An Introduction to Kriging

STAT 850 Final Project

Nikola Surjanovic

Simon Fraser University

Outline

1. Motivation
2. Gaussian Processes
3. Kriging
4. Application

Motivation

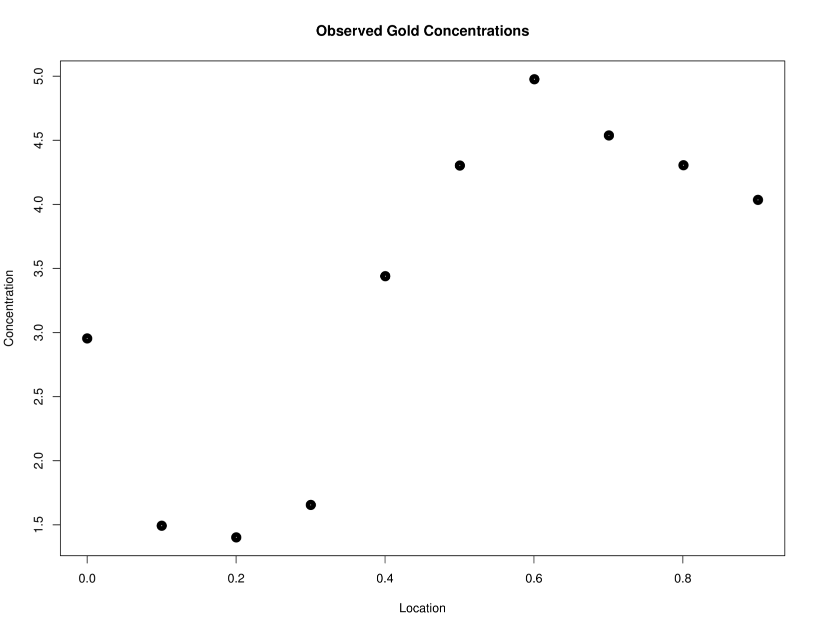
How it Began

Danie G. Krige’s Master’s thesis (“geostatistics”)

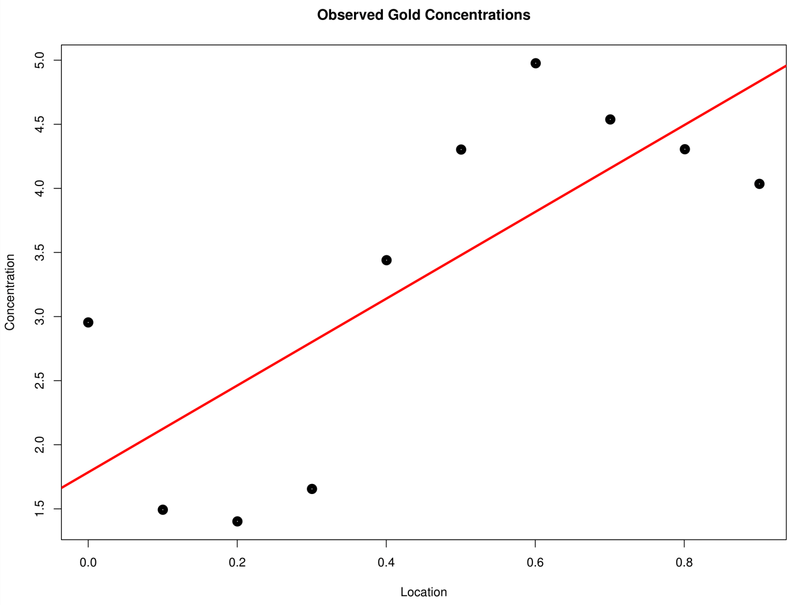
South Africa, Witwatersrand reef complex

Estimate “grade” of gold based on several boreholes

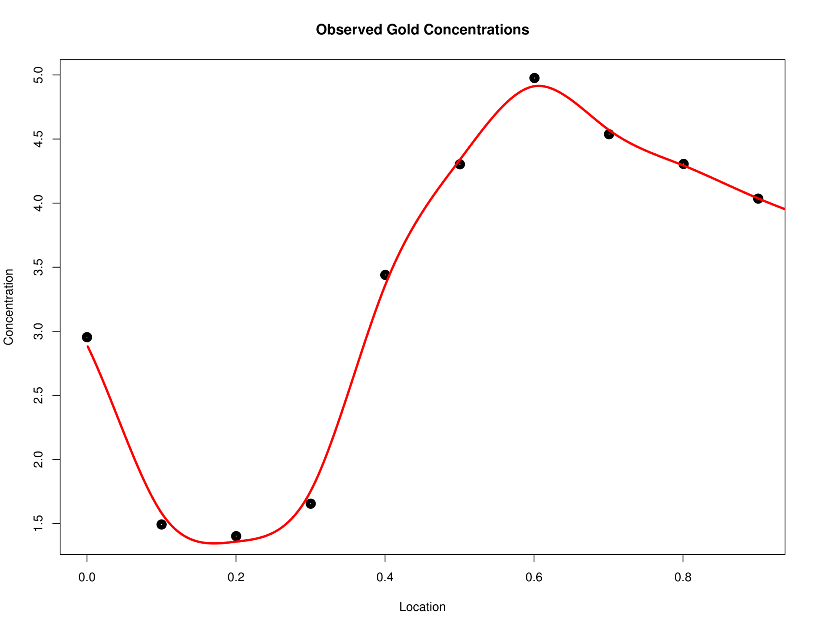
Distribution of Gold



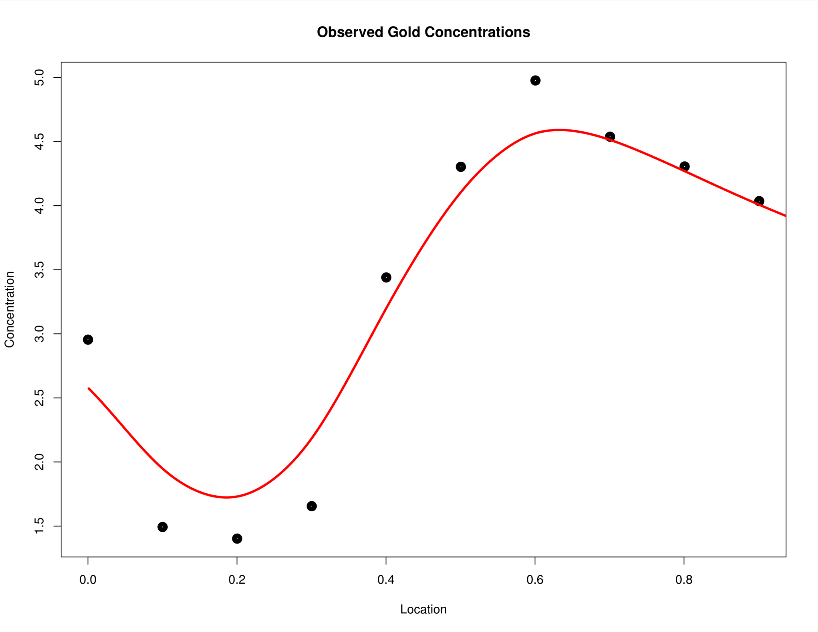
Linear Regression



Kriging Version 1

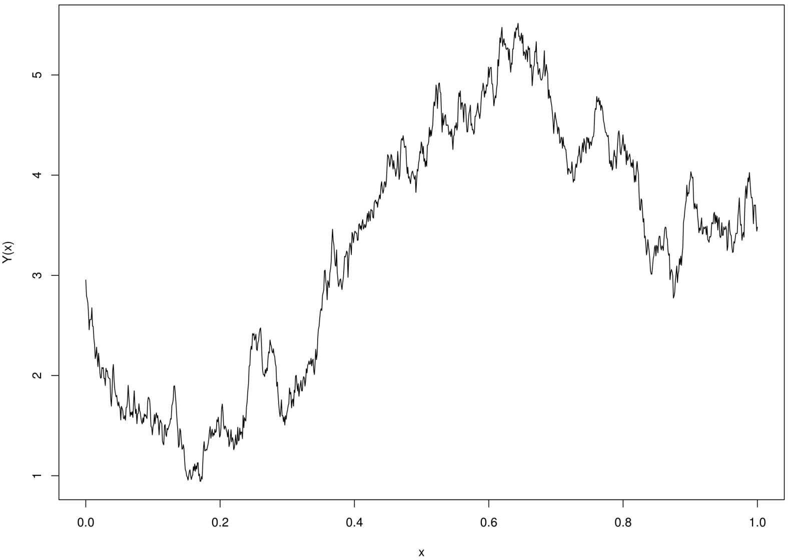


Kriging Version 2

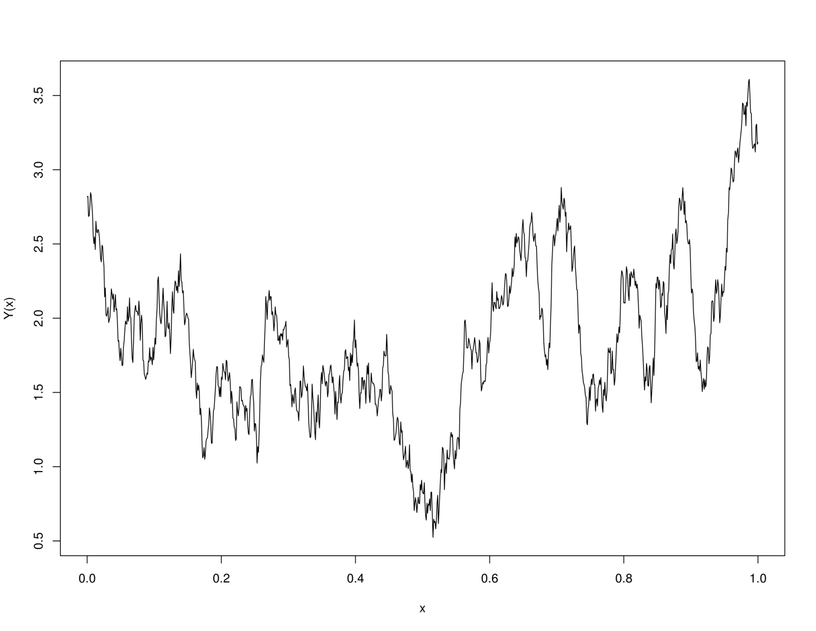


Gaussian Processes

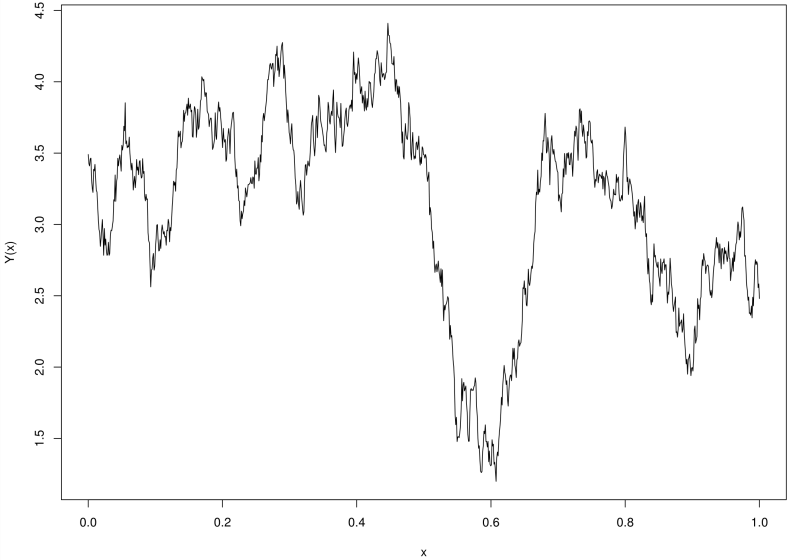
Example of a Stationary Gaussian Process Realization



Example of a Stationary Gaussian Process Realization



Example of a Stationary Gaussian Process Realization



Gaussian Process Definition

Stochastic process: a collection of random variables, $\{X_t : t \in T\}$

Gaussian process: For all finite k, and $t_1, \ldots, t_k \in T$, $(X_{t_1}, \ldots, X_{t_k})$ is multivariate normal

Example: $T = \mathbb{R}$ or $\mathbb{R}^2$

Spatial Gaussian Process Notation

Gaussian spatial process: $\{S(x) : x \in \mathbb{R}^2\}$, with $(S(x_1), \ldots, S(x_n))$ multivariate normal for all $x_1, \ldots, x_n \in \mathbb{R}^2$

Completely specified by:

- Mean function $\mu(x) = \mathsf{E}(S(x))$
- Covariance function $\gamma(x, x') = \mathsf{Cov}(S(x), S(x'))$

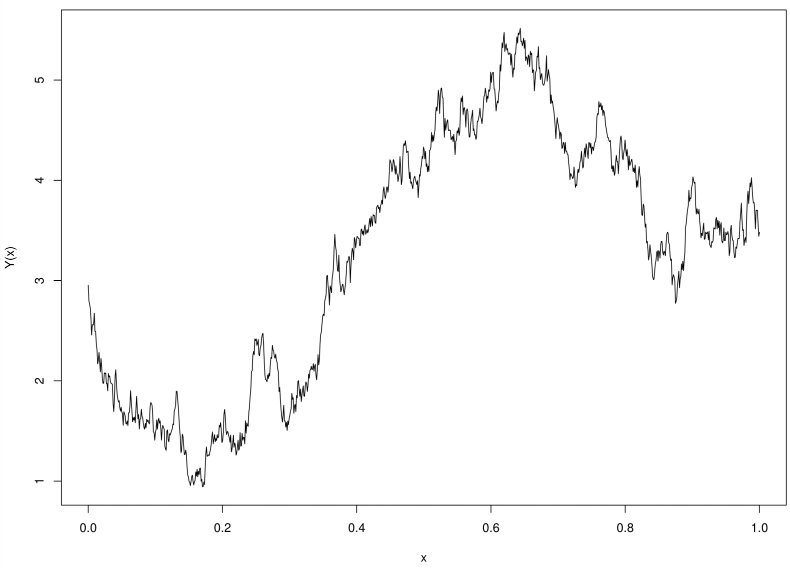
Stationary Gaussian Processes

A Gaussian process is stationary (and isotropic) if:

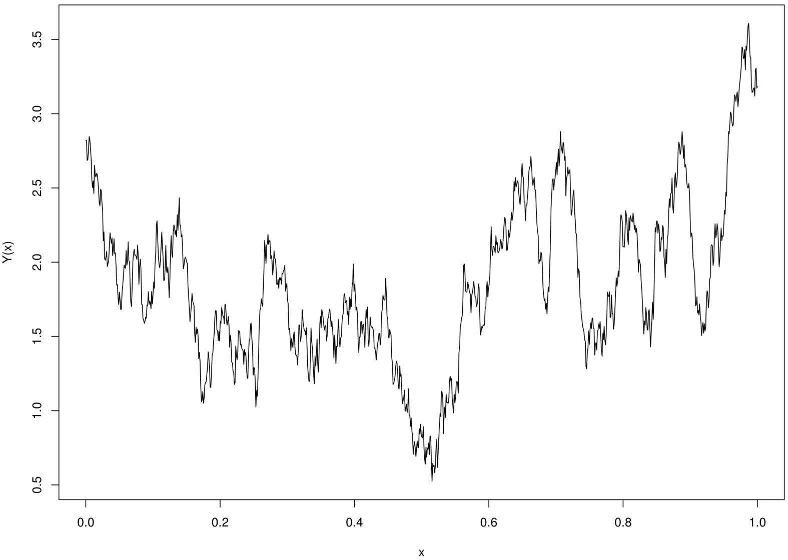
- The mean, $\mu(x) = \mu$, is constant
- The covariance, $\gamma(x, x') = \gamma(x - x') = \gamma(||x - x'||)$, depends only on the (Euclidean) distance between points

Note
Can be smooth!

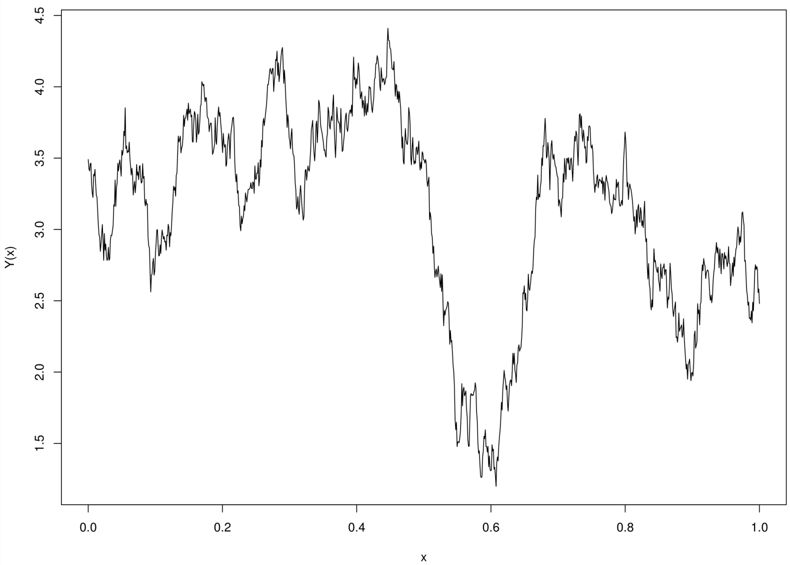
Example of a Stationary Gaussian Process Realization



Example of a Stationary Gaussian Process Realization



Example of a Stationary Gaussian Process Realization



Kriging

An Overview

Observe $y_1, y_2, \ldots, y_n$ (e.g. concentrations of gold) at locations $x_1, x_2, \ldots, x_n$

Idea: What if the distribution of gold is the result of a realization of a Gaussian process?

$$Y_i = S(x_i) + Z_i, \quad i = 1; \ldots, n,$$

where Z_i are i.i.d. $N(0, \tau^2)$

An Overview

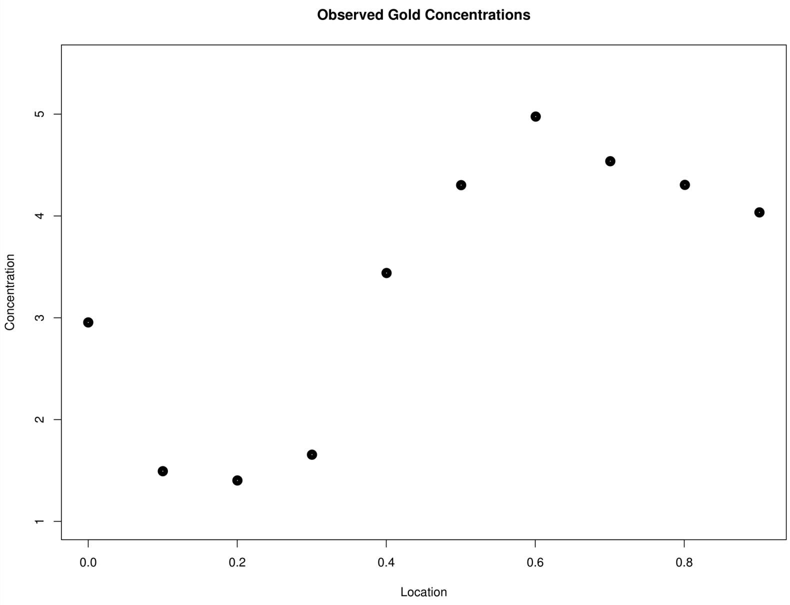
$$Y_i = S(x_i) + Z_i, \quad i = 1; \ldots, n, \quad Z_i \overset{iid}{\sim} N(0, \tau^2)$$

Measurement error: $\tau > 0$ (‘nugget effect’)

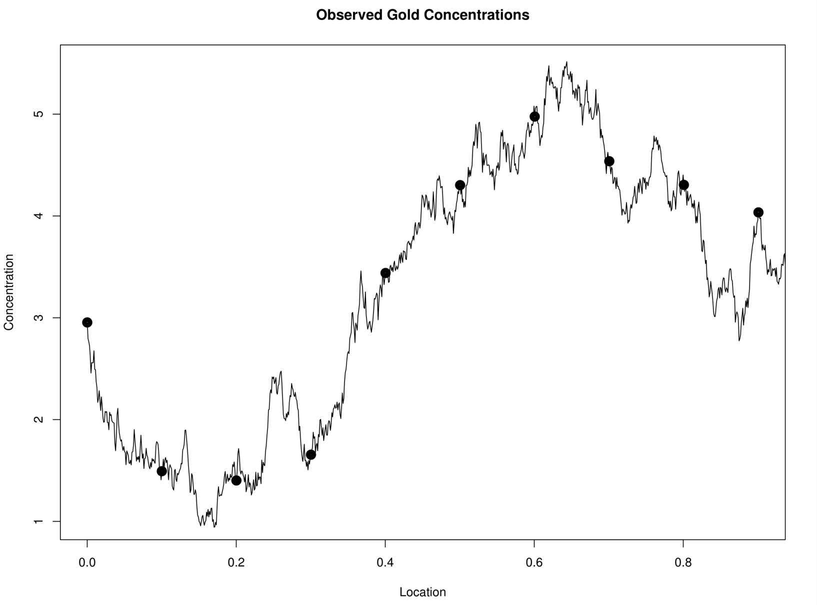
If $\tau = 0$, observed y_i come directly from a Gaussian process

Note
Useful models, but “rarely have any physical justification”!

Distribution of Gold



Distribution of Gold



Stationary (and isotropic): $\mathrm{Corr}(x, x') = \rho(u)$, where $u = ||x - x'||$

If stationary, assume we know μ and $\rho(u)$ for now

How do we predict gold concentrations at unobserved locations?

Making Predictions

How do we predict the gold concentration, $T = S(x)$, at an unobserved location x?

We have: $y = (y_1, \ldots, y_n)$. We want: $\widehat{T} = t(Y)$

Minimize "mean square prediction error": $\mathrm{E}\left((\widehat{T} - T)^2\right)$

Note
T, $\widehat{T}$, and Y are all random variables

Making Predictions

Minimize "mean square prediction error": $\mathsf{E}\left((\widehat{T} - T)^2\right)$

Use $\widehat{T} = \mathsf{E}(T|Y)$

Since $Y_i = S(x_i) + Z_i, \quad i = 1;\ldots,n, \quad Z_i \overset{iid}{\sim} N(0,\tau^2)$
$\implies$ Y is multivariate normal

Making Predictions

$$Y \sim N(\mu\mathbf{1}, \sigma^2 V),$$

where $\mathbf{1}$ is a vector of ones, and

$$\sigma^2 V = \sigma^2 R + \tau^2 I,$$

with I the $n \times n$ identity matrix, $R_{ij} = \rho(||x_i - x_j||)$, and $\sigma^2 = \gamma(0)$

Making Predictions

Similarly, (T, Y) is multivariate normal

Basic multivariate normal result, if $T = S(x)$:

$$E(T|Y) = \widehat{T} = \mu + r'V^{-1}(Y - \mu\mathbf{1}),$$

where $r' = (\rho(||x - x_1||), \ldots, \rho(||x - x_n||))$

Note
$\widehat{T}$ is linear in Y!

FALL 2019 - STAT 850 G100

LINEAR MODELS AND APPLICATIONS (4)

Class Number: 4607   *Delivery Method: In Person*

Simple vs. Ordinary Kriging

Simple Kriging

$$\widehat{T} = \mu + r'V^{-1}(Y - \mu\mathbf{1}), \text{ or } \widehat{T} = \hat{\mu} + r'V^{-1}(Y - \hat{\mu}\mathbf{1}),$$

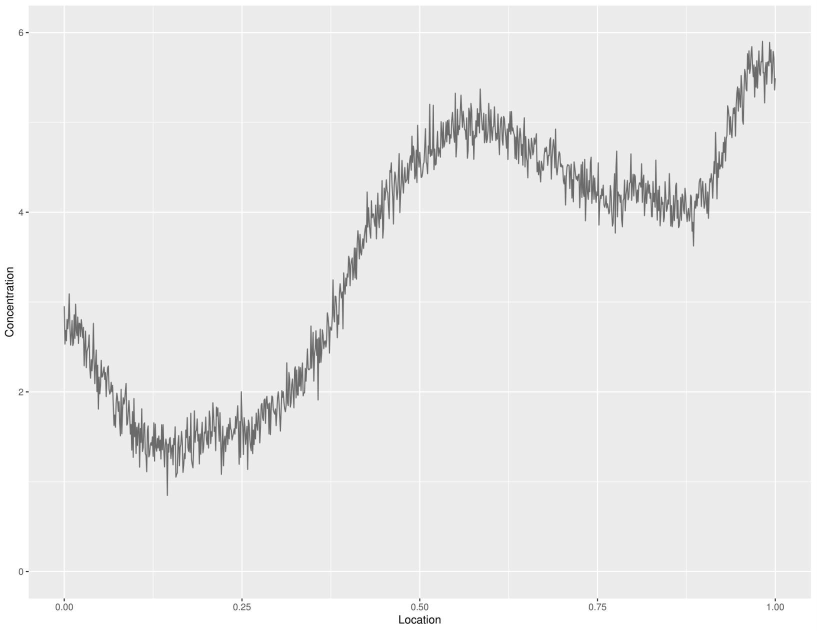
where $\hat{\mu} = \bar{Y}$

Ordinary Kriging

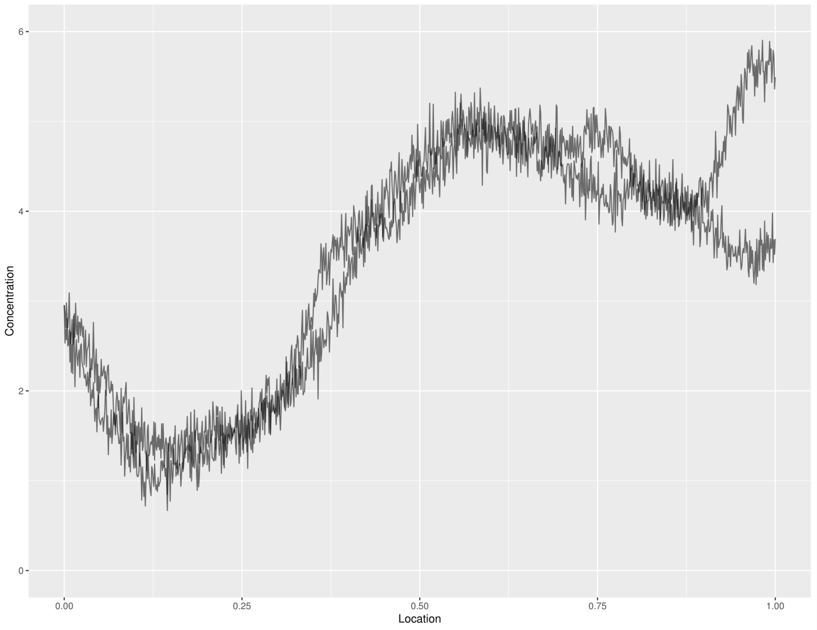
$$\widehat{T} = \hat{\mu} + r'V^{-1}(Y - \hat{\mu}\mathbf{1}),$$

where $\hat{\mu} = (\mathbf{1}'V^{-1}\mathbf{1})^{-1}\mathbf{1}'V^{-1}Y$ (generalized least squares)

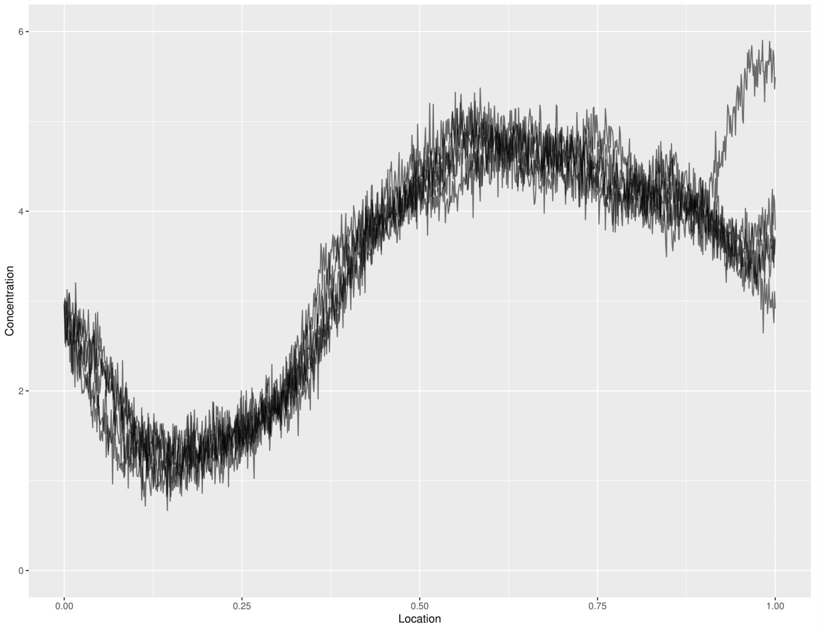
Visual Interpretation



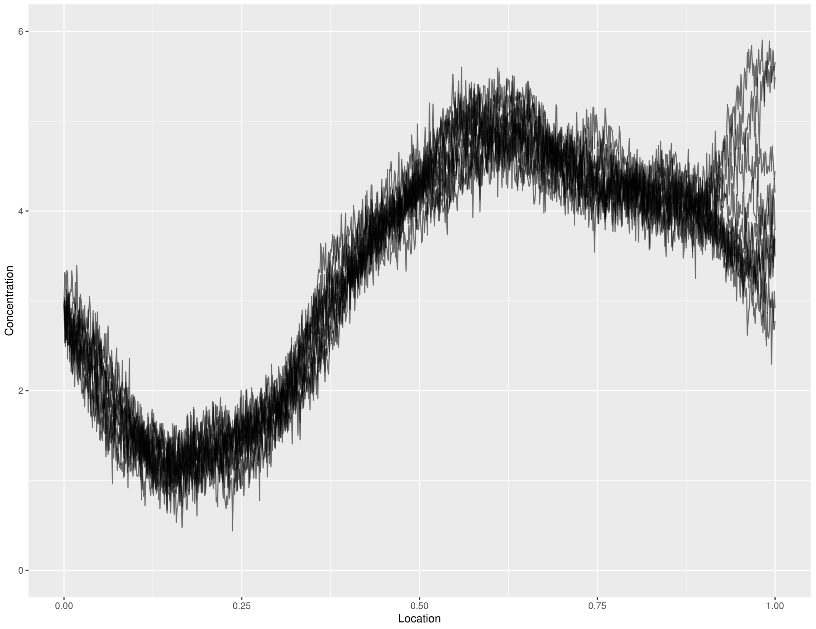
Visual Interpretation



Visual Interpretation



Visual Interpretation



Common Correlation Functions

Many families of correlation functions exist

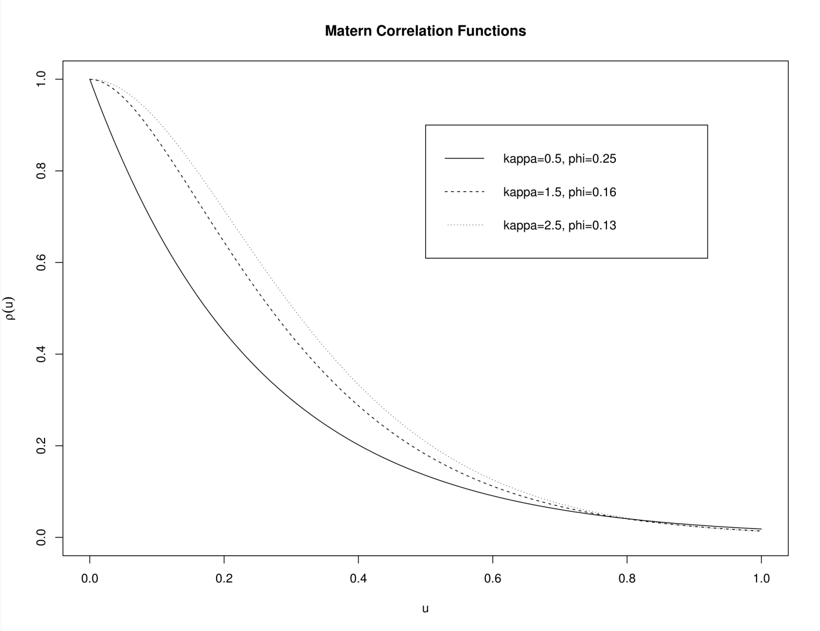
E.g. **Matérn correlation functions** (2 parameters: κ, ϕ)

$$\rho(u) = \{2^{\kappa-1}\Gamma(\kappa)\}^{-1}(u/\phi)^{\kappa} K_{\kappa}(u/\phi),$$

where K_{κ} is an order κ modified Bessel function

Choice of κ affects differentiability of Gaussian process

Matérn Correlation Functions



Estimation

Variograms can be used for estimation of $\rho(\cdot)$

Less subjective: maximum likelihood estimation of $\kappa, \phi, \tau^2, \sigma^2$

- Recall: Y is multivariate normal

Other considerations:

- Dealing with trends
- Trans-Gaussian kriging
- Non-linear target prediction

Application

Zinc Concentration Dataset

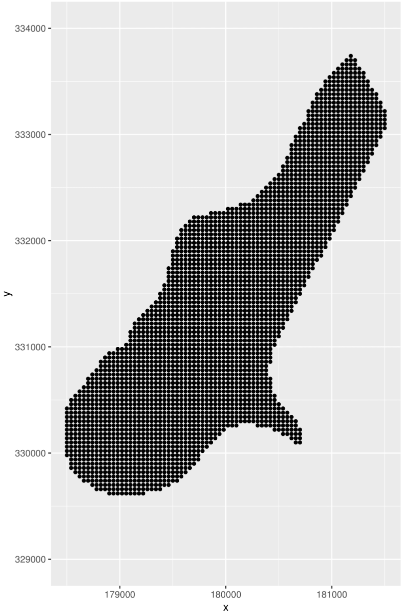
Goal: Predict concentrations of zinc

Location: Meuse river, data from the Netherlands

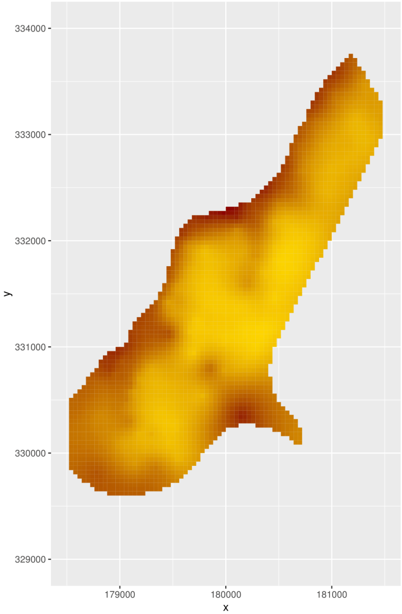
(Adapted from RPubs - An Introduction to Kriging in R by Nabil A.)

Concentration of Zinc

Points to be Evaluated



2D Kriging Predictions



References

Introduction to Kriging in R.
`https://rpubs.com/nabilabd/118172`. Accessed: 2019-11-13.

Noel Cressie. *Statistics for Spatial Data, Revised Edition*. Wiley. ISBN 978-0-471-00255-0.

Peter Diggle and Paulo Justiniano Ribeiro. *Model-based Geostatistics*. Springer. ISBN 978-0-387-48536-2.

Thank You!

Questions?

“Nugget” Explanation

Nugget explanation: measurement error or spatial variation on a small scale

Like modelling $Y(x)$ as a GP with $\rho(u)$ discontinuous at $u = 0$.

Differentiability and Continuity

Theorem 3.1 (Diggle and Ribeiro): "A stationary stochastic process with correlation function $\rho(u)$ is k times mean-square differentiable if and only if $\rho(u)$ is $2k$ times differentiable at $u = 0$."

Variograms

$$V(x, x') := \frac{1}{2}(S(x) - S(x')) \implies V(u) = \sigma^2(1 - \rho(u))$$

$$V_Y(u) = \tau^2 + \sigma^2(1 - \rho(u))$$