Universiteit Utrecht

Stochastic Hydrology

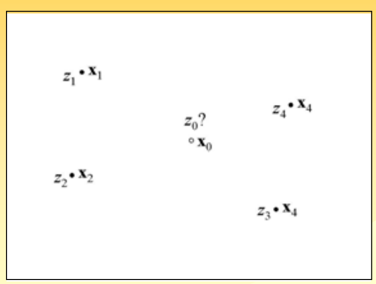
Geostatistics (1)

Marc F.P. Bierkens

Professor of Hydrology
Faculty of Geosciences

Universiteit Utrecht

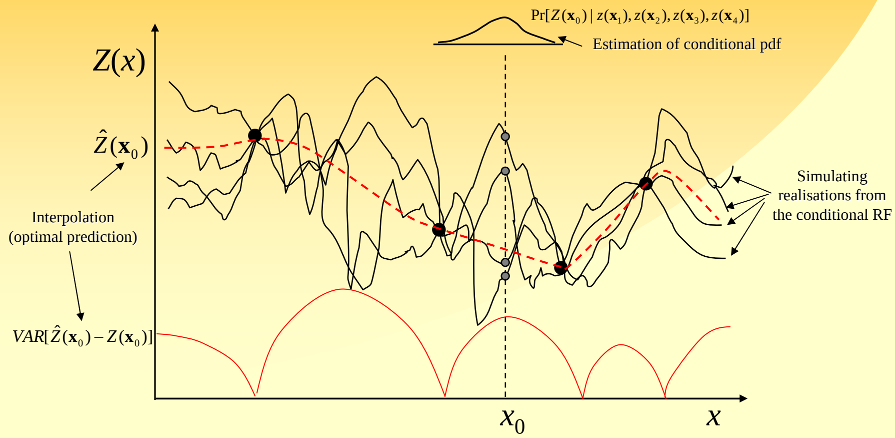
The focus of geostatistics



- Interpolation: prediction $\hat{Z}(\mathbf{x}_0)$ and prediction variance $VAR[\hat{Z}(\mathbf{x}_0) - Z(\mathbf{x}_0)]$
- Estimation of the conditional probability: $\Pr[Z(\mathbf{x}_0) \mid z(\mathbf{x}_1), z(\mathbf{x}_2), z(\mathbf{x}_3), z(\mathbf{x}_4)]$
- Simulation of realisations of the conditional RF $Z(\mathbf{x} \mid z(\mathbf{x}_1), z(\mathbf{x}_2), z(\mathbf{x}_3), z(\mathbf{x}_4))$

Universiteit Utrecht

The focus of geostatistics

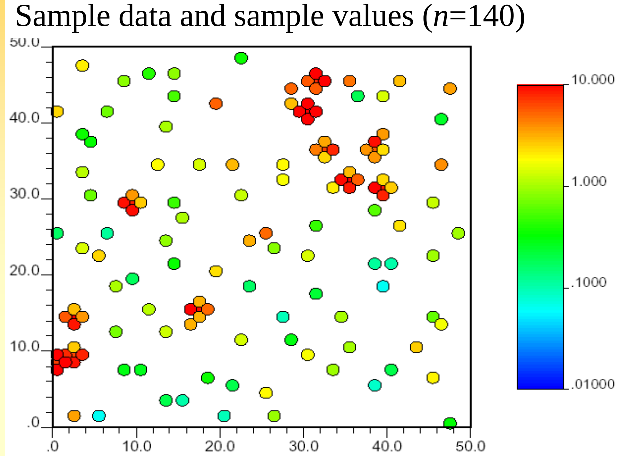


Universiteit Utrecht

Spatial descriptive statistics

Walker Lake data set

Sample data and sample values (n=140)



Universiteit Utrecht

Spatial descriptive statistics

Polygon declustering

$$m_z = \sum_{i=1}^{n} w_i z_i$$

$$\sigma_Z^2 = \sum_{i=1}^{n} w_i (z_i - m_z)^2$$

$$w_1 = \frac{A_1}{A_1 + A_2 + A_3 + A_4 + A_5}$$

$$w_2 = \frac{A_2}{A_1 + A_2 + A_3 + A_4 + A_5}$$

etc.

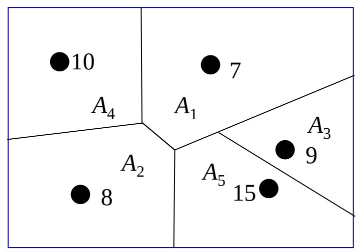
A_1= 0.30

A_2= 0.25

A_3= 0.10

A_4= 0.20

A_5= 0.15



$$m_z = 0.3 \cdot 7 + 0.25 \cdot 8 + 0.10 \cdot 9 + 0.20 \cdot 10 + 0.15 \cdot 15 = 9.25$$

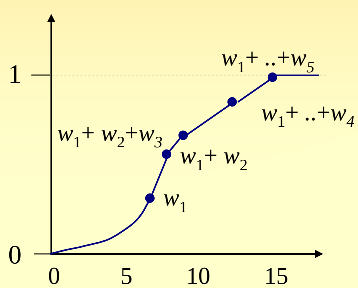
$$s_z^2 = 0.3 \cdot (-2.25)^2 + 0.25 \cdot (-1.25)^2 + 0.10 \cdot (-0.25)^2 + 0.20 \cdot (0.75)^2 + 0.15 \cdot (5.75)^2 = 6.99$$

Universiteit Utrecht

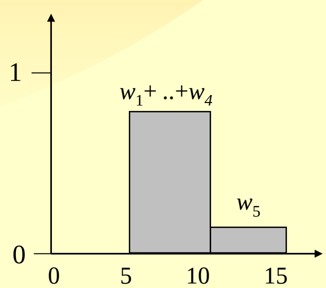
Spatial descriptive statistics

Declustered histogram and cumulative frequency distribution

Declustered cum. Freq. Distr.

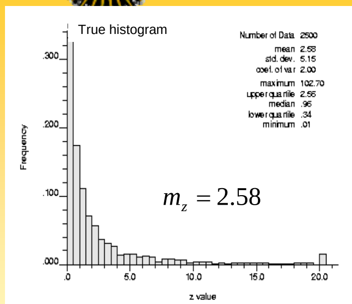


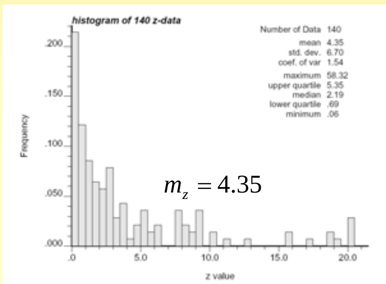
Declusterd histogram

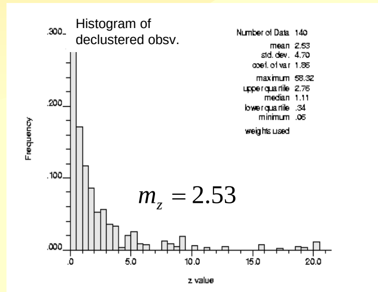


Universiteit Utrecht

Spatial descriptive statistics



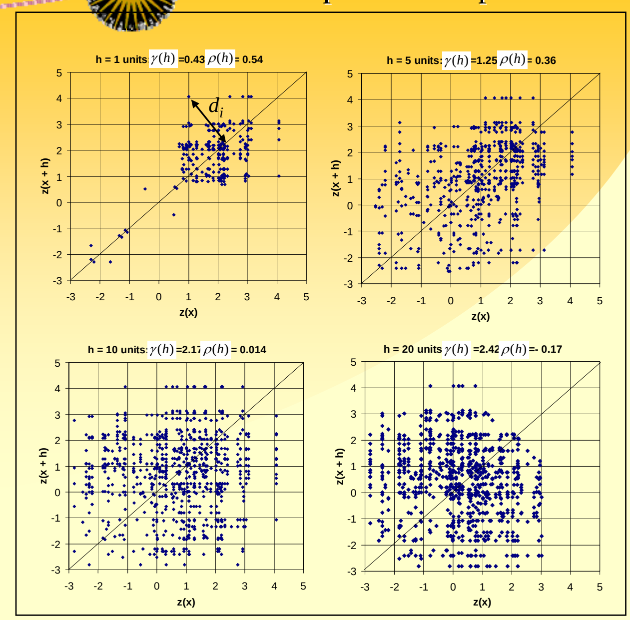




Universiteit Utrecht

Spatial descriptive statistics

Semivariance



Universiteit Utrecht

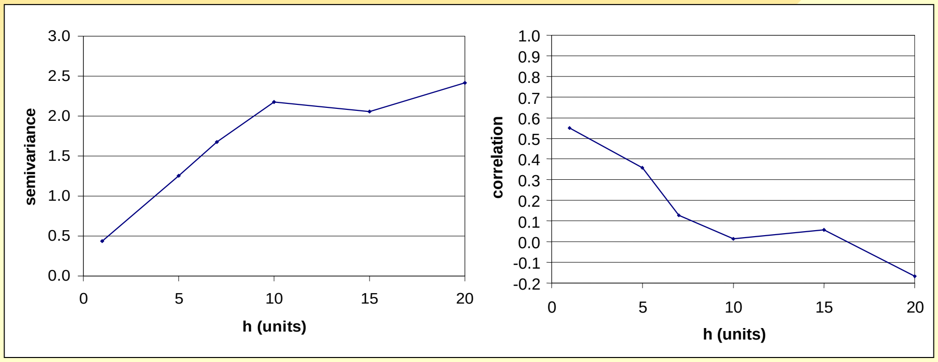
Spatial descriptive statistics

Semivariogram
Measure of difference
as function of distance

$$\gamma(\mathbf{h}) = \frac{1}{2N}\sum_{i=1}^{N} d_i^2(\mathbf{h}) = \frac{1}{2N}\sum_{i=1}^{N}\left[z(\mathbf{x}) - z(\mathbf{x}+\mathbf{h})\right]^2$$

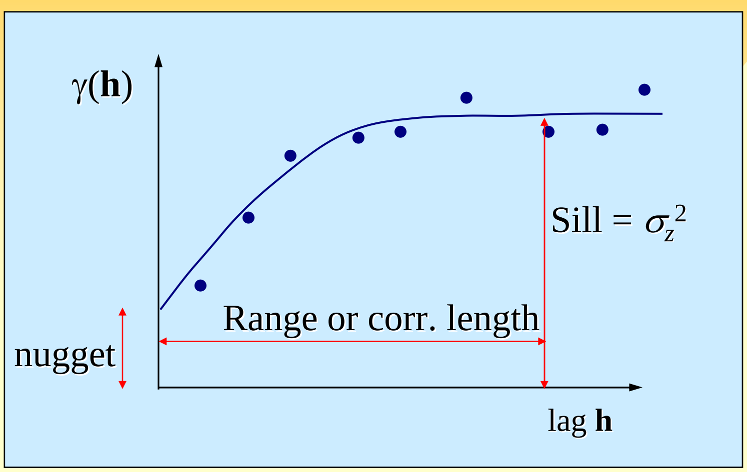
Correlogram
Meassure of similarity
as a function of distance

$$\rho(\mathbf{h}) = \frac{\frac{1}{N}\sum_{i=1}^{N} z(\mathbf{x})z(\mathbf{x}+\mathbf{h}) - m_{z(\mathbf{x})}m_{z(\mathbf{x}+\mathbf{h})}}{s_{z(\mathbf{x})}s_{z(\mathbf{x}+\mathbf{h})}}$$



Universiteit Utrecht

Spatial descriptive statistics



Universiteit Utrecht

Interpolation by Kriging

Simple Kriging

Wide sense stationary RF:

$\mu_z(\mathbf{x}) = \mu_z$ (constant)

Variance σ_Z^2 is constant and finite

$\mathrm{Cov}(Z(\mathbf{x}_1), Z(\mathbf{x}_2)) = C_Z(\mathbf{x}_1 - \mathbf{x}_2) = C_Z(\mathbf{h})$

Simple Kriging if:

- RF is wide sense stationary, and
- mean μ_z is known.

Universiteit Utrecht

Interpolation by Kriging

Simple Kriging

Linear Predictor:

$$\hat{Z}(\mathbf{x}_0) = \mu_Z + \sum_{i=1}^{n} \lambda_i (Z_i - \mu_Z)$$

Unbiased prediction means: $E[\hat{Z}(\mathbf{x}_0)] = E[Z(\mathbf{x}_0)]$

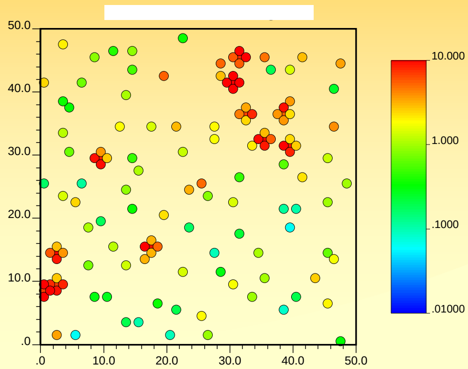
Simple kriging predictor is by definition unbiased!

Find the λ_i such that the error variance: $VAR[\hat{Z}(\mathbf{x}_0) - \hat{Z}(\mathbf{x}_0)]$ is minimal!

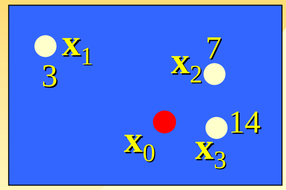
Universiteit Utrecht

Kriging

Simple kriging – step by step



Example data set and a trivial example



	0	1	2	3
0	0	3	1	1
1	3	0	3	4
2	1	3	0	1
3	1	4	1	0

Universiteit Utrecht

Simple kriging – step by step

Step 1: Estimate mean from the data

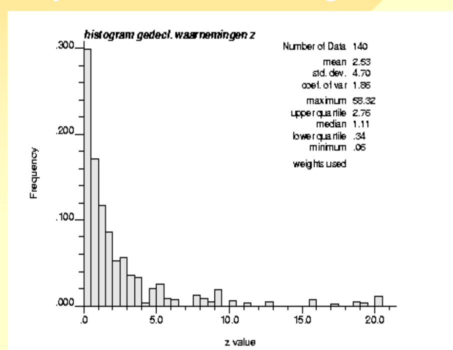
If locations have not been randomly chosen: declustering is advised

$$\hat{\mu}_z = m_z = \sum_{i=1}^{n} w_i z_i$$

Dataset: $\hat{\mu}_z = 2.53$

Trivial:

$$\hat{\mu}_z = 0.5 \cdot 3 + 0.2 \cdot 7 + 0.3 \cdot 14 = 7.1$$

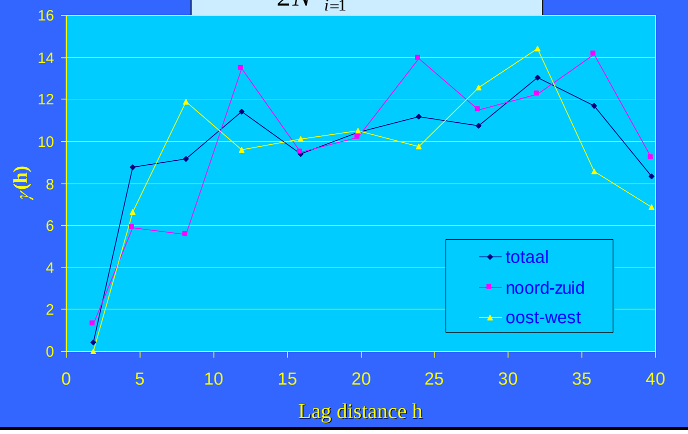


Universiteit Utrecht

Simple kriging – step by step

Step 2: Estimate the semivariogram

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N}\sum_{i=1}^{N}\left[z(\mathbf{x}) - z(\mathbf{x}+\mathbf{h})\right]^2$$

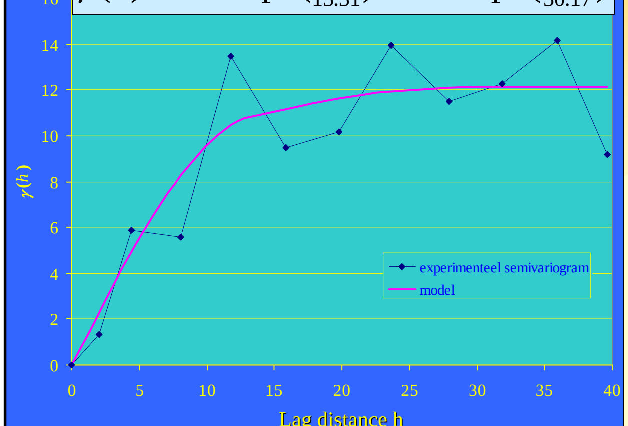


Universiteit Utrecht

Simple kriging – step by step

Step 3: Fit a semivariogram model

$$\gamma(h) = 8.73\,\mathrm{sph}(\tfrac{h}{13.31}) + 3.42\,\mathrm{sph}(\tfrac{h}{30.17})$$



Trivial example

$$\gamma(h) = 22.69\left[1 - e^{-\left(\frac{h}{2}\right)}\right]$$

Note:

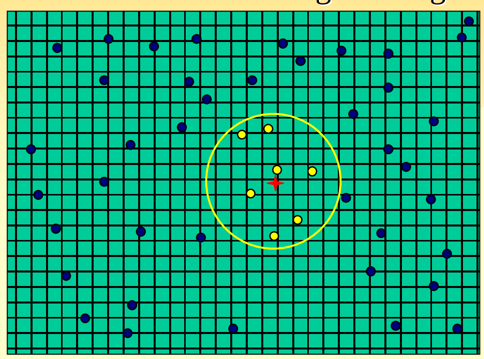
$$C_Z(h) = \sigma_Z^2 - \gamma_Z(h)$$

Universiteit Utrecht

Simple kriging – step by step

Step 4: Select observations within a search radius

Based on semivariogram range



Note: this step and all the next ones are repeated for every prediction location

Universiteit Utrecht

Simple kriging – step by step

Step 5: Solve the kriging equations for λ_i

$$\sum_{j=1}^{n} \lambda_j C(\mathbf{x}_i - \mathbf{x}_j) = C(\mathbf{x}_i - \mathbf{x}_0) \qquad i = 1,..,n$$

with $C(\mathbf{x}_i - \mathbf{x}_j) = \sigma_z^2 - \gamma(\mathbf{x}_i - \mathbf{x}_j)$

Trivial:

$$\lambda_1 C(\mathbf{x}_1 - \mathbf{x}_1) + \lambda_2 C(\mathbf{x}_1 - \mathbf{x}_2) + \lambda_3 C(\mathbf{x}_1 - \mathbf{x}_3) = C(\mathbf{x}_1 - \mathbf{x}_0)$$
$$\lambda_1 C(\mathbf{x}_2 - \mathbf{x}_1) + \lambda_2 C(\mathbf{x}_2 - \mathbf{x}_2) + \lambda_3 C(\mathbf{x}_2 - \mathbf{x}_3) = C(\mathbf{x}_2 - \mathbf{x}_0)$$
$$\lambda_1 C(\mathbf{x}_3 - \mathbf{x}_1) + \lambda_2 C(\mathbf{x}_3 - \mathbf{x}_2) + \lambda_3 C(\mathbf{x}_3 - \mathbf{x}_3) = C(\mathbf{x}_3 - \mathbf{x}_0)$$

Met $C(\mathbf{x}_i - \mathbf{x}_j) = 22.69 e^{-\frac{|\mathbf{x}_i - \mathbf{x}_j|}{2}}$

	0	1	2	3
0	0	3	1	1
1	3	0	3	4
2	1	3	0	1
3	1	4	1	0

Universiteit Utrecht

Simple kriging – step by step

Step 5: Solve the kriging equation for λ_i

Trivial:

$$\begin{aligned} 22.69\lambda_1 + 5.063\lambda_2 + 3.071\lambda_3 &= 5.063 \\ 5.063\lambda_1 + 22.69\lambda_2 + 13.76\lambda_3 &= 13.76 \\ 3.071\lambda_1 + 13.76\lambda_2 + 22.69\lambda_3 &= 13.76 \end{aligned}$$

Solution:

$$\lambda_1 = 0.0924 \qquad \lambda_2 = 0.357 \qquad \lambda_3 = 0.378$$

Universiteit Utrecht

Simple kriging – step by step

Step 6: Kriging prediction

$$\hat{Z}(\mathbf{x}) = \mu_z + \sum_{i=1}^{n} \lambda_i \left[Z(\mathbf{x}_i) - \mu_z \right]$$

Trivial:

$$\hat{Z}(\mathbf{x}) = 7.1 + 0.0924 \cdot (-4.1) + 0.357 \cdot (-0.1) + 0.378 \cdot 6.9 = 9.29$$

Universiteit Utrecht

Simple kriging – step by step

Step 7: Kriging variance

$$\mathrm{var}[\,\hat{Z}(\mathbf{x}_0) - Z(\mathbf{x}_0)\,] = \sigma_z^2 - \sum_{i=1}^{n} \lambda_i C(\mathbf{x}_i - \mathbf{x}_0)$$

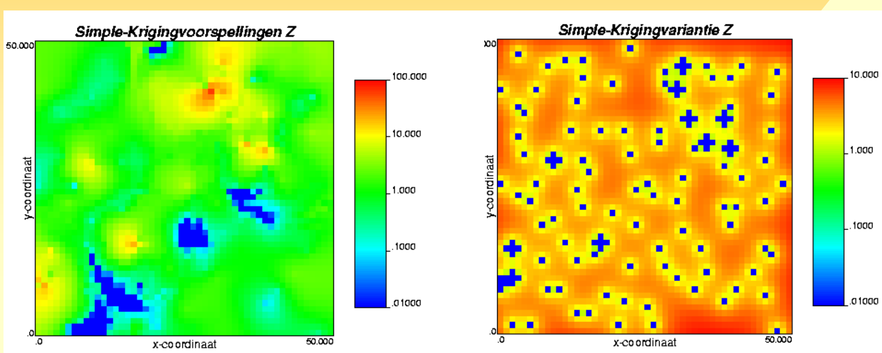
Trivial:

$\hat{Z}(\mathbf{x}_0) - Z(\mathbf{x}_0)$ 2˙0e3 + 0˙327 · 13˙7e + 0˙378 · 13˙7e =12˙11

var[Σ*(x0) − Σ*(x0)] =

Universiteit Utrecht

Kriging prediction and kriging variance



Universiteit Utrecht

Ordinary kriging

Intrinsic random function:

- Constant mean $\mu_z(\mathbf{x}) = \mu_z$
- Semivariogram a function of separation only

$$\gamma(Z(\mathbf{x}_1), Z(\mathbf{x}_2)) = \gamma(\mathbf{x}_1 - \mathbf{x}_2) = \gamma(\mathbf{h})$$

Universiteit Utrecht

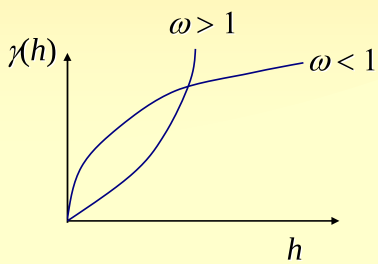
Ordinary kriging

Wide sense stationary RF: semivariogram has sill: $C_Z(h) = \sigma_Z^2 - \gamma_Z(h)$

Intrinsic RF: semivariogram does not have to have a sill

Example: power semivariogram

$$\gamma(h) = ch^{\omega} \quad 0 \le \omega < 2$$



- Wide sense stationary RF is also intrinsic, but not vice versa;
- An instrinsic RF is also wide sense stationary if the semivariogram has a sill.

Ordinary kriging

Ordinary Kriging is applied if:

- an instrinsic RF is assumed, or
- a wide sense stationary RF is assumed, but the mean μ_Z is not known

Universiteit Utrecht

Ordinary kriging: step by step

Step 1: Estimate the semivariogram

Step 2: Fit a semivariogram model

$$\gamma(h) = 8.73\,\mathrm{sph}\left(\tfrac{h}{13.31}\right) + 3.42\,\mathrm{sph}\left(\tfrac{h}{30.17}\right)$$

For trivia:

$$\gamma(h) = 22.69\left[1 - e^{-\left(\frac{h}{2}\right)}\right]$$

*Step 3: Select observations with search radius**

*This and next steps repeated for each prediction location

Universiteit Utrecht

Ordinary kriging: step by step

Step 4: Solve kriging system

$$\begin{cases} \sum_{j=1}^{n} \lambda_j \gamma(\mathbf{x}_i - \mathbf{x}_j) + \nu = \gamma(\mathbf{x}_i - \mathbf{x}_0) \quad i = 1,..,n \\ \sum_{i=1}^{n} \lambda_i = 1 \end{cases}$$

Trivial:

$$\begin{aligned} &\lambda_1\gamma(\mathbf{x}_1-\mathbf{x}_1)+\lambda_2\gamma(\mathbf{x}_1-\mathbf{x}_2)+\lambda_3\gamma(\mathbf{x}_1-\mathbf{x}_3)+\nu=\gamma(\mathbf{x}_1-\mathbf{x}_0)\\ &\lambda_1\gamma(\mathbf{x}_2-\mathbf{x}_1)+\lambda_2\gamma(\mathbf{x}_2-\mathbf{x}_2)+\lambda_3\gamma(\mathbf{x}_2-\mathbf{x}_3)+\nu=\gamma(\mathbf{x}_2-\mathbf{x}_0)\\ &\lambda_1\gamma(\mathbf{x}_3-\mathbf{x}_1)+\lambda_2\gamma(\mathbf{x}_3-\mathbf{x}_2)+\lambda_3\gamma(\mathbf{x}_3-\mathbf{x}_3)+\nu=\gamma(\mathbf{x}_3-\mathbf{x}_0)\\ &\lambda_1 + \lambda_2 + \lambda_3 = 1 \end{aligned}$$

Met $\gamma(\mathbf{x}_i - \mathbf{x}_j) = 22.69(1 - e^{-\frac{|\mathbf{x}_i - \mathbf{x}_j|}{2}})$

	0	1	2	3
0	0	3	1	1
1	3	0	3	4
2	1	3	0	1
3	1	4	1	0

Universiteit Utrecht

Ordinary kriging: step by step

Step 4: Solve the kriging system

Trivial:

$$\begin{aligned} &17.627\lambda_2 + 19.619\lambda_3 + \nu = 17.628\\ &16.627\lambda_1 + 8.930\lambda_3 + \nu = 8.930\\ &19.619\lambda_1 + 8.930\lambda_2 + \nu = 8.930\\ &\lambda_1 + \lambda_2 + \lambda_3 = 1 \end{aligned}$$

Solution:

$$\lambda_1 = 0.172 \quad \lambda_2 = 0.381 \quad \lambda_3 = 0.447 \quad \nu = 2.147$$

Universiteit Utrecht

Ordinary kriging: step by step

Step 5: Kriging prediction

$$\hat{Z}(\mathbf{x}_0) = \sum_{i=1}^{n} \lambda_i Z(\mathbf{x}_i)$$

Trivial:

$$\hat{Z}(\mathbf{x}_0) = 0.172 \cdot 3 + 0.381 \cdot 7 + 0.447 \cdot 14 = 9.44$$

Universiteit Utrecht

Ordinary kriging: step by step

Step 6: Kriging variance

$$\text{var}[\ \hat{Z}(\mathbf{x}_0) - Z(\mathbf{x}_0)] = \sum_{i=1}^{n} \lambda_i \gamma(\mathbf{x}_i - \mathbf{x}_0) + \nu$$

Trivial:

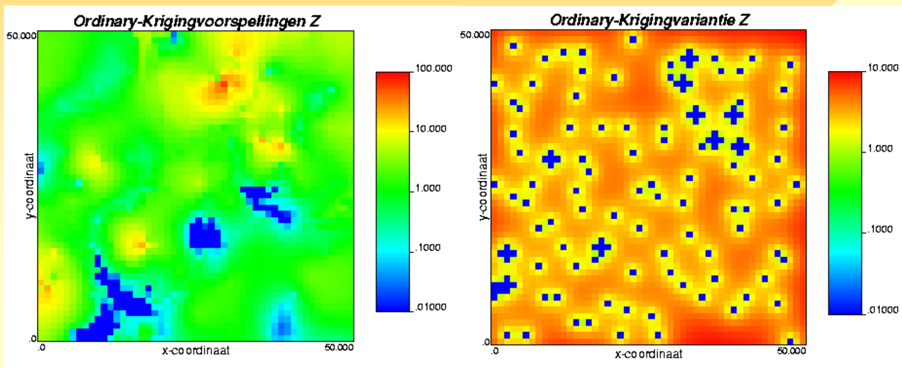
$$\hat{Z}\,172 \cdot 17.627 + 0.381 \cdot 8.930 + 0.447 \cdot 8.930 + 2.147 = 12.57$$

$$\text{var}[Z^*(\mathbf{x}_0) - Z^*(\mathbf{x}_0)] =$$

Universiteit Utrecht

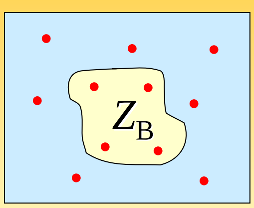
Ordinary kriging: step by step

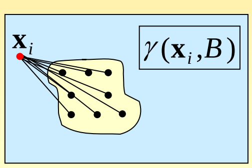
Kriging prediction and kriging variance

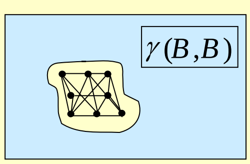


Universiteit Utrecht

(Ordinary) block kriging







$$\hat{Z}_B = \sum_{i=1}^{n} \lambda_i Z(\mathbf{x}_i)$$

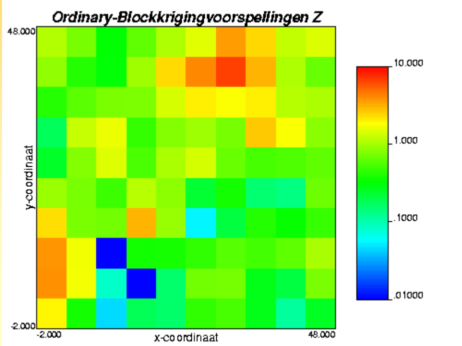
$$\begin{cases} \sum_{j=1}^{n} \lambda_j \, \gamma(\mathbf{x}_i - \mathbf{x}_j) + \nu = \gamma(\mathbf{x}_i, B) & i = 1,..,n \\ \sum_{i=1}^{n} \lambda_i = 1 & \end{cases}$$

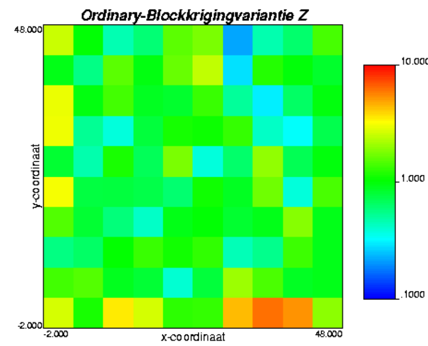
$$\text{var}[\hat{Z}_B - Z_B] = \sum_{j=1}^{n} \lambda_j \, \gamma(\mathbf{x}_i, B) + \nu - \gamma(B,B)$$

Ordinary block kriging can serve as declustering algoritme

Universiteit Utrecht

(Ordinary) block kriging





Universiteit Utrecht

Universiteit Utrecht

Stochastic Hydrology

Geostatistics (2)

Marc F.P. Bierkens

Professor of Hydrology
Faculty of Geosciences

Universiteit Utrecht

Estimating conditional probability

$$\Pr[Z(\mathbf{x}_0) \leq z \mid z(\mathbf{x}_1),\ldots,z(\mathbf{x}_n)]$$

Three ways:

- MultiGaussian kriging ⇐
- Indicator kriging
- Disjunctive kriging

Universiteit Utrecht

Estimating conditional probability

MultiGaussian kriging

Heavy assumptions:

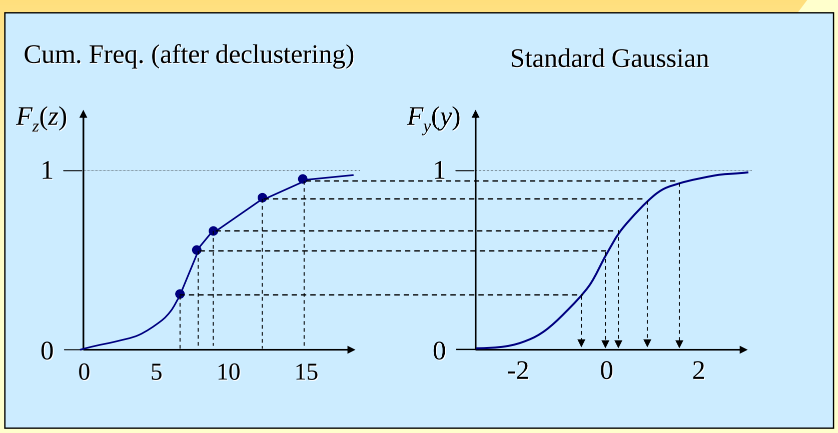
- $Z(\mathbf{x})$ is wide sense stationary
- Any set of values $Z(\mathbf{x}_1),\ldots, Z(\mathbf{x}_N)$ (or their transforms) are jointly Gaussian distributed.

Universiteit Utrecht

Conditional probability

MultiGaussian kriging step by step

Step 1: Normal-score transformation of data



Universiteit Utrecht

Conditional probability

MultiGaussian kriging step by step

2 Estimate semivariogram of transforms Y
Note: Semivariogram must have a sill!

3 Simple kriging Y-values with $\mu_Y = 0$: yields prediction $\hat{Y}(\mathbf{x}_0)$ and kriging-variance $\sigma^2_{\hat{Y}}(\mathbf{x}_0)$

4. Calculate cumulative probability as:

$$\Pr[Z(\mathbf{x}_0) \le z \mid z(\mathbf{x}_i), i = 1,..,n] = \frac{1}{\sqrt{2\pi\sigma^2_{\hat{Y}}\mathbf{x}_0}} \int_{-\infty}^{z} \exp\left(\frac{-[y_{ns}(z) - \hat{y}(\mathbf{x}_0)]^2}{\sigma^2_{\hat{Y}}(\mathbf{x}_0)} \right)$$

Universiteit Utrecht

Conditional probability

2. De multiGaussion method

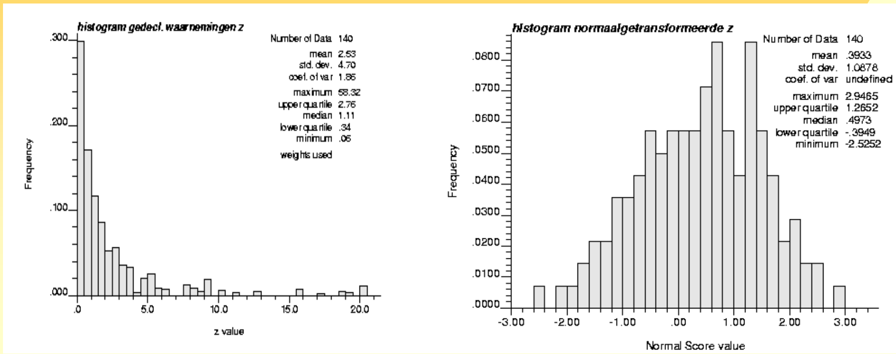
Two additional reasons for taking normal score transforms:

- Stabilisation of variance: better estimates of the semivariogram.
- Values have to be within a given range (e.g. only positive, between 0 en 1).

Universiteit Utrecht

Conditional probability

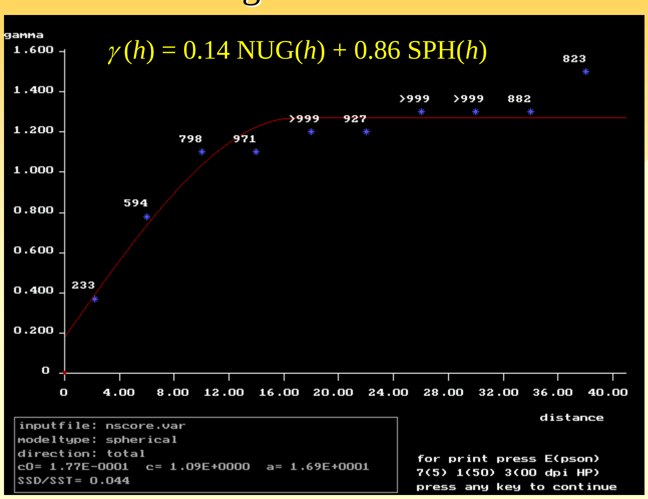
Example Walker Lake data set



Universiteit Utrecht

Conditional probability

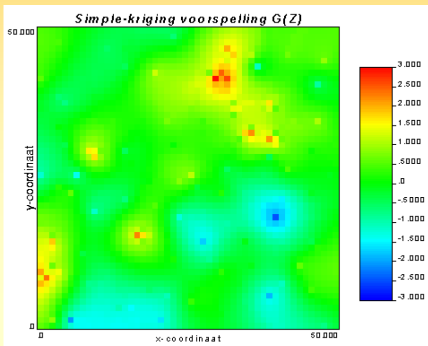
Semivariogram of ns-transforms

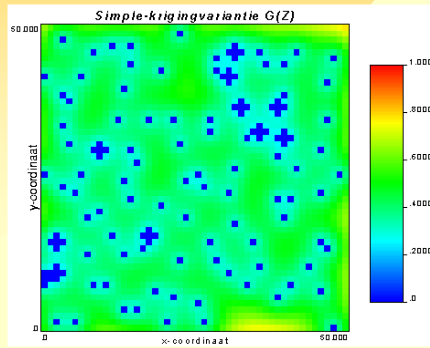


Universiteit Utrecht

Conditional probability

Simple kriging prediction *Y* and kriging variance

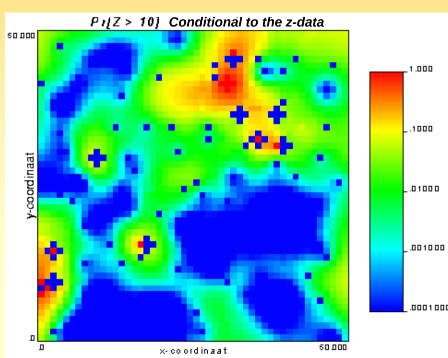




Universiteit Utrecht

Conditional probability

Example of exceedence probability: z_c = 10 -> y_c = G(10) = 1.757.



Universiteit Utrecht

Question: How do we obtain $\hat{Z}(\mathbf{x}_0)$

from $\hat{Y}(\mathbf{x}_0)$ and $\sigma^2_{\hat{Y}}(\mathbf{x}_0)$?

Coffee Break

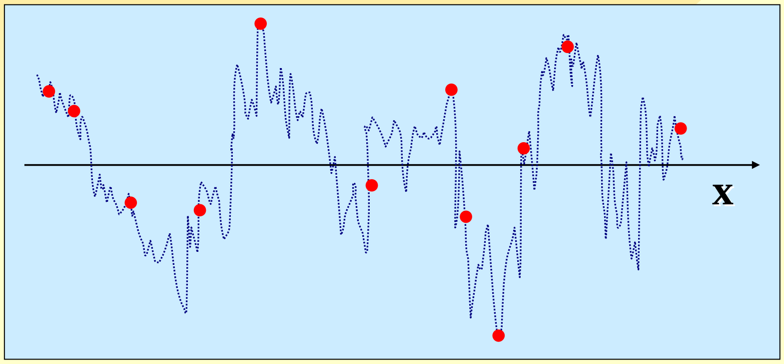
Universiteit Utrecht

Geostatistical Simulation

Simulating realisations of a (spatial) random function

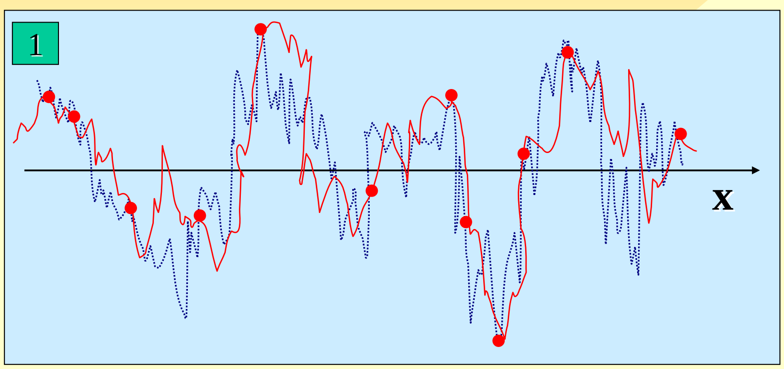
Universiteit Utrecht

Geostatistical Simulation



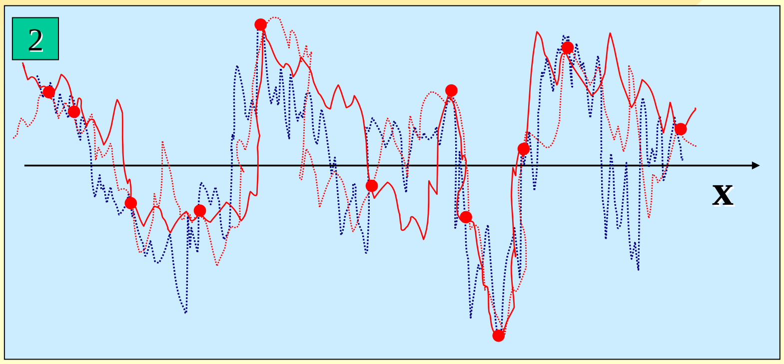
Universiteit Utrecht

Geostatistical Simulation



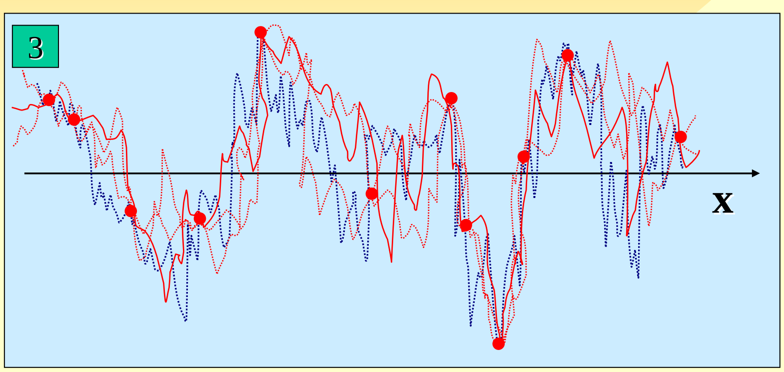
Universiteit Utrecht

Geostatistical Simulation



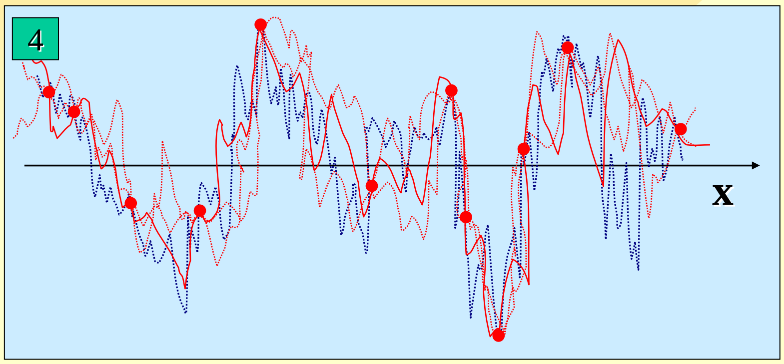
Universiteit Utrecht

Geostatistical Simulation



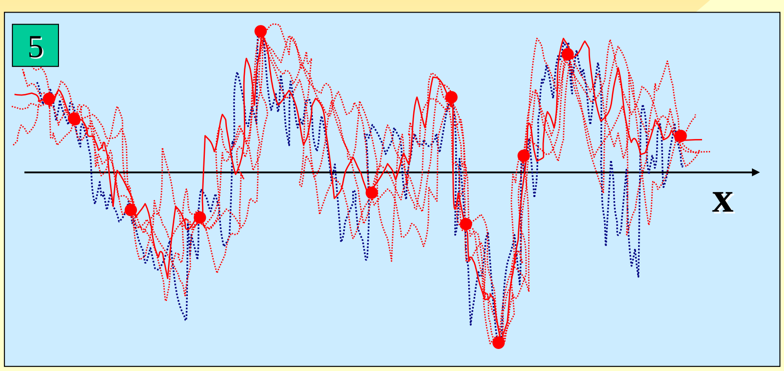
Universiteit Utrecht

Geostatistical Simulation



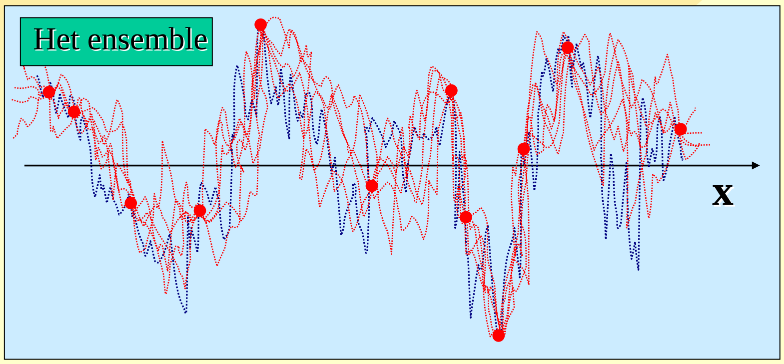
Universiteit Utrecht

Geostatistical Simulation



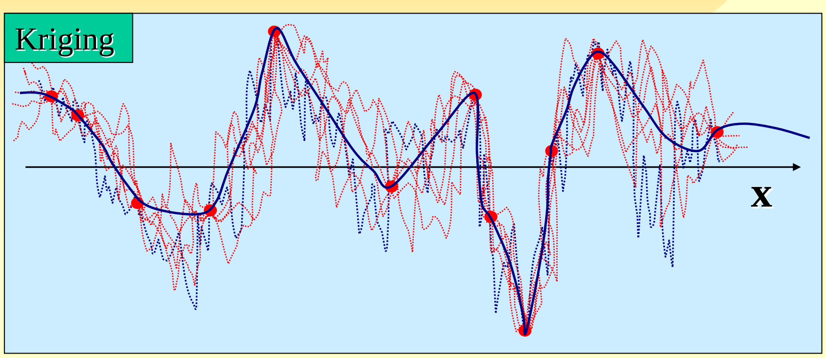
Universiteit Utrecht

Geostatistical Simulation



Universiteit Utrecht

Geostatistical Simulation

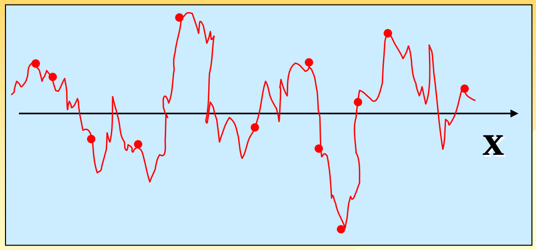


Universiteit Utrecht

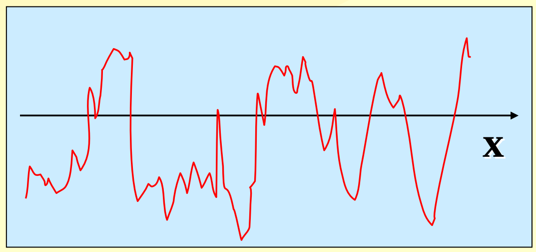
Geostatistical Simulation

Conditionel and unconditionel simulation

conditional



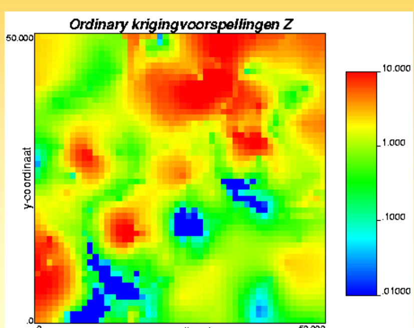
unconditional

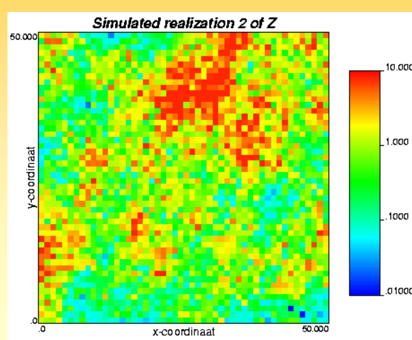


Universiteit Utrecht

Geostatistical Simulation

Conditionel simulation or kriging?





Suppose these are conductivity values:

- kriging predictions are too smooth
- flow is influenced by measurement locations

Universiteit Utrecht

Geostatistical Simulation

Conditionel simulation or kriging?

- If point-operation: kriging and conditional pdf

$$Y(\mathbf{x}_0) = g[Z(\mathbf{x}_0)]$$

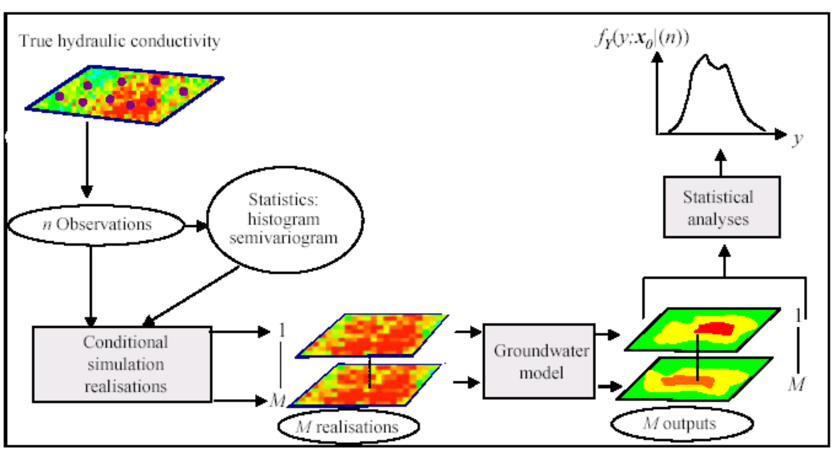
- If neighborhood operation: conditional simulation

$$Y(\mathbf{x}_0) = g[Z(\mathbf{x}_1),\ldots,Z(\mathbf{x}_N)]$$

Includes partial differential equations!

Universiteit Utrecht

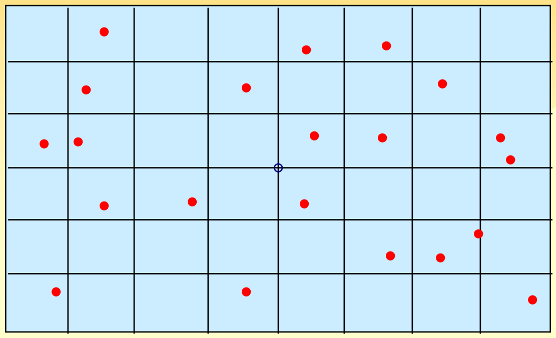
Monte Carlo Analysis or Monte Carlo Simulation



Universiteit Utrecht

Sequential Gaussian Simulation

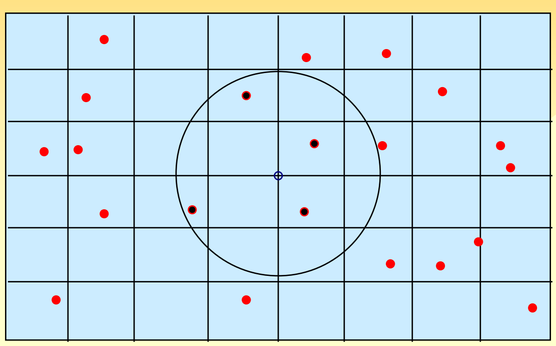
1. Define a simulation grid and visit grid points at random



Universiteit Utrecht

Sequential Gaussian Simulation

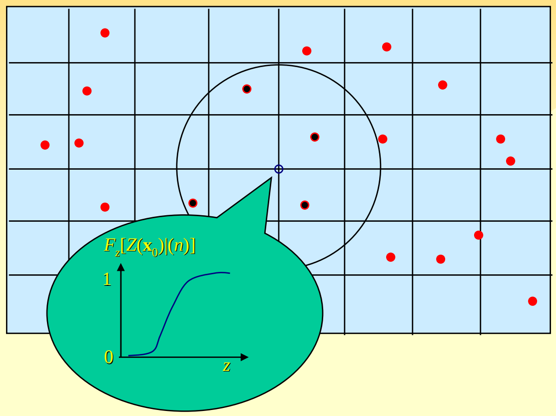
2. Select observations within search radius



Universiteit Utrecht

Sequential Gaussian Simulation

3. Estimate the conditional cpdf from kriging

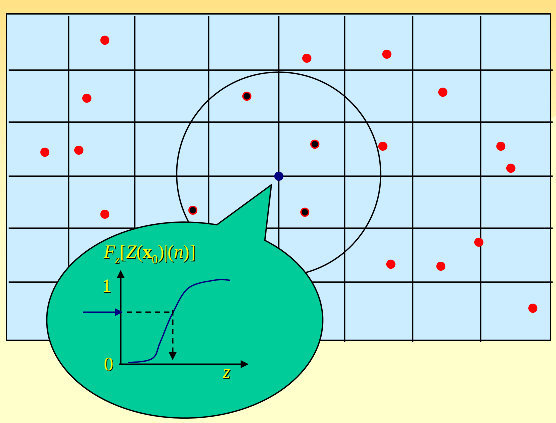


$F_z[Z(\mathbf{x}_0)|(n)]$ using MultiGaussian kriging

Universiteit Utrecht

Sequential Gaussian Simulation

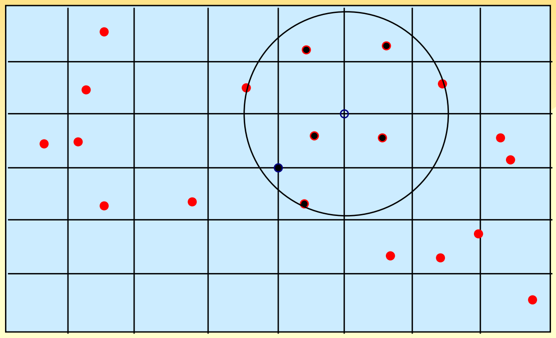
4. Draw a random variable from the cpdf and add it to the data set



Universiteit Utrecht

Sequential Gaussian Simulation

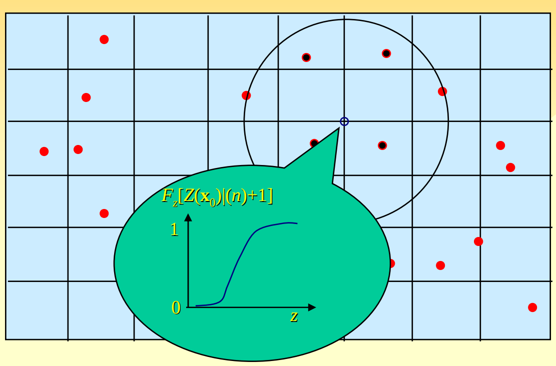
5. Go to the next grid location



Universiteit Utrecht

Sequential Gaussian Simulation

6. Estimate the conditional cpdf using kriging

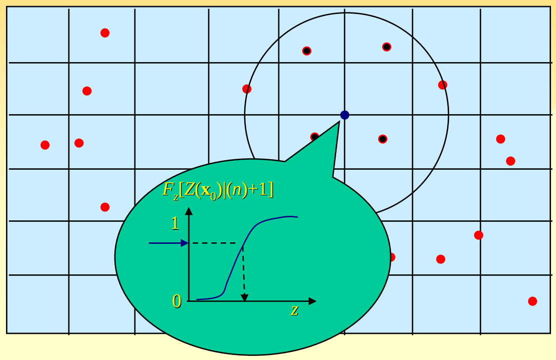


$F_z[Z(\mathbf{x}_0)|(n)+1]$ via MultiGaussian kriging
kriging using observations + previously simulated nodes

Universiteit Utrecht

Sequential Gaussian Simulation

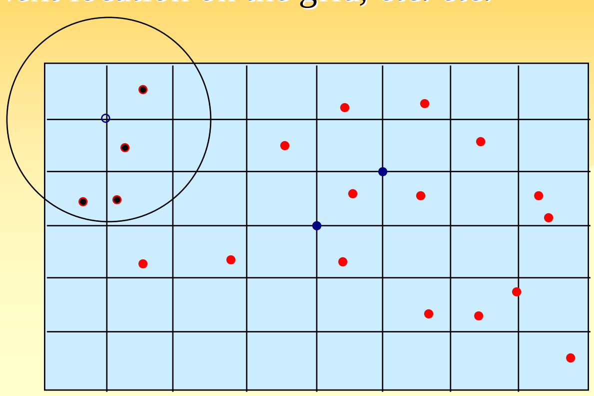
7. Draw a random variable from the cpdf and add it to the data set



Universiteit Utrecht

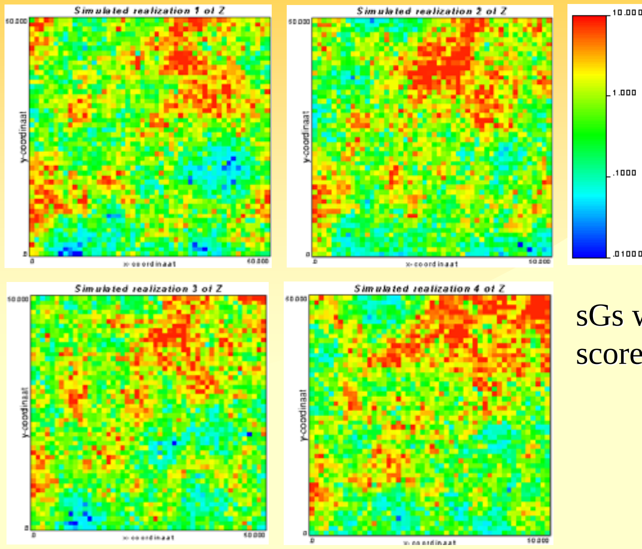
Sequential Gaussian Simulation

8. Next location on the grid, etc. etc.



Universiteit Utrecht

Sequential Gaussian Simulation
Walker Lake Example



sGs with normal-score transform