Universiteit Utrecht

Stochastic Hydrology

Forward stochastic modelling

Marc F.P. Bierkens

Professor of Hydrology
Faculty of Geosciences

Universiteit Utrecht

The problem of stochastic forward modelling

- Discharge calculated from water height over a weir subject to error
- Groundwater head from (partly) unknown conductivity
- Concentration of a lake using a decay rate that is not exacty known

Input or parameters *Z*, model output *Y*, function or model *g*():

$$Y = g(Z)$$

Given the pdf $f_Z(z)$ of *Z* or its moments (mean, variance),

What is the pdf $f_Y(y)$ of Y or its moments?

Universiteit Utrecht

The problem of stochastic forward modelling

Which method to use depends on:

- the form of $g()$;
- what is known about Z, the complete pdf or its moments.

Universiteit Utrecht

Classification of methods

Classification of problems:

- explicit functions of one random variable ⇐
- explicit functions of multiple random variables ⇐
- explicit vector functions
- explicit functions of random functions of time, space or space-time
- differential equations with a random parameter ⇐
- stochastic differential equations ⇐

Universiteit Utrecht

Explicit functions of one random variable

Explicit function: $Y = g(Z)$

Implicit function: $Y = g(Z, Y)$ Includes differential equations

Universiteit Utrecht

Explicit functions of one random variable

Method 1: Derived distribution

Goal:

- the probability density function $f_Y(y)$.

Requirements:

- the probability density $f_Z(z)$ of Z is known;
- the function $g(Z)$ is monotonous (only increasing or only decreasing), differentiable and can be inverted.

Cumulative distribution function: $F_Y(y) = F_Z(g^{-1}(y))$

Probability density function: $f_Y(y) = \left| \frac{d[g^{-1}(y)]}{dy} \right| f_Z(g^{-1}(y))$

Universiteit Utrecht

Explicit functions of one random variable

Method 1: Derived distribution: example

Discharge formula of a weir: $Q = aH^b$

Inverse: $g^{-1}(q) = \left(\frac{q}{a}\right)^{\frac{1}{b}} \longrightarrow F_Q(q) = F_H\left(\left(\frac{q}{a}\right)^{\frac{1}{b}}\right)$

Derivative of inverse: $\frac{dg^{-1}(y)}{dy} = \frac{1}{b}\left(\frac{q}{a}\right)^{\frac{1-b}{b}}$

$$\longrightarrow f_Q(q) = \frac{1}{b}\left(\frac{q}{a}\right)^{\frac{1-b}{b}} f_H\left(\left(\frac{q}{a}\right)^{\frac{1}{b}}\right)$$

Universiteit Utrecht

Explicit functions of one random variable

Method 2: Taylor approximation

Goal:

- the moments of Y, e.g. μ_Y and σ_Y^2.

Requirements:

- the moments of Z, e.g. μ_Z and σ_Z^2, are known;
- the variance σ_Z^2 should not be too large.
- the function g() is differentiable

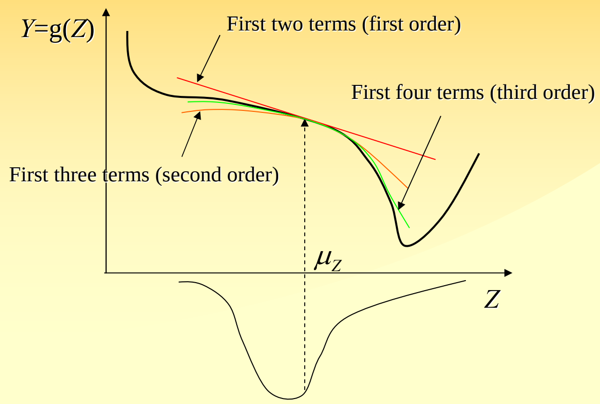
$$Y = g(Z) = g(\mu_Z) + \left(\left.\frac{dg(z)}{dz}\right|_{z=\mu_Z}\right)(Z-\mu_Z) +$$

$$\frac{1}{2}\left(\left.\frac{d^2g(z)}{dz^2}\right|_{z=\mu_Z}\right)(Z-\mu_Z)^2 + \frac{1}{6}\left(\left.\frac{d^3g(z)}{dz^3}\right|_{z=\mu_Z}\right)(Z-\mu_Z)^3 + \ldots$$

Universiteit Utrecht

Explicit functions of one random variable

Method 2: Taylor approximation



Universiteit Utrecht

Explicit functions of one random variable

Method 2: Taylor approximation

First order Taylor approximation:

$$\mu_Y = E[Y] \approx g(\mu_Z)$$

$$\sigma_Y^2 = E[(Y-\mu_Y)^2] \approx \left(\left. \frac{dg(z)}{dz} \right|_{z=\mu_Z} \right)^2 \sigma_Z^2$$

Second order Taylor approximation:

$$\mu_Y \approx g(\mu_Z) + \frac{1}{2} \left(\left. \frac{d^2 g(z)}{dz^2} \right|_{z=\mu_Z} \right) \sigma_Z^2$$

Universiteit Utrecht

Explicit functions of one random variable

Method 2: Taylor approximation: example

Discharge formula of a weir: $Q = aH^b$

$$\left(\left. \frac{dg(z)}{dz} \right|_{z=\mu_Z} \right) = abh^{b-1}\Big|_{h=\mu_H} = ab\mu_H^{b-1}$$

$$\left(\left. \frac{d^2g(z)}{dz^2} \right|_{z=\mu_Z} \right) = ab(b-1)h^{b-2}\Big|_{h=\mu_H} = ab(b-1)\mu_H^{b-2}$$

Universiteit Utrecht

Explicit functions of one random variable

Method 2: Taylor approximation: example

First order Taylor approximation:

$$\mu_Q \approx a\mu_H^b$$

$$\sigma_Q^2 \approx a^2 b^2 \mu_H^{2b-2} \sigma_H^2$$

Second order Taylor approximation:

$$\mu_Q \approx a\mu_H^b + \frac{1}{2} ab(b-1)\mu_H^{b-2}\sigma_H^2$$

Universiteit Utrecht

Explicit functions of one random variable

Method 3: Monte Carlo Simulation

Goal:
· the probability density function or its moments.
Requirements:
· the probability density of Z is known

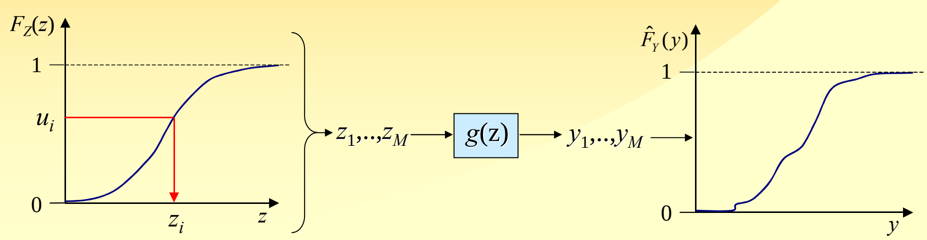
No requirements on pdf or form of $g(Z)$!

Universiteit Utrecht

Explicit functions of one random variable

Method 3: Monte Carlo Simulation

u_i from uniform[0,1]



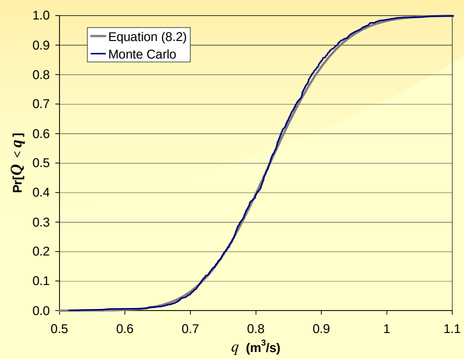
Universiteit Utrecht

Explicit functions of one random variable

Comparison of Monte Carlo Simulation and exact derived distribution

$$Q = aH^b$$

a=5 and b=1.5, with Q in m^3/d and with H in m following a Gaussian distribution with mean 0.3 m and standard deviation of 0.02 m; MC using 1000 realisations of H.



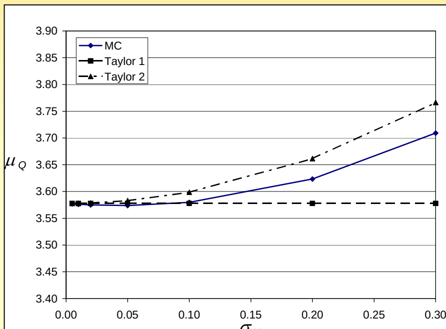
Universiteit Utrecht

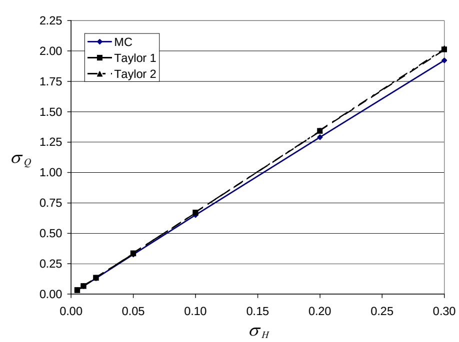
Explicit functions of one random variable

Comparison of Monte Carlo Simulation Taylor approximation

$$Q = aH^b$$

a=5 and b=1.5, with Q in m^3/d and with H in m following a Gaussian distribution with mean 0.8 m and varying standard deviation; MC using 1000 realisations of H.





Universiteit Utrecht

Explicit functions of multiple random variables

$$Y = g(Z_1,..,Z_m)$$

Method 1: linear modelling
Goal:
· the first two moments of Y .
Requirements:
· the function g() is linear: $Y = a + \sum_{i=1}^{m} b_i Z_i$

Mean: $\mu_Y = a + \sum_{i=1}^{m} b_i \mu_i$

Variance: $\sigma_Y^2 = \sum_{i=1}^{m} \sum_{j=1}^{m} b_i b_j \rho_{ij} \sigma_i \sigma_j$

Note: if the Z_i are multiGaussian then so is Y and we have the pdf!

Universiteit Utrecht

Explicit functions of multiple random variables

Method 2: Taylor approximation
Goal:
- the first two moments of Y .

Requirements:
- the function g() is differentiable
- the moments of the Z_i are known: $\mu_i, \sigma_i^2, \rho_{ij}$
- the variances σ_i^2 are not too large

Universiteit Utrecht

Explicit functions of multiple random variables

Method 2: Taylor approximation:

First order mean:

$$\mu_Y = g(\mu_1, \mu_2)$$

Second order mean:

$$\mu_Y = g(\mu_1,\mu_2) + \frac{1}{2}\frac{d^2 g}{dz_1^2}(\mu_1,\mu_2)\sigma_1^2 + \frac{d^2 g}{dz_1 dz_2}(\mu_1,\mu_2)\rho_{12}\sigma_1\sigma_2 + \frac{1}{2}\frac{d^2 g}{dz_2^2}(\mu_1,\mu_2)\sigma_2^2$$

First order variance:

$$\sigma_Y^2 = g(\mu_1,\mu_2) + \left(\frac{dg}{dz_1}(\mu_1,\mu_2)\right)^2 \sigma_1^2 +$$
$$2\left(\frac{dg}{dz_1}(\mu_1,\mu_2)\right)\left(\frac{dg}{dz_2}(\mu_1,\mu_2)\right)\rho_{12}\sigma_1\sigma_2 + \left(\frac{dg}{dz_2}(\mu_1,\mu_2)\right)^2 \sigma_2^2$$

Universiteit Utrecht

Explicit functions of multiple random variables

Method 2: Taylor approximation:

Example Consider the weir equation or rating curve $Q = Ah^B$, where A and B are stochastic variables with statistics $\mu_A, \sigma_A^2, \mu_B, \sigma_B^2$ and ρ_{AB}. The first order Taylor approximation of the mean becomes:

$$E[Q] \approx \mu_A h^{\mu_B} \qquad (8.42)$$

and the second order approximation:

$$E[Q] \approx \mu_A h^{\mu_B} + \frac{1}{2}(\mu_A h^{\mu_B}(\ln h)^2)\sigma_B^2 + (h^{\mu_B}\ln h)\rho_{AB}\sigma_A\sigma_B \qquad (8.43)$$

The variance from the first order Taylor analysis is given by:

$$\sigma_Q^2 \approx (h^{2\mu_B})\sigma_A^2 + (4\mu_A h^{2\mu_B}\ln h)\rho_{AB}\sigma_A\sigma_B + (2\mu_A^2 h^{2\mu_B}(\ln h)^2)\sigma_B^2 \qquad (8.44)$$

Universiteit Utrecht

Explicit functions of multiple random variables

Method 3: Monte Carlo simulation

Goal:
· the probability density function or its moments.
Requirements:
· the multivariate probability density of the Z_i is known

Universiteit Utrecht

Explicit functions of multiple random variables

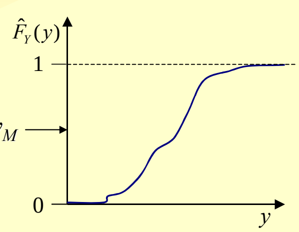
Method 3: Monte Carlo simulation

In practice: the multivariate probability density of the Z_i is assumed Gaussian
otherwise: transformation to Gaussians

Generation of M sets of dependent random variables from multivariate pdf (LU-decomposition

$$\{z_1^{(1)}, z_2^{(1)}, .., z_n^{(1)}\}$$
$$\{z_1^{(2)}, z_2^{(2)}, .., z_n^{(2)}\}$$
$$\vdots$$
$$\{z_1^{(M)}, z_2^{(M)}, .., z_n^{(M)}\}$$

$\longrightarrow$ $g(z_1,.., z_n)$ $\longrightarrow$ $y_1,..,y_M$ $\longrightarrow$



Universiteit Utrecht

Differential equation with random coefficients

Example: concentration from a lake C with time:

$$v\frac{dC}{dt} = -vKC + q_{in}$$

where v is the volume of the lake (assumed constant and known) q_{in} is the constant an known input load and K is a random decay coefficient.

Required: pdf of C or its moments

Method:

1) solve the differential equation (either analytically or numerically) assuming K deterministic.

2) analyse the solution, for the case that C is random.

Universiteit Utrecht

Differential equation with random coefficients

Solution $C(t=0) = 0$: $C(t) = \frac{q_{in}}{vK}\left(1 - e^{-Kt}\right)$

If moments of K are known: μ_K, σ_K^2

Perform Taylor approximation or Monte Carlo simulation

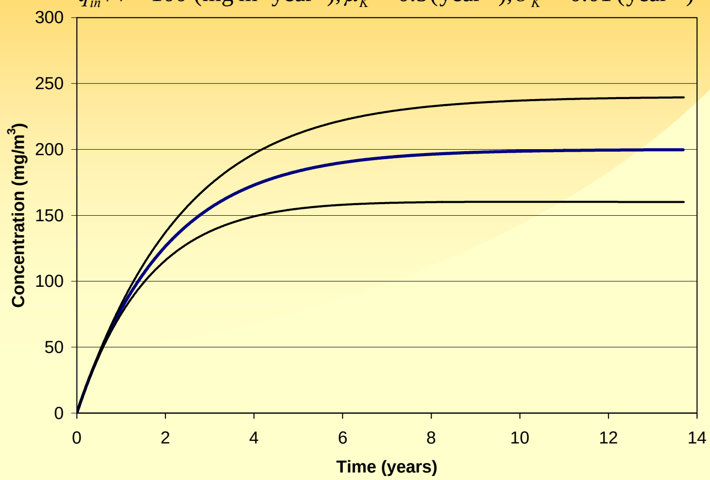
First order Taylor approximation of the variance:

$$\sigma_C^2(t) = \frac{q_{in}^2}{v^2}\frac{\left(e^{-\mu_K t}(1+\mu_K t) - 1\right)^2}{\mu_K^4}\sigma_K^2(t)$$

Universiteit Utrecht

Differential equation with random coefficients

$q_{in} / v = 100\ (\text{mg m}^{-3}\text{year}^{-1}), \mu_K = 0.5(\text{year}^{-1}), \sigma_K^2 = 0.01\ (\text{year}^{-2})$



Universiteit Utrecht

Stochastic differential equations

Examples:

1. Evolution of lake concentration with decay rate as a random function of time:

$$v\frac{dC}{dt} = -vK(t)C + q_{in}$$

2. Transient groundwater equation for two-dimensional flow with heterogeneous storage coefficient and transmissivity, described as random space functions:

$$S(x,y)\frac{\partial H}{\partial t} = \frac{\partial}{\partial x}\left(T(x,y)\frac{\partial H}{\partial x}\right) + \frac{\partial}{\partial y}\left(T(x,y)\frac{\partial H}{\partial y}\right)$$

Universiteit Utrecht

Stochastic differential equations

$$v\frac{dC}{dt} = -vK(t)C + q_{in}$$

$$S(x,y)\frac{\partial H}{\partial t} = \frac{\partial}{\partial x}\left(T(x,y)\frac{\partial H}{\partial x}\right) + \frac{\partial}{\partial y}\left(T(x,y)\frac{\partial H}{\partial y}\right)$$

Required, multivariate pdf of C or H or its moments:

$$\mu_C(t),\ \sigma_C^2(t),\ C_C(t_1,t_2)$$
$$\mu_H(x,y),\ \sigma_H^2(x,y),\ C_H(x_1,y_1,x_2,y_2)$$

We cannot solve for deterministic $S(x,y)$,$T(x,y)$ and $K(t)$, and then analyse the solution for randomness!

Universiteit Utrecht

Stochastic differential equations

Consider example 2: Concentration pollutant in a lake:

$$\frac{dC}{dt} = -K(t)C + q_{in}c_{in}$$

General solution:

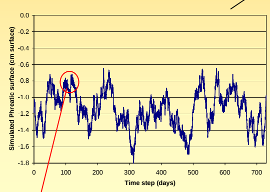
$$C(t) = C(t_0) + q_{in}c_{in}(t-t_0) - \int_0^t CK(t)dt$$

Universiteit Utrecht

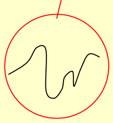
Stochastic differential equations

How to solve the stochastic integral: $\int_0^t CK(t)dt$

Coloured noise



limited variation

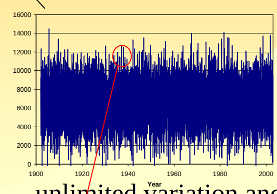


$\underset{\tau \to 0}{VAR}[z(t) - z(t+\tau)] \to 0$

Smooth at very small scales

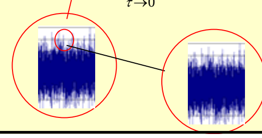
Normal calculus

White noise



unlimited variation and infinite variance

$\underset{\tau \to 0}{VAR}[z(t) - z(t+\tau)] \to \infty$



Itô calculus

Universiteit Utrecht

Stochastic differential equations

We consider only coloured noise

Solution methods:

a) (Semi-)Analytical derivation of the moments $\mu_C(t), \sigma_C^2(t), C_C(t_1,t_2)$ as a function of the moments $\mu_K(t), \sigma_K^2(t), C_K(t_1,t_2)$

-> strict conditions on RF *K* (and *H* and *S*): stationary and multiGaussian

-> in groundwater flow: uniform flow in infinite domains!

-> small variance $\sigma_K^2 \ (\sigma_T^2, \sigma_S^2)$

Examples:

- Small pertrurbation analysis
- Moment differential equations (usually solved numerically)

Universiteit Utrecht

Stochastic differential equations

Solution methods:

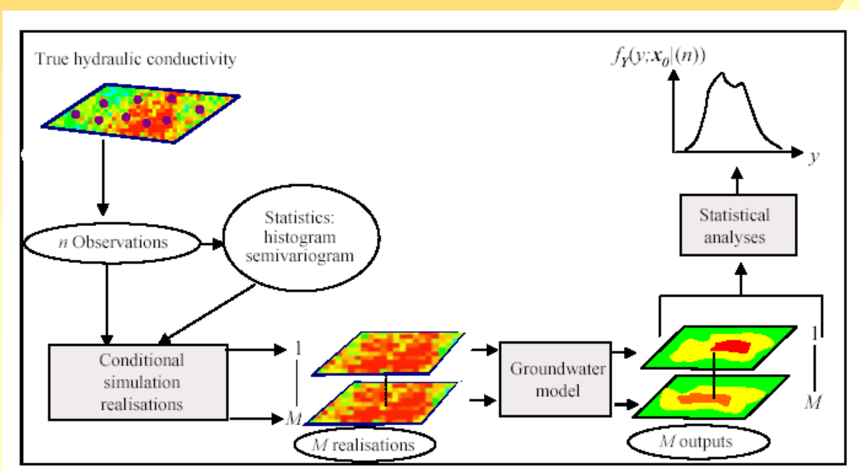
b) Monte Carlo simulation

- required: multivariate pdf of parameters/input
- no additional assumption needed!
- Conditioning on data possible!

Universiteit Utrecht

Stochastic differential equations

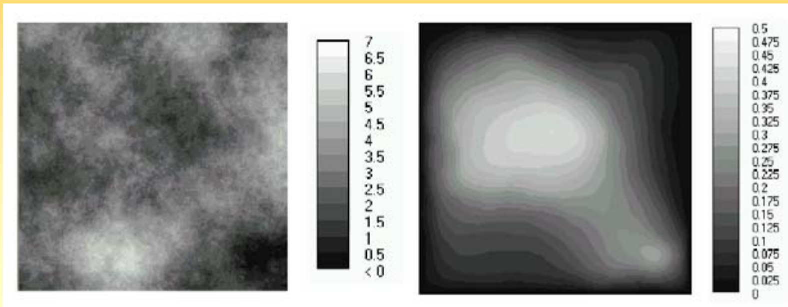
Monte Carlo simulation: Groundwater example



Universiteit Utrecht

Stochastic differential equations

Monte Carlo simulation: Groundwater example



Reality. Left: Log-hydraulic conductivity Y=ln$K$ (K in m/d); Right: Hydraulic head (m)

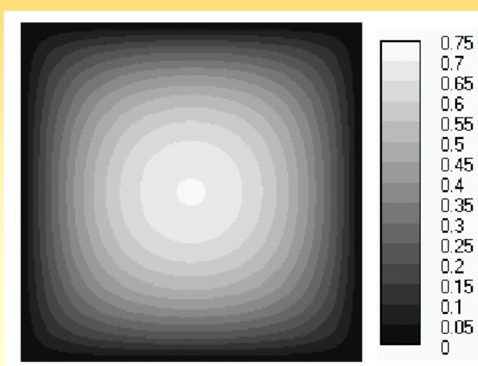
$K_G = \exp(\mu_Y) = 10$ m/d; $\sigma_Y^2 = 2$

Hydraulic head at boundaries = 0, precipiation is 1 mm/day.

Universiteit Utrecht

Stochastic differential equations

Monte Carlo simulation: Groundwater example



Solution assuming homogenous $K = K_G = 10$ m/d.

Universiteit Utrecht

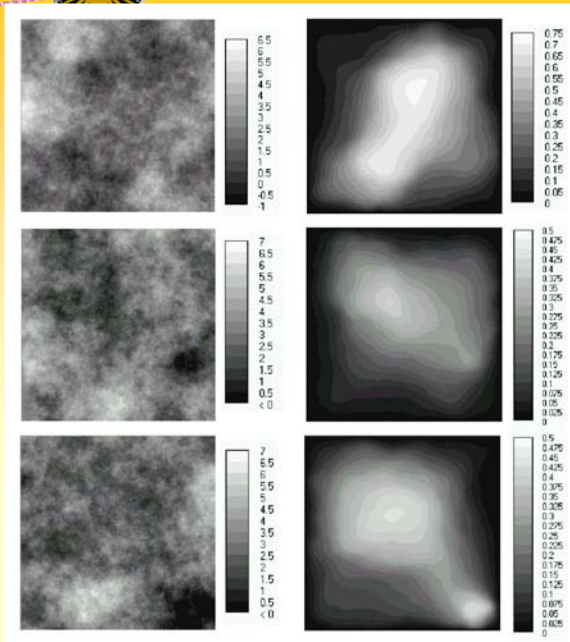
Monte Carlo simulation: Groundwater example

Realisations with increasing number of conditioning data

0

9

36



Universiteit Utrecht

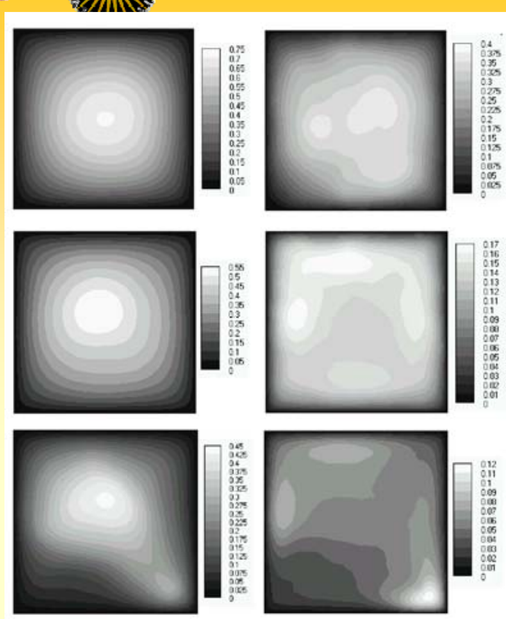
Monte Carlo simulation: Groundwater example

Mean (left) and variance of 1000 realisations with increasing number of conditioning data

0

9

36



Universiteit Utrecht

Monte Carlo simulation: Groundwater example

Relationship between the domain average standard deviation of hydraulic head and the number of observations used for conditioning (zero observations means unconditional simulation)