

PEMILIHAN METODE ESTIMASI DATA CURAH HUJAN TIDAK LENGKAP

Wahidah Sanusi¹, Sudarmin², Sahlan Sidjara¹

¹ Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Makassar, Indonesia

² Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Makassar, Indonesia

Email: wahidah.sanusi@unm.ac.id

Abstrak. Estimasi data curah hujan tidak lengkap perlu dilakukan untuk menghindari kesimpulan penelitian yang bias. Lima metode digunakan dalam penelitian ini untuk mengestimasi data hujan tidak lengkap pada 5 stasiun hujan di kabupaten Maros dan 7 stasiun di kabupaten Gowa. Metode tersebut adalah metode Rata-rata Aritmetika (RA), Rasio Normal (RN), Rasio Normal Termodifikasi (RNT), Koefisien Korelasi Terboboti (KKT) dan Pembobot Invers Jarak (PIJ). Pemilihan metode yang sesuai berdasarkan kriteria *Mean Absolute Error*, *Root Mean Square Error*, indeks-S, dan Koefisien Korelasi. Data yang digunakan adalah data curah hujan bulanan dari 12 stasiun hujan tersebut untuk periode 1988-2017. Berdasarkan persentase data tidak lengkap pada setiap stasiun kajian, hasil penelitian menunjukkan bahwa metode RA sesuai digunakan untuk stasiun M5, G5 dan, G7. Metode RNT sesuai untuk stasiun M4 dan G1. Metode PIJ sesuai untuk stasiun M1, M2, M3, G3 dan G6. Sementara untuk stasiun G2 dan G4 lebih sesuai digunakan metode RN. Selain itu metode PIJ dapat juga digunakan untuk mengestimasi data tidak lengkap di stasiun G2. Hasil penelitian ini selanjutnya akan diaplikasikan untuk melengkapi data curah hujan pada seluruh stasiun kajian.

Kata Kunci: curah hujan, data tidak lengkap, rasio normal, rasio normal termodifikasi, koefisien korelasi terboboti, pembobot invers jarak.

1. PENDAHULUAN

Salah satu masalah yang sering dialami dalam suatu penelitian adalah tidak lengkapnya data yang akan digunakan. Hal tersebut juga menjadi masalah dalam analisis iklim. Keberadaan data tidak lengkap tersebut dapat menyebabkan penelitian menjadi bias, bahkan mungkin hasil yang diperoleh akan meragukan. Dalam analisis iklim, seperti analisis curah hujan, keberadaan data tidak lengkap mungkin disebabkan oleh beberapa faktor, seperti

kerusakan alat pengukur hujan, pemindahan lokasi alat pengukur hujan, atau kesalahan petugas pencatat jumlah curah hujan, baik kesalahan dalam membaca alat tersebut maupun kesalahan dalam penginputan data. Berdasarkan hal tersebut, maka penganggaran data tidak lengkap merupakan hal penting yang harus dilakukan sebelum menganalisis data hujan lebih lanjut ((Laili & Sutikno 2013, Sanusi, 2016, Suhaila, dkk 2008)

Beberapa metode interpolasi yang dapat digunakan dalam mengestimasi data tidak lengkap antara lain metode Regresi, Rata-rata Aritmetika, Rasio Normal, Rasio Normal Termodifikasi, Pembobot Invers Jarak, Koefisien Korelasi Terboboti, dan Spline (Caldera dkk, 2016; De Silva, 2007; Jamaluddin, dkk 2008; Sanusi, dkk, 2017)). Dalam tulisan ini digunakan metode Rata-rata Aritmetika, Rasio Normal, Rasio Normal Termodifikasi, Koefisien Korelasi Terboboti dan Pembobot Invers Jarak. Pemilihan metode yang sesuai dengan data yang ada berdasarkan kriteria *Mean Absolute Error*, *Root Mean Square Error*, indeks-S, dan Koefisien Korelasi.

2. Metode Penelitian

a. Metode estimasi data hilang

Metode estimasi yang paling sederhana adalah rata-rata hitung (RA). Sementara Paulhus dan Kohler (Jamaluddin dkk. 2008) telah memperkenalkan metode rasio normal (RN) dengan rumus yang diberikan pada persamaan (1),

$$P_x = \frac{1}{N} \sum_{\substack{j=1 \\ j \neq x}}^N \left[\frac{\mu_x}{\mu_j} \right] P_j \quad (1)$$

dengan μ_x dan μ_j masing-masing nilai rata-rata curah hujan yang ada pada stasiun target x (stasiun yang mempunyai data hilang) dan stasiun tetangga j . Selanjutnya, Jamaluddin dkk. (2008) memperkenalkan metode rasio normal termodifikasi (RNT) dan metode koefisien korelasi terboboti (KKT) masing-masing diberikan pada persamaan (2) dan (3) berikut ini:

$$P_x = \sum_{\substack{j=1 \\ j \neq x}}^N \left[\frac{(m_j - 2)r_{jx}^2 (1 - r_{jx}^2)^{-1}}{\sum_{\substack{j=1 \\ j \neq x}}^N (m_j - 2)r_{jx}^2 (1 - r_{jx}^2)^{-1}} \right] P_j \quad (2)$$

$$P_x = \sum_{\substack{j=1 \\ j \neq x}}^N \left[\frac{r_{jx}}{\sum_{\substack{j=1 \\ j \neq x}}^N r_{jx}} \right] P_j \quad (3)$$

dengan r_{jx} merupakan nilai koefisien korelasi data hujan antara stasiun target x dan stasiun tetangga j dan m_j merupakan panjang data yang digunakan untuk menghitung koefisien korelasi. Metode lainnya adalah Pembobot Invers Jarak (PIJ) yang didefinisikan sebagai berikut:

$$P_x = \sum_{j=1}^N \left[\frac{\frac{1}{d_j}}{\sum_{j=1}^N \frac{1}{d_j}} \right] P_j \quad (4)$$

dengan d_j merupakan jarak antara stasiun target x dan stasiun tetangga j (Caldera dkk, 2016).

Hal penting lainnya dalam estimasi data tidak lengkap adalah pemilihan stasiun tetangga. Pemilihan tersebut selain berdasarkan pada kedekatan menurut letak geografisnya, juga tingginya nilai koefisien korelasi data hujan antara stasiun target dan stasiun calon tetangga. Sementara itu,

beberapa peneliti telah menyarankan bahwa penggunaan tiga atau empat stasiun tetangga adalah cukup untuk mengestimasi data hilang (Jamaluddin dkk. 2008).

b. Sumber data

Data penelitian yang digunakan adalah data jumlah curah hujan bulanan (mm) dari 5 stasiun hujan di kabupaten Maros dan 7 stasiun hujan di kabupaten Gowa masing-masing dari tahun 1988 hingga 2017. Pemilihan stasiun-stasiun hujan tersebut berdasarkan kelengkapan dan panjang data. Data tersebut diperoleh dari Badan Meteorologi, Klimatologi dan Geofisika Wilayah IV propinsi Sulawesi Selatan dan Dinas Sumber Daya Air, Cipta Karya dan Tata Ruang propinsi Sulawesi Selatan. Nama stasiun hujan dan persentase data lengkap masing-masing stasiun diberikan pada Tabel 1. Pada tabel tersebut terlihat bahwa persentase data tidak lengkap antara 3.6% dan 14.4%.

Tabel 1 Nama stasiun hujan dan persentase data tidak lengkap

Nama Stasiun	Data tidak lengkap (%)
Hasanuddin Maros (M1)	3.6
Batu Bassi Maros (M2)	9.7
Bonti-bonti Maros (M3)	7.2
Tanralili Maros (M4)	10.3
Manrimisi Maros (M5)	14.4
Pagentungan Gowa (G1)	3.9
Barembeng Gowa (G2)	11.9
Bonto Sallang Gowa (G3)	11.1
Tinggi Mae Gowa (G4)	9.7
Sungguminasa Gowa (G5)	8.3
Tete Batu Gowa (G6)	6.1
Sanro Bone Gowa (G7)	13.6

c. Prosedur estimasi data tidak lengkap

Langkah-langkah yang dilakukan sebagai berikut:

- 1) Menyusun data curah hujan seluruh stasiun yang akan digunakan dalam penelitian dengan mengambil data yang tersedia pada waktu yang sama.
- 2) Menghitung nilai koefisien korelasi curah hujan antar stasiun.
- 3) Menentukan stasiun-stasiun tetangga untuk setiap stasiun target dengan mempertimbang nilai koefisien korelasi yang tertinggi antara stasiun target dan tetangga.
- 4) Pertama-tama, data dari stasiun target diambil sebanyak 5% secara acak, di mana data ini diandaikan sebagai data tidak lengkap.
- 5) Menghitung estimasi data hilang dengan metode RA, RN, RNT, KKT dan PIJ dengan menggunakan persamaan (1) – (4).
- 6) Menghitung nilai empat uji statistik, yaitu *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), indeks-S, dan Koefisien Korelasi (KK). Rumus keempat uji statistik tersebut dapat dilihat dalam Sanusi, dkk (2017).
- 7) Memilih metode estimasi yang terbaik berdasarkan nilai terkecil dari MAE dan RMSE dan nilai terbesar dari Indeks-S dan KK.
- 8) Mengulang langkah (4) – (8) untuk andaian data tidak lengkap, masing-masing 10% dan 15%.
- 9) Ulangi prosedur ini pada setiap stasiun hujan kajian.

3. Hasil dan Pembahasan

Berdasarkan Tabel 1, maka dalam penelitian ini diujicobakan persentase data tidak lengkap masing-masing 5%, 10% dan 15% yang hasilnya diberikan pada Tabel 2. Tabel tersebut menunjukkan bahwa untuk 5% data tidak lengkap, metode PIJ dan RN pada umumnya memberikan nilai MAE dan RMSE yang minimum serta nilai Indeks-S dan KK yang maksimum. Sementara untuk 10% dan 15%, selain metode PIJ, pada umumnya metode RNT lebih sesuai digunakan untuk mengestimasi data yang tidak lengkap. Namun demikian masing-masing metode RA dan RN dapat juga menjadi alternatif metode estimasi bagi stasiun G4, G5 dan G7.

Berdasarkan persentase data tidak lengkap dalam Tabel 1, maka secara ringkas diberikan metode estimasi yang dapat digunakan untuk masing-masing stasiun hujan seperti yang diperlihatkan dalam Tabel 3.

Tabel 2 Nilai masing-masing kriteria berdasarkan metode estimasi dan persentase data tidak lengkap

Stasiun	Metode/ Kriteria	5%					10%					15%				
		RA	RN	RNT	KKT	PIJ	RA	RN	RNT	KKT	PIJ	RA	RN	RNT	KKT	PIJ
M1	MAE	109.42	114.69	80.32	104.53	63.62	63.02	66.44	47.74	57.11	101.26	71.69	82.33	65.15	67.75	100.77
	RMSE	122.78	126.84	94.23	113.86	67.57	91	96.13	67.2	83.34	123.64	92.92	110.26	87.52	88.06	124.24
	Indeks S	0.96	0.96	0.98	0.97	0.99	0.97	0.97	0.99	0.98	0.94	0.97	0.96	0.98	0.98	0.95
	KK	0.96	0.96	0.97	0.96	0.98	0.97	0.97	0.98	0.97	0.9	0.95	0.95	0.96	0.96	0.91
M2	MAE	149.06	175.93	146.96	149.68	40.87	116.99	105.2	112.5	115.98	80.2	96.2	88.91	87.78	93.28	68.41
	RMSE	201.86	193.97	172.7	191.98	67.04	151.94	122.86	146.95	149.08	92.74	128.6	119.66	126.57	128.83	119.93
	Indeks S	0.94	0.95	0.96	0.95	0.97	0.89	0.94	0.91	0.9	0.97	0.86	0.89	0.87	0.86	0.92
	KK	0.93	0.93	0.94	0.93	0.97	0.9	0.91	0.88	0.89	0.96	0.83	0.83	0.8	0.82	0.87
M3	MAE	173.34	113.79	136.17	163.78	34.26	135.25	107.55	96.43	122.67	47.38	125.97	125.36	99.21	119.61	113.81
	RMSE	232.15	156.33	197.38	223.2	38.33	156.28	120.95	110.96	139.84	67.17	168.05	166.27	137.08	160.64	203.02
	Indeks S	0.91	0.96	0.93	0.92	0.99	0.95	0.97	0.98	0.96	0.99	0.89	0.9	0.93	0.9	0.85
	KK	0.99	0.99	0.99	0.99	0.98	0.98	0.97	0.98	0.98	0.99	0.93	0.92	0.93	0.93	0.75
M4	MAE	88.35	100.42	102.97	96.08	109.41	83.3	78.62	78.35	81.49	66.31	128.58	133.67	114.12	126.03	76.45
	RMSE	123.68	116.13	123.5	123.74	156.97	106.12	105.79	100.78	104.01	118.81	162.52	165.3	158.39	161.6	87.77
	Indeks S	0.96	0.96	0.95	0.95	0.9	0.92	0.92	0.93	0.92	0.96	0.91	0.91	0.92	0.91	0.96
	KK	0.93	0.93	0.94	0.94	0.84	0.87	0.86	0.88	0.87	0.94	0.85	0.85	0.85	0.85	0.95
M5	MAE	184.42	169.61	180.85	182.78	112.35	365.07	401.6	364.58	364.77	116.25	127.77	141.83	129.4	128.44	271.77
	RMSE	221.37	206.36	218.74	220.12	133.56	1021.9	1066.52	1020.09	1021.15	140.2	173.04	190.94	174.22	173.54	407.78
	Indeks S	0.65	0.64	0.66	0.66	0.92	0.35	0.36	0.35	0.35	0.79	0.9	0.85	0.89	0.9	0.71
	KK	0.42	0.44	0.44	0.42	0.88	0.66	0.67	0.66	0.66	0.67	0.83	0.83	0.83	0.83	0.64
G1	MAE	58.29	57.91	48.72	52.53	70.15	101.6	109.06	84.26	96.51	143.1	107.05	145.96	79.76	101.1	95.44
	RMSE	66.87	83.42	62.86	64.32	115.19	139.86	152.35	100.87	127.56	181.93	156.59	202.23	124.67	147.82	128.23
	Indeks S	0.99	0.98	0.99	0.99	0.93	0.97	0.97	0.99	0.98	0.89	0.94	0.9	0.97	0.95	0.92
	KK	0.98	0.97	0.98	0.98	0.93	0.99	0.99	0.99	0.99	0.95	0.96	0.96	0.96	0.96	0.86
G2	MAE	162.09	163.17	151.81	157.53	79.62	126.95	159.55	113.42	121.26	132.06	144.81	117.91	146.87	144.64	134.42
	RMSE	187.16	218.67	168.8	178.76	84.12	160.83	225.24	139.38	151.55	179.29	180.74	146.79	189.71	182.62	185.95
	Indeks S	0.92	0.88	0.94	0.93	0.87	0.94	0.88	0.96	0.95	0.77	0.84	0.87	0.84	0.84	0.93
	KK	0.92	0.92	0.92	0.92	0.92	0.98	0.98	0.99	0.99	0.69	0.76	0.76	0.77	0.77	0.89
G3	MAE	118.77	103.28	119.16	120.71	53.44	172.82	162.32	169.88	171.6	43.37	89.58	86.19	88.83	88.66	48.58
	RMSE	130.33	124.12	131.97	132.66	77.93	255.44	303.2	247.59	252.77	50.36	116.52	123.36	118.1	116.15	72.27
	Indeks S	0.79	0.76	0.8	0.79	0.96	0.8	0.67	0.81	0.8	0.99	0.92	0.89	0.93	0.93	0.97
	KK	0.66	0.65	0.67	0.65	0.99	0.81	0.82	0.83	0.82	0.98	0.88	0.88	0.9	0.89	0.95
G4	MAE	119.62	77.48	111.24	117.46	96.89	140.39	97.33	138.7	141.14	97.96	96.4	114.23	94.87	95.73	157.1
	RMSE	157.16	93.78	151.55	154.72	120.89	172.61	140.59	172.27	173.72	161.3	131.52	152.62	131.43	131.9	204.77
	Indeks S	0.84	0.91	0.84	0.84	0.77	0.79	0.81	0.8	0.79	0.78	0.84	0.74	0.84	0.84	0.85
	KK	0.85	0.85	0.82	0.84	0.83	0.68	0.7	0.71	0.69	0.72	0.71	0.72	0.7	0.7	0.79
G5	MAE	132.71	118.74	134.1	132.77	156.22	96.02	158.84	95.5	96.63	140.96	100.74	90.94	107.2	102.57	107.67
	RMSE	179.52	137.19	186.65	180.35	224.35	150.8	198.62	154.28	152.37	175.7	134.17	124.38	139.58	134.36	143.45
	Indeks S	0.75	0.78	0.74	0.75	0.76	0.93	0.83	0.93	0.93	0.9	0.89	0.86	0.89	0.89	0.92
	KK	0.69	0.72	0.66	0.69	0.62	0.88	0.88	0.88	0.88	0.83	0.86	0.86	0.87	0.87	0.88
G6	MAE	125.28	165.36	105.87	119.1	50.25	111.14	122.94	100.9	106.45	92.02	75.46	92.7	63.66	71.07	49.08
	RMSE	172.95	212.92	148.59	164.59	54.52	153.12	172.96	139.96	147.79	119.51	95.59	123.39	88.05	90.17	63.82
	Indeks S	0.9	0.85	0.93	0.91	0.99	0.94	0.92	0.96	0.95	0.96	0.96	0.93	0.97	0.97	0.98
	KK	0.96	0.96	0.97	0.97	0.99	0.93	0.93	0.94	0.94	0.95	0.93	0.93	0.94	0.94	0.96
G7	MAE	133	134.86	150.85	140.67	129.8	219.37	125.08	222.45	219.43	167.88	77.23	71	81.42	79.34	175
	RMSE	139.89	173.73	157.72	146.94	159.59	297.6	167.03	306.95	301.59	195.47	94.32	95.1	101.16	97.41	236.03
	Indeks S	0.95	0.89	0.94	0.95	0.84	0.73	0.84	0.72	0.72	0.88	0.93	0.89	0.92	0.93	0.82
	KK	0.96	0.97	0.95	0.96	0.72	0.73	0.72	0.7	0.71	0.85	0.94	0.94	0.93	0.93	0.84

Tabel 3 Metode estimasi yang sesuai persentase data tidak lengkap

Stasiun	Metode	Stasiun	Metode
M1	PIJ	G2	RN atau PIJ
M2	PIJ	G3	PIJ
M3	PIJ	G4	RN
M4	RNT	G5	RA
M5	RA	G6	PIJ
G1	RNT	G7	RA

4. Kesimpulan

Hasil penelitian ini menunjukkan bahwa untuk mengestimasi data hujan yang tidak lengkap diperlukan metode estimasi yang sesuai dengan mempertimbangkan persentase data tidak lengkap dalam kumpulan data curah hujan tersebut. Hasil penelitian ini selanjutnya digunakan untuk melengkapi data masing-masing stasiun hujan kajian.

5. Ucapan terima kasih

Penulis mengucapkan terima kasih kepada Kementerian Riset, Teknologi dan Pendidikan Tinggi atas bantuan dana penelitian melalui hibah Penelitian Dasar Unggulan Perguruan Tinggi (PDUPT) 2018 dengan nomor kontrak: 36/UN36.9/PL/2018. Terima kasih pula kami ucapkan kepada BMKG wilayah IV propinsi Sulawesi Selatan dan Dinas Sumber Daya Air, Cipta Karya dan Tata Ruang propinsi Sulawesi Selatan atas kesediannya menyiapkan data curah hujan untuk penelitian ini.

Daftar Pustaka

- Caldera, H.P.G.M., Piyathisse, V.R.P.C. & Nandalal, K.D.W. 2016. A Comparison of Methods of Estimating Missing Daily Rainfall Data. *Engineer*, 49(04): 1-8.
- De Silva, R.P., Dayawansa, N.D.K. & Ratnasiri, D. 2007. A comparison of methods used in estimating missing rainfall data. *Agricultural Sciences* 3(2): 101-108
- Jamaluddin, S., Deni, S. M. & Jemain, A. A. 2008. Revised spatial weighting methods for estimation of missing rainfall data. *Asia-Pacific Journal of Atmospheric Sciences*, 44(2): 93 – 104

- Laili, Z. N. & Sutikno, 2013. Perbandingan Uji Homogenitas Runtun Data Curah Hujan Sebagai Pra-Pemrosesan Kajian Perubahan Iklim. *Jurnal Sains dan Seni POMITS*, 2 (2): D-255-D-259.
- Sanusi, W. 2016. Analisis homogenitas data curah hujan kota Makassar. *Indonesian Journal of Fundamental Sciences*.
- Sanusi,W, Wan Zin, W. Z, Mulbar, U, Danial, M. & Side, S. 2017. Comparison of the methods to estimate missing values in monthly precipitation data. *IJASEIT*, 7(6): 2168-2174.
- Suhaila, J. Deni, S. M. & Jemain, A. A. 2008. Detecting inhomogeneity of rainfall series in Peninsular Malaysia. *Asia Pasific Journal of Atmospheric Sciences*, 44: 369-380.