

**ANALISA STATISTIKA UNTUK CURAH HUJAN HARIAN PADA DAS
KAMPAR BERDASARKAN AIC
(AKAIKE INFORMATION CRITERION)**

Citra Dewi Simbolon, Bambang Sujatmoko, Mardani Sebayang
Jurusan Teknik Sipil S1, Fakultas Teknik Universitas Riau
Kampus Binawidya Jl. H.R. Subrantas Km 12,5 Simpang Baru Pekanbaru 28293
email : citra.simbolon@gmail.com

ABSTRACT

The applicability of the frequency distributions had been widely used for the analysis of daily rainfall because the rain had different characteristic according to space and time. Distributions were often used for example gamma distribution, exponential, normal, log normal, Gumbel, and log the third person. This research aims to election statistics for rainfall distribution based on AIC and to determine which one was most appropriate distribution based on the Chi-kuadrat test and Kormologrov Smirnov test. The distributions in this research were searched maximum likelihood value, then evaluated based on the value of the Akaike Information Criterion (AIC). Maximum likelihood estimate of the value was used to maximize the distribution parameters in AIC. Distribution which has the smallest AIC value that will be used for the Chi Square test and Kolmogorov-Smirnov test. Research result shows all stations in the Kampar river basin, generate normal distribution that has the smallest AIC value. For Chi Square test and Kolmogorov-Smirnov test showed a normal distribution can be accepted in the watershed Kampar.

Keywords: Akaike Information Criterion, Chi Square, exponential distribution gamma distribution, Gumbel distribution, normal distribution, log normal distribution, log the third person distribution, Kolmogorov-Smirnov

PENDAHULUAN

Hujan memiliki karakteristik sesuai dengan ruang dan waktunya. Karakteristika tersebut menyebabkan pengamatan untuk setiap daerah pengaliran sungai akan memiliki sifat-sifat yang tersendiri dan berlaku untuk daerah itu sendiri. Karakteristika hujan dalam pengembangan sumber daya air adalah kedalaman intensitas, lama hujan, daerah tangkapan dan arah gerak hujan. (Suripin, 2004)

Hujan sangat berpengaruh pada keadaan sungai. Besar kecilnya curah hujan dapat menyebabkan banjir ataupun kekeringan. Untuk itu diperlukan penelitian tentang karateristik hujan secara intensif. Sungai Kampar yang ada di Riau Daratan memiliki potensi yang unik yang bisa dikembangkan untuk kepentingan Pengelolaan Sumber Daya Air dan ataupun kepentingan penelitian. Namun karakteristik hujan yang sesuai dengan ruang dan waktu menyebabkan Pengelolaan Sumber Daya Air pada Sungai Kampar memerlukan analisis data hidrologi.

Analisis data dalam jumlah yang banyak telah banyak dilakukan dalam analisa hidrologi. Salah satu hal penting dalam analisis hidrologi adalah menafsirkan probabilitas kemungkinan suatu kejadian yang akan datang berdasarkan data hidrologi (Triadmojo, 2008). Analisis dari sebuah data hidrologi yang kompleks memunculkan adanya sebuah sajian sederhana yang disebut model hidrologi. Dalam hidrologi terdapat beberapa macam model hidrologi diantaranya model matematika (Sri Harto, 1993).

Model matematika telah banyak digunakan dalam analisis hidrologi. Rantai markov telah digunakan untuk mensimulasi curah hujan basah dan kering (Khalilullah, 2011). Distribusi gamma dan eksponensial digunakan untuk analisa curah hujan (Suhaila, 2008). Distribusi Gamma juga digunakan untuk studi efek variabilitas curah hujan selama musim panen, untuk mempelajari efek pemanasan global, dalam teori antrian dan teori keandalan (reliabilitas), menghitung jarak antara waktu tiba di fasilitas pelayanan (misalnya bank dan loket kereta api), dan lamanya waktu sampai rusaknya suku cadang dan alat listrik (Walpole, 1995).

Untuk seleksi distribusi yang paling tepat untuk data curah hujan di DAS Kampar maka digunakan Akaike Information Criterion (AIC). AIC adalah sebuah ukuran relatif dari model statistika. AIC dikembangkan oleh Hirotugu Akaike dan pertama kali dipublikasikan pada tahun 1974. Terutama pada pemilihan model regresi terbaik, AIC mempunyai tujuan untuk peramalan (*forecasting*), yaitu dapat menjelaskan kecocokan model dengan data yang ada (*insample forecasting*) dan nilai yang terjadi di masa mendatang (*out of sample forecasting*).

Distribusi Frekuensi

Dalam penelitian ini, enam distribusi akan di uji. Enam distribusi tersebut beserta nilai estimasi maksimum likelihoodnya akan dijelaskan sebagai berikut.

1. Distribui Gamma

Variabel acak kontinu X, mempunyai distribusi Gamma, dengan parameter α dan β , jika fungsi padat peluangnya diberikan oleh: (Sebayang, 2004)

$$f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} \quad x \geq 0$$

Maksimum likelihood untuk gamma $\hat{\alpha} = \frac{\bar{x}^2}{s^2}$ dan $\hat{\beta} = \frac{s^2}{\bar{x}}$.

2. Distribusi Eksponensial

Variabel acak kontinu X mempunyai distribusi Eksponensial dengan parameter β , jika fungsi padat peluangnya berbentuk (Sebayang, 2004) :

$$f(x) = \frac{1}{\beta} e^{-\frac{x}{\beta}}, \text{ untuk } x \geq 0$$

Maksimum likelihood untuk eksponensial $\hat{\beta} = \bar{X}$

3. Distribusi normal

Distribusi normal disebut juga distribusi Gauss, dengan fungsi densitas peluangnya berbentuk (Triatmojo, 2008)

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{(2\sigma)^2}\right]$$

Maksimum likelihood untuk normal $\mu = \bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ dan $s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$

4. Distribusi log normal

Fungsi densitas probabilitasnya dapat dituliskan dalam bentuk sebagai berikut: (Triatmojo, 2008)

$$p(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(\log x - \mu)^2}{2\sigma^2}\right]$$

Estimasi maksimum likelihood untuk distribusi log normal adalah $x = \exp \mu$

$$\text{dan } \sigma = \sqrt{\frac{\sum_{i=1}^N (\ln X_i - \mu)^2}{N-1}}$$

dimana:

$$\mu = \frac{\sum_{i=1}^N \ln X_i}{N}$$

5. Distribusi gumbel

Distribusi Gumbel mempunyai fungsi probabilitas sebagai berikut: (Triatmojo, 2008)

$$F(x) = e^{-e^{\frac{x-\alpha}{\beta}}}$$

estimasi maksimum likelihood dari distribusi gumbel adalah

$$\bar{\beta} = \bar{x} - \frac{\sum_{i=1}^n x_i e^{-\frac{x_i}{\beta}}}{\sum_{i=1}^n e^{-\frac{x_i}{\beta}}}$$

dan

$$\bar{\alpha} = -\bar{\beta} \log \left[\frac{1}{n} \sum_{i=1}^n e^{-\left(\frac{-x_i}{\beta}\right)} \right]$$

6. Distribusi log person III

Fungsi densitas probabilitas distribusi log Person III mempunyai bentuk sebagai berikut: (Triatmodjo, 2008)

$$p(x) = \frac{x^{\alpha-1} e^{-\frac{x}{\beta}}}{\beta^{\alpha} \Gamma(\alpha)}$$

estimasi maksimum likelihood untuk distribusi ini dihitung dengan rumus:

$$\hat{\alpha} = \frac{\bar{x}^2}{s^2} \text{ dan } \hat{\beta} = \frac{s^2}{\bar{x}}$$

Pemilihan distribusi

Menentukan model terbaik dalam data yaitu menggunakan Akaike Information Criteria (AIC). Pada suatu model dikatakan baik apabila nilai AIC nya paling kecil.

$$AIC = -2 \log L + 2k$$

dengan :

L = maksimum Likelihood

k = jumlah parameter

Uji Kesesuaian Ditribusi

a. Uji Chi-Kuadrat

Uji Chi-Kuadrat bertujuan untuk menentukan apakah persamaan distribusi yang dipilih dapat mewakili distribusi statistik sampel data yang telah

dianalisis. Pengambilan keputusan uji Chi-Kuadrat menggunakan parameter χ_h^2 (Suripin, 2004).

- b. Uji Smirnov Kolmogorov
 Uji kecocokan Smirnov Kolmogorov sering disebut juga uji kecocokan non parametrik, karena pengujiannya tidak menggunakan fungsi distribusi tertentu. Prosedur dasarnya mencakup perbandingan antara frekuensi kumulatif eksperimental dan distribusi teoritis yang diasumsikan.

METODE PENELITIAN

Lokasi dan Data Penelitian

Penelitian dilakukan pada DAS Kampar, secara geografis mencakup kawasan seluas 24.548 km² dan terletak pada 100°10"-103°15" BT dan 0°41 "LU-0°35"LS, dengan panjang 580 km dan lebar 100-300 serta kedalaman 6-10 m. Data curah hujan yang digunakan adalah data dari 4 stasiun yang ada di DAS Kampar yaitu stasiun Pasar Kampar, stasiun Lipat Kain, stasiun Koto Baru dan stasiun Muara Lembu. Panjang data curah hujan yang digunakan adalah 28 tahun dari tahun 1983 sampai tahun 2010.

Analisa Penelitian

- Untuk uji seleksi menggunakan Akaike Information Criterion
 AIC adalah suatu ukuran yang relatif dari suatu model statistik. Nilai-nilai AIC digunakan untuk pemilihan model yang sesuai. AIC tidak menghasilkan suatu model test dalam menguji suatu hipotesis batal. Metode AIC didasarkan pada metode *maximum likelihood estimation* (MLE).
- Uji kesesuaian menggunakan uji Chi Kuadrat dan uji Smirnov-Kolmogorov
 Uji Chi-Kuadrat adalah teknik statistik yang digunakan untuk menguji hipotesis bila dalam populasi terdiri atas dua atau lebih kelas dimana data berbentuk nominal dan sampelnya besar. Uji kecocokan Smirnov – Kolmogorov sering disebut juga uji kecocokan non parametrik, karena pengujiannya tidak menggunakan fungsi distribusi tertentu namun dengan memperhatikan kurva dan penggambaran data pada kertas probabilitas.

HASIL DAN PEMBAHASAN

Uji Outlier

Hasil perhitungan uji *outlier* setiap stasiun dapat dilihat pada tabel berikut:

Tabel 1 Rekapitulasi Uji *Outlier*

Stasiun	Data curah hujan		Uji <i>outlier</i>		Keterangan
	Min (mm)	Maks (mm)	x_l (mm)	x_h (mm)	
Koto Baru	70,5	190	54,223	219,299	Semua data digunakan
Muara Lembu	62	139	50,887	163,729	Semua data digunakan
Lipat Kain	75	150	68,492	154,403	Semua data digunakan
Pasar Kampar	76	211	57,096	202,221	Data maksimum tidak digunakan

(sumber:hasil perhitungan, 2013)

Dari tabel diatas dapat dilihat bahwa untuk setiap stasiun nilai minimum curah hujan lebih kecil dari batas terendah nilai uji *outlier*. Untuk nilai maksimum curah hujan lebih kecil dari batas tertinggi nilai uji *outlier* pada stasiun Koto Baru, Muara Lembu dan Lipat kain. Sedangkan pada stasiun Pasar Kampar nilai maksimum melebihi batas tertinggi uji *outlier* sehingga nilai maksimum tersebut tidak digunakan pada analisa.

Perhitungan Nilai Statistik

Hasil perhitungan nilai statistik untuk setiap stasiun dapat dilihat pada tabel berikut:

Tabel 2. Rekapitulasi Perhitungan Nilai Statistik

Stasiun	Nilai statistik				
	\bar{x} (mm)	S (mm)	C_v	C_s	C_k
Koto Baru	113,236	32,822	0,289	0,884	3,264
Muara Lembu	93,679	22,104	0,236	0,595	2,655
Lipat Kain	104,143	17,237	0,166	0,804	3,912
Pasar Kampar	103,425	24,406	0,236	0,891	4,071

(sumber:hasil perhitungan, 2013)

Dari rekapitulasi uji statistik dapat dilihat bahwa Stasiun Koto Baru, Stasiun Muara Lembu, Stasiun Lipat Kain dan Stasiun Pasar Kampar memiliki nilai *skewness* 0,884; 0,595; 0,804; dan 0,891. Nilai *skewness* tersebut positif menunjukkan bahwa distribusi data lebih condong ke kanan dari data di sekitar rata-rata. Nilai *kurtosis* pada Stasiun Koto Baru, Stasiun Lipat Kain dan Stasiun Pasar Kampar adalah 3,264; 3,912; dan 4,071 yang berarti distribusi berbentuk tinggi relatif. Sedangkan Stasiun Muara Lembu memiliki nilai *kurtosis* lebih kecil dari 3 yaitu 2,655 yang berarti distribusi berbentuk rata relatif.

Nilai AIC Untuk Setiap Stasiun

Nilai AIC untuk setiap stasiun pada DAS Kampar dapat dilihat pada tabel di bawah ini.

Tabel 3 Rekapitulasi Nilai AIC Untuk Setiap Stasiun

	Koto Baru	Muara Lembu	Lipat Kain	Pasar Kampar
AIC _{Gamma}	-0,108	0,056	3,390	-0,0609
AIC _{Eksponensial}	-2,107	-1,943	-2,035	-2,060
AIC _{Normal}	-3,140	-2,583	-2,557	-2,769
AIC _{Log Normal}	4,071	4,764	6,913	4,302
AIC _{Gumbel}	1,072	2,011	3,039	1,923
AIC _{Log Person III}	3,381	3,415	3,392	3,389

(sumber:hasil perhitungan, 2013)

Dari tabel diatas dapat dilihat bahwa distribusi normal memiliki nilai AIC yang kecil untuk setiap stasiun pada DAS Kampar. AIC bernilai minus menunjukkan bahwa kecocokan distribusi dan data adalah besar. Distribusi

normal yang mempunyai nilai AIC terkecil akan digunakan dalam uji Chi Kuadrat dan uji Smirnov-Kolmogorov.

Uji Chi Kuadrat Dan Uji Smirnov-Kolmogorov

Hasil perhitungan uji Chi Kuadrat Dan Smirnov-Kolmogorov setiap Stasiun dapat dilihat pada tabel berikut:

Tabel 4 Rekapitulasi Uji Chi Kuadrat Dan Smirnov-Kolmogorov Untuk Distribusi Normal

Stasiun	Chi kuadrat		Smirnov-Kolmogorov		keterangan
	χ^2	χ^2_{cr}	Δ_{maks}	Δ_{cr}	
Koto Baru	3,976	7,8150	0,133	0,241	Distribusi normal diterima
Muara Lembu	3,119	7,8150	0,109	0,241	Distribusi normal diterima
Lipat Kain	3,904	7,8150	0,039	0,241	Distribusi normal diterima
Pasar Kampar	4,277	7,8150	0,045	0,250	Distribusi normal diterima

(sumber:hasil perhitungan, 2013)

Dari tabel diatas dapat dilihat bahwa distribusi normal yang nilai AIC kecil untuk setiap stasiun pada DAS Kampar diterima dalam uji Chi Kuadrat dan Smirnov-Kolmogorov. Dapat dilihat dari nilai χ^2 lebih kecil dari nilai χ^2_{cr} untuk uji Chi Kuadrat dan nilai Δ_{maks} lebih kecil dari nilai Δ_{cr} untuk uji Smirnov-Kolmogorov.

KESIMPULAN DAN SARAN

Kesimpulan

1. Dari hasil uji outlier menunjukkan bahwa stasiun Koto Baru, Muara Lembu dan lipat Kain semua data digunakan untuk analisa. Sedangkan pada stasiun Pasar Kampar, nilai maksimum tidak digunakan karena lebih besar dari ambang batas atas nilai uji *outlier*.
2. Berdasarkan nilai AIC distribusi yang memiliki nilai AIC terkecil untuk setiap stasiun pada DAS Kampar adalah distribusi normal karena distribusi normal memiliki kecocokan yang besar antara model dan data DAS Kampar.
3. Nilai AIC distribusi normal untuk stasiun Koto Baru adalah -3,140, pada stasiun Muara Lembu -2,583, untuk stasiun Lipat Kain -2,557 dan untuk stasiun Pasar Kampar -2,769.
4. Pada pengujian Chi Kuadrat dan Smirnov-Kolmogorov, distribusi normal untuk semua stasiun DAS Kampar diterima.

Saran

saran yang dapat dikemukakan dalam studi ini adalah untuk variasi penelitian selanjutnya disarankan agar mengkaji lebih lanjut data curah hujan dengan menggunakan distribusi yang lain seperti distribusi gamma campuran dan eksponensial campuran.

DAFTAR PUSTAKA

- Enders.** 1995. Bank of England 2004 [online]. Available at :
<<http://www.bankofengland.co.uk/five.pdf>> [Diakses tanggal 5 Pebruari 2012]
- Engelhardt, Bain.** 1991. Introduction To Probability And Mathematical Statistics. California. Duxburry Press
- Harto, Sri.** 1993. *Analisis Hidrologi*. Jakarta : Gramedia Pustaka Utama.
- Hu, Shuhua.** 2008. *Akaike's Information Criteria (AIC)*. Available at
<http://coopunit.forestry.uga.edu/coop_wkshop/inference_effects/aic_reg.pdf> [Diakses 14 Desember 2012]
- Khalilullah, M.** 2011. *Simulasi Data Curah Hujan Harian Pada Das Kampar Menggunakan Stokastik Rantai Markov*. Skripsi. Pekanbaru. Teknik Sipil FT UR.
- Misbahussurur, Ahmad.** Estimasi parameter Distribusi Gamma Dengan Metode Maksimum Likelihood. Available at <<http://lib.uin-malang.ac.id/thesis/fullchapter/05510023-ahmad-misbahussurur.ps>> [diakses tanggal 12 februari 2013]
- Montgomery, D. C.** *Engineering statistics*. New york : John wiley n sons, inc.
- Rizki, Dwiki.** 2011. *Uji Chi-Kuadrat*. Available at:
<http://ilerning.com/index.php?option=com_content&view=article&id=693:uji-chi-kuadrat-edit-mar&catid=39:hipotesis&Itemid=70> [diakses tanggal 13 April 2013]
- Sebayang, Mardani.** 2004. *Probabilitas dan Statistik dalam Ilmu Rekayasa*. Pekanbaru. Teknik Sipil FT UR.
- Soemarto, C. D.** 1999. *Hidrologi Teknik*. Jakarta : Erlangga.
- Suripin.** 2004. *Sistem Drainase Perkotaan yang Berkelanjutan*. Yogyakarta: Andi.
- Triatmodjo, B.** 2008. *Hidrologi Terapan*. Yogyakarta : Beta offset
- Walpole, Ronald E.** 1995. *Ilmu Peluang dan Statistika Untuk Insinyur dan Ilmuan*. Bandung: ITB.