

MEMBANGKITKAN DATA SINTETIK (SYNTHETIC DATA GENERATING)

Masalah: kekurangan data

Tujuan:

- Untuk memperoleh data deret berkala buatan (*artificially generating time series*) atau data sintetik yang berasal dari data historis.
- Untuk memperpanjang rekaman data sehingga mempunyai beberapa alternatif dalam hal analisis teknis maupun ekonomis dari suatu proyek sumber daya air.

Berbeda dengan data simulasi yaitu data keluaran perhitungan model.

Deret berkala mengandung: trend, periodik dan stokastik

Komponen trend dan periodik bersifat pasti (*deterministic*) dan tdk tergantung waktu sedangkan komponen stokastik bersifat stasioner dan tergantung waktu, artinya sifat statistik dari sampel tidak berbeda dari sifat populasinya.

Metode stokastik dapat membangkitkan data berkala tahunan atau bulanan, metode yang digunakan adalah Penggunaan tabel bilangan acak dan proses Markov.

1. tabel bilangan acak

Digunakan tabel bilangan acak dengan memilih bilangan secara sembarang, banyaknya bilangan acak yang diambil tergantung dari jumlah nilai deret berkala buatan yang akan dibangkitkan. Bentuk persamaannya

$$X = \bar{X} \pm S. k$$

Dengan X debit yang dibangkitkan, \bar{X} debit rerata dari data historis dan k wilayah luas di bawah kurva normal. k dihitung dengan mencari nilai acak dari tabel secara sembarang, kemudian dicari nilai peluang, misal nilai acak dari tabel 3291 maka peluang $1 - 0,3291 = 0.6709$ maka nilai ini dicari di tabel wilayah luas di bawah kurva normal diperoleh $+0,44$. Dengan cara ini perkiraan debit banjir hanya disarankan sampai periode ulang sebesar 2 kali lama ketersediaan data (data historis).

2. proses Markov

Menggunakan model auto-regresif tahunan, model yang paling sederhana adalah Model Markov-Chain:

$$X_i = \Gamma(X_{i-1}) + (1-\Gamma) \bar{X} + S. t. (1-\Gamma^2)^{1/2}$$

dengan,

X_i = debit tahunan pada tahun ke t

X_{i-1} = debit tahunan pada tahun ke t-1

t = variat acak dari distribusi normal dengan rata-rata 0 dan deviasi standar 1

\bar{X} = debit rerata tahunan dari data historik

Γ = koefisien Markov-Chain,

Cara I: nilainya berkisar antara 0.20 – 0.30 umumnya digunakan 0.25

Cara II: Dihitung sebagai koefisien korelasi serial lag-1

$$\Gamma_1 = \frac{\sum_{i=1}^n (X_i)(X_{i+1}) - \frac{1}{n-1} \left(\sum_{i=1}^{n-1} X_i \right) \left(\sum_{i=2}^{n-1} X_i \right)}{\left[\sum X_i^2 - \frac{1}{n-1} \left(\sum_{i=1}^{n-1} X_i \right)^2 \right]^{0.5} \left[\sum_{i=2}^{n-1} X_i^2 - \frac{1}{n-1} \left(\sum_{i=2}^{n-1} X_i \right)^2 \right]^{0.5}}$$

persamaan di atas dapat disederhanakan menjadi

$$\Gamma_1 = \frac{f_1 - \frac{1}{n-1} (f_2)(f_3)}{\sqrt{f_4} \sqrt{f_5}}$$