Random functions

Marc F.P. Bierkens

Professor of Hydrology
Faculty of Geosciences

Introduction

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

$\rightarrow$ 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 $Z(x,y,z,t)$

$\rightarrow$ Random Function
Average daily discharge of the river Meuse at Eysden (m$^3$/s)

Exhaustive TDR data set (2000 observations) Tarrawarra Catchment Australia

Exhaustive TDR data set (2000 observations) Tarrawarra Catchment Australia


Concept of a random function

$Z(x,y,t)$:

$Z(t)$:

ensemble

realisation
Random functions of different domains

- Random time function (RTF), random process or stochastic process:
  \[ Z(t) \]

- Random space function or random field:
  \[ Z(x), x = (x, y) \text{ or } x = (x, y, z) \]

- Random space-time function or space-time random field:
  \[ Z(x, t) \]

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.
Random functions are dependent RVs

Univariate Statistics of random functions

Mean:

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

Variance:

\[ \sigma_Z^2(t) = \mathbb{E}[\{Z(t) - \mu_Z(t)\}^2] = \int_{-\infty}^{\infty} \{z - \mu_Z(t)\}^2 f_Z(z; t) \, dz. \]
Bivariate Statistics of random functions

Bivariate pdf

\[ f(z_1, z_2; t_1, t_2) = \lim_{\varepsilon_1, \varepsilon_2 \to 0} \frac{\Pr[z_1 < Z(t_1) \leq z_1 + \varepsilon_1, z_2 < Z(t_2) \leq z_2 + \varepsilon_2]}{\varepsilon_1 \varepsilon_2} \]

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

Covariance:

\[ \text{COV}[Z(t_1), Z(t_2)] = \mathbb{E}[\{Z(t_1) - \mu_z(t_1)\}\{Z(t_2) - \mu_z(t_2)\}] = \iint_{-\infty}^{\infty} \{z_1 - \mu_z(t_1)\}\{z_2 - \mu_z(t_2)\} f(z_1, z_2; t_1, t_2) \, dz_1 \, dz_2 \]

Multivariate pdf of a random function

Also called: multi-point pdf

\[ f(z_1, z_2, \ldots, z_N; t_1, t_2, \ldots, t_N) = \]

\[ \lim_{\varepsilon_1, \ldots, \varepsilon_N \to 0} \frac{\Pr(z_1 < z(t_1) \leq z_1 + \varepsilon_1, z_2 < z(t_2) \leq z_2 + \varepsilon_2, \ldots, z_N < z(t_N) \leq z_2 + \varepsilon_N)}{\varepsilon_1 \varepsilon_2 \ldots \varepsilon_N} \]

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

1) Value: 
   a) discrete valued $D$
   b) continuous valued $Z$

2) Domain type: 
   a) time $t$
   b) space $x$
   c) space-time $(x,t)$

3) Domain definition: 
   a) continuous $x$, $t$
   b) discrete (at finite number of points) $x_k$, $t_k$
   c) random points (point process) $X_j, T_k$
Strict stationary random functions

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

\[ t' \quad \text{time} \]

Strict stationary random functions

\[ f(z_1, z_2, \ldots, z_N; x_1, x_2, \ldots, x_N) = f(z_1, z_2, \ldots, z_N; x_1 + h, x_2 + h, \ldots, x_N + h) \quad \forall x, h \]

\[ |h| = \sqrt{h_x^2 + h_y^2} \]

Strict stationary RF

Strict Stationary and isotropic RF

space
Ergodic random functions

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

Ergodic random functions

e.g. ergodicity in the mean:

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

Requirements:
• random function is stationary
• time interval/area is large compared to correlation scale
Second order stationary RF

Bivariate pdf invariant under translation:

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

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 \( C(\tau) \)
Covariance function

\[ C_z(\tau) = \sigma_z^2 \]

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

Correlation function

\[ \rho_z(\tau) = \frac{C_z(\tau)}{\sigma_z^2} \]

\[ \rho_z(0) = 1 \]
Random functions (2)

Marc F.P. Bierkens

Professor of Hydrology
Faculty of Geosciences
Covariance function

\[ C_Z[h(x_1), h(x_2)] = C_Z(x_2 - x_1) = C_Z(h) = C_Z(h_x, h_y) \]

Isotropy: both correlation scales are the same

\[ |h| = h = \sqrt{h_x^2 + h_y^2} \]
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

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

Influence of correlation
Covariance functions

Exponential

Spherical

Gaussian

Hole effect (wave)

White noise

Measures of correlation

Time and space

- Correlation length or (effective) range:
  Lag $\tau$ 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 fluctuation:
    $$\theta = \int_{-\infty}^{\infty} \rho(\tau) d\tau = 2I_{Z(t)}$$
Measures of correlation

Space (2D and 3D)

• Integral scale:

\[
2D: \quad I_{Z(x)} = \left[ \frac{4}{\pi} \int_{0}^{\infty} \int_{0}^{\infty} \rho(h_1, h_2) \, dh_1 \, dh_2 \right]^{1/2}
\]

\[
3D: \quad I_{Z(x)} = \left[ \frac{6}{\pi} \int_{0}^{\infty} \int_{0}^{\infty} \int_{0}^{\infty} \rho(h_1, h_2, h_3) \, dh_1 \, dh_2 \, dh_3 \right]^{1/3}
\]

Conditional random functions

\[ Z(t) \]

\[ t_1 \quad t_2 \quad t \]
Conditional random functions

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

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

Conditional mean and variance \( t=t_1 \):

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

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

Spectrum

Many hydrological time series have a seasonal character

![Spectrum Graph](image)
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

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 \omega in the process Z(t)

\sigma_Z^2 \sim \sum_i Var[C_i]
Spectrum

Spectrum and covariance are related:

\[ C_Z(\tau) = \int_{-\infty}^{\infty} G_Z(\omega) \cos(\omega \tau) d\omega \]

\[ G_Z(\omega) = \frac{2}{\pi} \int_{0}^{\infty} C_Z(\tau) \cos(\omega \tau) d\omega \]

Wiener-Khinchine relations

- exponential
- harmonic
- hole type
- white noise
Spectrum

Spectrum can be used for:

• Analysis of periodicity in data
• Filtering (lev out certain wavelengths)
• Solving stochastic differential equations

Spectrum Rhine data set