Spatial statistics

Paul D. Sampson
Peter Guttorp

Course description

Lectures (10)

Practica (8)

Need to submit solutions to the problems given in at least four of the practica. Can be done in groups of 2-3.

Homework

Need to submit eight homework problems. To be done individually. A data analysis can, with permission, replace three problems.

Office hours

PG Tu 12-1

PDS Th 11-12

PASI

Outline

1. Kriging (9/28)
2. Spatial covariance (10/3)
3. Nonstationary structures I: deformations (10/10)
4. Nonstationary structures II: linear combinations etc. (10/17)
5. Space-time models (10/24)
6. Markov random fields (10/31)
7. Misalignment and use of deterministic models (11/7)
8. Design of monitoring network (11/14)
9. Extremes (11/16)
10. Statistical climatology (11/28)

Kriging

The geostatistical model

Gaussian process $Z(s), s \in D \subseteq R^2$

$\mu(s)=EZ(s)$ Var $Z(s) < \infty$

Z is strictly stationary if

$$(Z(s_1),\dots,Z(s_k)) \stackrel{d}{=} (Z(s_1+h),\dots,Z(s_k+h))$$

Z is weakly stationary if

$$\mu(s) \equiv \mu \quad \mathrm{Cov}(Z(s_1),Z(s_2)) = C(s_1 - s_2)$$

Z is isotropic if weakly stationary and

$$C(s_1 - s_2) = C_0(\|s_1 - s_2\|)$$

The problem

Given observations at n locations

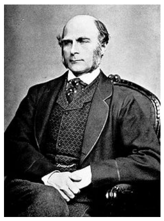
$$Z(s_1),\ldots,Z(s_n)$$

estimate

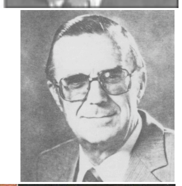
$Z(s_0)$ **(the process at an unobserved location)**

or $\int_A Z(s)d\nu(s)$ **(an average of the process)**

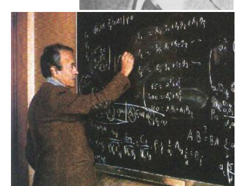
In the environmental context often time series of observations at the locations.

Some history

Regression (Bravais, Galton, Bartlett)



Mining engineers (Krige 1951, Matheron, 60s)

Spatial models (Whittle, 1954)

Forestry (Matérn, 1960)


Objective analysis (Gandin, 1961)

More recent work:


Stein (1999)

Gelfand et al. (2010)

A Gaussian formula

If $\quad \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \sim \mathbf{N}\left(\begin{pmatrix} \mu_{\mathbf{X}} \\ \mu_{\mathbf{Y}} \end{pmatrix}, \begin{pmatrix} \Sigma_{\mathbf{XX}} & \Sigma_{\mathbf{XY}} \\ \Sigma_{\mathbf{YX}} & \Sigma_{\mathbf{YY}} \end{pmatrix} \right)$

then $\quad (\mathbf{Y} \mid \mathbf{X}) \sim \mathbf{N}(\mu_{\mathbf{Y}} + \Sigma_{\mathbf{YX}} \Sigma_{\mathbf{XX}}^{-1} (\mathbf{X} - \mu_{\mathbf{X}}),$

$$\Sigma_{\mathbf{YY}} - \Sigma_{\mathbf{YX}} \Sigma_{\mathbf{XX}}^{-1} \Sigma_{\mathbf{XY}})$$

Simple kriging

Let $X = (Z(s_1),\ldots,Z(s_n))^T$, $Y = Z(s_0)$, so that

$\mu_X=\mu 1_n$, $\mu_Y=\mu$,

$\Sigma_{XX}=[C(s_i-s_j)]$, $\Sigma_{YY}=C(0)$, and

$\Sigma_{YX}=[C(s_i-s_0)]$.

Then

$$p(X) \equiv \hat{Z}(s_0) = \mu + \left[C(s_i - s_0)\right]^T \left[C(s_i - s_j)\right]^{-1} \left(X - \mu 1_n\right)$$

This is the best unbiased linear predictor when μ and C are known (simple kriging).

The prediction variance is

$$m_1 = C(0) - \left[C(s_i - s_0)\right]^T \left[C(s_i - s_j)\right]^{-1} \left[C(s_i - s_0)\right]$$

Some variants

Ordinary kriging (unknown μ)

$$p(X) \equiv \hat{Z}(s_0) = \hat{\mu} + \left[C(s_i - s_0)\right]^T \left[C(s_i - s_j)\right]^{-1} \left(X - \hat{\mu} 1_n\right)$$

where

$$\hat{\mu} = \left(1_n^T \left[C(s_i - s_j)\right]^{-1} 1_n\right)^{-1} 1_n^T \left[C(s_i - s_j)\right]^{-1} X$$

Universal kriging ($\mu(s)=A(s)\beta$ for some spatial variable A)

$$\hat{\beta} = (\left[A(s_i)\right]^T \left[C(s_i - s_j)\right]^{-1} \left[A(s_i)\right])^{-1}$$

$$\left[A(s_i)\right]^T \left[C(s_i - s_j)\right]^{-1} X$$

$$p(X) \equiv \hat{Z}(s_0) = \hat{\mu}(s_0) + \left[C(s_i - s_0)\right]^T \left[C(s_i - s_j)\right]^{-1} \left(X - [\hat{\mu}(s_i)]\right)$$

where $\hat{\mu}(s) = A(s)\hat{\beta}$

Still optimal for known C.

Universal kriging variance

simple kriging variance

$$E\left(\hat{Z}(s_0) - Z(s_0)\right)^2 = \boxed{m_1} +$$

$$\boxed{\begin{aligned}&\left(A(s_0) - [A(s_i)^T[C(s_i - s_j)]^{-1}[C(s_i - s_0)]\right)^T\\ &\times\left(\left[A(s_i)\right]^T\left[C(s_i - s_j)\right]^{-1}\left[A(s_i)\right]\right)^{-1}\\ &\times\left(A(s_0) - [A(s_i)^T[C(s_i - s_j)]^{-1}[C(s_i - s_0)]\right)\end{aligned}}$$

variability due to estimating μ

The (semi)variogram

$$\gamma(\|h\|) = \frac{1}{2}\mathrm{Var}(Z(s+h) - Z(s)) = C(0) - C(\|h\|)$$

Intrinsic stationarity

Weaker assumption (C(0) needs not exist)

Kriging predictions can be expressed in terms of the variogram instead of the covariance.

The exponential variogram

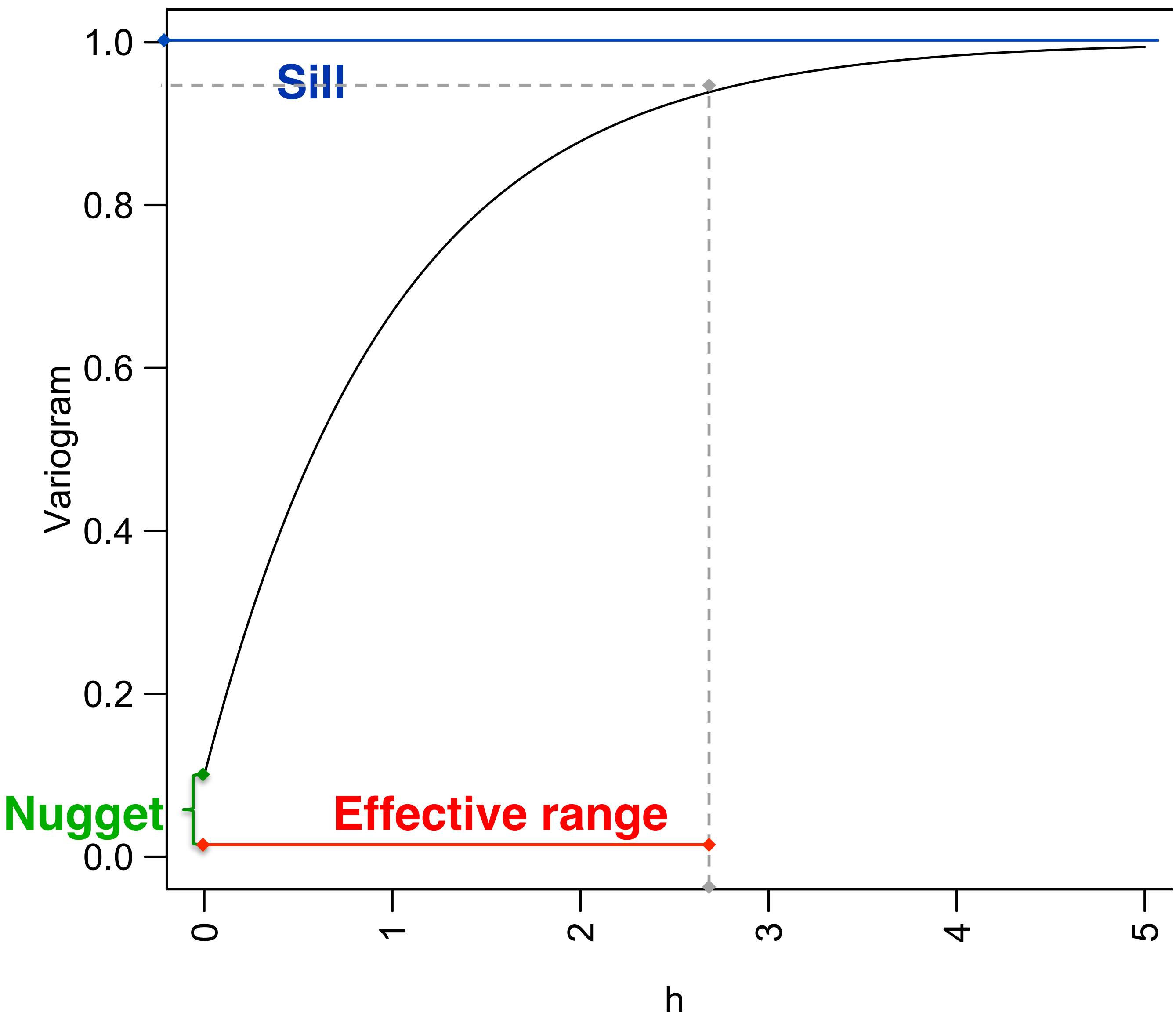
A commonly used variogram function is $\gamma(h) = \sigma^2 (1 - e^{-h/\varphi})$.

Corresponds to a Gaussian process with continuous but not differentiable sample paths.

More generally,

$$\gamma(h) = (\sigma^2 - \tau^2)\left(1 - e^{-h/\varphi}\right) + \tau^2$$

has a nugget τ^2, corresponding to measurement error and spatial correlation at small distances.



1.0
0.8
0.6
0.4
0.2
0.0
Variogram
Sill
Nugget
Effective range
0
1
2
3
4
5
h

Ordinary kriging

$$\hat{Z}(s_0) = \sum_{i=1}^{n} \lambda_i Z(s_i)$$

where

$$\lambda^T = \left(\gamma + 1 \frac{1 - 1^T \Gamma^{-1} \gamma}{1^T \Gamma^{-1} 1} \right)^T \Gamma^{-1}$$

$$\gamma = (\gamma(s_0 - s_1), \ldots, \gamma(s_0 - s_n))^T$$

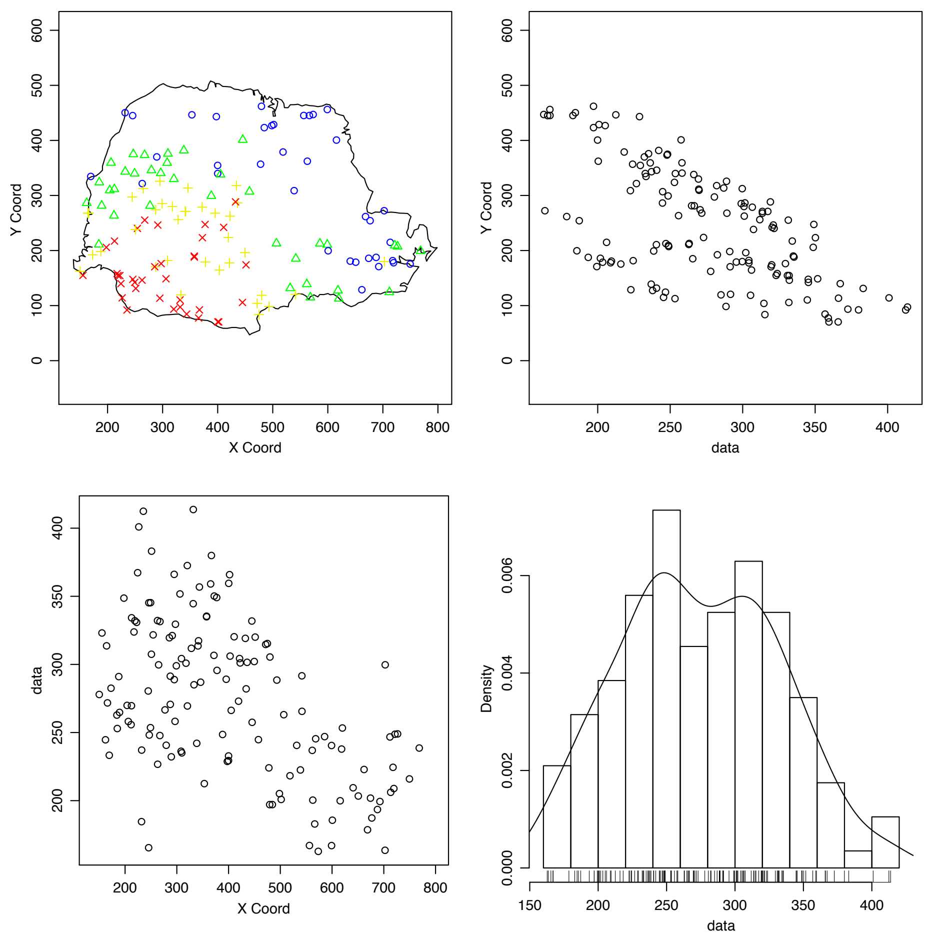
$$\Gamma_{ij} = \gamma(s_i - s_j)$$

and kriging variance

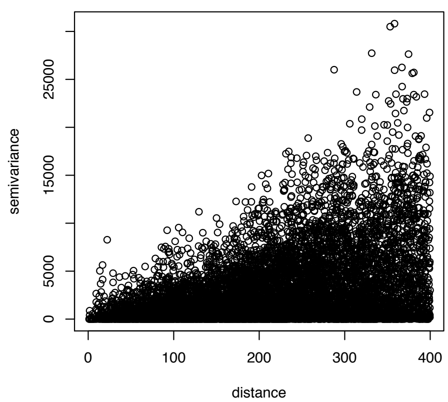
$$m_1(s_0) = 2\sum_{i=1}^{n} \lambda_i \gamma(s_0 - s_i) - \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j \gamma(s_i - s_j)$$

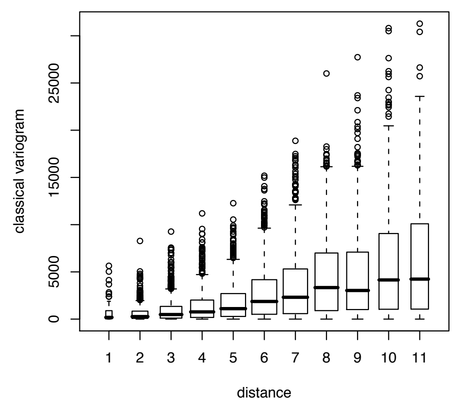
An example

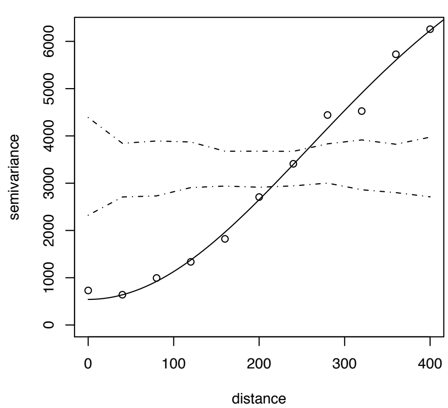
Precipitation data from Parana state in Brazil (May-June, averaged over years)



Variogram plots

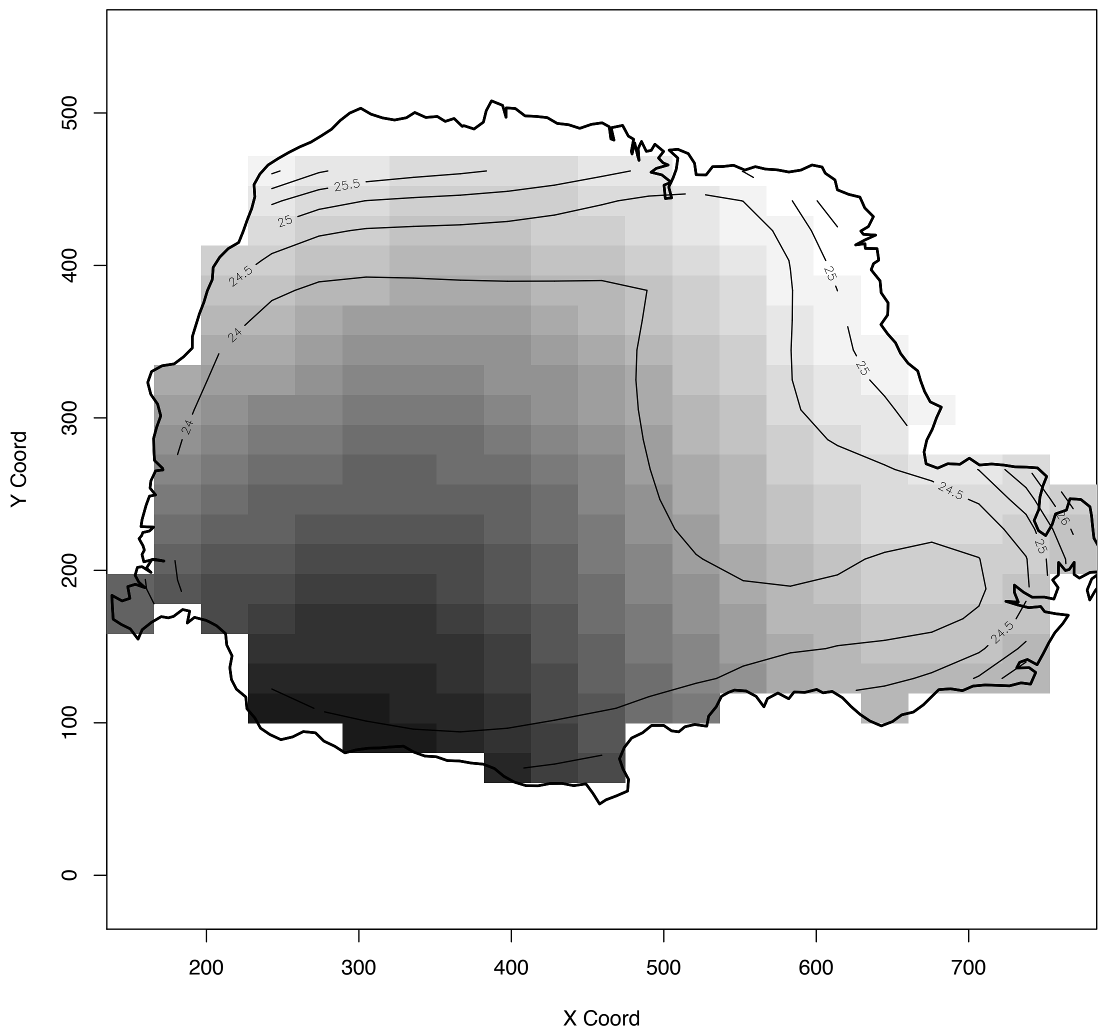






$$\gamma(h) = 542 + 8141\left(1 - \exp\left(-\left(\frac{h}{365}\right)^2\right)\right)$$

Kriging surface



Ozone data set

Built in data set in “maps” library in R

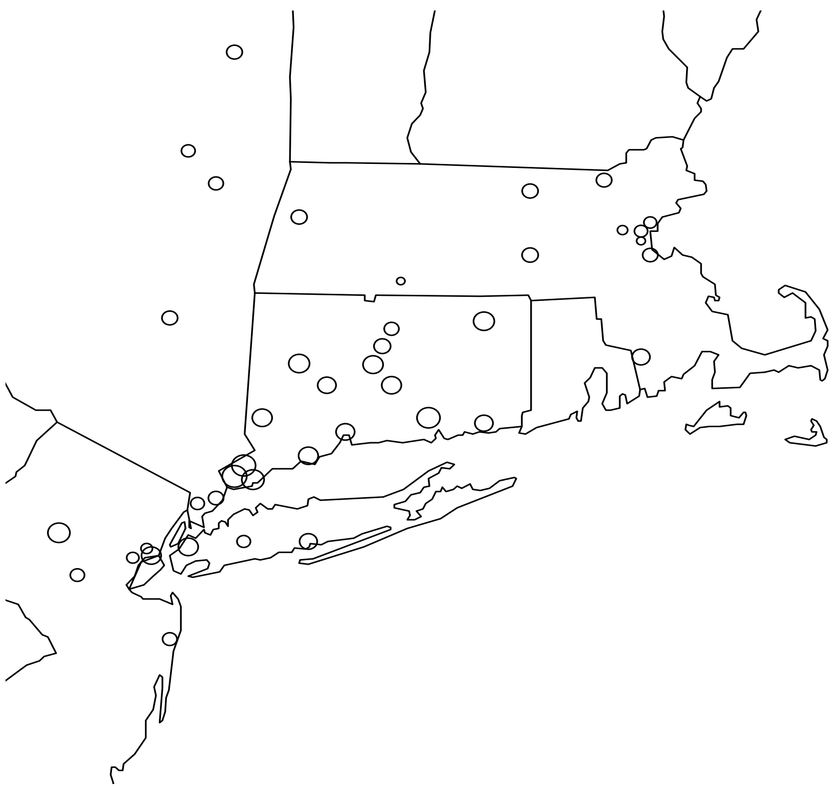
NW US ozone data

1974 June-August median daily maximum ground level ozone data from 41 stations in New Jersey, New York, Connecticut and Massachusetts

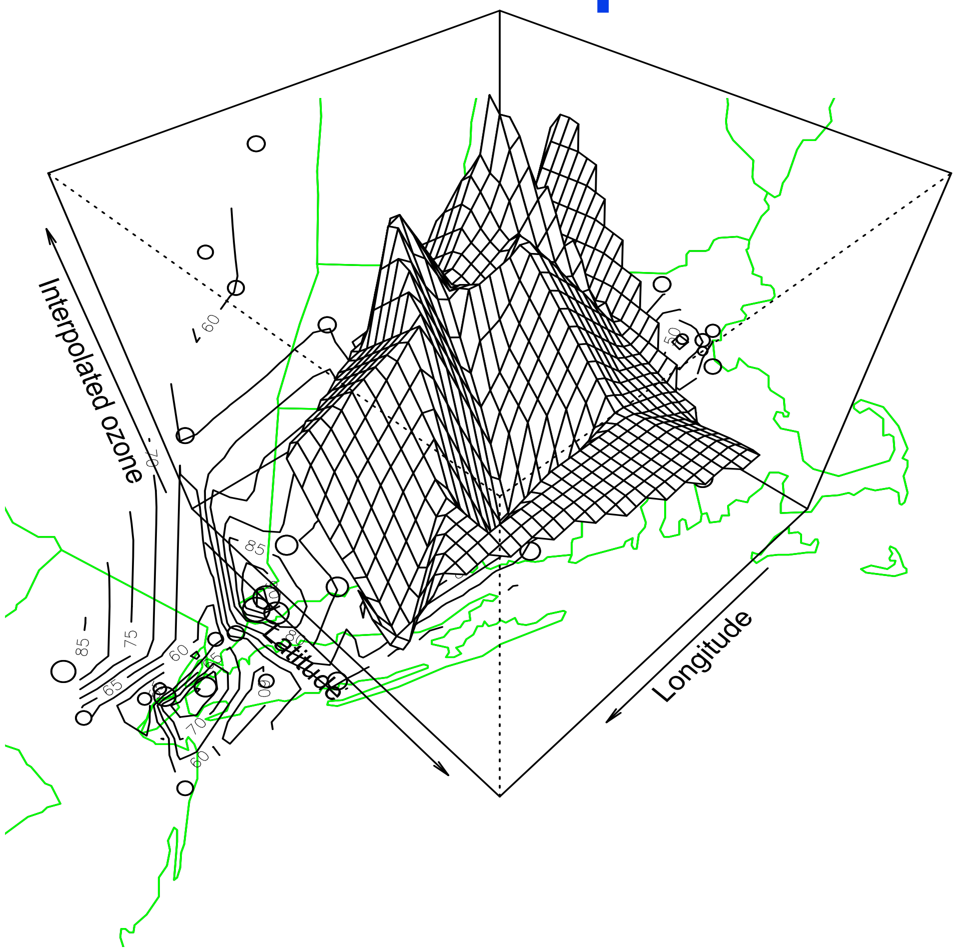
Contour plot using bilinear interpolation

Kriging with exponential covariance function and nugget

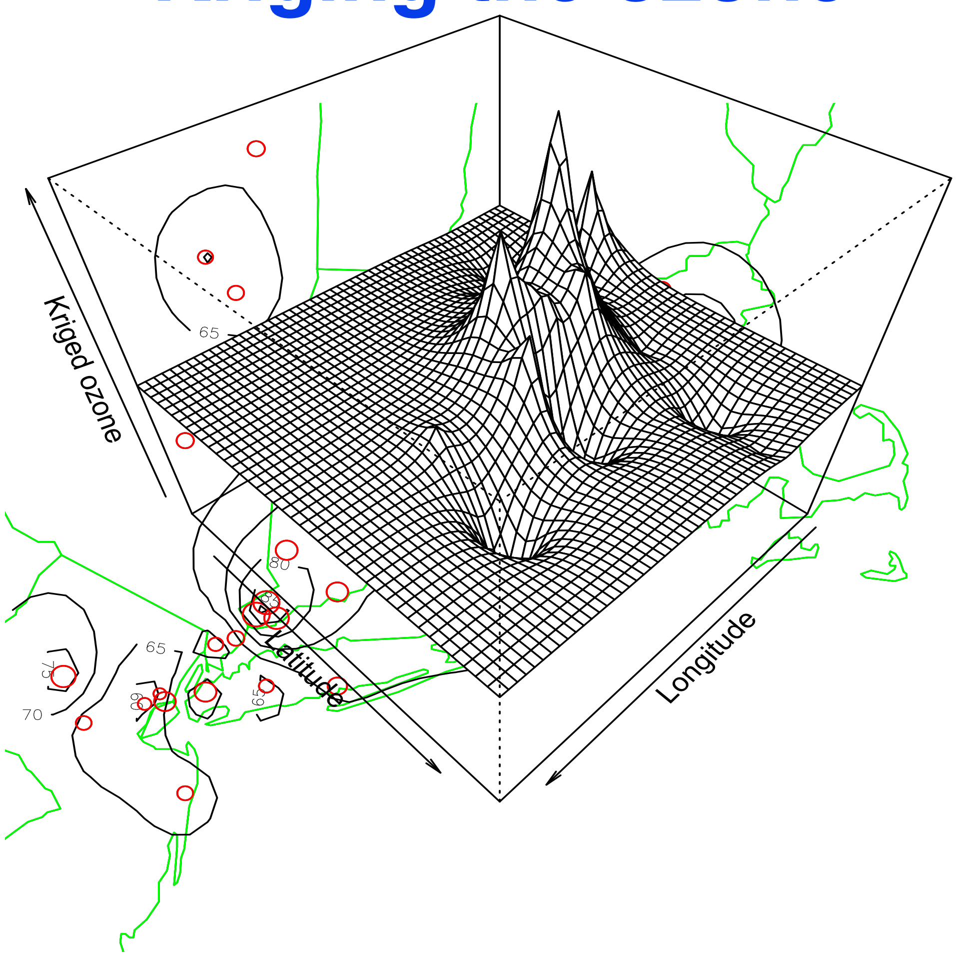
Data



Bilinear interpolation



Kriging the ozone



How many contours?

A Gaussian prediction $\hat{Z}(s)$ falls between contour lines between a and b with probability

$$\begin{aligned} p(s) &= \Pr(a < \hat{Z}(s) < b) \\ &= \Phi((1-r)q) - \Phi(-rq) \end{aligned}$$

where q=(b-a)/σ and r= $(\hat{Z}(s)-a)/(b-a)$

If q=.5 the probability is at most 0.2 that a statement about the level of Z(s) is correct (Polfeldt, 1999).

If q=2 it is at most 2/3

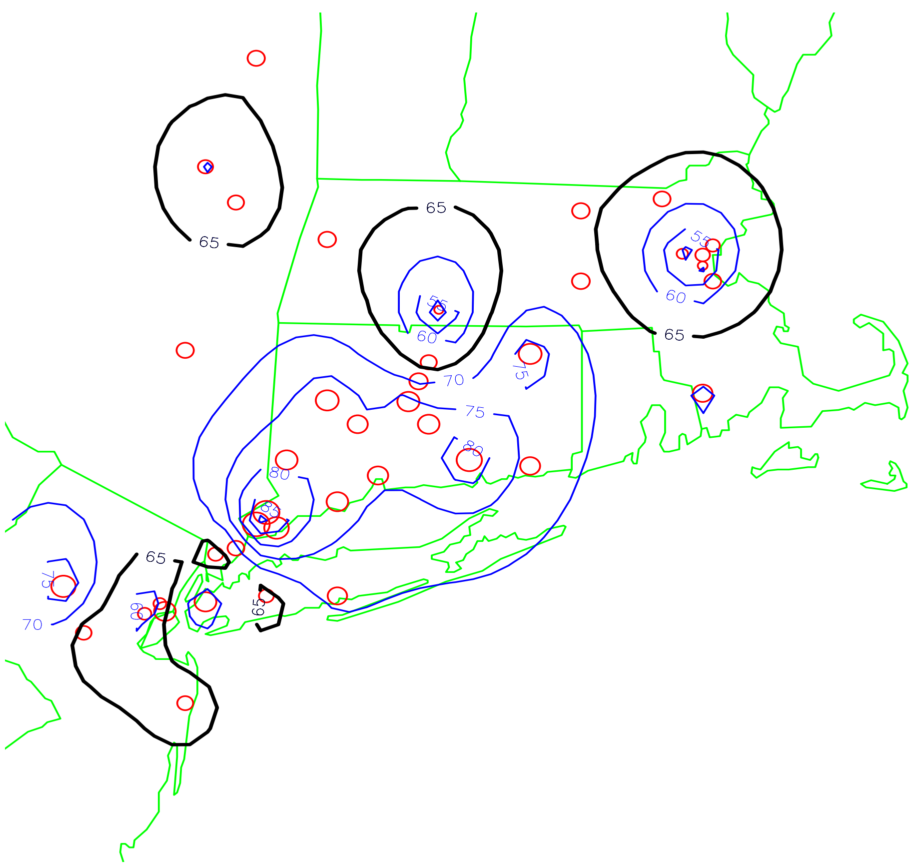
If q=4 it is at most 0.95.

Consequences

Points close to contour lines are always very uncertain as to whether they should be above or below the line.

If the contour lines are well separated there are high probabilities of correctness in the middle between them.

Revisit kriging contours



Bayesian kriging

Instead of estimating the parameters, we put a prior distribution on them, and update the distribution using the data.

Model:

$$(\boxed{\mathbf{X}}\,|\,\boxed{\theta}) \sim N(\beta, \sigma^2\,\boxed{\mathbf{C}(\phi)} + \boxed{\tau^2\mathbf{I}})$$

$\boxed{(Z(s_1)\ldots Z(s_n))^T}$

$\boxed{\theta=(\beta,\sigma^2,\phi,\tau^2)^T}$

Matrix with i,j-element $C(s_i - s_j;\ \phi)$ (correlation)

measurement error

Prior and posterior

Prior: $f(\theta) = f(\beta)f(\sigma^2)f(\phi)f(\tau^2)$

Posterior: $f(\theta \mid X = x) \propto f(\theta)f(x \mid \theta)$

$$f(\phi \mid X = x) \propto f(\phi)\iiint f(x \mid \theta)f(\beta)f(\sigma^2)f(\tau^2)\,d\beta\,d\sigma^2 d\tau^2$$

Predictive distribution:

$$f(Z(s_0) \mid X = x) = \int \underbrace{\boxed{p(x;\theta)}}_{\text{kriging predictor}} f(\theta \mid x)\,d\theta$$

Specifically

Exponential isotropic correlation function:
$$C(h;\phi) = \exp(-\phi h)$$

Default correlation model in geoR.

Prior on β defaults to flat, but can also be normal or fixed

Prior on σ^2 defaults to reciprocal, but can be scaled inverse chisquare or flat

Prior on ϕ can be exponential, uniform, reciprocal, squared reciprocal, or user specified (discrete)

More priors

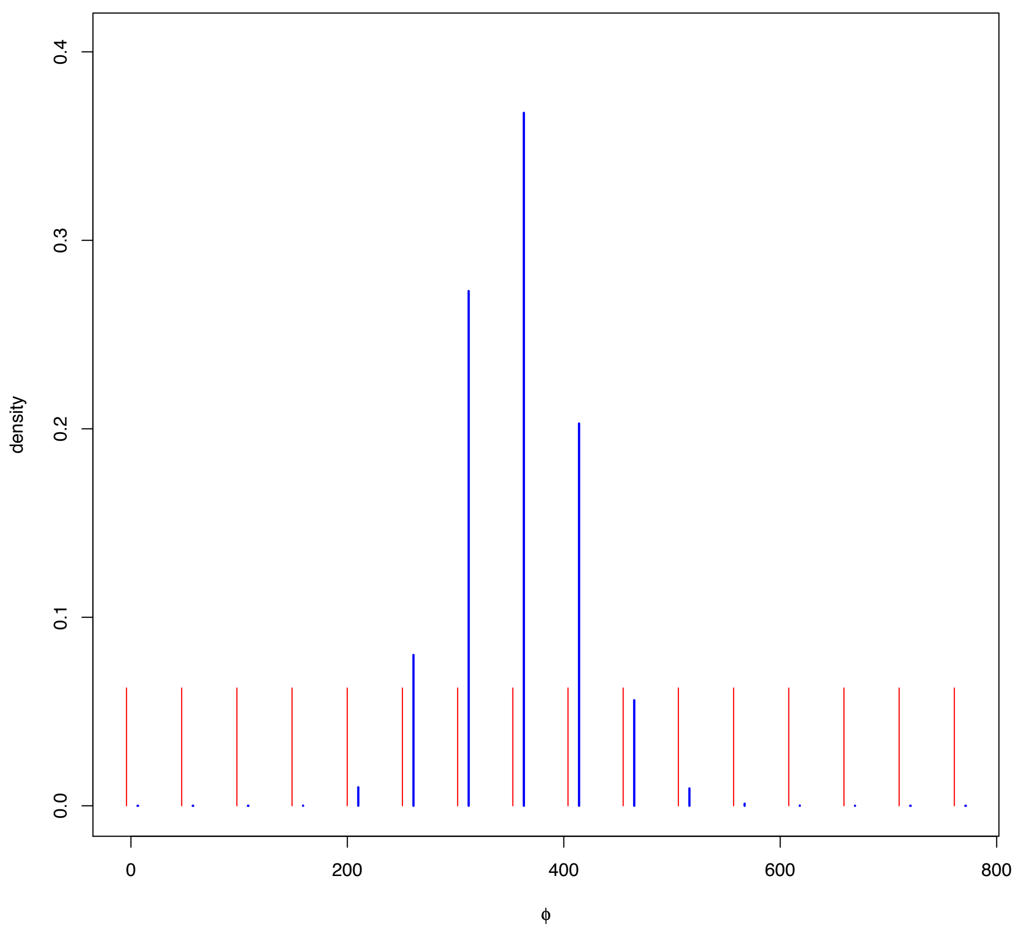
A prior is assigned to τ/σ. Defaults to fixed=0, but can also be uniform or user specified (discrete).

These choices are made for computational reasons. For example, the posterior distribution of σ^2 is inverse chisquared.

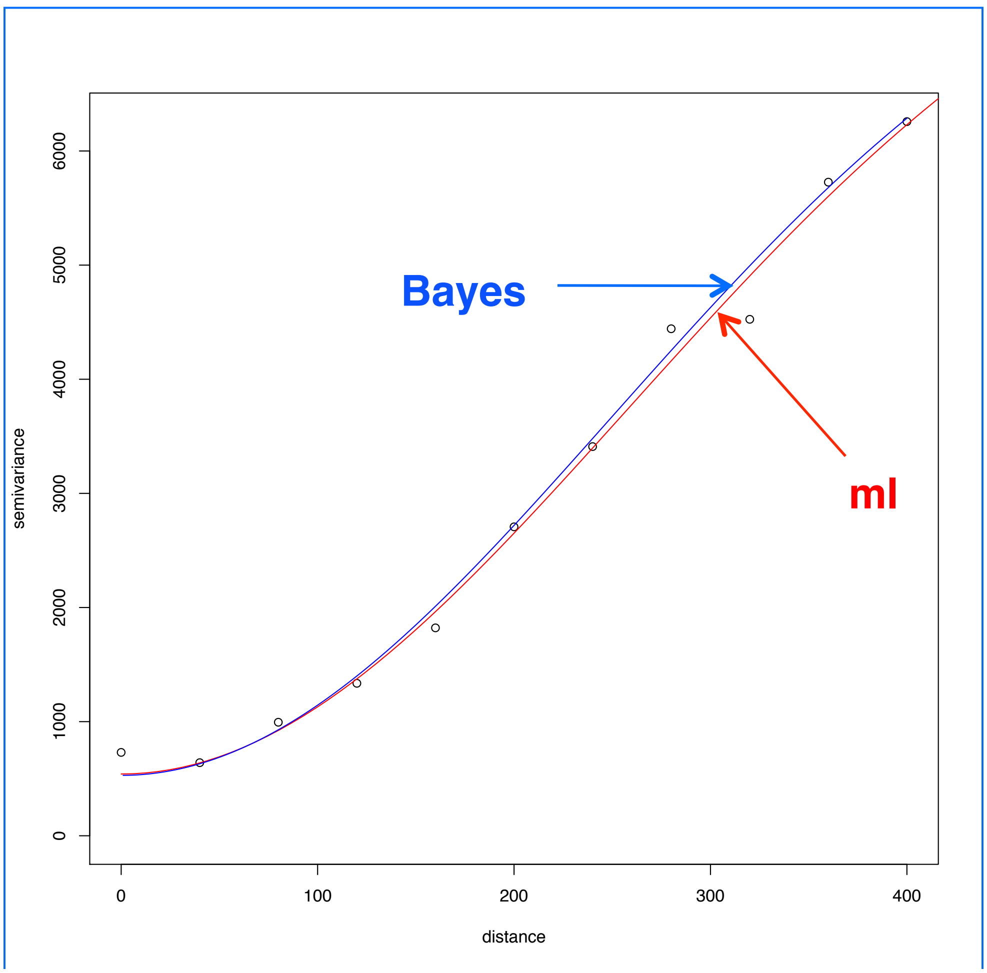
For details, see

http://www.leg.ufpr.br/geoR/geoRdoc/bayeskrige.pdf

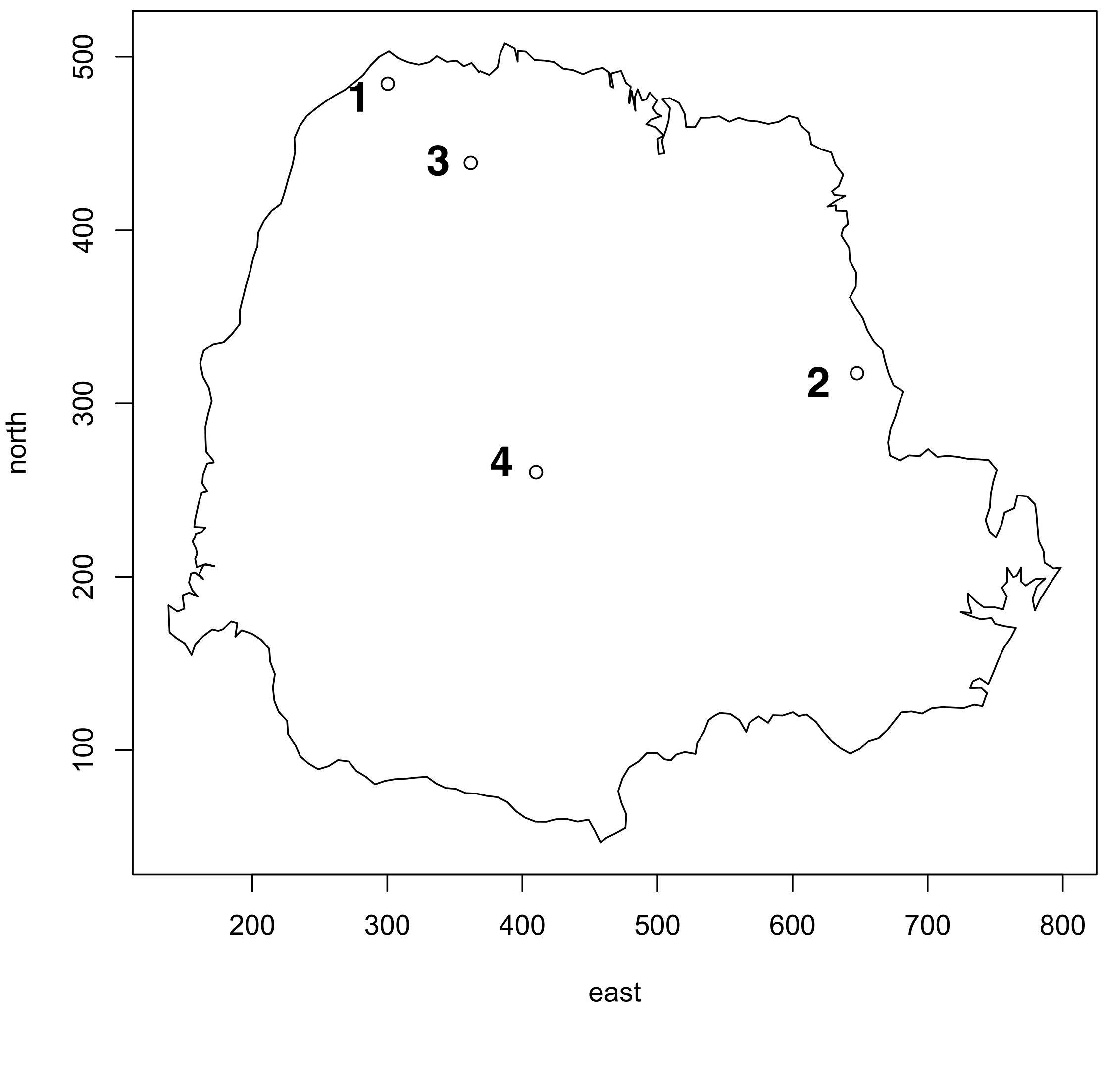
Prior/posterior of ϕ



Estimated variogram



Prediction sites



Predictive distribution