostatistics review

Introduction

Definitions

Isotropy

Covariance functions

The variogram

Empirical variogram

Estimation

Nugget effect

Kriging

Spatial mixed models

Field experiments

Observational data

References

Exercise

The empirical semi-variance between two points is defined as

$$\frac{1}{2} \{y(\mathbf{s}_i) - y(\mathbf{s}_j)\}^2$$

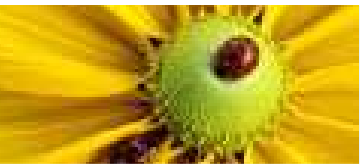
The variogram at lag  $\mathbf{h}$  can be defined as

$$E \left[ \frac{1}{2} \{y(\mathbf{s}_i) - y(\mathbf{s}_j)\}^2 \right]$$

with the expectation taken across all possible pairs  $\mathbf{s}_i, \mathbf{s}_j$  with  $\mathbf{h} = \mathbf{s}_i - \mathbf{s}_j$ .

The empirical variogram at lag  $\mathbf{h}$  is estimated by the mean taken across all point pairs at lag  $\mathbf{h}$ , or within a given neighbourhood of  $\mathbf{h}$ :

- the amount of information (number of point pairs) in the variogram at any distance depends on the structure of the sample
- number of available point pairs decreases as lag distance increases in any direction
- by convention, curtail the use empirical variogram at distances greater than half the range of the original sample area



# Estimating the spatial correlation function

Introduction

Geostatistics review

Introduction

Definitions

Isotropy

Covariance functions

The variogram

Empirical variogram

Estimation

Nugget effect

Kriging

Spatial mixed models

Field experiments

Observational data

References

Exercise

In traditional geo-statistics, the form of the spatial correlation function is estimated from the variogram

- either by eye
- or by weighted least squares

Fitting the variogram to a standard parametric model allows interpolation of correlation to spatial lags not present in the sample.

Neither of these estimation methods accounts for the correlations inherent in the empirical variogram, as each observed sample contributes to many neighbouring points in the variogram (up to  $n - 1$  points).

This correlation makes the empirical variogram much smoother than might be expected.

This method does not take account of sampling uncertainty in the empirical variogram.



# Examples of variograms

- Introduction
- Geostatistics review
- Introduction
- Definitions
- Isotropy
- Covariance functions
- The variogram
- Empirical variogram
- Estimation**
- Nugget effect
- Kriging
- Spatial mixed models
- Field experiments
- Observational data
- References
- Exercise

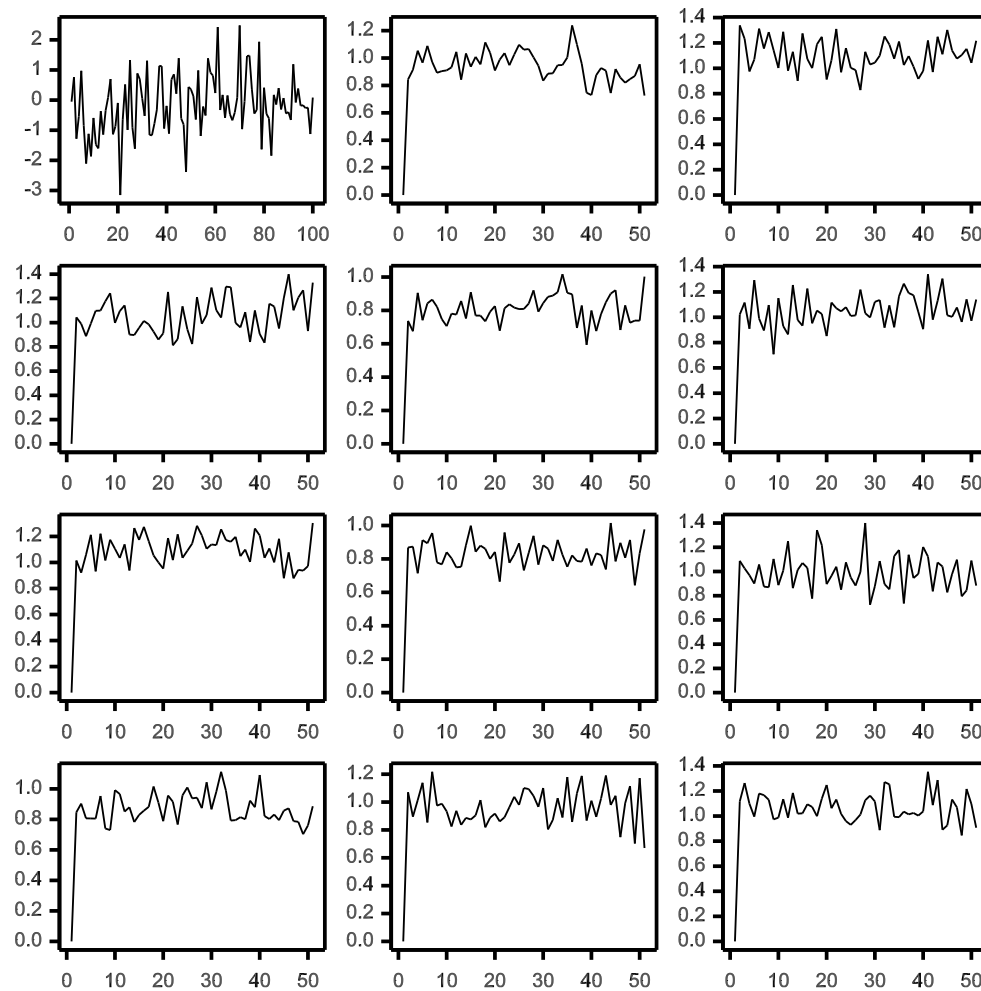
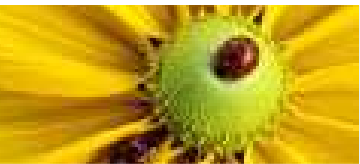


Figure 3: Data and variograms for pure noise  $\mathbf{y} \sim N(0, \mathbf{I}_{100})$ . Top left: example data. Other panes: variograms for independent samples.



## Examples of variograms (2)

- Introduction
- Geostatistics review
- Introduction
- Definitions
- Isotropy
- Covariance functions
- The variogram
- Empirical variogram
- Estimation**
- Nugget effect
- Kriging
- Spatial mixed models
- Field experiments
- Observational data
- References
- Exercise

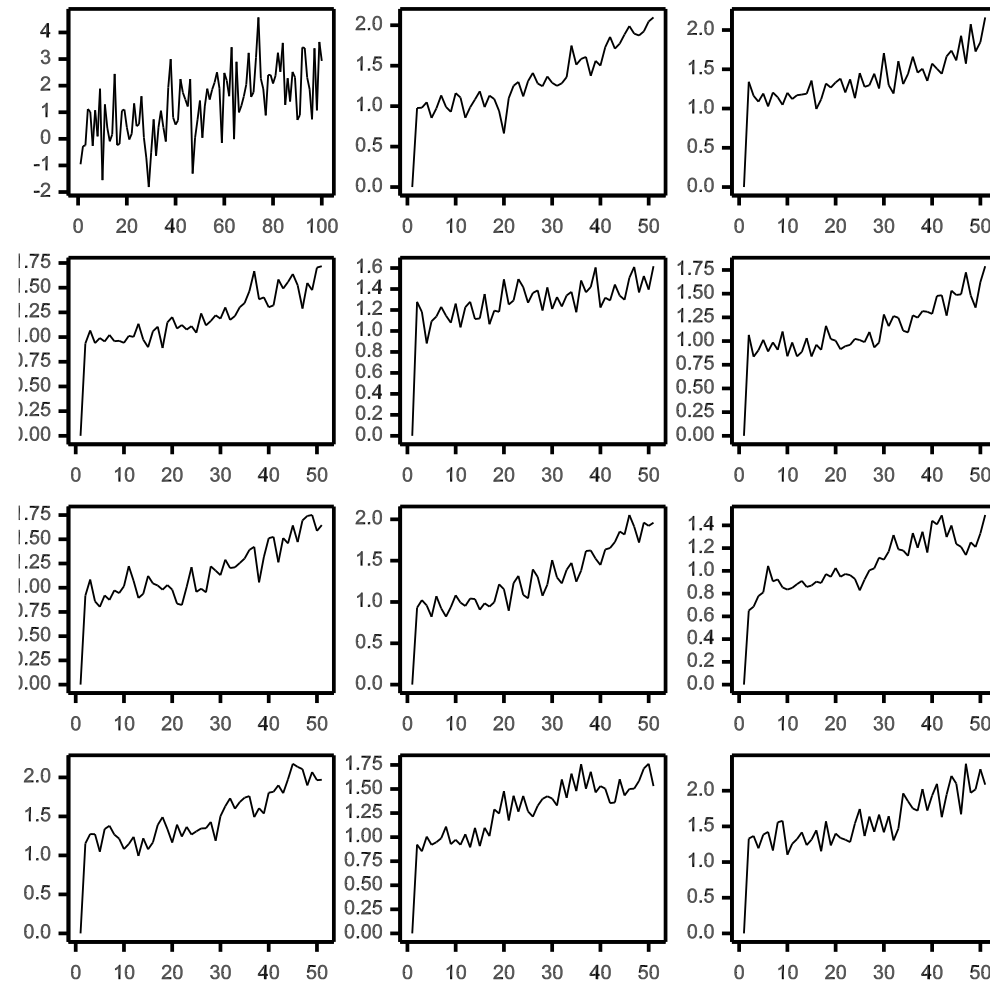


Figure 4: Data and variograms for trend plus noise  $\mathbf{y} \sim N(0.025\mathbf{x}, \mathbf{I}_{100})$  for  $\mathbf{x} = (1 \dots 100)'$ . Top left: example data. Other panes: variograms for independent samples.



## Examples of variograms (3)

- Introduction
- Geostatistics review
- Introduction
- Definitions
- Isotropy
- Covariance functions
- The variogram
- Empirical variogram
- Estimation**
- Nugget effect
- Kriging
- Spatial mixed models
- Field experiments
- Observational data
- References
- Exercise

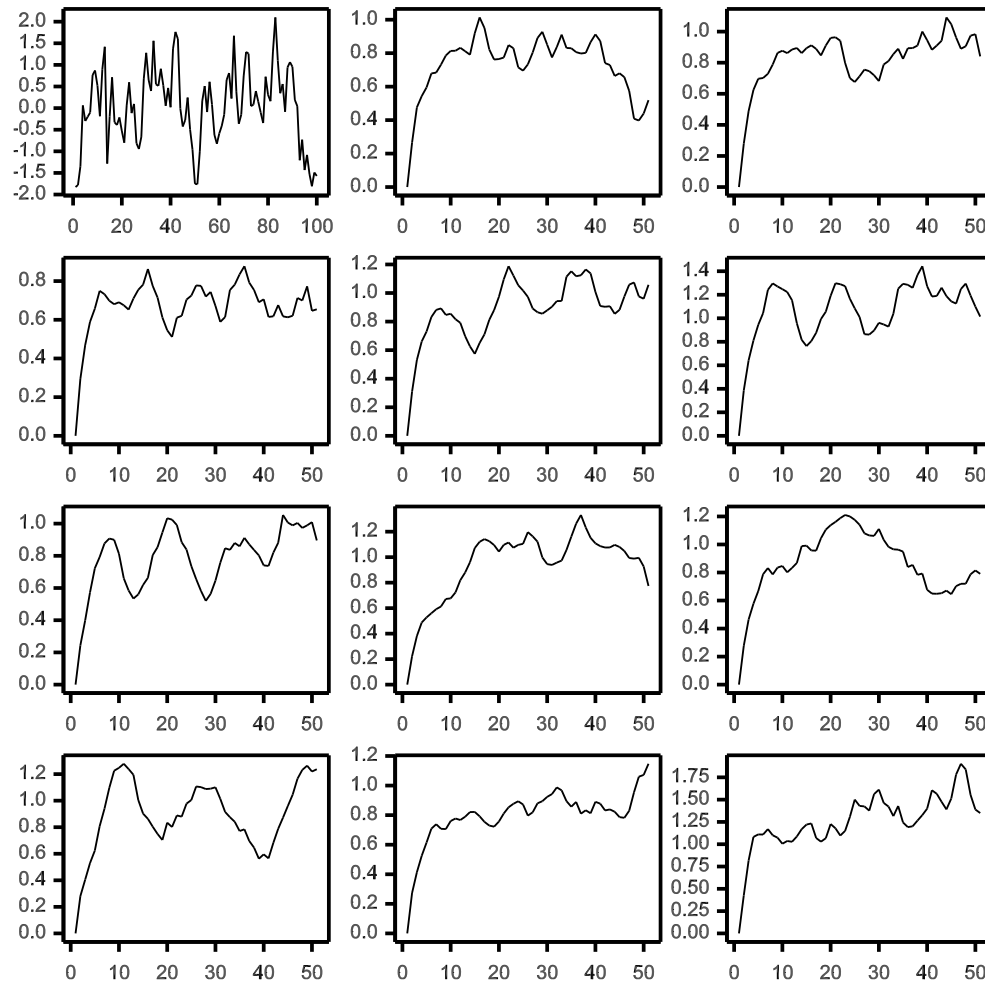
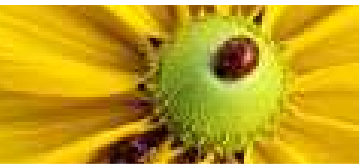


Figure 5: Data and variograms for auto-regressive error with correlation 0.7 at lag 1,  $y \sim N(0.025x, C)$ . Top left: example data. Other panes: variograms for independent samples.



## Examples of variograms (4)

- Introduction
- Geostatistics review
- Introduction
- Definitions
- Isotropy
- Covariance functions
- The variogram
- Empirical variogram
- Estimation**
- Nugget effect
- Kriging
- Spatial mixed models
- Field experiments
- Observational data
- References
- Exercise

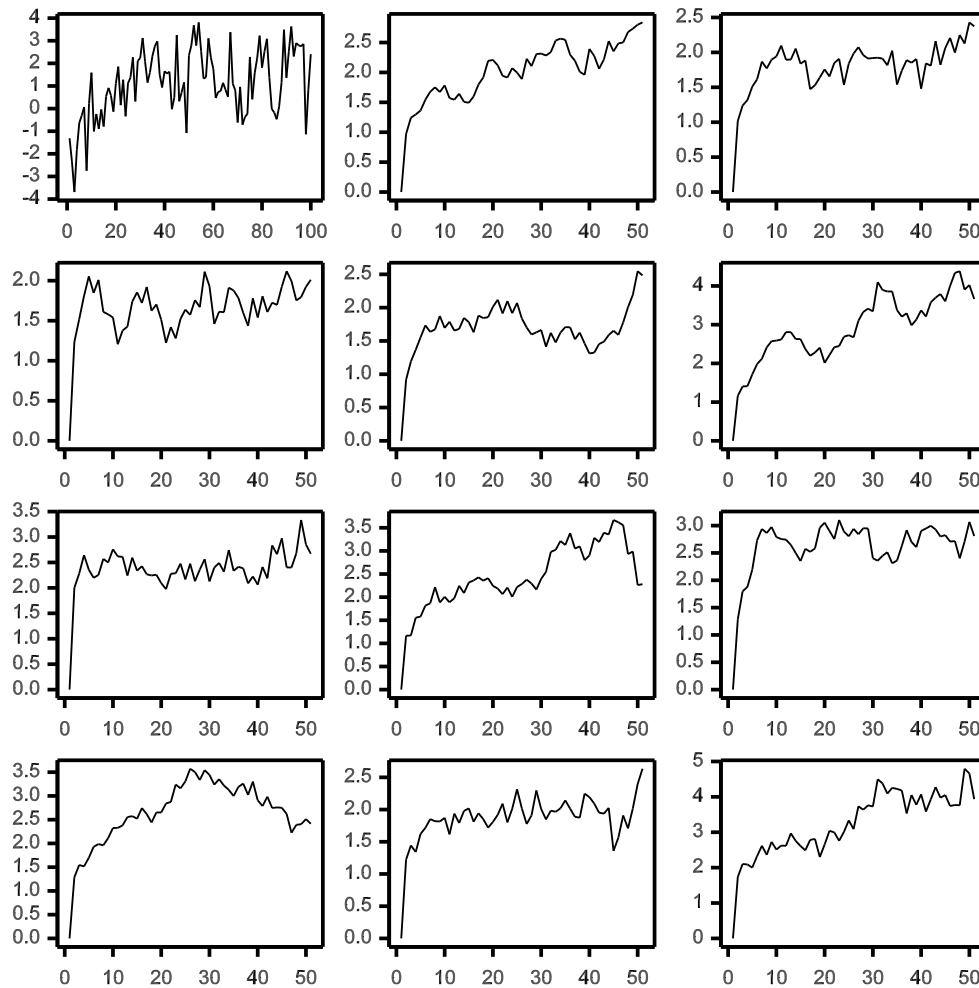


Figure 6: Data and variograms for trend plus auto-regressive error plus noise  $\mathbf{y} \sim N(0.025\mathbf{x}, \mathbf{C} + \mathbf{I}_{100})$  for  $\mathbf{x} = (1 \dots 100)'$ . Top left: example data. Other panes: variograms for independent samples.



# Nugget effect

Introduction

Geostatistics review

Introduction

Definitions

Isotropy

Covariance functions

The variogram

Empirical variogram

Estimation

Nugget effect

Kriging

Spatial mixed models

Field experiments

Observational data

References

Exercise

The nugget effect describes a non-zero value of the variogram at the origin (lag 0), contrary to the definition  $\gamma(0) = 0$ .

In general, the use of a nugget effect can be controversial, as it is estimated from the variogram and the smallest lag distance may be far from the origin.

The nugget variance is estimated by extrapolation of the variogram to the origin, which is done (in the absence of replication) without real information on the true value.

Causes of a nugget effect include

- small range spatial process (range smaller than sampling distance)
- measurement error

In general, we cannot distinguish the two sources of variation.

We can make some progress by taking replicate measurements

- separate samples within a single location - gives information on small range variation
- repeated measurement of a single sample (technical replication) gives information on measurement error

# Kriging

Kriging is the term used for prediction at locations unobserved in the sample,  
 $\tilde{y}(s_p) = \tilde{y}_p$ .

The principle of prediction used is BLUP = best (minimum MSE) linear unbiased prediction.

A linear prediction must take the form

$$\tilde{y}_p = \lambda_0 + \sum_{i=1}^n \lambda_i y(\mathbf{s}_i)$$

hence, with second-order stationarity,

$$E \left( \lambda_0 + \sum_{i=1}^n \lambda_i y(\mathbf{s}_i) \right) = \lambda_0 + \sum_{i=1}^n \lambda_i \mu.$$

Unbiasedness requires  $E(\tilde{y}_p) = \mu$  ie.

$$\mu = \lambda_0 + \sum_{i=1}^n \lambda_i \mu \Rightarrow \lambda_0 = \left(1 - \sum_{i=1}^n \lambda_i\right) \mu$$

which is solved by  $\lambda_0 = 0$  and  $\sum_{i=1}^n \lambda_i = 1$ .

Introduction

Geostatistics review

Introduction

Definitions

Isotropy

Covariance functions

The variogram

Empirical variogram

Estimation

Nugget effect

Kriging

Spatial mixed models

Field experiments

Observational data

References

Exercise



## Kriging (2)

Introduction

Geostatistics review

Introduction

Definitions

Isotropy

Covariance functions

The variogram

Empirical variogram

Estimation

Nugget effect

Kriging

Spatial mixed models

Field experiments

Observational data

References

Exercise

Consider the square error on the predictor:

$$\begin{aligned}(\tilde{y}_p - y_p)^2 &= (\tilde{y}_p - \mu + \mu - y_p)^2 \\ &= \left( \sum_{i=1}^n \lambda_i y_i - \mu + \mu - y_p \right)^2 \\ &= \left( \sum_{i=1}^n \lambda_i (y_i - \mu) \right)^2 + 2(\mu - y_p) \sum_{i=1}^n (\lambda_i (y_i - \mu)) + (\mu - y_p)^2 \\ &= \sum_{i,j=1}^n \lambda_i \lambda_j (y_i - \mu)(y_j - \mu) + 2 \sum_{i=1}^n \lambda_i (y_i - \mu)(\mu - y_p) + (\mu - y_p)^2\end{aligned}$$

using  $\sum_{i=1}^n \lambda_i = 1$ , hence

$$\begin{aligned}E(\tilde{y}_p - y_p)^2 &= \sum_{i,j=1}^n \lambda_i \lambda_j \text{cov}(y_i, y_j) - 2 \sum_{i=1}^n \lambda_i \text{cov}(y_i, y_p) + \text{var}(y_p) \\ &= \text{var}(y_p) - 2\boldsymbol{\lambda}\mathbf{c} + \boldsymbol{\lambda}'\mathbf{V}\boldsymbol{\lambda}\end{aligned}$$

where  $\mathbf{c} = (\text{cov}(y_1, y_p) \dots \text{cov}(y_n, y_p))'$  and  $\mathbf{V}$  is the sample variance-covariance matrix.

## Kriging (3)

Introduction

Geostatistics review

Introduction

Definitions

Isotropy

Covariance functions

The variogram

Empirical variogram

Estimation

Nugget effect

Kriging

Spatial mixed models

Field experiments

Observational data

References

Exercise

We need to maximise the mean squared error subject to the constraint  $\boldsymbol{\lambda}'\mathbf{1} = 1$ , which we do using Lagrange multipliers, hence we minimize

$$\text{var}(y_p) - 2\boldsymbol{\lambda}\mathbf{c} + \boldsymbol{\lambda}'\mathbf{V}\boldsymbol{\lambda} - 2\phi(\boldsymbol{\lambda}'\mathbf{1} - 1)$$

Differentiate with respect to parameters  $\boldsymbol{\lambda}$  and  $\phi$  and set equal to zero:

$$\begin{aligned} -2\mathbf{c} + 2\mathbf{V}\boldsymbol{\lambda} - 2\phi\mathbf{1} &= \mathbf{0} \\ \Rightarrow \mathbf{V}\boldsymbol{\lambda} - \phi\mathbf{1} &= \mathbf{0} \end{aligned}$$

and

$$\boldsymbol{\lambda}'\mathbf{1} = 0$$

or, in matrix form,

$$\begin{bmatrix} \mathbf{V} & \mathbf{1} \\ \mathbf{1}' & 0 \end{bmatrix} \begin{pmatrix} \boldsymbol{\lambda} \\ \phi \end{pmatrix} = \begin{pmatrix} \mathbf{c} \\ 1 \end{pmatrix}$$

and by inverting the partitioned matrix:

$$\boldsymbol{\lambda} = \mathbf{V}^{-1} \left( \mathbf{I} - \frac{\mathbf{1}\mathbf{1}'\mathbf{V}^{-1}}{\mathbf{1}'\mathbf{V}\mathbf{1}} \right) \mathbf{c} + \frac{\mathbf{V}^{-1}\mathbf{1}}{\mathbf{1}'\mathbf{V}\mathbf{1}}$$

## Kriging (4)

Introduction

Geostatistics review

Introduction

Definitions

Isotropy

Covariance functions

The variogram

Empirical variogram

Estimation

Nugget effect

Kriging

Spatial mixed models

Field experiments

Observational data

References

Exercise

The kriging variance is defined as

$$\begin{aligned}\text{var}(\tilde{y}_p - y_p) &= \text{var}(\boldsymbol{\lambda}'\mathbf{y} - y_p) \\ &= \text{var}(\boldsymbol{\lambda}'\mathbf{y}) - 2\text{cov}(\boldsymbol{\lambda}'\mathbf{y}, y_p) + \text{var}(y_p) \\ &= \boldsymbol{\lambda}'\mathbf{V}\boldsymbol{\lambda} - 2\boldsymbol{\lambda}'\mathbf{c} + \text{var}(y_p)\end{aligned}$$

which is exactly the mean squared error that has been minimized across all possible estimators. Using

$$\begin{aligned}\boldsymbol{\lambda}'\mathbf{V}\boldsymbol{\lambda} &= \mathbf{c}'\mathbf{V}^{-1}\mathbf{c} - \frac{\mathbf{c}'\mathbf{V}^{-1}\mathbf{1}\mathbf{1}'\mathbf{V}^{-1}\mathbf{c}}{\mathbf{1}'\mathbf{V}^{-1}\mathbf{1}} + \frac{1}{\mathbf{1}'\mathbf{V}^{-1}\mathbf{1}} \\ 2\boldsymbol{\lambda}'\mathbf{c} &= 2\mathbf{c}'\mathbf{V}^{-1}\mathbf{c} - 2\frac{\mathbf{c}'\mathbf{V}^{-1}\mathbf{1}\mathbf{1}'\mathbf{V}^{-1}\mathbf{c}}{\mathbf{1}'\mathbf{V}^{-1}\mathbf{1}} + 2\frac{\mathbf{1}'\mathbf{V}^{-1}\mathbf{c}}{\mathbf{1}'\mathbf{V}^{-1}\mathbf{1}}\end{aligned}$$

the kriging variance can be written as

$$\begin{aligned}\text{var}(\tilde{y}_p - y_p) &= \text{var}(y_p) - 2\boldsymbol{\lambda}'\mathbf{c} + \boldsymbol{\lambda}'\mathbf{V}\boldsymbol{\lambda} \\ &= \text{var}(y_p) - \mathbf{c}'\mathbf{V}^{-1}\mathbf{c} + \frac{\mathbf{c}'\mathbf{V}^{-1}\mathbf{1}\mathbf{1}'\mathbf{V}^{-1}\mathbf{c}}{\mathbf{1}'\mathbf{V}^{-1}\mathbf{1}} - 2\frac{\mathbf{1}'\mathbf{V}^{-1}\mathbf{c}}{\mathbf{1}'\mathbf{V}^{-1}\mathbf{1}} + \frac{1}{\mathbf{1}'\mathbf{V}^{-1}\mathbf{1}} \\ &= \text{var}(y_p) - \mathbf{c}'\mathbf{V}^{-1}\mathbf{c} + \frac{(\mathbf{1} - \mathbf{1}'\mathbf{V}^{-1}\mathbf{c})^2}{\mathbf{1}'\mathbf{V}^{-1}\mathbf{1}}\end{aligned}$$

# Spatial linear mixed model

Introduction

Geostatistics review

Spatial mixed models

Kriging variance

Field experiments

Observational data

References

Exercise

Consider a linear mixed model for the spatial sample  $\mathbf{y}$  of the form

$$\mathbf{y} = \mu \mathbf{1} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

where  $\text{var}(\mathbf{u}) = \sigma^2 \mathbf{G}$  and  $\text{var}(\mathbf{e}) = \sigma^2 \mathbf{R}$ , and

$$\mathbf{V} = \text{var}(\mathbf{y}) = \sigma^2 (\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})$$

is a stationary variance structure.

Consider a BLUP for a location not observed within the original sample:

$$y_p = \mu + \mathbf{z}_p \mathbf{u}_p + e_p$$

We know that, when using REML,  $\tilde{\mathbf{u}} = E(\mathbf{u} | \mathbf{L}'_2 \mathbf{y}) = \mathbf{G}\mathbf{Z}'\mathbf{P}\mathbf{y}$  and  $\tilde{\mathbf{e}} = E(\mathbf{e} | \mathbf{L}'_2 \mathbf{y}) = \mathbf{R}\mathbf{P}\mathbf{y}$  where  $\mathbf{L}'_2 \mathbf{X} = \mathbf{0}$ .

So consider  $E(y_p | \mathbf{L}'_2 \mathbf{y})$ :

$$E(y_p | \mathbf{L}'_2 \mathbf{y}) = E[y_p] + \text{cov}(y_p, \mathbf{L}'_2 \mathbf{y}) [\text{var}(\mathbf{L}'_2 \mathbf{y})]^{-1} \mathbf{L}'_2 \mathbf{y}$$

and use

$$\text{cov}(y_p, \mathbf{L}'_2 \mathbf{y}) = \sigma^2 (\mathbf{z}_p \mathbf{G}_{po} \mathbf{Z}' + \mathbf{R}_{po}) \mathbf{L}_2$$

where  $\sigma^2 \mathbf{G}_{po} = \text{cov}(\mathbf{u}_p, \mathbf{u})$  and  $\sigma^2 \mathbf{R}_{po} = \text{cov}(e_p, \mathbf{e})$ .

## Spatial linear mixed model (2)

Hence

$$\begin{aligned} E(y_p | \mathbf{L}'_2 \mathbf{y}) &= \mu + (\mathbf{z}_p \mathbf{G}_{po} \mathbf{Z}' + \mathbf{R}_{po}) \mathbf{L}_2 (\mathbf{L}'_2 \mathbf{H} \mathbf{L}_2)^{-1} \mathbf{L}'_2 \mathbf{y} \\ &= \mu + (\mathbf{z}_p \mathbf{G}_{po} \mathbf{Z}' + \mathbf{R}_{po}) \mathbf{P} \mathbf{y} \\ &= \mu + \mathbf{z}_p \mathbf{G}_{po} \mathbf{G}^{-1} \tilde{\mathbf{u}} + \mathbf{R}_{po} \mathbf{R}^{-1} \tilde{\mathbf{e}} \end{aligned}$$

but  $\mu$  is unknown and is replaced by its BLUE  $\hat{\mu} = \frac{\mathbf{1}' \mathbf{V}^{-1} \mathbf{y}}{\mathbf{1}' \mathbf{V}^{-1} \mathbf{1}}$  to get prediction

$$E(y_p | \mathbf{L}'_2 \mathbf{y}) = \hat{\mu} + \mathbf{z}_p \mathbf{G}_{po} \mathbf{G}^{-1} \tilde{\mathbf{u}} + \mathbf{R}_{po} \mathbf{R}^{-1} \tilde{\mathbf{e}}$$

We want to compare this with the kriging estimate written in terms of  $\mathbf{c}$ , so note that

$$E(y_p | \mathbf{L}'_2 \mathbf{y}) = \hat{\mu} + \mathbf{c}' \mathbf{P} \mathbf{y} / \sigma^2$$

where here  $\mathbf{P} = \sigma^2 \left( \mathbf{V}^{-1} - \frac{\mathbf{V}^{-1} \mathbf{1} \mathbf{1}' \mathbf{V}^{-1}}{\mathbf{1}' \mathbf{V}^{-1} \mathbf{1}} \right)$ , so

$$\begin{aligned} E(y_p | \mathbf{L}'_2 \mathbf{y}) &= \frac{\mathbf{1}' \mathbf{V}^{-1} \mathbf{y}}{\mathbf{1}' \mathbf{V}^{-1} \mathbf{1}} + \mathbf{c}' \mathbf{P} \mathbf{y} / \sigma^2 \\ &= \left( \frac{\mathbf{1}' \mathbf{V}^{-1}}{\mathbf{1}' \mathbf{V}^{-1} \mathbf{1}} + \mathbf{c}' \left[ \mathbf{I} - \frac{\mathbf{V}^{-1} \mathbf{1} \mathbf{1}'}{\mathbf{1}' \mathbf{V}^{-1} \mathbf{1}} \right] \mathbf{V}^{-1} \right) \mathbf{y} = \boldsymbol{\lambda}' \mathbf{y} \end{aligned}$$

## Spatial linear mixed model (3)

Introduction

Geostatistics review

Spatial mixed models

Kriging variance

Field experiments

Observational data

References

Exercise

Given a known spatial covariance function, the prediction from the linear mixed model is therefore the same prediction as that made from kriging.

Advantages of the REML / mixed model formulation:

- readily extended to account for more fixed effects
- parameters of a given covariance structure estimated directly from the data (adjusted for any fixed effects)

In contrast, in the presence of fixed effects, the estimated empirical variogram is biased and does not provide a good basis for model selection:

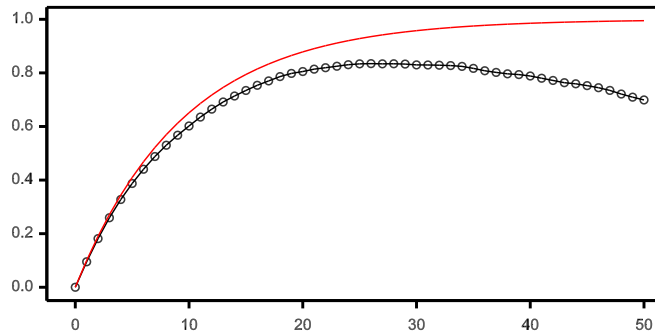


Figure 7: Estimated (black o) and theoretical (red -) variogram from a model with trend plus exponential spatial correlation of 0.9 at lag 1.



## Spatial linear mixed model (4)

[Introduction](#)

[Geostatistics review](#)

[Spatial mixed models](#)

[Kriging variance](#)

[Field experiments](#)

[Observational data](#)

[References](#)

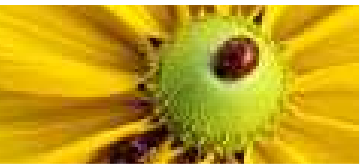
[Exercise](#)

The role of the variogram in the linear mixed model differs from geo-statistics:

- in geo-statistics, the correlation function is often estimated from the variogram
- in the linear mixed model, parameters of a correlation function are estimated directly from the data

Given the bias of the empirical variogram in the presence of fixed effects, it arguably has no useful role in estimation.

Several authors find it useful to use the variogram as a diagnostic tool, either to suggest a suitable correlation function or to find lack of fit.



## Spatial linear mixed model (5)

Introduction

Geostatistics review

Spatial mixed models

Kriging variance

Field experiments

Observational data

References

Exercise

Within the mixed model, there is a partition of random effects between  $u$  and  $e$  (the residual term, which must be present).

For random terms with  $Z \neq I$ , there is only one allocation.

If there is only one random term with  $Z = I$ , then this must be the residual term.

If there are several terms with  $Z = I$ , then in theory they may be combined as a composite residual term.

In practice, most software requires the residual to consist of a single term.

Special cases:

- spatial model with replicate measurements but no independent error
  - ◆ then  $Z \neq I_n$  so the model has no residual
- spatial model plus independent error with no replication - either term could be used as the residual



# Kriging variance

For a general model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\tau} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

and a prediction

$$\tilde{y}_p = \mathbf{x}_p \hat{\boldsymbol{\tau}} + \mathbf{z}_p \tilde{\mathbf{u}}_p + \tilde{e}_p$$

it is helpful to consider the form of the kriging variance

$$\begin{aligned} \text{var}(\tilde{y}_p - y_p) &= \text{var}(\mathbf{x}_p(\hat{\boldsymbol{\tau}} - \boldsymbol{\tau}) + \mathbf{z}_p(\tilde{\mathbf{u}}_p - \mathbf{u}_p) + \tilde{e}_p - e_p) \\ &= \text{var}(\mathbf{x}_p(\hat{\boldsymbol{\tau}} - \boldsymbol{\tau}) + \mathbf{z}_p \mathbf{G}_{p0} \mathbf{G}^{-1}(\tilde{\mathbf{u}} - \mathbf{u}) + \mathbf{R}_{p0} \mathbf{R}^{-1}(\tilde{\mathbf{e}} - \mathbf{e}) \\ &\quad + \mathbf{z}_p \mathbf{G}_{p0} \mathbf{G}^{-1}(\mathbf{u} - \mathbf{u}_p) + \mathbf{R}_{p0} \mathbf{R}^{-1} \mathbf{e} - e_p) \\ &= \begin{pmatrix} \mathbf{x}_p \\ \mathbf{z}_p \mathbf{G}_{p0} \mathbf{G}^{-1} \\ \mathbf{R}_{p0} \mathbf{R}^{-1} \end{pmatrix} \text{var} \begin{pmatrix} \hat{\boldsymbol{\tau}} - \boldsymbol{\tau} \\ \tilde{\mathbf{u}} - \mathbf{u} \\ \tilde{\mathbf{e}} - \mathbf{e} \end{pmatrix} \begin{pmatrix} \mathbf{x}_p \\ \mathbf{z}_p \mathbf{G}_{p0} \mathbf{G}^{-1} \\ \mathbf{R}_{p0} \mathbf{R}^{-1} \end{pmatrix}' \\ &\quad + \sigma^2 \mathbf{z}_p (\mathbf{G}_{pp} - \mathbf{G}_{p0} \mathbf{G}^{-1} \mathbf{G}_{0p}) \mathbf{z}_p' + \sigma^2 (\mathbf{R}_{pp} - \mathbf{R}_{p0} \mathbf{R}^{-1} \mathbf{R}_{0p}) \end{aligned}$$

This variance can be interpreted as the sum of uncertainty in the estimates plus uncertainty due to spatial interpolation (prediction).

Introduction

Geostatistics review

Spatial mixed models

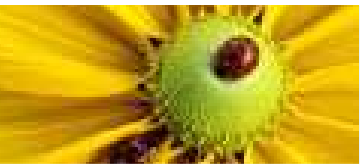
Kriging variance

Field experiments

Observational data

References

Exercise



## Kriging variance (2)

Introduction

Geostatistics review

Spatial mixed models

Kriging variance

Field experiments

Observational data

References

Exercise

Special cases:

- if prediction is at a point within the data set,  $\text{cov}(u_p, \mathbf{u}) = \mathbf{1}$  and  $\text{cov}(e_p, \mathbf{e}) = \mathbf{1}$ , then the fit is exact with no uncertainty in the prediction
- if  $\text{cov}(u_p, \mathbf{u}) = \mathbf{0}$  and  $\text{cov}(e_p, \mathbf{e}) = \mathbf{0}$ , ie. independent random effects, then  $\mathbf{G}_{po} = \mathbf{0}$  and  $\mathbf{R}_{po} = \mathbf{0}$  hence

$$\begin{aligned}\tilde{y}_p &= \mathbf{x}_p \hat{\boldsymbol{\tau}} \\ \text{var}(\tilde{y}_p - y_p) &= \mathbf{x}_p \text{var}(\hat{\boldsymbol{\tau}}) \mathbf{x}'_p + \sigma^2 \mathbf{z}_p \mathbf{G}_{pp} \mathbf{z}'_p + \sigma^2 \mathbf{R}_{pp}\end{aligned}$$

ie. uncertainty due to fixed effects plus uncertainty due to unobserved random effects.

This use of prediction can be used to take into account uncertainty in future random effects when making predictions.



# Spatial analysis of field experiments

Introduction

Geostatistics review

Spatial mixed models

Field experiments

Field experiments

Separability

2D variogram

Example

FE summary

Observational data

References

Exercise

- Aim is to model spatial variation to get improved SEDs
- Spatial models may be in conflict with randomization-based mixed model
- e.g. small blocks can be effective approximation to smooth spatial trend
- Danger of over-modelling variance pattern

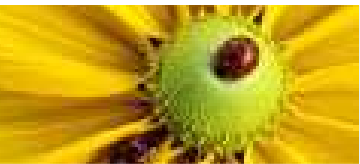
Procedure suggested here based on Gilmour *et al.* (1997).

Modelling variation in field experiments as a sum of:

- design-based blocking factors (to be retained as far as possible)
- local natural trend (local spatial trend)
- global natural trend (smooth trend across trial)
- extraneous variation

Extraneous variation often associated with harvesting or sowing patterns, plot trimming, etc

Use of diagnostics (inc variogram) to detect different sources of error



# Spatial analysis of field experiments

Introduction

Geostatistics review

Spatial mixed models

Field experiments

Field experiments

Separability

2D variogram

Example

FE summary

Observational data

References

Exercise

- Field experiments often laid out as a grid (rows  $\times$  columns)
- Field operations usually carried out in either row or column directions (sometimes alternating directions, *i.e.* up/down)
- Shape of field has sometimes determined directions of operations on long-term basis
- So: many factors that may affect variation are closely associated with either rows or columns

We therefore often assume that the spatial process is separable between rows and columns, *i.e.* row processes act independently of column processes.

- this seems realistic if we expect management practices to be the dominant form of spatial variation
- but unrealistic if spatial variation is due to underlying environmental conditions



# Separability

Introduction

Geostatistics review

Spatial mixed models

Field experiments

Field experiments

Separability

2D variogram

Example

FE summary

Observational data

References

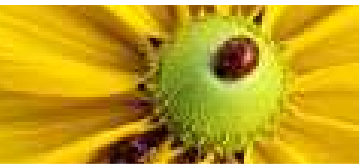
Exercise

- For data  $y_{ij}$  from row  $i$  and column  $j$  of the grid:
- let  $\mathbf{y} = (y_{11} \ y_{12} \ y_{13} \ \dots \ y_{rc})'$  be the vector of data ordered as columns within rows
- let  $C_r(\phi_r)$  be the covariance structure across rows within any column
- let  $C_c(\phi_c)$  be the covariance structure across columns within any row

Separability implies that

$$\text{var}(\mathbf{y}) = \sigma^2 C_r(\phi_r) \otimes C_c(\phi_c)$$

- simple model: two one-dimensional processes to consider
- separability usually assumed rather than tested
- departures from separability may be detected by variogram
- two dimensional variogram required to account for separate row & column processes



## Two-dimensional variogram

Introduction

Geostatistics review

Spatial mixed models

Field experiments

Field experiments

Separability

2D variogram

Example

FE summary

Observational data

References

Exercise

For a regular  $r \times c$  grid, a two dimensional variogram can be defined.

Let  $\tilde{e}_{ij}$  be the residual from row  $i$  and column  $j$ .

Then vertices of the two-dimensional variogram are calculated as

$$\tilde{V}(s, t) = \frac{1}{2(r-s)(c-t)} \sum_{i=1}^{r-s} \sum_{j=1}^{c-t} [ \tilde{e}_{ij} - \tilde{e}_{i+s, j+t} ]^2$$

- should be able to see row/column variance pattern in edges of variogram ( $s = 0$  or  $t = 0$ )
- variogram should show no trend or systematic pattern
- variogram should show regular behaviour consistent with separability
- GenStat command `f2dres [row=; column=] data=` (also easy way of getting 1D variograms for regular spacing)



## Example: analysis of uniformity data

Introduction

Geostatistics review

Spatial mixed models

Field experiments

Field experiments

Separability

2D variogram

Example

FE summary

Observational data

References

Exercise

- Uniformity trial to investigate impact of field operations
- Trial laid out as regular array 25 rows x 6 columns
- Management practices aligned with rows and columns:
  - ◆ trials sown by traversing columns in alternate directions
  - ◆ using cone seeder which sows two plots at once with left or right side
  - ◆ harvesting done in similar (but different) pattern to sowing
- Gilmour *et al.* suggest starting with AR1  $\otimes$  AR1 model then using diagnostic plots to identify lack of fit
  - ◆ plot of residuals against rows/columns for individual rows/columns or together
  - ◆ two-dimensional variogram
- procedure seems to work well in practice

# GenStat specification

Introduction

Geostatistics review

Spatial mixed models

Field experiments

Field experiments

Separability

2D variogram

Example

FE summary

Observational data

References

Exercise

```
"
  Model 1: initial investigation of spatial pattern
  =====
"
vcomp row.column
vstructure [row.column] factor=row,column; model=ar,ar
reml [prin=model,comp,dev] yield
vdis [prin=moni]

" Save residuals and look at patterns against rows/columns "
vkeep [residuals=res]
trellis [group=column; nrow=3; ncol=2] res; row; method=line
trellis res,res; row; method=point,mean

trellis [group=row] res; column; method=line
trellis res,res; column; method=point,mean

f2dres [row=row; col=column; title='Variogram for AR1 x AR1'] res; variogram=var1
```

- need to specify row.column residual term to apply variance model
- for separable model, use direct product structure





# Uniformity data

Introduction

Geostatistics review

Spatial mixed models

Field experiments

Field experiments

Separability

2D variogram

Example

FE summary

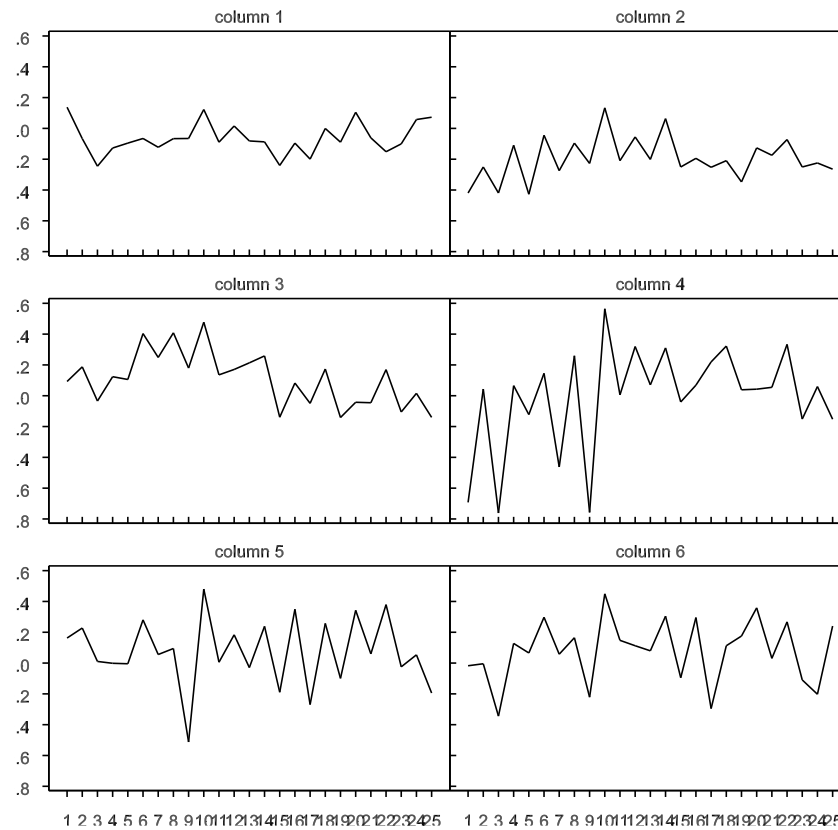
Observational data

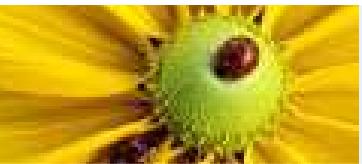
References

Exercise

Initial fit gives:  $\sigma^2 = 0.06$ ,  $\phi_r = 0.25$ ,  $\phi_c = 0.44$

Plot of residuals against row number within each column:





# Uniformity data

Introduction

Geostatistics review

Spatial mixed models

Field experiments

Field experiments

Separability

2D variogram

Example

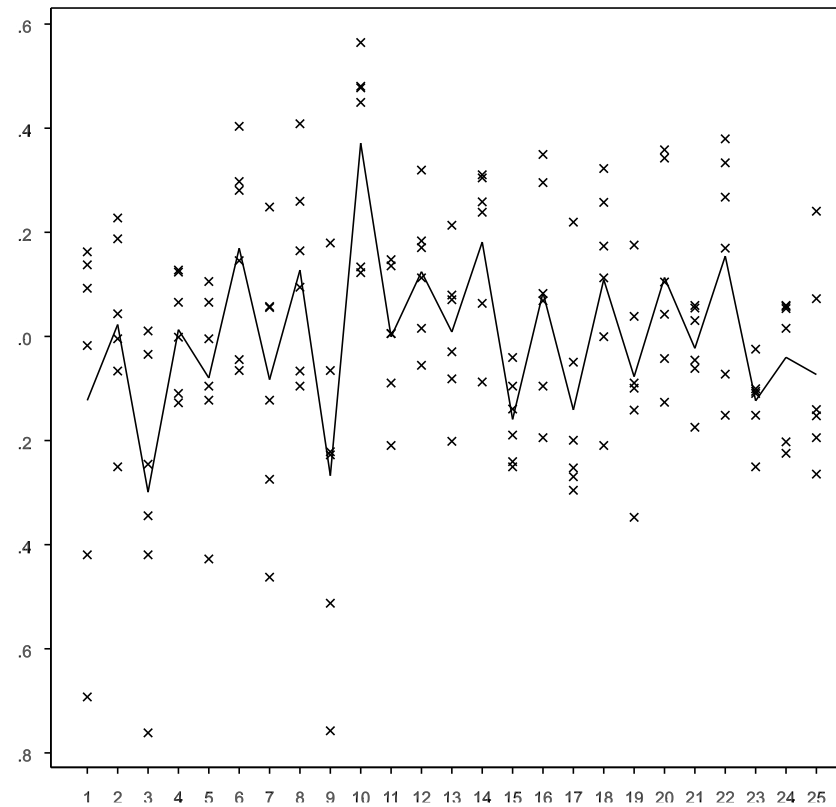
FE summary

Observational data

References

Exercise

Plot of residuals against row number with mean:



Strong alternating pattern due to field operations.



# Uniformity data

Introduction

Geostatistics review

Spatial mixed models

Field experiments

Field experiments

Separability

2D variogram

Example

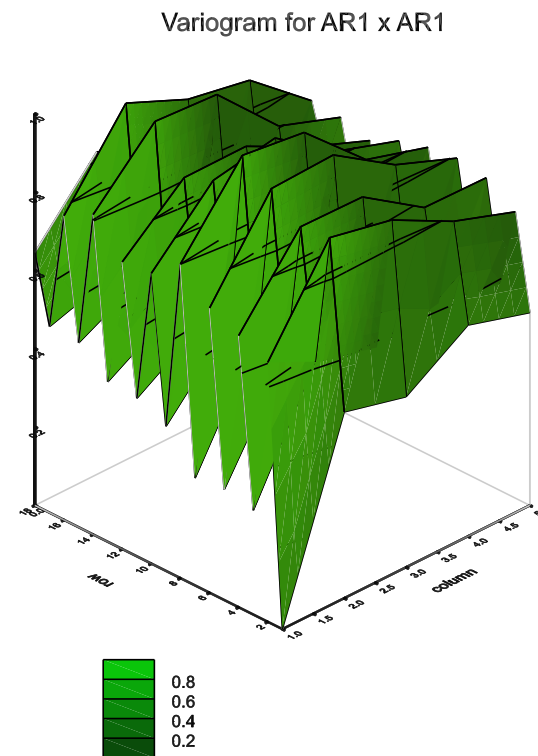
FE summary

Observational data

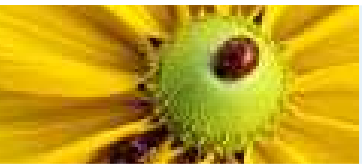
References

Exercise

## Two-dimensional variogram from initial model



Alternating pattern dominates row direction.



# Uniformity data

Introduction

Geostatistics review

Spatial mixed models

Field experiments

Field experiments

Separability

2D variogram

Example

FE summary

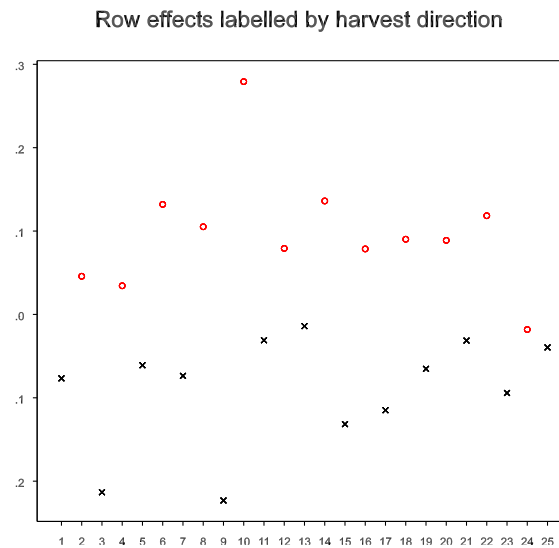
Observational data

References

Exercise

Second model = add random row effects:  $\sigma_r^2 = 0.19$ ,  $\sigma^2 = 0.04$ ,  $\phi_r = 0.36$ ,  $\phi_c = 0.11$

Plot estimated effects against factors describing field operations:



Alternating pattern appears to be due to harvest direction (possible problem with plot trimming): add fixed term for harvest direction to model



# Uniformity data

Introduction

Geostatistics review

Spatial mixed models

Field experiments

Field experiments

Separability

2D variogram

Example

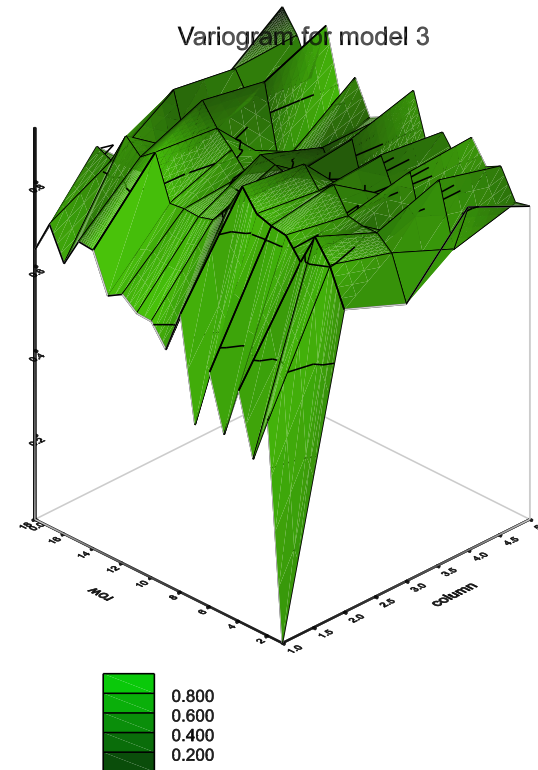
FE summary

Observational data

References

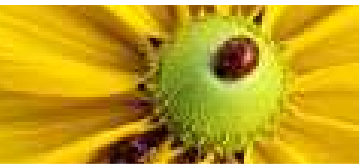
Exercise

Variogram now much improved:



Might try adding row and column random effects to take account of remaining systematic pattern in variogram (and residual plots)

Use comparison of AIC or BIC to formally select model.



# Uniformity data

- Introduction
- Geostatistics review
- Spatial mixed models
- Field experiments
- Field experiments
- Separability
- 2D variogram
- Example**
- FE summary
- Observational data
- References
- Exercise

Use AIC/BIC to refine model: all models include harvest direction as fixed effects

Variance component		Residual	AR correlation		$-2RL$	$N_v$	AIC	BIC
Row	Column	Variance	Row	Column				
0	0	0.042	0	0	-313.83	1	-311.83	-308.83
0.005	0.012	0.027	0	0	-350.21	3	<u>-344.21</u>	<u>-335.21</u>
0	0	0.043	0.35	0.17	-331.42	3	-325.42	-316.42
0.003	0	0.040	0.36	0.10	-332.55	4	-324.55	-312.55
0	0.012	0.032	0.10	0.19	-348.64	4	-340.64	-328.64
0.004	0.012	0.027	0.07	0.06	-350.76	5	-340.76	-325.76

Note duality between main effects and correlations

- including row main effect decreases column AR parameter
- including column main effect decreases row AR parameter
- exactly the same phenomena as noted for longitudinal data: random main effects induce correlation within same level of factor



# Field experiments: summary

Introduction

Geostatistics review

Spatial mixed models

Field experiments

Field experiments

Separability

2D variogram

Example

FE summary

Observational data

References

Exercise

Suggested procedure is iterative:

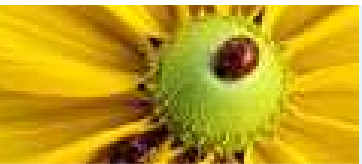
- try model → check residuals → try new model
- then formalize choice using AIC/BIC
- difficult to justify formally, seems to work in practice

In this example, there was no evidence of global trend

- can take out trend using polynomials or smoothing splines

Real danger of over-modelling:

- be wary of uncertainty in variograms / residual plots
- use of BIC gives some protection
- be aware of role of design-based factors in model (pragmatism)



# Linear mixed models for observational data

Introduction

Geostatistics review

Spatial mixed models

Field experiments

Observational data

Example

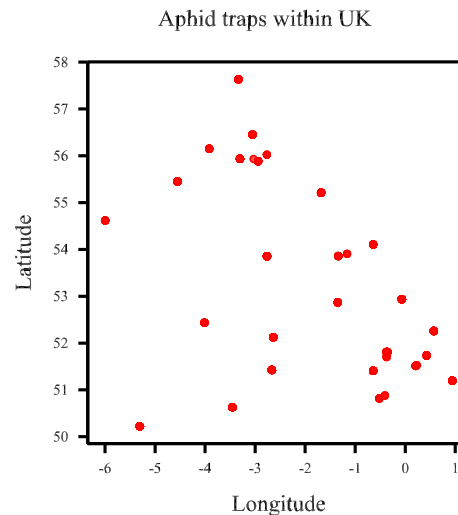
References

Exercise

Aim: to quantify influence of environmental factors on aphid occurrence in traps.

Scope of project:

- environmental factors: climate, land use and pollution
- trap positions across Western Europe
- data for 35 years (mainly UK in first 10 years)
- here: pilot study using a subset of trap data from the UK from 1965-1998, climate and land use data only







# Spatial analysis

Introduction

Geostatistics review

Spatial mixed models

Field experiments

Observational data

Example

References

Exercise

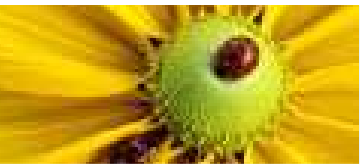
## Aims of analysis:

- use (linear) regression to model relationship between aphid variables and environment
- investigate spatial trend in aphid variables not accounted for by environmental variables

Note: geo-statistics usually works on a single (large) spatial sample: we have small samples within each year.

## Features of analysis:

- mixed models allow joint estimation of regression coefficients and spatial correlation
- assuming spatial correlation is the same, can accumulate information across years
- can allow additional variation due to sites and/or years
- can allow correlation across years (independent of spatial correlation)
- REML estimation of variance parameters



# Model

Introduction

Geostatistics review

Spatial mixed models

Field experiments

Observational data

Example

References

Exercise

The proposed model uses

- fixed terms for regression on environmental variables
- variance components for site and year
- direct product correlation structure with independence across years, spatial correlation ( $\mathbf{C}$ ) across sites within years

Model is written:

$$y_{ij} = \mathbf{X}\boldsymbol{\tau} + t_i + s_j + v_{ij} + \eta_{ij}$$

where

- $y_{ij}$  is the data in year  $i$  ( $i=1 \dots m$ ) at trap  $j$  ( $j=1 \dots n$ ), ordered as sites within years
- $\mathbf{X}\boldsymbol{\tau}$  represents regression terms (fixed)
- $t_i$  is the effect for year  $i$ , with  $\text{var}(\mathbf{t}) = \sigma_Y^2 \mathbf{I}_m$
- $s_j$  is the effect for trap  $j$ ,  $\text{var}(\mathbf{s}) = \sigma_S^2 \mathbf{I}_n$
- $v_{ij}$  is spatially correlated variation,  $\text{var}(\mathbf{v}) = \sigma_v^2 \mathbf{I}_m \otimes \mathbf{C}$
- $\eta_{ij}$  is additional independent sampling error,  $\text{var}(\boldsymbol{\eta}) = \sigma^2 \mathbf{I}_{mn}$



# Random model

Introduction

Geostatistics review

Spatial mixed models

Field experiments

Observational data

Example

References

Exercise

$$y_{ij} = \mathbf{X}\boldsymbol{\tau} + t_i + s_j + v_{ij} + \eta_{ij}$$

Independent year effects ( $t$ ) allow for effect of season that applies across all sites

- does not help prediction in new years, might be explained by year-level variables

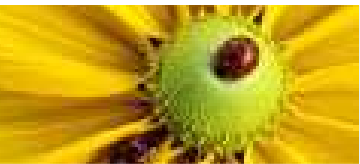
Independent site effects ( $s$ ) allow for consistent effect at each site across years

- does not help prediction at new sites, might be explained by site-level variables

Spatial correlation ( $v$ ) allows for smooth spatial trend within each year

- can predict at new sites within years where data available, explained by site.year variables

Random error ( $\eta$ ) - sampling variation.



## Models for spatial correlation

Introduction

Geostatistics review

Spatial mixed models

Field experiments

Observational data

Example

References

Exercise

Correlation across sites is based on distance between sites (exponential model).

For data at sites A (at location  $(x_a, y_a)'$ ) and B (at location  $(x_b, y_b)'$ ) the correlation within year is modelled either as

- $C(A, B) = \phi_x^{|x_a - x_b|} \phi_y^{|y_a - y_b|}$  (anisotropic model)
- $C(A, B) = \phi^{(|x_a - x_b| + |y_a - y_b|)}$  (isotropic model, city-block distances)
- $C(A, B) = \phi^{\sqrt{(|x_a - x_b|^2 + |y_a - y_b|^2)}}$  (isotropic model, Euclidean distances)

where in all cases  $0 \leq \phi, \phi_x, \phi_y \leq 1$ .

For these models, the correlation decreases smoothly with distance, the parameter  $\phi$  is the correlation at a distance of 1 unit in given direction (1 unit = 160km).

First model: look at variation without taking account of environmental variables

- leaving out fixed effects changes values of variance parameters
- not generally advisable - here, may give insight into baseline size and sources of variation due to environmental and site/year factors
- can assess proportion of natural variation accounted for by environmental variables

# Modelling strategy

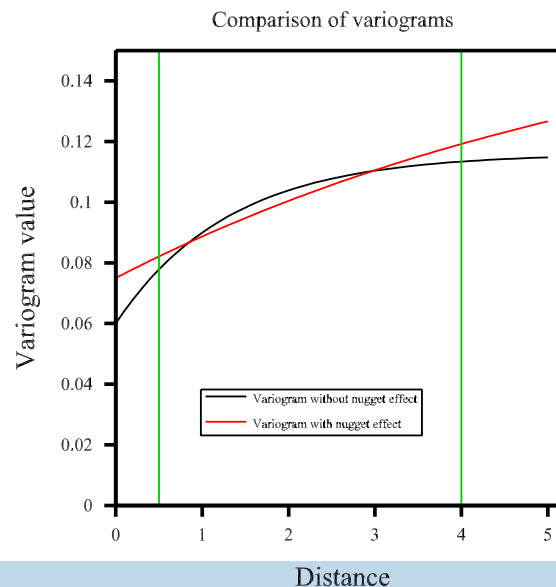
- Introduction
- Geostatistics review
- Spatial mixed models
- Field experiments
- Observational data
- Example**
- References
- Exercise

Fit various models and compare using residual log-likelihood (nested models) or information criteria (non-nested).

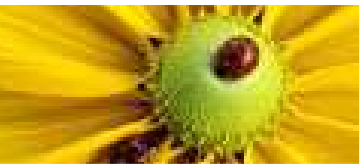
## Summary of best variance models without fixed terms

Variable	Estimated variance parameters					
	$\sigma_Y^2$	$\sigma_S^2$	$\sigma_v^2$	$\phi_y$	$\phi_x$	$\sigma^2$
Log(Mp total)	0.061	0.119	0.113	0.69	0.45	-
Log(Rp total)	-	0.059	0.097	0.86	0.86	0.015*
Mp 1st flight	236.8	136.6	669.3	0.35	0.35	-
Rp 1st flight	481.3	158.1	602.2	0.18	0.18	-

\* Interpretation of nugget effect questionable in this context



# Modelling strategy



Introduction

Geostatistics review

Spatial mixed models

Field experiments

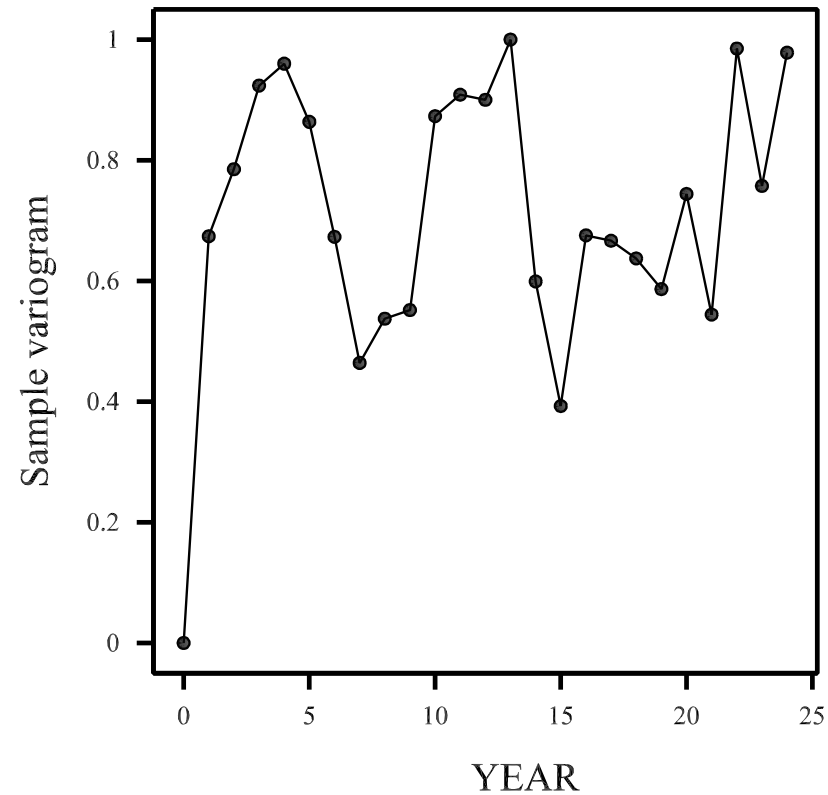
Observational data

Example

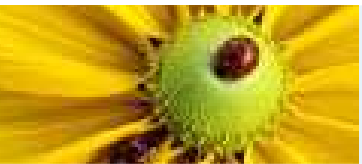
References

Exercise

Log( $R_p$  total) variogram across years



Empirical variogram is probably consistent with either model - experience of the data supports presence of large sampling variation.



Introduction

Geostatistics review

Spatial mixed models

Field experiments

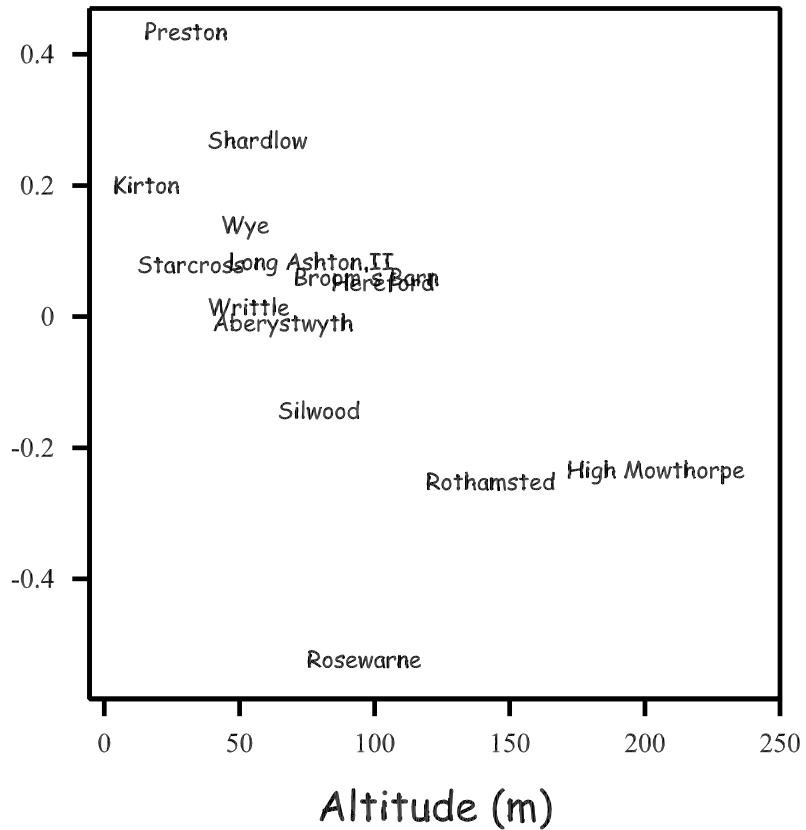
Observational data

Example

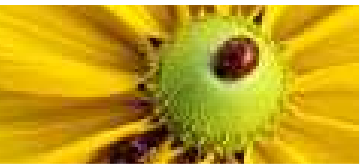
References

Exercise

### Trap random effects for $\log(R_p \text{ total})$



Examination of the trap effects suggests that there is some dependence on environmental effects, here altitude.



# Mixed model with environmental covariates

Introduction

Geostatistics review

Spatial mixed models

Field experiments

Observational data

Example

References

Exercise

Limited set of covariates:

- site position: longitude (X), latitude (Y), altitude (Z)
- climate: Jan/Feb temperature, Jan/Feb rainfall
- landscape: area of potatoes, oilseed rape or sugar beet in 50 km<sup>2</sup>

Modelling strategy

- fit all covariates, reselect random model
- drop unimportant fixed variables (Wald tests)
- not robust procedure with respect to variable selection, return to this later
- chosen variables: X, Y, Y<sup>2</sup>, Z, osr





# Model summary

- Introduction
- Geostatistics review
- Spatial mixed models
- Field experiments
- Observational data
- Example**
- References
- Exercise

Comparison of variance models excluding and including environmental variables:

-/+ covariates	Variable	Estimated variance parameters					
		$\sigma_Y^2$	$\sigma_S^2$	$\sigma_v^2$	$\phi_y$	$\phi_x$	$\sigma^2$
-	Log(Mp total)	0.061	0.119	0.113	0.69	0.45	-
+	Log(Mp total)	0.053	0.003	0.113	0.70	0.46	-
-	Log(Rp total)	-	0.059	0.097	0.86	0.86	0.015
+	Log(Rp total)	-	0.010	0.087	0.83	0.83	0.014
-	Mp 1st flight	236.8	136.6	669.3	0.35	0.35	-
+	Mp 1st flight	-	-	639.0	0.32	0.32	-
-	Rp 1st flight	481.3	158.1	602.2	0.18	0.18	-
+	Rp 1st flight	102.4	-	649.1	0.19	0.19	-

- independent site and year variation greatly reduced
- site.year spatial variation not reduced by addition of covariates
- site.year level covariates (climate and landuse) turn out to have year+site additive structure

# Prediction/kriging

Introduction

Geostatistics review

Spatial mixed models

Field experiments

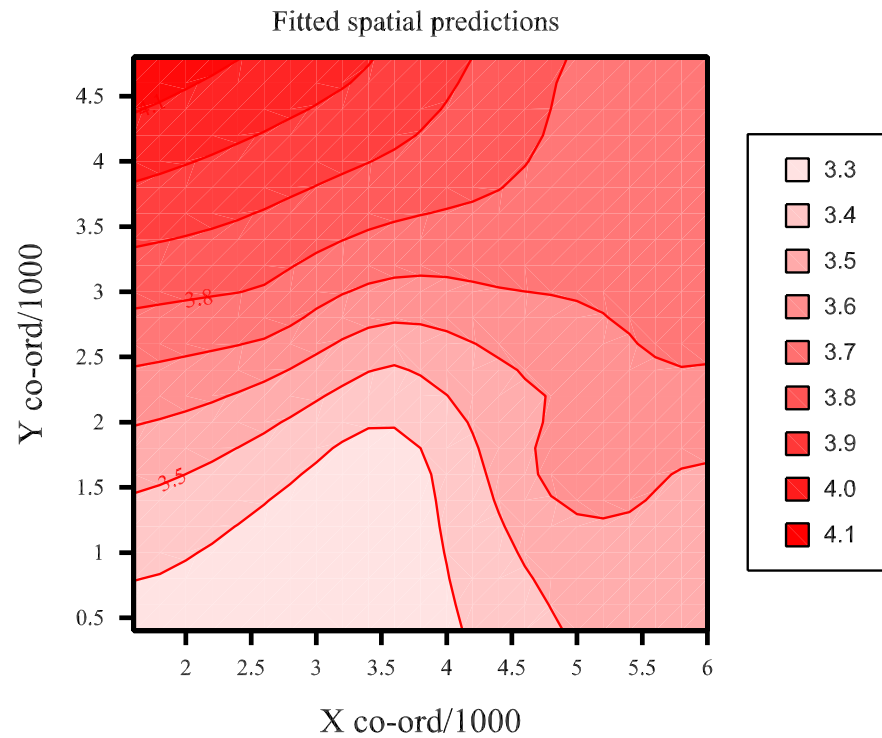
Observational data

Example

References

Exercise

Predictions at unobserved locations (kriging) are made using all terms:



This surface is more complex than could be predicted using the fixed terms only.



## References / Further reading

Introduction

Geostatistics review

Spatial mixed models

Field experiments

Observational data

References

References

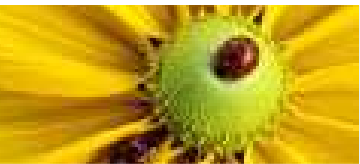
Exercise

CRESSIE, N. A. C. (1993) *Analysis of spatial data* John Wiley & Sons, New York.

GILMOUR, A. R. AND CULLIS, B. R. AND VERBYLA, A. P. (1997) Accounting for natural and extraneous variation in the analysis of field experiments. *Journal of Agricultural, Biological, and Environmental Statistics*, **2**, 269–273.

HASKARD, K.A. (2007) *An isotropic Matern spatial covariance model: REML estimation and properties*. PhD thesis, University of Adelaide.

LARK, R.M., CULLIS, B.R, & WELHAM, S.J. (2006) On spatial prediction of soil properties in the presence of a spatial trend: the empirical best linear unbiased predictor (E-BLUP) with REML. *European Journal of Soil Science*, **57**, 787-799.



# Exercise

Introduction

Geostatistics review

Spatial mixed models

Field experiments

Observational data

References

Exercise

Exercise

Field experiment with spatial trend:

- Laid out as grid of 16 rows  $\times$  6 columns
- Tests 8 varieties at 6 seedrates with 2 reps
- Rep 1 = columns 1-3, rep 2 = columns 4-6
- data held in spreadsheet yield.xls

Try to find a sensible model for this data.