KAJIAN PANJANG DATA HISTORIS YANG REPRESENTATIF PADA MODEL STOKASTIK

Setiarso Gunawan[1], Sri Eko Wahyuni[2], Suharyanto[2]

ABSTRACT

Stochastic models are models to generate new data series based on historical data and have similar statistical parameter with statistic historical data. Methods of forecasting are developed base on statistic and mathematic science. The historical data are observed data or sample data. The limited data is become main constrain for extrapolation of data. The mean error of generated data should be lower than 5%, its mean data of generated have the validation rate on 95 %. Three samples location for study are Catchment of Bengawan Solo in Bojonegoro, Catchment of Serang in Kedungombo - Grobogan and Catchment of Citarum in Cirata - Bandung. The synthetic data and then is used to calculate the statistic parameter. Error of generated data is measured with relative error. The relative error is result of divided and subtract statistic parameter of generated data and the statistic parameter of historical data longest and statistic parameter of generated data. The result of data length analysis is relative error and historical length of the data. The analyzed result indicate that historical data are studied have representative historical data about 30 years length of data.

Keywords **:** *stochastic, historical data,synthetics data, representative data length and relative error.*

PENDAHULUAN

Panjang data historis sebagai masukan untuk mendapatkan gambaran sebenarnya fenomena hidrologi yang terjadi, sangat penting ditetapkan batas minimumnya. Panjang data historis yang memadai dan representatif pada suatu DPS sangat berpengaruh pada proses dan hasil perencanaan bangunan pengembangan sumber daya air (Soemarto, 1987).

Pembangkitan data historis menjadi data baru dilakukan dengan 3 (tiga) model stokastik, yaitu model Markov untuk data aliran historis tahunan, model Thomas-Fiering dan model Box-Jenkins (ARIMA) untuk data aliran historis bulanan. Hasil pembangkitan digunakan untuk mengkaji hubungan antara kemiripan sifat statistik data hasil bangkitan dengan panjang data historis yang dipakai untuk membangkitkan data tersebut.

Untuk menentukan panjang data historis yang representatif dilakukan pengkajian terhadap kesalahan relatif antara data hasil bangkitan terhadap data populasinya. Hubungan antara kesalahan relatif dan panjang data tersebut ditunjukkan dengan persamaan regresi garis yang menggambarkan tingkat keeratan hubungan. Panjang data historis yang memenuhi tingkat kesalahan 5 % ditentukan melalui substitusi persamaan regresi garis tersebut. Ketersediaan data historis dapat menimbulkan variasi data hasil pembangkitan yang signifikan.

[1] Magister Teknik Sipil FT. Universitas Diponegoro Jl. Hayam Wuruk Semarang

[2] Jurusan Teknik Sipil FT. Universitas Diponegoro Jl. Prof. Soedarto, SH Tembalang Semarang

Informasi yang didapat dari data pendek amat sedikit, sehingga sering diperlukan pembangkitan data (*data generation*). Filosofi pembangkitan data adalah membuat rangkaian data baru berdasarkan data historis yang umumnya pendek untuk mendapatkan data yang lebih panjang.

Sebagian besar fenomena hidrologi dikatagorikan sebagai proses stokastik, dengan disertai kondisi yang terdapat ketidakpastian pada prosesnya. Sehingga untuk memperoleh data baru diperlukan penelitian tentang stokastik yang berkaitan dengan data historis yang representatif untuk pembangkitan. Data historis tersebut harus diaplikasikan pada model stokastik yang memenuhi persyaratan secara teknis dan statistik. Pengumpulan data harus dipakai tiga atau lebih lokasi penelitian yang tersebar. Dipilih lokasi DPS yang luasnya lebih dari 100 km^2 dan tersebar, yaitu DPS Bengawan Solo di stasiun hidrometri Bojonegoro (Jawa Timur), DPS Serang di stasiun hidrometri waduk Kedungombo (Jawa Tengah) dan DPS Citarum di stasiun hidrometri waduk Cirata (Jawa Barat). Pada data sintetik hasil pembangkitan kemudian dilakukan uji dan analisis statistik guna menentukan panjang data historis yang representatif.

Gambar 1. Lokasi Penelitian

TINJAUAN PUSTAKA

Umum

Stokastik adalah suatu hal tentang ketidakpastian. Fenomena hidrologi merupakan daur stokastik. Daur dan stokastik dalam proses hidrologi ini sama pentingnya dalam penelaahan hidrologi (Tao, 1976). Data daur yang deterministik dapat dijelaskan dengan persamaan matematika eksplisit, sedangkan data acak yang stokastik tidak ada persamaan eksplisitnya karena tiap-tiap data bersifat unik (Raudkivi, 1979).

Pada deret acak terdapat koefisien korelasi untuk semua lag mempunyai kesalahan baku akibat jumlah sampel sebesar $\pm 1/\sqrt{n}$. Untuk lag satu, uji yang pasti menunjukkan bahwa r_1 benar-benar bukan nol adalah bila nilainya berada di luar rentang untuk batas konviden 95% (Salas et. al., 1980).

Model Hidrologi

Hidrologi adalah ilmu yang membahas tentang air yang ada di bumi, yaitu kejadian, sirkulasi dan penyebaran, sifat-sifat fisis dan kimiawi serta reaksinya terhadap lingkungan, termasuk hubungannya dengan kehidupan. Hidrologi teknik mencakup bidang yang berhubungan langsung dengan perencanaan, perancangan, dan pelaksanaan proyek-proyek teknik bagi pengaturan dan pemanfaatan air (Linsley et.al., 1982).

Model hidrologi secara umum dapat dijabarkan sebagai sebuah sajian sederhana *(simple representation)* dari sebuah sistem hidrologi yang kompleks (Sri Harto dan Sudjarwadi, 1989). Sedangkan menurut Eko Wahyuni (1999) adalah suatu model yang dapat mensimulasi berbagai proses hidrologi yang terjadi pada suatu daerah tangkapan dengan menggambarkan proses fisis yang sesungguhnya dari siklus hidrologi, atau menirukan peristiwa hidrologi yang terjadi.

Penentuan parameter hidrologi yang terkait dengan pembuatan model dapat dilakukan dengan menentukan tujuan model tersebut dibuat. Model stokastik mempunyai keterkaitan dengan parameter berpengaruh berupa musim, distribusi debit, sistem daerah aliran sungai yang akan dipakai sebagai model dan autokorelasi data itu sendiri.

Di dalam penyelesaian masalah-masalah hidrologi stokastik diperlukan suatu model stokastik. Menurut Mediana (1988), model stokastik sangat diperlukan dalam hubungannya dengan ketersediaan data yang terbatas.

Pembangkitan Data

Pembangkitan data debit sintetik pertama kali dilakukan oleh Sudler tahun 1927 (Kottegoda, 1980), dan dilengkapi dengan penggunaan tabel angka acak dengan distribusi dikenal sebagai Metode Monte Carlo (Raudkivi, 1979). Pada umumnya model stokastik berdasarkan model Autoregresif (*Markov Chain*). Untuk data bulanan atau selang waktu yang lebih pendek, distribusi datanya dianggap makin menjauhi normal (Gauss) dan lebih mendekati distribusi Gamma (Pearson). Beberapa teknik untuk merubah deret data agar mendekati normal telah banyak dipakai. Pengubahan ini dikenal sebagai pemutihan (*prewhitening*). Distribusi normal dapat digunakan untuk data tahunan dan bulanan (Mediana, 1988).

Model Stokastik

Model stokastik menurut Soemarto (1987) adalah model hidrologi dengan basis matematik yang menghasilkan suatu urutan nilai yang merupakan hasil dari proses acak dengan cara memasukkan probabilitas. Simulasi model ini akan memberikan data sintetik dengan nilai tengah (*mean*), tingkat keragaman (*variance*), kesalahan standar (*standart deviation*) yang tetap terpelihara. Sedangkan model stokastik menurut Eko Wahyuni (2001) adalah model hidrologi yang selalu berkisar dengan waktu, mewakili suatu urutan peristiwa dan selalu dipengaruhi oleh peristiwa sebelumnya.

- **Model Markov**

Metode paling sederhana dan banyak dipakai untuk membangkitkan data hidrologi adalah model Autoregresif atau disebut juga rantai Markov menurut nama ahli matematika A.A. Markov (1856-1922). Model tersebut dapat ditulis sebagai berikut:

$$(y_t - \mu = \beta_1 (y_{t-1} - \mu) + \beta_2 (y_{t-2} - \mu) + \ldots\ldots\ldots$$
$$+ \beta_k (y_{t-k} - \mu) + \varepsilon_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

di mana:

μ = rata-rata dari populasi, yang diperkirakan sama dengan rata-rata per sampel

β = parameter yang didapat dari koefisien korelasi antara satu variabel dan variabel sebelumnya

ε_t = bilangan acak dengan rata-rata nol dan deviasi standar tertentu

Clarke (1973) menyatakan bahwa makin besar daerah pengaliran sungai (DPS), makin besar korelasi di antara jumlah aliran tahunan.

- **Model Thomas-Fiering**

Thomas dan Fiering mengembangkan model untuk membangkitkan aliran sungai bulanan. Secara implisit, model ini mengijinkan adanya ketidakstasionairan dalam data aliran bulanan (Clarke, 1973). Adanya persistensi disebabkan oleh efek penyimpanan air sebagai lengas tanah dan air tanah (Raudkivi, 1979).

Secara umum, persamaannya dituliskan sebagai berikut:

$$Q_{i+1} = \bar{Q}_{i+1} + b_i (Q_i - \bar{Q}_i) + t_i S_{i+1} \sqrt{1 - r_i^2} \quad \text{...} (2)$$

dimana :

Q = debit bulanan (m^3/dt)
ii = indeks, dari 1 sampai 12
b_i = (r_i.S_i+1)/Si
r = koefisien korelasi antara debit bulan ke i dan bulan ke i+1
S = deviasi standar
t = bilangan acak, biasanya merupakan variabel bebas berdistribusi normal dengan rata-rata 0 dan varian 1.

Model Thomas-Fiering ini biasanya dipakai untuk sungai dengan aliran *perennial*.

- **Model Box-Jenkins**

Box dan Jenkins mengembangkan model yang dapat dipakai untuk data yang tak stasionair. Suatu variabel y_t dihubungkan dengan variabel acak bebas ε_t dan nilai-nilai y_{t-1}, y_{t-2},... yang mendahuluinya dalam deret dinyatakan dengan persamaan:

$$\phi(B)\, \nabla^d y_t = \theta(B)\, \varepsilon_t \quad \text{........................} (3)$$

dengan ϕ(B) dan θ(B) merupakan polinomial berderajat p dan q.

Operator *backward* B didefinisikan sebagai $By_t = y_{t-1}$. Operator *difference* ∇didefinisikan sebagai : $\nabla y_t = y_t - y_{t-1} = (1 - B)\, y_t$; yang mana operator ini digunakan untuk menghilangkan kecenderungan atau *trend*. Pembedaan (d*ifferencing*) tingkat 1 (d = 1) adalah : $u_t = x_t - x_{t-1}$

Identifikasi Model pada Model Stokastik

Identifikasi adalah merupakan langkah dalam pembuatan model deret berkala di mana pola dalam statistik seperti fungsi autokorelasi, fungsi autokorelasi parsial dan sebagainya dihubungkan terhadap model tentatif untuk data. Identifikasi dapat dilakukan dengan perhitungan autokorelasi, yaitu suatu asosiasi atau ketergantungan bersama (*mutual dependence*) antara nilai – nilai suatu deret berkala yang sama pada periode waktu yang berlainan, atau selang waktu (*time lag*) yang berbeda dan perhitungan autokorelasi parsialnya untuk menunjukkan hubungan antara nilai suatu variabel saat ini dengan nilai sebelumnya dari variabel yang sama dengan menganggap pengaruh dari semua kelambatan waktu lainnya adalah konstan.

Estimasi Parameter Model Stokastik

Parameter adalah merupakan karakteristik dari suatu populasi seperti nilai rata-rata atau deviasi. Penentuan parameter model dilakukan dengan menentukan tujuan model stokastik dibuat (Makridakis et. al., 1998). Model stokastik mempunyai keterkaitan dengan parameter berpengaruh berupa musim, distribusi debit, sistem daerah aliran sungai yang akan dipakai sebagai model dan autokorelasi data itu sendiri.

Parameter model stokastik berupa konstanta, variabel bebas yang berupa koefisien – koefisien deret berkala (*time series*): koefisien autoregresif dan koefisien *Moving Average*, bilangan acak dan koefisien musiman.

Uji Diagnostik

Pemeriksaan diagnostik adalah merupakan salah satu tahap pembuatan model stokastik di mana nilai taksiran kesalahan model diuji tentang kebebasan, nilai tengah nol, ketepatan ragamnya dan lain-lain dan untuk mendapatkan kevalidan suatu model.

Untuk mendapatkan suatu model yang representatif diperlukan model yang dapat diukur ketepatannya dengan data historisnya. Ketepatan dapat diukur menggunakan dimensi seperti rata-rata kesalahan kuadrat (*mean square error*, *MSE*); rata-rata kesalahan presentase absolut (*mean absolute percent error*, *MAPE*) atau rata-rata kesalahan presentase atau bias (*mean percent error*, *MPE*). Untuk mendapatkan model yang representatif diperlukan uji statistik yaitu Uji chi – kuadrat, Uji-t , F-test dan Uji D – W.

Analisis Distribusi

Analisis distribusi sesungguhnya merupakan prakiraan (*forecasting*) dalam arti probabilitas untuk terjadinya suatu peristiwa hidrologi. Distribusi kemungkinan teori *probability distribution* yang biasa digunakan adalah Normal (*Gauss*), Log Normal, Pearson III (*Gamma*), Log Pearson tipe III dan Gumbel.

Tabel 1. Jenis Distribusi

No	Jenis Disstribusi	Syarat
1	Normal (Gauss)	$C_s = 0$ $C_k = 3$
2	Log Normal	C_s (In x) = 0 C_k (In x) = 3
3	Pearson Tipe III (Gamma)	$C_s > 0$ $C_k = 1{,}5\ C_s^2 + 3$
4	Log Pearson Tipe III	C_s (In x) = 0 C_k (In x) = 1,5 (C_s (In x)2) +3
5	Gumbell	$C_s = 1{,}14$ $C_k = 5{,}4$

Uji kecocokan data atau distribusi digunakan cara Uji Chi Kuadrat. Distribusi yang dipilih dan dianggap tidak cocok apabila harga X^2 melewati harga X^2 kritik.

Kesalahan Relatif dan Panjang Data

Kesalahan relatif adalah prosentase selisih besaran parameter statistik hasil pembangkitan data historis populasi dan besaran parameter statistik hasil pembangkitan data historis dengan panjang data tertentu dibagi besaran parameter hasil pembangkitan data historis populasi.

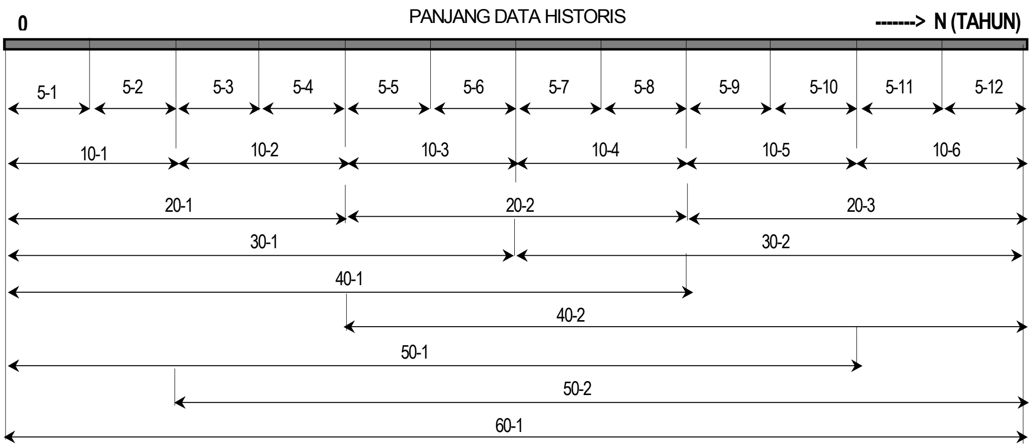
Sri Harto (1986) menunjukkan adanya hubungan antara panjang data yang tersedia dengan kesalahan yang terjadi. Semakin pendek data historis yang digunakan, maka semakin besar kesalahan relatif yang terjadi. Menurut Warsini (1987) suatu analisis hidrologi dengan data historis kurang dari 20 tahun akan memberikan kesalahan relatif lebih besar dari 3 %.

Penelitian Solomon (1970) menunjukkan kesalahan relatif limpasan tahunan rata-rata rencana dengan panjang data di Ontario dan Canada, kesalahan relatif 5 % terjadi pada panjang data historis 30 tahun; sedangkan penelitian Bayu Arianto (1988) pada DPS Brantas Hulu menunjukkan kesalahan relatif 5 % untuk panjang data antara 20 - 40 tahun sesuai dengan periode ulang hujan rancangannya. Hubungan antara kesalahan relatif dengan panjang data historis untuk dibangkitkan belum ditemukan.

ANALISIS MODEL STOKASTIK DAN PEMBAHASAN

Pembuatan model stokastik terdiri dari 3 model yang diteliti yaitu model Markov, Thomas-Fiering dan Box-Jenkins yang meliputi tahap-tahap pembentukan model yaitu prosedur dan validasi untuk model Markov dan Thomas-Fiering serta identifikasi, estimasi parameter model, dan validasi untuk model Box-Jenkins.

Pembangkitan direncanakan dengan skenario sebagai berikut :



Gambar 2. Skenario Analisis panjang Data Historis untuk Seri Simulasi Model Stokastik

Model Markov

- **Prosedur Model Markov**

1. Debit rata-rata tahunan dari data aliran historis dihitung dengan:

   $\mu = \sum q_i$ / n

2. Deviasi standar tahunan dihitung dengan:

   $\sigma = \sum (q_i - \mu)^2$ / (n-1)

3. Autokorelasi lag 1 dihitung dengan tahap-tahap berikut:

$$f_1 = \sum_{i=1}^{n-1} x_i - x_{i+1}$$

$$f_2 = \sum_{i=1}^{n-1} x_i$$

$$f_3 = \sum_{i=2}^{n} x_i \quad ;$$

$$f_4 = \sum_{i=1}^{n-1} x_1^2 - \left(\frac{1}{n-1}\sum_{i=1}^{n-1} x_i\right)^2$$

$$f_5 = \sum_{i=2}^{n} x_1^2 - (\frac{1}{n-1}\sum_{i=2}^{n} x_i)^2$$

Nilai autokorelasi lag 1 (r_1) adalah :

$$r_1 = \frac{f_1 - \frac{1}{n-1}(f_2 \times f_3)}{\sqrt{f_4} \times \sqrt{f_5}}$$

4. Debit pembangkitan dihitung dengan:
   $q_i = \mu + \rho (q_{i-1} - \mu) + t_i \sqrt{(1-\rho^2)}$

- **Validasi Model Markov**

Uji kesamaan nilai tengah antara data aliran historis dengan data sintetik hasil pembangkitan model stokastik Markov sebagai berikut :

1. Menentukan hipotesis sebagai berikut:
   H_0 : $\mu_1 = \mu_2$ (tidak ada perbedaan nyata)
   H_1 : $\mu_1 \neq \mu_2$ (ada perbedaan nyata)
2. Statistik penguji berdistribusi student-t dihitung dengan:

$$t = \frac{\overline{x_1} - \overline{x_2}}{\partial \left| \frac{1}{n_1 + n_2} \right|^{0.5}}$$

3. Derajat kebebasan dihitung dengan:
   DK = $n_1 + n_2 - 2$
4. Daerah kritik dua sisi, yaitu $-t_{cr}(k;\alpha)$ dan $t_{cr}(k_j\alpha)$ dengan derajat kebebasan DK dan tingkat kepercayaan α, diperoleh dari tabel distribusi –t, yaitu: $t_{cr}(98;0,025) = 1,960$
5. $H_0 : \mu_1 = \mu_2$ diterima, karena $-t_{cr} < t < t_{cr}$

Uji kesamaan varian antara data aliran historis dengan data sintetik hasil pembangkitan model stokastik sebagai berikut:

1. Menentukan hipotesis sebagai berikut:
   H_0 : $\mu_1 = \mu_2$ (tidak ada perbedaan nyata)
   H_1 : $\mu_1 \neq \mu_2$ (ada perbedaan nyata)
2. Statistik penguji berdistribusi F dihitung dengan:

$$F = \frac{n_1 . S_1{}^2 .(n_2 - 1)}{n_2 . S_2{}^2 .(n_1 - 1)}$$

3. Derajat kebebasan (DK) dihitung dengan:
   $DK_1 = n_1 - 1$
4. Luas daerah distribusi F, yaitu Fcr (DK_1; DK_2;α) dengan kebebasan data aliran historis DK_1 dan data sintetik DK_2 serta dengan tingkat kepercayaan α, diperoleh dari tabel distribusi F, yaitu:
   F_{cr} (DK_1;DK_2;α) = 1,59
5. $H_0 : \mu_1 = \mu_2$ diterima karena $F < F_{cr}$ (k_1 ; k_2 ; $\alpha/2$)

Model Thomas-Fiering

- **Prosedur Model Thomas-Fiering**

Prosedur pembangkitan data debit bulanan model Thomas-Fiering dapat ditulis sebagai berikut :

1. Debit rata-rata tiap bulan dari data aliran historis dihitung dengan:
   $q_j = \sum q_{ij} / n$
2. Deviasi standar tiap bulan dihitung dengan :
   $\sigma_j = \sum (q_{i,j} - q_i)^2 / (n-1)$
3. Koefisien korelasi serial tiap bulan data debit historis dihitung dengan:
   $\rho_j = \sum (q_{i,j} - q_i)(q_{i,j} - q_{i-j}) / (\sum (q_{i,j} - q_i)^2 (q_{i,j} - q_{i-j})^2)$
4. Koefisien regresi bulanan dihitung dengan:
   $\beta_j = \rho_j \, \sigma_j / \sigma_{j-1}$
5. Koefisien variasi bulanan dihitung dengan:

$Cv_j = \sigma_j / q_j$

6. Koefisien asimetri bulanan dihitung dengan:
   $\gamma_j = 3Cv_j / Cv_j^3$
7. Bilangan acak berdistribusi seragam antara 0 dan 1 dibuat dengan program komputer Excel 2000 dan diubah menjadi bilangan acak berdistribusi normal baku dengan dengan nilai tengah 0 dan variasi 1, dan mengoreksi menjadi distribusi Gamma.
8. Data sintetik model Thomas-Fiering dibangkitkan dengan:
   $$q_{i,j} = \overline{q_i} + \beta_j.\left(q_{i,j} - \overline{q_{j-1}}\right) + t_{i,\gamma,j}.\sigma_j\sqrt{\left(1-\rho_j^2\right)}$$

- **Validasi Model Thomas-Fiering**

Uji kesamaan nilai tengah antara data aliran historis dengan data sintetik hasil pembangkitan model stokastik Thomas-Fiering sebagai berikut :

1. Menentukan hipotesis sebagai berikut:
   H_0 : $\mu_1 = \mu_2$ (tidak ada perbedaan nyata)
   H_1 : $\mu_1 \neq \mu_2$ (ada perbedaan nyata)
2. Statistik penguji berdistribusi student-t dihitung dengan:
   $$t = \frac{\left|\overline{x_1} - \overline{x_2}\right|}{\sigma\left|\frac{1}{n_1} + \frac{1}{n_2}\right|^{0.5}}$$
3. Derajat kebebasan dihitung dengan:
   DK = $n_1 + n_2 - 2$
4. Daerah kritik dua sisi, yaitu $-t_{cr}(k;\alpha)$ dan $t_{cr}(k;\alpha)$ dengan derajat kebebasan DK dan tingkat kepercayaan α, diperoleh dari tabel distribusi –t, yaitu: $t_{cr}(98;0,025) = 1,960$
5. H_0 : $\mu_1 = \mu_2$ diterima karena $-t_{cr} < t < t_{cr}$

Uji kesamaan varian antara data aliran historis dengan data sintetik hasil pembangkitan model stokastik sebagai berikut:

1. Menentukan hipotesis sebagai berikut:
   H_0 : $S_1 = S_2$ (tidak terdapat perbedaan nyata)
   H_1 : $S_1 \neq S_2$ (terdapat perbedaan nyata)
2. Statistik penguji berdistribusi F dihitung dengan:
   $$F = \frac{n_1.S_1^2.(n_2 - 1)}{n_2.S_2^2.(n_1 - 1)}$$
3. Derajat kebebasan (DK) dihitung dengan:
   $DK_1 = n_1 - 1$
4. Luas daerah distribusi F, yaitu Fcr (DK_1; DK_2;α) dengan kebebasan data aliran historis DK_1 dan data sintetik DK_2 serta dengan tingkat kepercayaan α, diperoleh dari tabel distribusi F, yaitu:
   F_{cr} (DK_1; DK_2;α) = 1,55
5. H_0 : $\mu_1 = \mu_2$ diterima karena $F < F_{cr}$ (k_2 ; k_1 ; $\alpha/2$).

Model Box-Jenkins (ARIMA)

- **Identifikasi Model**

Indentifikasi jenis model ditentukan dengan menghitung autokorelasi dan autokorelasi parsial. Langkah-langkah menghitung autokorelasi adalah sebagai berikut :

1. Menghitung Covarian (COV) dengan rumus :
   $$COV\left(X_t, X_{t-1}\right) = \sum\frac{\left(X_t - \overline{X}\right)\left(X_{t-1} - \overline{X}\right)}{N}$$
2. Menghitung Auto Correlation (AC) dengan rumus :
   $$R_t = COR\left(X_t, X_{t-1}\right) = \frac{COV\left(X_t, X_{t-1}\right)}{STD\left(X_t\right)STD\left(X_{t-1}\right)}$$
   $$= \frac{COV\left(X_t, X_{t-1}\right)}{\sqrt{VAR\left(X_t\right)}\sqrt{VAR\left(X_{t-1}\right)}}$$

3. Memplotkan Auto Correlation (AC) ke dalam grafik Correlogram untuk mengidentifikasi model Moving Average atau ARIMA (0,0,q) dan Auto Regressive and Moving Average atau ARIMA(p,0,q).

Langkah-langkah menghitung autokorelasi parsial adalah sebagai berikut :

1. Menghitung Partial Auto Correlation (PAC) dengan memakai metode Invers Matriks pada persamaan:

$$\left.\begin{array}{l} R_1 = A_1 + A_2R_1 + \ldots + A_pR_{p-1} \\ R_2 = A_1R_1 + A_2 + \ldots + ApR_{p-2} \\ R_3 = A_1R_2 + A_2R_1 + \ldots + ApR_{p-3} \\ . \qquad\qquad\qquad\qquad . \\ R_p = A_1R_{p-1} + A_2R_{p-2} + \ldots + A_p \end{array}\right\} \text{Yule-Walker}$$

Dari persamaan tersebut kemudian disusun menjadi matriks yaitu :

Matrik Elemen: $\begin{vmatrix} 1 & R_1 & R_2 & R_p \\ R_1 & 1 & R_1 & R_{p-1} \\ R_2 & R_1 & 1 & R_{p-2} \\ R_p & R_{p-1} & R_{p-2} & 1 \end{vmatrix} \begin{vmatrix} A_1 \\ A_2 \\ A_3 \\ A_p \end{vmatrix} = \begin{vmatrix} R_1 \\ R_2 \\ R_3 \\ R_p \end{vmatrix}$

Penyelesaian persamaan dengan Matriks:

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots\ldots + a_{1n}x_n = b_1$$
$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots\ldots + a_{2n}x_n = b_2$$
$$a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \ldots\ldots + a_{nn}x_n = b_n$$

DeterminanA= $|A|$;AdjoinA=C^T;$A^{-1} = \frac{C^T}{|A|}$

2. Memplotkan Partial Auto Correlation (PAC) ke dalam Correlogram untuk mengidentifikasi model Auto Regresive atau ARIMA (p,0,0) atau Auto Regressive Moving Average ARIMA(p,0,q).

- **Validasi Model**

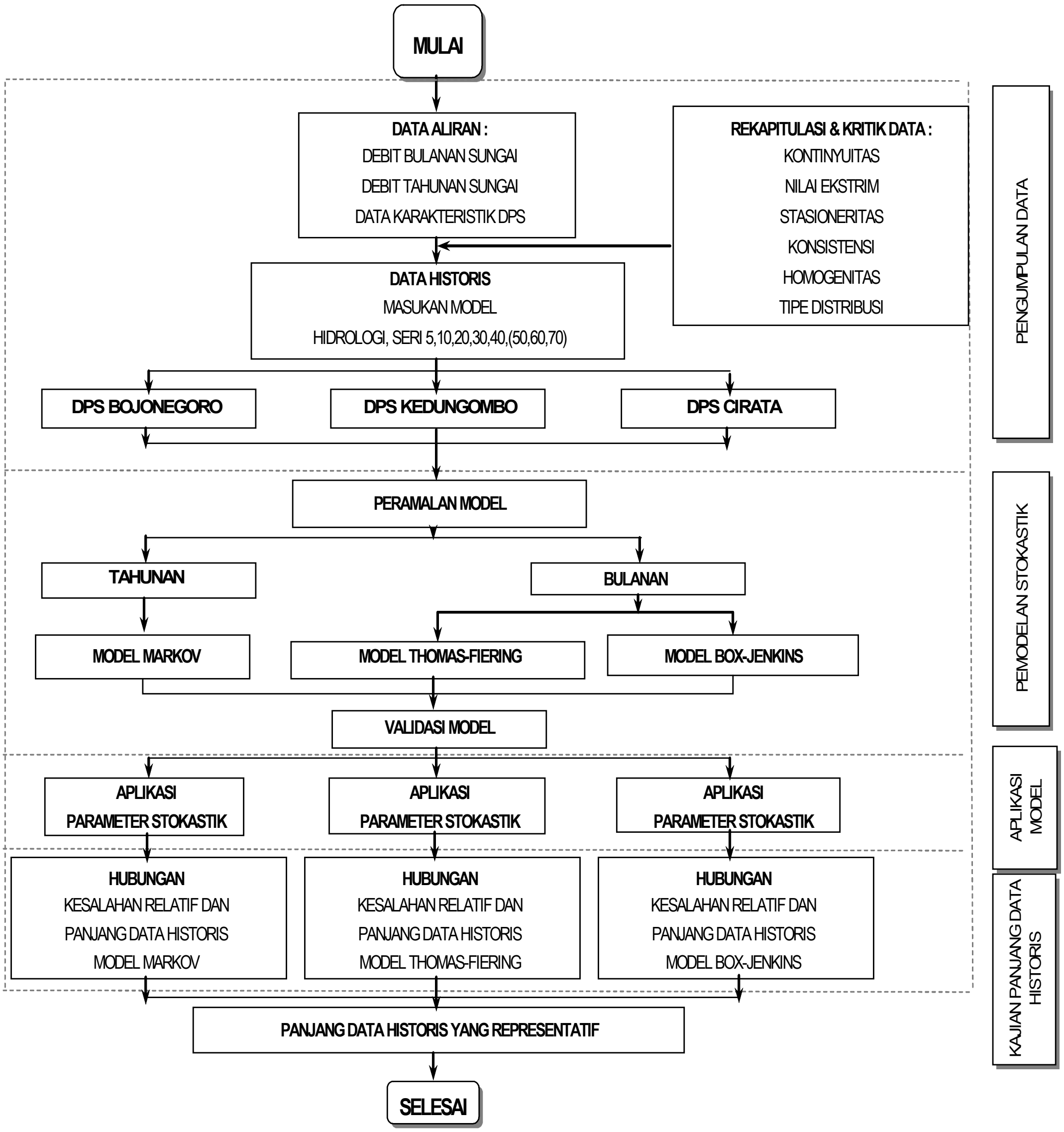
  Validasi model Box-Jenkins ditunjukkan dengan mencek residu model merupakan data acak.

Panjang Data Historis Yang Representatif

Kesalahan relatif dengan tingkat resiko 5 % digunakan untuk menarik garis plot hubungan antara ke duanya untuk mendapatkan panjang data historis yang representatif. Untuk mendapatkan panjang data aliran historis minimum yang representatif pada model stokastik dilakukan proyeksi dari kesalahan relatif dengan nilai 5 % ke absis data historis.

Untuk memperoleh persamaan regresi hubungan antara panjang data sebagai absis dan kesalahan relatif sebagai ordinat, digunakan analisis korelasi-regresi Logaritmik dengan persamaan umum: y = k (x-a)b . Menurut program paket Excel 2000 untuk mendapatkan persamaan regresi tersebut dipakai perintah Trendline - Regression Type – Logaritmic dengan persamaan y = a Ln(x) + b.

Untuk mengetahui nilai koefisien regresi dan persamaan regresi dipakai perintah *Display R square* dan *Display Equation* pada *Option.* Pada model Markov untuk panjang data – kesalahan relatif langsung (untuk nilai tengah) diperoleh persamaan y = - 6,3473 Ln (x) + 25,196 dengan R^2 = 0,60 dan untuk panjang data – kesalahan relatif langsung (untuk deviasi standar) diperoleh persamaan y = - 25,447 Ln (x) + 94,241 dengan R^2 = 0,46. Angka 0,60 dan 0,46 ini menunjukkan bahwa tingkat hubungan ke dua variabel panjang data dan kesalahan relatif langsung tersebut tidak terlalu kuat (yang baik sekitar 0,70), hal itu disebabkan pada seri-seri 5 tahunan mempunyai variasi yang besar.



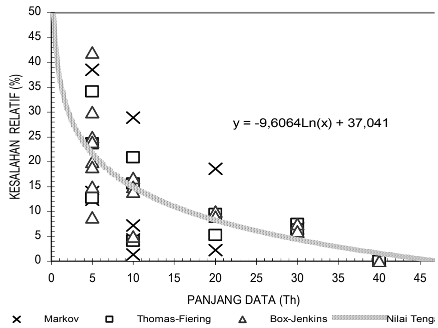
Gambar 3. Bagan Alir Kajian Panjang Data Historis yang Representatif pada Model Stokastik

Rekapitulasi hasil analisis model Markov, Thomas-Fiering dan Box-Jenkins berkaitan dengan parameter statistik, kesalahan relatif dan panjang data yang representatif ditampilkan pada Tabel di bawah ini.

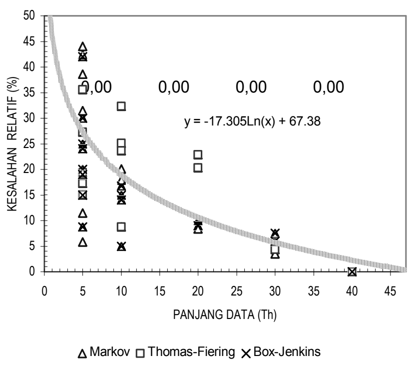
Tabel 2. Rekapitulasi Hasil Analisis Model Markov, Thomas-Fiering dan Box-Enkins

No	Daerah Pengaliran Sungai (Dps)	Lokasi/ Stasiun Hidro-Metri	Data Historis Tahunan			Pembangkitan Tahunan		Kesalahan Relatif 5%			Panjang Data Yang Repre Sentatif (Tahun)
			Panjang Data (Tahun)	Nilai Tengah	Deviasi Stan-Dar	Nilai Tengah	Deviasi Standar	Pnj. Data Markov (Tahun)	Pnj. Data Thomas-Fiering (Tahun)	Pnj. Data Box-Jenkins (Tahun)	
1	Bengawan Solo	Bojonegoro	40	325,44	86,63	330,92	89,97	24 - 26	25 - 36	24 - 30	27,3
2	Serang	Kedungombo	47	24,35	6,24	24,91	6,96	27 - 29	27 - 40	24 - 31	29,3
3	Citarum	Cirata	73	74,10	20,83	74,90	23,18	34 - 41	35 - 41	25 - 39	35,3
								Panjang Data yang Representatif			30

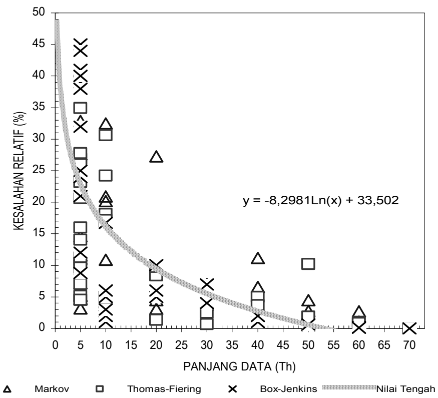
Hubungan kesalahan relatif dan panjang data yang representatif dapat dilihat pada gambar berikut ini.



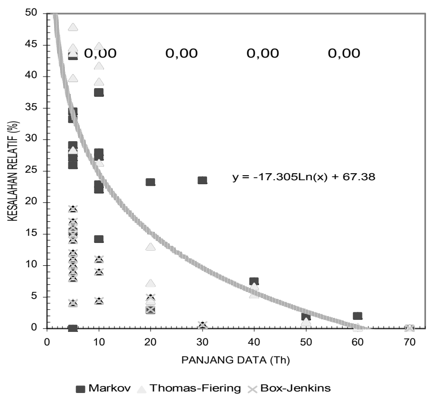
Gambar 4. Grafik Hubungan Kesalahan Relatif Nilai Tengah dan Panjang Data DPS Serang



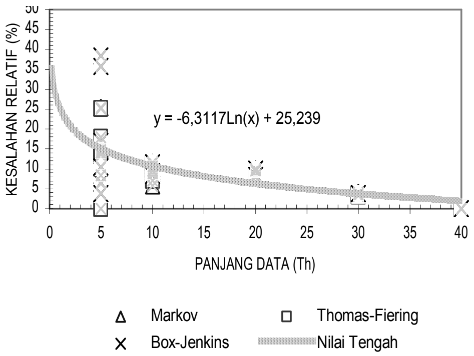
Gambar 5. Grafik Hubungan Kesalahan Relatif Deviasi Standar dan Panjang Data DPS Serang



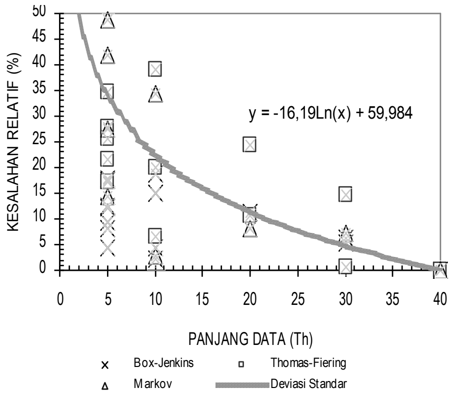
Gambar 6. Grafik Hubungan Kesalahan Relatif Nilai Tengah dan Panjang Data DPS Citarum



Gambar 7. Grafik Hubungan Kesalahan Relatif Deviasi Standar dan Panjang Data DPS Citarum



Gambar 8. Grafik Hubungan Kesalahan Relatif Nilai Tengah dan Panjang Data DPS Bengawan Solo



Gambar 9. Grafik Hubungan Kesalahan Relatif Deviasi Standar dan Panjang Data DPS Bengawan Solo

KESIMPULAN DAN SARAN

Berdasarkan pembahasan dan hasil penelitian yang dilakukan pada lokasi penelitian dengan pembatasan dan asumsi seperti yang diuraikan pada bab terdahulu, dapat diambil kesimpulan bahwa:

1. Panjang data aliran historis dengan seri 5, seri 10, seri 20, seri 30, seri 40, seri 50, seri 57, seri 60, seri 70 dan seri 73 sesudah dibangkitkan dengan model Markov (untuk data tahunan), Thomas-Fiering dan Box-Jenkins (untuk data bulanan) menghasilkan data sintetik yang mempunyai nilai statistik (nilai tengah, deviasi standar, variasi, *skewness*, kurtosis, dan lainnya) yang berlainan.
2. Pada data aliran historis seri dan DPS yang sama sesudah dibangkitkan dengan 3 model stokastik tersebut, menghasilkan nilai statistik yang mirip atau mendekati satu dengan lainnya.
3. Ditemukan kesalahan relatif langsung yang tidak mendekati satu dengan lainnya pada seri tahun yang sama (misalnya seri 5-1, 5-2, ..., 5-8) terutama kesalahan relatif deviasi standar. Diduga hal ini disebabkan adanya variasi data yang lebar pada seri tahun yang kecil.
4. Besaran atau nilai panjang data aliran historis yang representatif masing-masing lokasi penelitian adalah : Bojonegoro sekitar 24 tahun, Kedungombo sekitar 27 tahun dan Cirata sekitar 35 tahun.
5. Kesalahan relatif 5 % pada nilai panjang data historis di Bojonegoro yang mempunyai data 40 tahun menghasilkan data historis minimum rata-rata 27 tahun, Kedungombo (47 tahun) menghasilkan 29 tahun, dan Cirata (73 tahun) menghasilkan 35 tahun.
6. Hasil analisis dari penelitian ini ditemukan panjang data historis minimum yang representatif yang terjadi adalah 24 tahun dengan kriteria luas DPS lebih besar dari 250 km2 dengan panjang data aliran historis sebesar 40 tahun sebagai data untuk pembangkitan, sedangkan panjang data historis rata-rata representatif yang terjadi adalah 30 tahun.
7. Direkomendasikan data historis yang representatif untuk pembangkitan model stokastik harus sama atau lebih dari 30 tahun.

DAFTAR PUSTAKA

Box, E.G.P dan G.M.Jenkins, 1982, *Time Series Analisys Forecasting and Control*, Holdel – Day, San Fransisco.

Cryer, J.D., 1990, *Time-Series Analysis*, Duxbury Press, Boston, USA.

Dixon, W.J. dan Massey F.J., 1983, *Introduction to Statistical Analysis*, Fourth Edition, McGraw-Hill, Inc, USA.

Eko Wahyuni, S., 1998, *Pembangkitan Data dengan Model Stokastik AR*, Makalah Seminar Kelompok Sipil Hidro, Universitas Diponegoro, Semarang.

Hoff, J.C., 1987, *A practical Guide to Box-Jenkins Forecasting*, Lifetime Learning Publication Belmont, California.

Makridakis, S., Wheelwright, S.C. dan McGee, V.E., 1998, *Forecasting: Methods and Application*, Second Edition, John Wiley & Sons, Inc, USA.

Meyn, S.P., dan R.L. Tweedie, 1993, *Markov Chain and Stochastic Stability,* Springer-Verlag, Great Britain.

Papoulis, A., 1991, *Probability, Random Variables, and Stochastic Processes*, Third Edition, McGraw-Hill, Inc, USA.

Reddy, P.J., 1987, *Stochastic Hydrology*, Laxmi Publication, New Delhi, Madras, Jalandhar.

Salas, J.D., Delleur, J.W., Yevjevic, V. dan Lane, W.L., 1980, *Applied Modelling of Hydrologic Time Series*, Water Resources Publication, Littleton.

Triatmodjo, B., 1991, *Model Deret Berkala untuk Turunan Data Sungai Serang*, Jurnal Teknik Hidraulik No.6/VI, HATHI, Bandung, 12-19.

Whey, P., 1997, Probability, Random Variables & Random Processes, McGraw-Hill, Companies, Inc.

Wood, E.F. and O'Connel P.E., 1985, *Real-time Forecasting, Hydrological Forecasting*, John Wiley & Sons, Chichester, England.

Yevjevic, V., 1982, *Stochastic Procceses in Hydrology*, Water Resources Publication, Littleton, Colorado, USA.