Jur. Ris. & Apl. Mat. Vol. 1 (2017), no. 1, 1-52
Jurnal Riset dan Aplikasi Matematika
e-ISSN: 2581-0154
URL: journal.unesa.ac.id/index.php/jram

Prediksi Curah Hujan dengan Model Deret Waktu dan Prakiraan Krigging pada 12 Stasiun di Bogor Periode Januari 2013 - Desember 2014

Mahfudhotin

S2 Matematika, MIPA, Institut Teknologi Bandung
mahfudhotin@s.itb.ac.id

Abstrak

Curah hujan merupakan salah satu faktor iklim yang berpengaruh di berbagai bidang sehingga pemerintah membangun stasiun hujan untuk mengukur curah hujan di lokasi tertentu di Indonesia yang dianggap memiliki potensi. Akan tetapi curah hujan di luar daerah stasiun hujan tidak diketahui secara pasti, sehingga perlu dilakukan prediksi curah hujan dengan menggunakan analisis deret waktu dengan metode Box-Jenkins yang dikenal dengan *Autoregressive Integrated Moving Average* (ARIMA), maupun analisis krigging untuk melihat kebergantungan spasial lokasi. Identifikasi model dilakukan dengan meihat plot ACF dan PACF data. Data yang digunakan adalah data curah hujan di Bogor periode 10 harian dari bulan Januari 2013 - Desember 2014 sehingga diperoleh model deret waktu terbaik untuk 12 stasiun yang terdiri dari ARIMA(1,1,1), ARIMA(1,1,0), dan ARIMA(3,1,0). Krigging dilakukan untuk memprakirakan 5 waktu ke depan.
Kata Kunci: Analisis Deret Waktu (*time series*),*Autoregressive Integrated Moving Average* (ARIMA), *Autocorrelation Function* (ACF), *Partial Autocorrelation Function*(PACF), Metode Krigging, Prediksi , Prakiraan

1. Pendahuluan

Bogor merupakan salah satu kota besar di Indonesia yang terletak di dataran tinggi., dimana dibangun stasiun hujan, antara lain Bendungan Empang, Cibodas, Cihideung, Ciriung, Kec. Ciawi, Klapanunggal, Lanud Atang Sanjaya, Perk. Cikasungka, Perk. Gunung Mas, Perk. Pondok Gedeh, Stamet Citeko, dan Staklim Darmaga.

Analisis deret waktu (*time series*) adalah suatu observasi yang dibangun berurutan dalam waktu. Sebelum memilih suatu model deret waktu, ada dua asumsi yang harus diperhatikan yaitu asumsi homoskedastisitas dan heterokedastisitas. Paper ini model yang digunakan adalah

Tanggal Masuk: 04-11-16; Revisi: 11-09-17; Diterima: 28-10-2017

model dengan asumsi homoskedastisitas, dimana nilai volatilitas berubah-ubah terhadap waktu (variansi datanya tidak stasioner).

Tahapan pemodelan deret waktu dilakukan dengan pemeriksaan kestasioneran. Hal ini dilakukan karena seringkali data deret waktu bersifat tidak stasioner. Identifikasi model dilakukan dengan memperhatikan ACF dan PACF. Estimasi parameter model deret waktu dilakukan melalui Metode Likelihood Maksimum (MLE). Lebih lanjut, untuk melihat perilaku galat dilakukan uji diagnostik. Seleksi model dapat itentukan melalui *Akaike Information Criteria* (AIC) dan *Bayesian Information Criterion* (BIC).

$$AIC = -2log(\hat{\theta}) + 2k$$

$$BIC = -2log(\hat{\theta}) + klog(n)$$

dengan $log(\hat{\theta})$ adalah nilai maksimum fungsi log likelihood dari suatu model yang diestimasi, n adalah banyaknya pengamatan, dan k merupakan banyaknya parameter.

Kriging adalah salah satu metode interpolasi spasial yang memanfaatkan nilai spasial pada lokasi tersampel untuk memprakirakan nilai lokasi lain yang belum atau tidak tersampel. Hal ini dapat diidentikasi dengan adanya kebergantungan spasial.

2. Kajian Teori

Model ARIMA

Jika suatu data deret waktu mengikuti model ARMA dengan *mean* dan variansi yang tidak konstan maka perlu dilakukan diferensi sehingga menjadi *Autoregressive Integrated Moving Average* (ARIMA). Pada model ini memiliki kesamaan sifat ACF dan PACF dari model ARMA, perbedaannya terletak pada bentuk data yang sudah didiferensi. Sebagai contoh model ARIMA (1,1,1) adalah :

$$\begin{aligned}
W_t &= \phi_1 W_{t-1} + \phi_2 W_{t-2} + \cdots + \phi_p W_{t-p} \\
Y_t - Y_{t-1} &= \phi_1(Y_{t-1} - Y_{t-2}) + \phi_2(Y_{t-2} - Y_{t-3}) + \cdots + \phi_p(Y_{t-p} - Y_{t-p-1}) \\
&\quad + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \\
Y_t &= (1+\phi_1)Y_{t-1} + (\phi_2 - \phi_1)Y_{t-2} + (\phi_3 - \phi_2)Y_{t-3} + \cdots + (\phi_p - \phi_{p-1})Y_{t-p} \\
&\quad - \phi_p Y_{t-p-1} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}.
\end{aligned}$$

Prediksi Model Deret Waktu

Untuk membangun sebuah model time series adalah memprediksi besar nilai yang akan terjadi pada waktu yang akan datang. Misalkan pengamatan $Y_1, Y_2, \cdots, Y_{t-1}$, akan diprediksi nilai dari Y_{t+l} dimana ℓ menyatakan lama waktu masa depan yang akan diprediksi. Perhatikan persamaan untuk menghitung prediksi dari model ARIMA ditulis sebagai berikut:

$$\begin{aligned}\tilde{Y}_t(\ell) &= \phi_1\tilde{Y}_t(\ell-1) + \phi_2\tilde{Y}_t(\ell-2) + \phi_3\tilde{Y}_t(\ell-3) + \cdots \\ &+ \phi_{p+1}\tilde{Y}_t(\ell-p-1) - \theta_1 E(e_{t+\ell-1}|Y_1, Y_2, \cdots, Y_t) \\ &- \theta_2 E(e_{t+\ell-2}|Y_1, Y_2, \cdots, Y_t) - \cdots - \theta_q E(e_{t+\ell-q}|Y_1, Y_2, \cdots, Y_t)\end{aligned}$$

Pandang kasus model ARIMA $(1, 1, 1)$. Persamaan model deret waktunya ditulis

$$\tilde{Y}_t(\ell) = (1+\phi)\tilde{Y}_t(\ell-1) - \phi\tilde{Y}_t(\ell-2) + \theta_0$$

Krigging

Untuk memprakirakan data pengamatan yang tidak teramati dengan menggunakan analisis semivariogram dan kriging. Bentuk variogram yang digunakan adalah model Cubical,

$$v = \begin{cases} c_0 + c(7r^2 - 8.75r^3 + 3.5r^5 - 0.75r^7) & r < 1 \\ c_0 + c & r\text{yang lain} \end{cases}$$

Menurut Issac dan Srivastava (1989), Kriging adalah metode geostatistik yang digunakan untuk mengestimasi nilai dari sebuah titik atau blok sebagai kombinasi linier dari nilai yang terdapat disekitar titik yang akan diestimasi. Estimator kriging $Z(\hat{s_0})$ pada lokasi s_0,

$$\begin{aligned}(Z(\hat{s_0})) &= \lambda_1 Z(s_1) + \lambda_2 Z(s_2) + ... + \lambda_n Z(s_n) \\ &= \sum_{i=1}^{n} \lambda_i Z(s_i)\end{aligned}$$

dengan λ_i adalah bobot kriging.

Pada umumnya, jenis kriging yang digunakan adalah ordinary kriging. Ordinary kriging adalah metode geostatistika yang digunakan untuk memprakirakan data pada lokasi tertentu. Ordinary kriging menduga suatu variabel pada suatu titik tertentu dilakukan dengan mengamati data yang sejenis pada suatu daerah.

Untuk menaksir nilai disuatu titik s_0 yaitu $Z(\hat{s_0})$ digunakan kombinasi linear dari $Z(s_1), Z(s_2), ..., Z(s_n)$ ditulis menjadi :

$$Z(\hat{s_b g0}) = \sum_{i=1}^{n} \lambda_i Z(s_i), \lambda_i \in \Re$$

dengan λ_i adalah bobot kriging dan s_i adalah lokasi sampel dengan $i = 1, 2, ..., n$.

3. **Pembahasan**

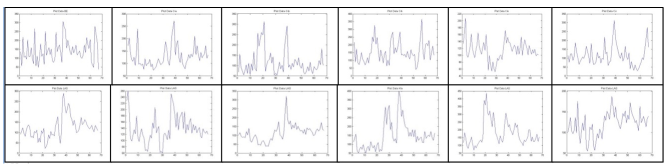
Model Deret Waktu dan Krigging Curah Hujan

Data yang digunakan adalah data curah hujan 10 harian dari tanggal Januari 2013 sampai Desember 2014 di 12 stasiun(kecamatan) Bogor yaitu sejumlah 67 buah data. Lebih lanjut, akan dicari model deret waktu yang sesuai dengan data selanjutnya diprediksi data di 5 waktu selanjutnya dan melakukan prakiraan curah hujan pada daerah yang data curah hujannya tidak diketahui menggunakan informasi data curah hujan pada lokasi disekitarnya.

Stat Deskriptif	BE	CIB	CIH	CIR	CIA	KLA	LAS	CIK	GM	PG	CIT	DAR
Banyak Data	67	67	67	67	67	67	67	67	67	67	67	67
Jumlah	9,563	7,795	10,405	7,690	8,693	10,048	11,896	7,732	7717	8128	8248	9261
Mean	142.73	116.34	155.30	114.78	129.75	149.97	177.55	115.40	115.1791	121.3134	123.1045	138.2239
Median	132	101	143	102	121	134	152	114	116	124	115	134
Minimum	37	53	41	29	73	51	69	52	38	55	23	60
Maximum	306	312	365	313	271	399	436	207	321	187	289	260
SD	62.278	60.327	67.020	55.696	42.026	78.352	78.177	28.886	47.11621	28.70604	52.26278	43.87511
Variance	3,878.533	3,639.289	4,491.637	3,102.085	1,766.192	6,138.969	6,111.706	834.426	2219.937	824.0366	2731.398	1925.025
Standard Error	7.608	7.370	8.188	6.804	5.134	9.572	9.551	3.529	5.756161	3.507001	6.384915	5.360197
Skewness	0.645	1.624	0.839	1.248	1.333	1.224	1.281	0.283	1.301259	-0.29749	0.967717	0.718212
Kurtosis	-0.029	2.099	0.643	2.109	1.697	1.274	1.343	1.026	4.606922	-0.01366	1.469242	0.628538
25th Percentile	102	76	104	79	100	95	125	101	86	103.5	92.5	115
50th Percentile	132	101	143	102	121	134	152	114	116	124	115	134
75th Percentile	174	127	193	137	143	182	218	129	136	139	138	154.5

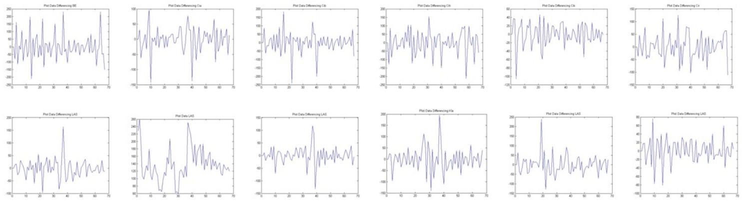
Gambar 1: Statistika Deskriptif Data Curah Hujan 12 Lokasi

Data curah hujan memiliki rentang nilai yang cukup besar terjadi pada lokasi Perk. Gunung Mas dan Perk. Pondok Gedeh. Pada tahap awal pemodelan, 72 data dibagi menjadi 67 data untuk membangun data dan 5 data untuk validasi model. Hal yang penting dan perlu diperhatikan dalam menggambarkan suatu deret waktu adalah kestasioneran. Parameter yang digunakan adalah mean dan variansi. Kestasioneran dari parameter model deret waktu sangatlah diperhatikan karena jika terjadi perubahan tiap waktu maka parameter yang ditaksir kurang sesuai dengan data kenyataan.



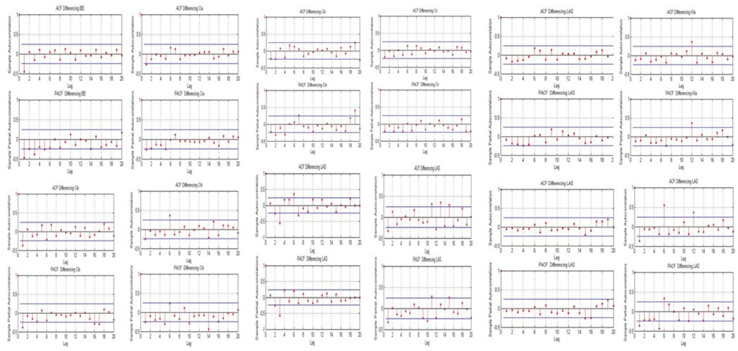
Gambar 2: Plot Data 12 Lokasi Curah Hujan di Bogor

Plot curah hujan menunjukkan trend data yang tidak stasioner dengan kecenderungan mean dan variansi yang naik kemudian turun secara tajam. Kestasioneran data curah hujan diperoleh melalui differencing dimana defferencing yang digunakan adalah $Y_t = \log(P_t) - \log(P_{t-1})$.



Gambar 3: Return data 12 Lokasi Curah Hujan di Bogor

Hasil differencing memperlihatkan trend data stasioner yakni mean dan variansi cenderung konstan.



Gambar 4: Plot ACF dan PACF Data Curah Hujan 12 Stasiun di Bogor

Berdasarkan ACF dan PACF, model deret waktu dapat ditentukan dan mendapatkan parameter yang diestimasi menggunakan metode maksimum likelihood. Stasiun Bendungan Empang model ARIMA(1,1,1), $Y_t = 0.254 + 0.0349\phi + 1\theta$; Cibodas model ARI(1,1), $Y_t = 1.146 - 0.384\phi$; Cihideung model ARIMA(1,1,1), $Y_t = -0.368 + 0.5648\phi + 0.984\theta$; Ciriung model ARIMA(1,1,1), $Y_t = 0.026 + 0.605\phi + 1\theta$; Kec. Ciawi model ARIMA(1,1,1), $Y_t = -0.150 + 0.474\phi + 0.888\theta$; Klapanunggal model ARIMA(1,1,1), $Y_t = 0.286 + 0.748\phi + 1\theta$; Lanud Atang Sanjaya model ARI(1,1), $Y_t = 0.516 - 0.0653\phi$; Perk Cikasungka model

ARIMA(1,1,1), $Y_t = -0.194 + 0.569\phi + 1\theta$; Perk Gunung Mas model ARIMA(1,1,1), $Y_t = 0.1569 + 0.697\phi + 1\theta$; Perk Pondok Gedeh model ARIMA(1,1,1), $Y_t = 0.348 + 0.428\phi + 1\theta$; Stamet Citeko model ARI(3,1), $Y_t = 0.157 - 0.066\phi_1 - 0.210\phi_2 - 0.578\phi_3$; Staklim Darmaga model ARIMA(1,1,1), $Y_t = -2.229 - 0.415\phi - 0.091\theta$; Seleksi model dilakukan dengan menghitung nilai AIC dan BIC.

Tabel 1: AIC dan BIC

Stasiun	Model	AIC	BIC
Bendungan Empang	ARIMA(1,1,1)	745.573	755.57
Cibodas	ARIMA(1,1,0)	732.14	737.139
Cihideung	ARIMA(1,1,1)	746.55	756.55
Ciriung	ARIMA(1,1,1)	701.596	711.59
Kec. Ciawi	ARIMA(1,1,1)	678.33	688.332
Klapanunggal	ARIMA(1,1,1)	715.800	725.799
Lanud Atang Sanjaya	ARIMA(1,1,0)	717.327	722.327
Perk Cikasungka	ARIMA(1,1,1)	621.083	621.083
Perk Gunung Mas	ARIMA(1,1,1)	655.843	665.842
Perk Pondok Gedeh	ARIMA(1,1,1)	615.780	625.779
Stamet Citeko	ARIMA(3,1,0)	646.39	661.395
Staklim Darmaga	ARIMA(1,1,1)	677.07	682.074

Berdasarkan 2 memberikan informasi nilai AIC dan BIC. Menurut Cryer (2008), BIC cenderung digunakan karena secara eksperimen terbukti BIC menghasilkan model lebih akurat dibandingkan AIC. Sehingga nilai BIC lebih dipertimbangkan dalam pemilihan model. Nilai BIC yang terkecil diperoleh model yang paling sesuai dengan data curah hujan untuk 12 stasiun di Bogor sehingga didapatkan model deret waktu yang sesuai dan parameternya.

4. **Prediksi Model Deret Waktu**
   Setelah identifikasi model, estimasi parameter dan uji diagnostik maka dilakukan prediksi berdasarkan model terpilih. Prediksi nilai curah hujan 5 waktu kedepan dari 12 lokasi yaitu November - Desember 2014.

Gambar 5: Perbandingan data prediksi dan data asli bberdasarkan model deret waktu

Bendungan Empang		Cibodas		Cihideung		Ciriung	
Forecast	Nilai Real	Forecast	Nilai Real	Forecast	Nilai Real	Forecast	Nilai Real
119.0163	53	93.012	51	112.43	120	118.92	135
135.8203	71	81.23	73	146.87	173	128.01	191
141.2488	121	91.34	86	146.23	156	146.029	120
143.0025	139	102.23	92	138.007	129	92.02	77
143.5691	126	95.2	82	149.92	183	98.43	102
Kec. Ciawi		**Klapanunggal**		**Lanud Atang Sanjaya**		**Perk. Cikasungka**	
Forecast	Nilai Real	Forecast	Nilai Real	Forecast	Nilai Real	Forecast	Nilai Real
102.4203	114	85.5231	121	117.9870	123	152.0050	162
117.3850	134	99.9947	143	129.0400	133	120.8840	123
123.7075	152	111.3010	134	142.0000	121	112.0600	102
126.3787	123	120.1343	152	144.3800	111	117.8500	128
127.5072	134	127.0356	133	107.9300	103	127.9300	122
Perk. Gunung Mas		**Perk. Pondok Gedeh**		**Stamet Citeko**		**Staklim Darmaga**	
Forecast	Nilai Real	Forecast	Nilai Real	Forecast	Nilai Real	Forecast	Nilai Real
126.9300	118	130.5200	124	78.3400	91	126.9500	119
165.0400	203	102.8300	112	89.2800	113	113.9300	113
177.4800	193	109.7700	137	103.7700	127	138.4400	104
189.4200	156	152.7300	129	108.9400	116	148.5600	129
148.3800	134	117.6600	112	109.7800	103	158.5400	123

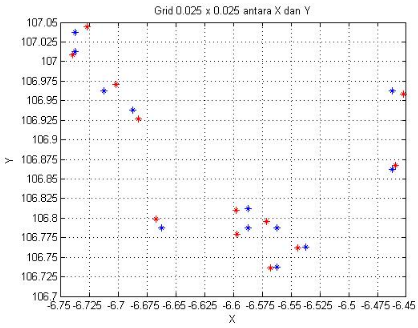
5. **Pemilihan Model Semivariogram**

   Data curah hujan pada 12 lokasi di Bogor tidak hanya bergantung pada data di waktu sebelumnya tetapi juga bergantung pada data curah hujan di lokasi sekitarnya. Sehingga dapat dilakukan analisis semivariogram untuk data curah hujan tersebut.

   Data yang dipakai untuk memodelkan semivariogram hanya 10 lokasi dan 2 lokasi lainnya yakni Perkebunan Cikasungka dan Kecamatan Ciawi digunakan untuk prakiraannya. Data curah hujan dari prediksi 5 waktu kedepan yang telah dihitung sebelumnya akan dimodelkan dengan model semivariogram yaitu Sperikal, Eksponensial, Gaussian, Cubical, dan Cardinal Sine.

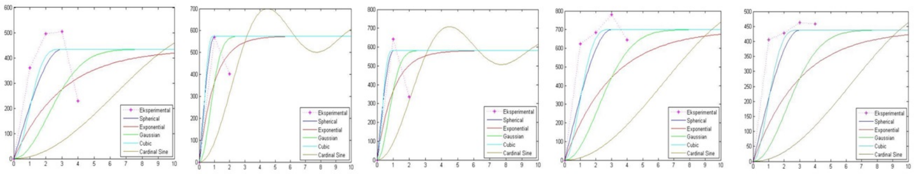
   Untuk mempermudah perhitungan jarak antar lokasi, yang pada analisa ini dibuat dengan ukuran $0,25x0,25$ satuan sehingga diperoleh koordinat baru yang menjadikan koordinat sebelumnya berada pada titik tengah di tiap-tiap grid. Grafik grid menunjukkan adanya ketergantungan secara implisit dimana terlihat titik-titik yang berdekatan.

Gambar 6: Gambar Grid Curah Hujan bogor



Berikut adalah model semivariogram dari data 5 prediksi curah hujan pada 10 lokasi,

Gambar 7: Kecocokan Semivariogram Prakiraan Waktu ke 68 sampai dengan 72 (dari kiri ke kanan)



Berdasarkan 7 Model semivariogram yang cocok adalah model semivariogram Cubical. Selain itu, nilai Mean Square Error (MSE) menunjukkan nilai yang kecil dibandingkan model semivariogram yang lain sebagai berikut,

Tabel 2: MSE model semivariogram Cubical Isotropik Waktu ke 68 sampai dengan 72

Waktu	Ke−68	Ke−69	Ke−70	Ke−71	Ke−72
MSE	12674.47	14997.44	3724.404	24608.66	10365.5652

Ordinary Krigging digunakan untuk memprakirakan data pada 2 lokasi yaitu daerah Perkebunan Cikasungka dan Kecamatan Ciawi. Terlihat bahwa pada kolom 'krigging' menunjukkan prakiraan data 68 sampai dengan 72 berdasarkan lokasi (spasial) dengan mengabaikan kebergantungan deret waktu.

Gambar 8: Data prediksi dengan deret waktu, data asli, dan hasil prakiraan dengan metode kriging

Perk. Cikasungka			Kec. Ciawi		
Forecast	Nilai Real	Kriging	Forecast	Nilai Real	Kriging
152.0050	162	149.02	102.4203	114	103.56
120.8840	123	110.92	117.3850	134	114.8
112.0600	102	92.01	123.7075	152	93.03
117.8500	128	134.2	126.3787	123	134.9
127.9300	122	113.3	127.5072	134	114.93

Kesimpulan

Dari analisis data menggunakan model deret waktu dan metode krigging didapatkan hasil prediksi dan prakiraan masing-masing. Sehingga dapat disimpulkan bahwa,

1. Berdasarkan deret waktu data curah hujan, terlihat bahwa sebagian besar 12 stasiun di Bogor pada periode pertama, kedua sampai kelima, prediksi deret waktu nya lebih mendekati data asli (sebenarnya). Khususnya Prediksi Stasiun Cikasungka lebih dekat dengan data sebenarnya.

2. Model semivariogram yang paling baik digunakan untuk data prediksi deret waktu adalah model Cubical dibandingkan dengan model Sperikal, Eksponensial, Gaussian, dan Cardinal Sine. Hal ini juga ditunjukkan pada MSE yang kecil dari model Cubical terhadap model yang lain.

3. Untuk kecamatan Ciawi, pada periode pertama, metode kriging memberikan pendekatan yang paling baik.

Daftar Pustaka

1. Amstrong, Margaret. Basic Linear Geostatistic. Springer, 1998.

2. BMKG.go.id. 14 Mei 2015

3. Cryer, Jonathan D. and Kung-Sik Chan. (2008). Time Series Analysis with Application in R, second edition, Lowa City : Springer. Berndt, E. K.

4. Isaaks, Edward H., and Srivastava, R.M., 1989, Applied Geostatistics, Oxford University Press, Oxford.

6. Mukhaiyar,Utriweni. (2015). Catatan Kuliah Topik Statistika II

7. Wei, W.W.S. (1994). Time Series Analysis: Univariate and Multivariate Methods, Second Edition, Addison Wesley.