Universiteit Utrecht

Stochastic Hydrology

Random functions

Marc F.P. Bierkens

Professor of Hydrology
Faculty of Geosciences

Universiteit Utrecht

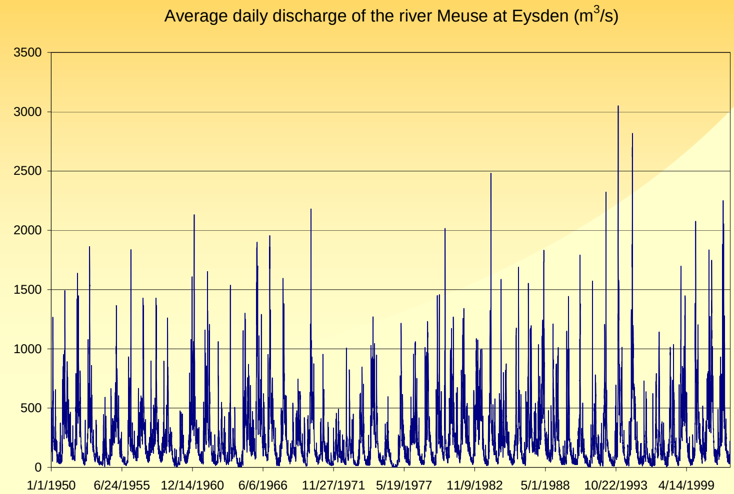
Introduction

Up to now: *unknown value* of a variable *Z* without considering time and spatial co-ordinates *t*,*x*,*y*,*z*
→ Random Variable

Now consider the *unknown variation* of a variable *Z* in time, space or space-time: and spatial co-ordinates Z(*t*), Z(*x*,*y*,*z*) or Z(*x*,*y*,*z*,*t*)
→ Random Function

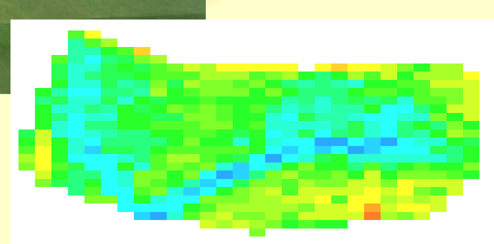
Universiteit Utrecht

Z(t):



Universiteit Utrecht

Z(x,y) :



Exhaustive TDR data set (2000 observations) Tarrawarra Catchment Australia

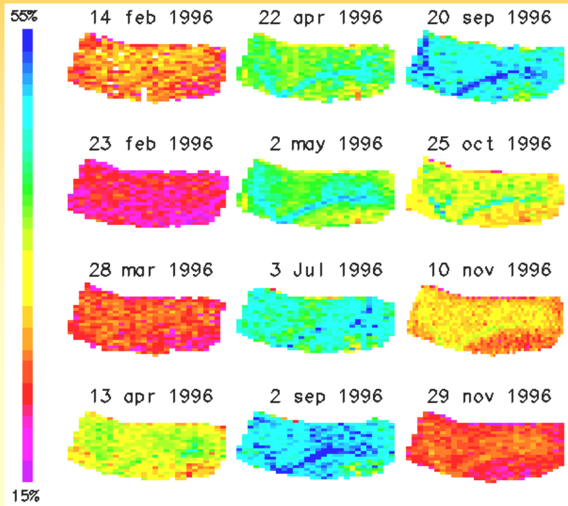
(Western and Grayson, 1998, Water Resources Research, 34(10): pp2765-2768)

Universiteit Utrecht

Exhaustive TDR data set (2000 observations) Tarrawarra Catchment Australia

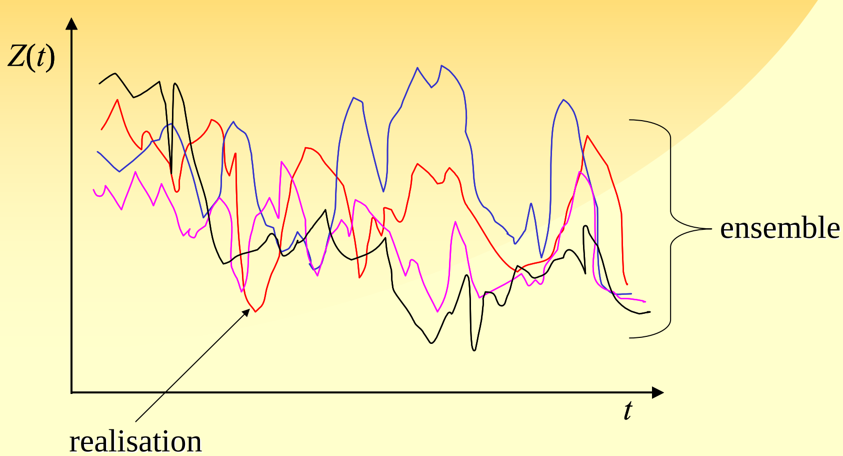
(Western and Grayson, 1998, Water Resources Research, 34(10): pp2765-2768)

$Z(x,y,t)$:



Universiteit Utrecht

Concept of a random function



Universiteit Utrecht

Random functions of different domains

- Random time function (RTF), random process or stochastic process:

$$Z(t)$$

- Random space function or random field:

$$Z(\mathbf{x}), \mathbf{x} = (x, y) \text{ or } \mathbf{x} = (x, y, z)$$

- Random space-time function or space-time random field:

$$Z(\mathbf{x}, t)$$

Universiteit Utrecht

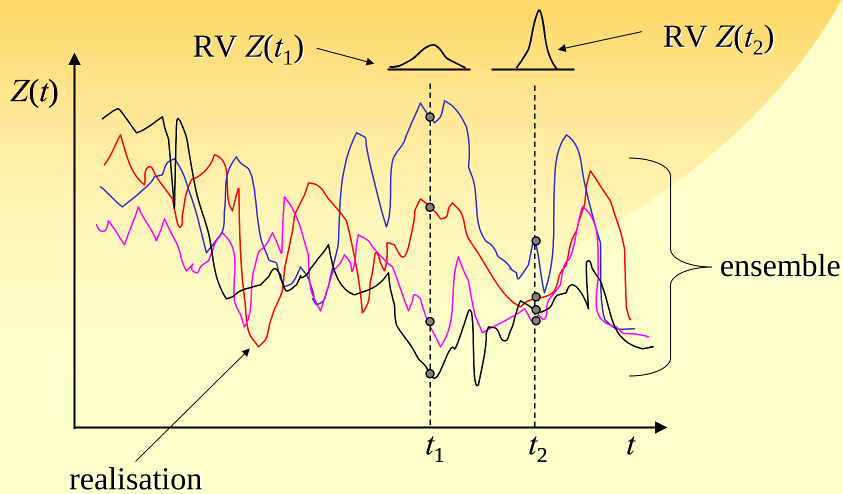
Random functions are dependent RVs

Another way of viewing a random function is as:

a collection of random variables (one at every location in space or point in time) that are all mutually statistically dependent.

Universiteit Utrecht

Random functions are dependent RVs



Universiteit Utrecht

Univariate Statistics of random functions

Mean:

$$\mu_Z(t) = \mathrm{E}[Z(t)] = \int_{-\infty}^{\infty} z\, f_Z(z;t)dz$$

Variance:

$$\sigma_Z^2(t) = \mathrm{E}[\{Z(t) - \mu_Z(t)\}^2] = \int_{-\infty}^{\infty} \{z - \mu_Z(t)\}^2 f_Z(z;t)dz.$$

Universiteit Utrecht

Bivariate Statistics of random functions

Bivariate pdf

$$f(z_1, z_2; t_1, t_2) = \lim_{\partial z_1, \partial z_2 \to 0} \frac{\Pr[z_1 < Z(t_1) \le z_1 + \partial z_1, z_2 < Z(t_2) \le z_2 + \partial z_2]}{\partial z_1 \partial z_2}$$

Gives the probability density of values of the random function at two points in time, space or space-time.

Covariance:

$$\mathrm{COV}[Z(t_1), Z(t_2)] = \mathrm{E}[\{Z(t_1) - \mu_Z(t_1)\}\{Z(t_2) - \mu_Z(t_2)\}] =$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \{z_1 - \mu_Z(t_1)\}\{z_2 - \mu_Z(t_2)\} f(z_1, z_2; t_1, t_2) dz_1 dz_2$$

Universiteit Utrecht

Multivariate pdf of a random function

Also called: multi-point pdf

$$f(z_1, z_2, \ldots.\, z_N \,;\, t_1, t_2, \ldots.\, t_N) =$$

$$\lim_{\partial z_1, \ldots, \partial z_N \to 0} \frac{\Pr(z_1 < z(t_1) \le z_1 + \partial z_1, z_2 < z(t_2) \le z_2 + \partial z_2, \ldots, z_N < z(t_N) \le z_2 + \partial z_N}{\partial z_1 \partial z_2 \cdots \partial z_N}$$

Gives the probability density of values of the random function at a set of points in time, space or space-time.

Universiteit Utrecht

Types of random functions

1) Value: a) discrete valued D

   b) continuous valued Z

2) Domain type: a) time t

   b) space $\mathbf{x}$

   c) space-time $(\mathbf{x},t)$

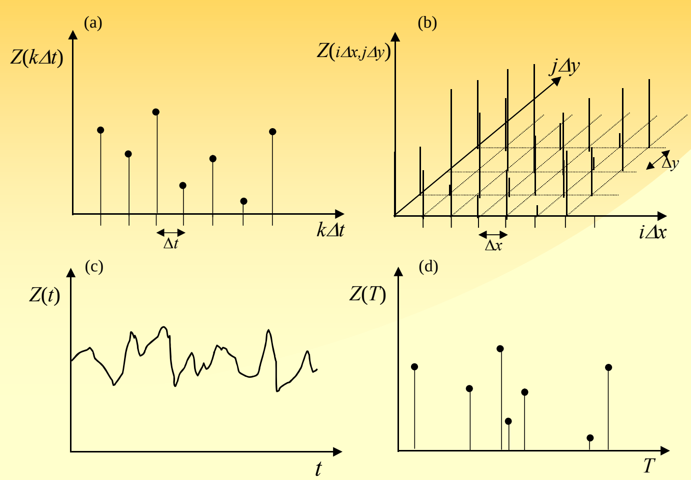
3) Domain definition: a) continuous $\mathbf{x}$, t

   b) discrete (at finite number of points) $\mathbf{x}_i$, t_k

   c) random points (point process) $\boldsymbol{X}_j$, T_k

Universiteit Utrecht

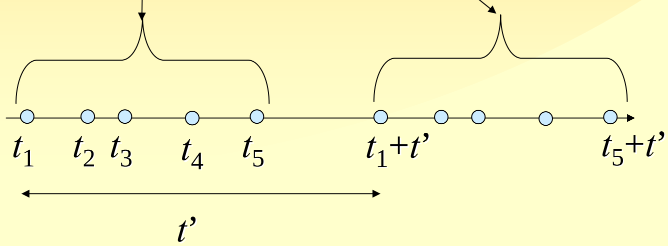
Types of random functions



Universiteit Utrecht

Strict stationary random functions

$$f(z_1, z_2, \ldots, z_N;\ t_1, t_2, \ldots\ t_N) =$$
$$f(z_1, z_2, \ldots, z_N\ ;\ t_1 + t', t_2 + t', \ldots,\ t_N + t') \quad \forall t_i, t'$$

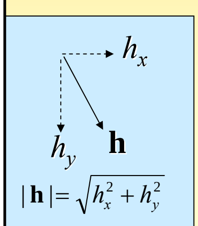


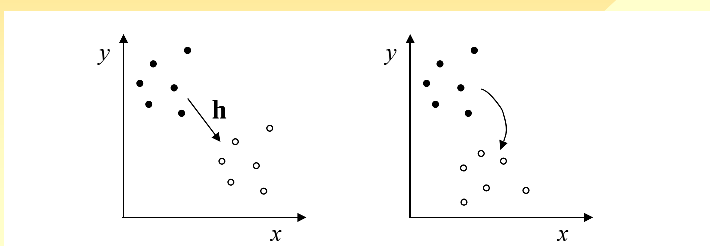
time

Universiteit Utrecht

Strict stationary random functions

$$f(z_1, z_2, \ldots, z_N;\ \mathbf{x}_1, \mathbf{x}_2, \ldots\ \mathbf{x}_N) =$$
$$f(z_1, z_2, \ldots, z_N\ ;\ \mathbf{x}_1 + \mathbf{h}, \mathbf{x}_2 + \mathbf{h}, \ldots,\ \mathbf{x}_N + \mathbf{h}) \quad \forall \mathbf{x}_i, \mathbf{h}$$



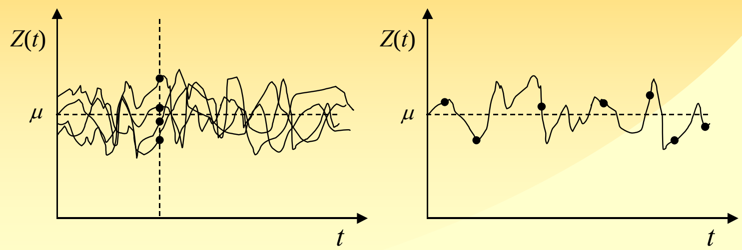


Strict stationary RF

Strict Stationary and isotropic RF

space

Ergodic random functions



Statistics over time (space/space-time) converge to the statistics over the ensemble!
WE CAN THUS ESTIMATE STATISTICS FROM DATA!

Universiteit Utrecht

Ergodic random functions

e.g. ergodicity in the mean:

$$\lim_{T\to\infty}\frac{1}{T}\int_T z(t)dt = \int_{-\infty}^{\infty} z\, f_Z(z;t)dz = \mu_Z$$

Requirements:

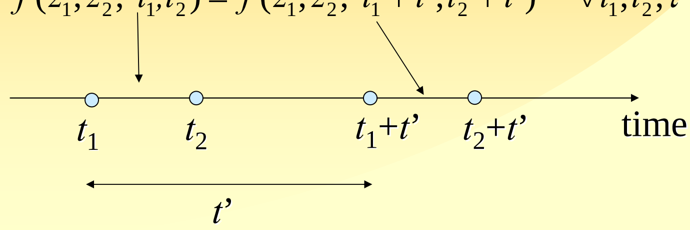
- random function is stationary
- time interval/area is large compared to correlation scale

Universiteit Utrecht

Second order stationary RF

Bivariate pdf invariant under translation:

$$f(z_1, z_2;\ t_1, t_2) = f(z_1, z_2;\ t_1 + t', t_2 + t') \quad \forall t_1, t_2, \tau$$



Universiteit Utrecht

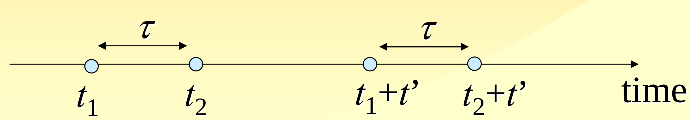
Wide sense stationary RF

Time:

$\mu_Z(t) = \mu_Z$ is constant

$\sigma_Z^2(t) = \sigma_Z^2$ is constant and finite

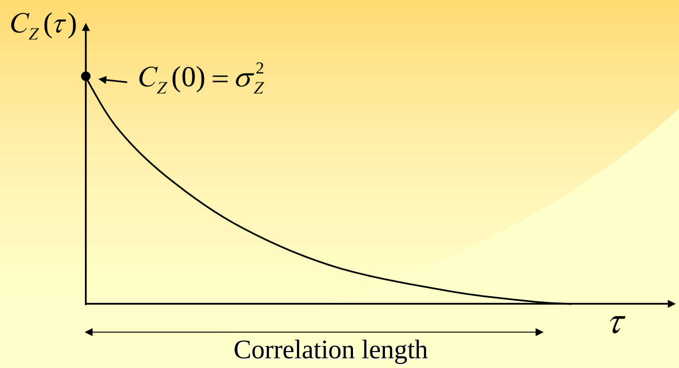
$\text{COV}[Z(t_1), Z(t_2)] = C_Z(t_2 - t_1) = C_Z(\tau)$



Covariance function $C(\tau)$

Universiteit Utrecht

Covariance function

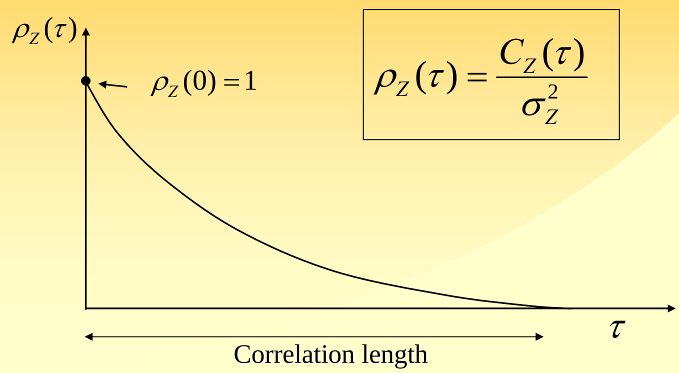


Note: $C(\tau) = C(-\tau)$

time

Universiteit Utrecht

Correlation function



$$\rho_Z(\tau) = \frac{C_Z(\tau)}{\sigma_Z^2}$$

time

Universiteit Utrecht

Universiteit Utrecht

Stochastic Hydrology

Random functions (2)

Marc F.P. Bierkens

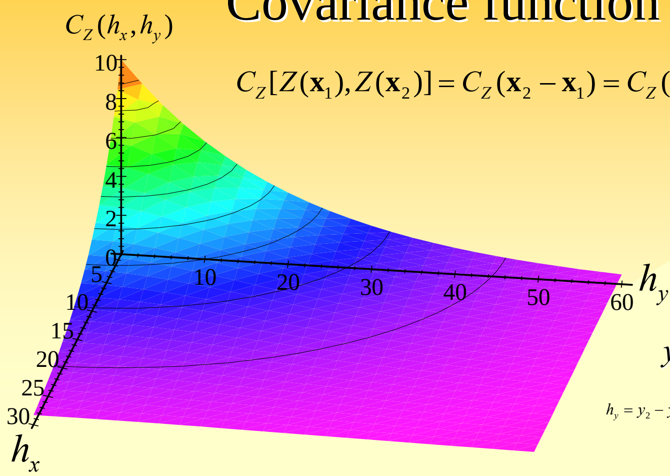
Professor of Hydrology
Faculty of Geosciences

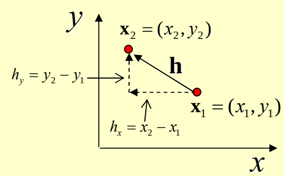
Universiteit Utrecht

space

Covariance function

$$C_Z[Z(\mathbf{x}_1), Z(\mathbf{x}_2)] = C_Z(\mathbf{x}_2 - \mathbf{x}_1) = C_Z(\mathbf{h}) = C_Z(h_x, h_y)$$





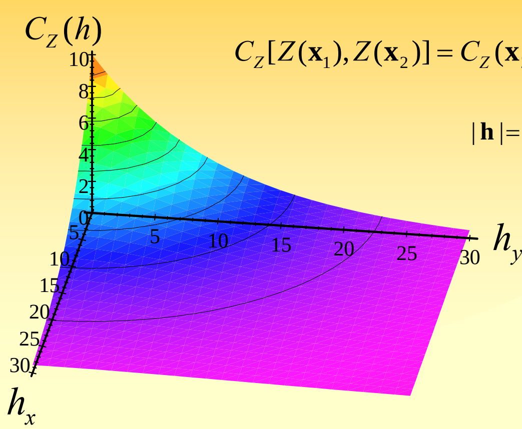
Universiteit Utrecht

space

Covariance function

$$C_Z[Z(\mathbf{x}_1), Z(\mathbf{x}_2)] = C_Z(\mathbf{x}_2 - \mathbf{x}_1) = C_Z(|\mathbf{h}|) = C_Z(h)$$

$$|\mathbf{h}| = h = \sqrt{h_x^2 + h_y^2}$$



Isotropy: both correlation scales are the same

Universiteit Utrecht

Estimation of Covariance functions

Time

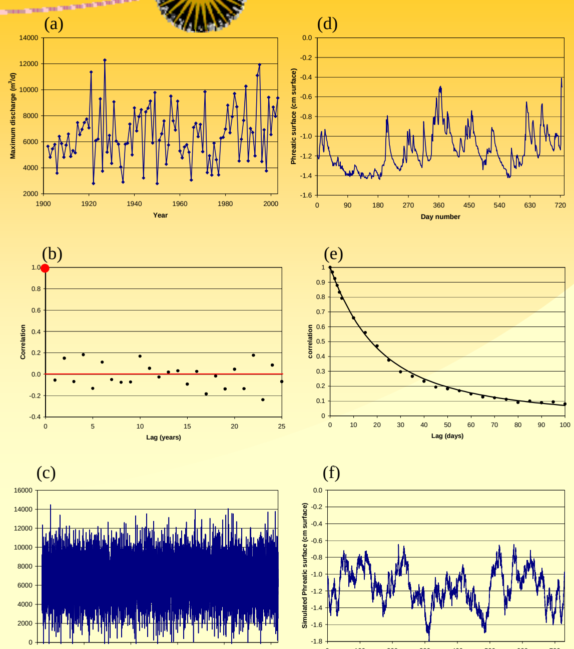
$$\hat{C}_Z(k\Delta t) = \frac{1}{n-k}\sum_{i=1}^{n-k}(z_i - \hat{\mu}_Z)(z_{i+k} - \hat{\mu}_Z)$$

Space

$$\hat{C}_Z(\mathbf{h}) = \frac{1}{n(\mathbf{h})}\sum_{i=1}^{n(\mathbf{h})}[z(\mathbf{x}) - \hat{\mu}_Z][z(\mathbf{x}_i + \mathbf{h} \pm \Delta\mathbf{h}) - \hat{\mu}_Z]$$

Universiteit Utrecht

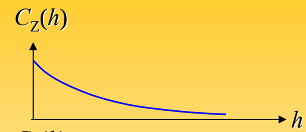
Influence of correlation



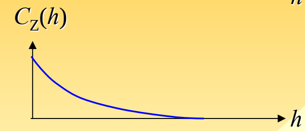
Universiteit Utrecht

Covariance functions

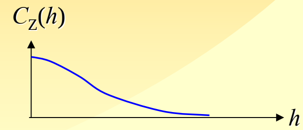
Exponential



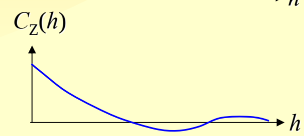
Spherical



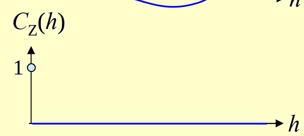
Gaussian



Hole effect (wave)



White noise



Universiteit Utrecht

Measures of correlation

Time and space

- Correlation length or (effective) range:

  Lag τ at which $\rho(\tau)=0$ (spherical) or $\rho(\tau)<0.05$ (exponential)

Time

- Integral scale:

$$I_{Z(t)} = \int_0^\infty \rho(\tau)\,d\tau$$

- Scale of flucuation:

$$\theta = \int_{-\infty}^{\infty} \rho(\tau)\,d\tau = 2I_{Z(t)}$$

Universiteit Utrecht

Measures of correlation

Space (2D and 3D)

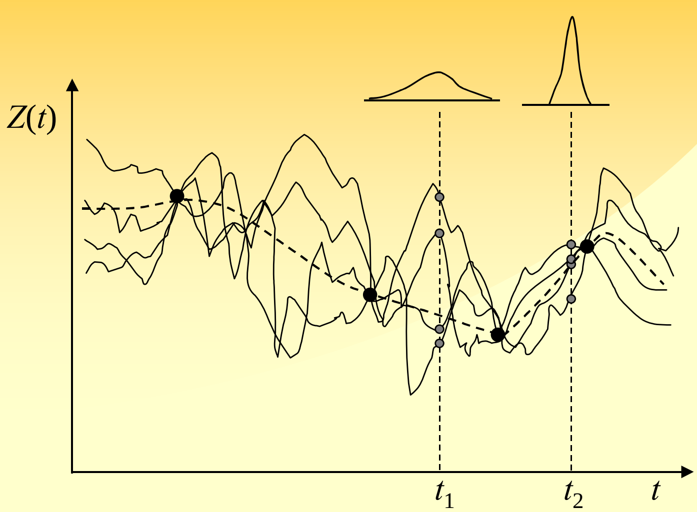
- Integral scale:

2D: $$I_{Z(\mathbf{x})} = \left[\frac{4}{\pi} \int_0^\infty \int_0^\infty \rho(h_1, h_2) dh_1 dh_2 \right]^{1/2}$$

3D: $$I_{Z(\mathbf{x})} = \left[\frac{6}{\pi} \int_0^\infty \int_0^\infty \rho(h_1, h_2, h_3) dh_1 dh_2 dh_3 \right]^{1/3}$$

Universiteit Utrecht

Conditional random functions



Universiteit Utrecht

Conditional random functions

Conditional random function: $Z(t \mid y_1,..,y_m)$

Conditional density function at $t=t_1$: $f_Z(z;t_1 \mid y_1,..,y_m)$

Conditional mean and variance $t=t_1$:

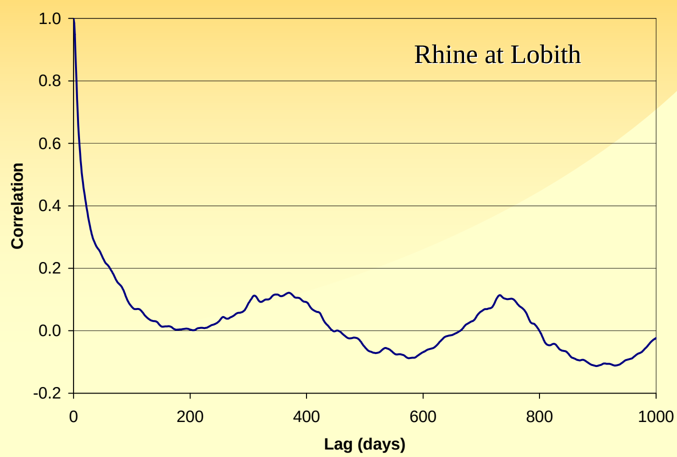
$$\mu_{Z|Y} = E[Z(t) \mid y_1,..,y_m] = \int_{\infty}^{\infty} f_Z(z;t_1 \mid y_1,..,y_m)dz$$

$$VAR[(t \mid y_1,..,y_m)] = \int_{\infty}^{\infty} (z-\mu_{Z|Y})^2 f_Z(z;t_1 \mid y_1,..,y_m)dz$$

Universiteit Utrecht

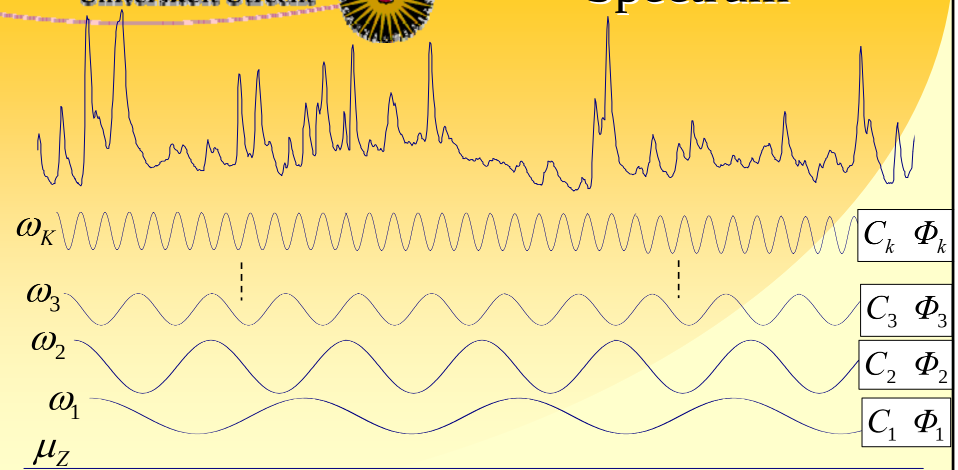
Spectrum

Many hydrological time series have a seasonal character



Universiteit Utrecht

Spectrum



$$Z(t) = \mu_Z + C_1 \cos(\omega t + \Phi_1) + C_2 \cos(2\omega t + \Phi_2) + \ldots + C_K \cos(K\omega t + \Phi_K)$$

Stationary random functions

Universiteit Utrecht

Spectrum

Then we have: $\sigma_Z^2 = \int_0^\infty G_Z(\omega) d\omega$

$G_Z(\omega)$: (one-sided) spectral density function

$G_Z(\omega)$: The magnitude of the harmonic signal with wavelength ω in the process Z(t)

$$\sigma_Z^2 \sim \sum_i Var[C_k]$$

Spectrum

Spectrum and covariance are related:

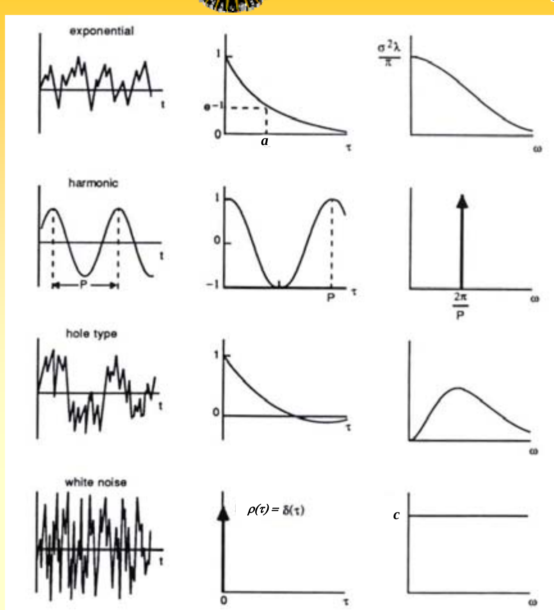
$$C_Z(\tau) = \int_0^\infty G_Z(\omega)\cos(\omega\tau)\,d\omega$$

$$G_Z(\omega) = \frac{2}{\pi}\int_0^\infty C_Z(\tau)\cos(\omega\tau)\,d\omega$$

Wiener-Khinchine relations

Universiteit Utrecht

Spectrum



Universiteit Utrecht

Spectrum

Spectrum can be used for:

- Analysis of periodicity in data
- Filtering (leve out certain wavelengths)
- Solving stochastic differential equations

Universiteit Utrecht

Spectrum Rhine data set