

METODE CLUSTER UNTUK MELENGKAPI DATA CURAH HUJAN PADA SUMBER AIR WADUK SAGULING (THE CLUSTER METHOD FOR COMPLETING RAINFALL DATA FOR WATER RESOURCES SAGULING RESERVOIR)

Lilik S. Supriatin¹, Sinta B. Sipayung

Pusat Sains dan Teknologi Atmosfer, Lapan

¹e-mail: lilik_lapan@yahoo.com

ABSTRACT

In this paper we discuss statistical analysis of the cluster method was applied to fill the rainfall data that is incomplete. As the case study was the rainfall data from 11 observation stations of the rainfall that was the source of water for the Saguling reservoir in producing electricity energy. The eleven observation stations of this rainfall were Cicalengka, Paseh, Chinchona, Ciparay, Ujung Berung, Bandung, Cililin, Montaya, Sukawana, Saguling, and Cisondari. The results show that the first, it is not all cluster or the data group from the observation station of the rainfall that statistically had the resemblance (in the same cluster) will have the high correlation coefficient. Only 35% from 120 groups data of the rain gauge station that was produced by the analysis cluster had the correlation coefficient 40% -70%, around 7% from 120 groups had the correlation coefficient 70% - 90%, in fact were received by groups cluster that had the value of the coefficient of the negative correlation. Was based on the analysis of the correlation coefficient ($r > 0.43$), the method cluster to fill up the incomplete rainfall data had the accuracy 42,5%.The second that the cluster method could not be used to fill up the data of the daily rainfall that was empty to the group (cluster) that had the low correlation coefficient. The third that the cluster method only could be used to fill up the empty rainfall data if results of the analysis cluster also show that the correlation coefficient that was high ($r > 0.5$) between two climatology stations so as the rainfall data that was empty could be from a climatology station expected from the other climatology station that including in cluster that was same.

Keywords: *Rainfall, Cluster, Data*

ABSTRAK

Makalah ini membahas analisis statistik dengan metode *cluster* (pengelompokkan) yang diaplikasikan untuk mengisi data curah hujan yang tidak lengkap (kosong/*blank*). Sebagai studi kasus adalah data curah hujan dari 11 stasiun pengamatan curah hujan yang merupakan sumber air untuk berlangsungnya kinerja waduk Saguling dalam menghasilkan energi listrik. Sebelas stasiun pengamatan curah hujan yang datanya digunakan adalah Cicalengka, Paseh, Chinchona, Ciparay, Ujung Berung, Bandung, Cililin, Montaya, Sukawana, Saguling, dan Cisondari. Hasil analisis *cluster* menunjukkan bahwa, pertama, tidak semua *cluster* atau kelompok data dari stasiun pengamatan curah hujan yang secara statistik memiliki kemiripan (dalam satu *cluster*) akan memiliki koefisien korelasi yang tinggi pula. Hanya 35% dari 120 kelompok data stasiun penakar hujan yang dihasilkan oleh analisis *cluster* memiliki koefisien korelasi 40% -70%, sekitar 7% dari 120 kelompok memiliki koefisien korelasi 70% - 90%, bahkan terdapat kelompok-kelompok *cluster* yang memiliki nilai koefisien korelasi negatif. Berdasarkan analisis koefisien korelasi ($r > 0,43$), metode *cluster* untuk mengisi data curah hujan yang tidak lengkap memiliki ketepatan 42,5%. Kedua,

metode *cluster* tidak dapat digunakan untuk mengisi data curah hujan harian yang kosong pada kelompok (*cluster*) yang memiliki koefisien korelasi yang rendah. Ketiga, metode *cluster* hanya dapat digunakan untuk mengisi data curah hujan yang kosong jika hasil dari analisis *cluster* juga menunjukkan koefisien korelasi yang tinggi ($r > 0,5$) antara dua stasiun klimatologi sehingga data curah hujan yang kosong dari suatu stasiun klimatologi dapat diduga dari stasiun klimatologi lain yang termasuk dalam *cluster* yang sama.

Kata Kunci : *curah hujan, cluster, data*

1 PENDAHULUAN

Data curah hujan seringkali ditemukan dalam kondisi tidak lengkap (terputus) pada suatu periode waktu atau tidak tercatatnya data curah hujan. Kondisi ini dapat disebabkan oleh tiga hal, yaitu instrumen pengukur yang tidak bekerja dengan baik (rusak), lokasi stasiun klimatologi yang digusur untuk peruntukan lain, dan *human eror*. Tidak tercatatnya data curah hujan pada suatu stasiun klimatologi dapat dilengkapi dengan memanfaatkan data curah hujan dari tempat lain yang berdekatan.

Banyaknya data curah hujan harian yang kosong menjadi hambatan untuk estimasi kemampuan produksi energi listrik pada sebuah Pembangkit Listrik Tenaga Air (PLTA), khususnya PLTA Saguling. Kebutuhan data curah hujan harian yang lengkap amat penting untuk keberlangsungan kinerja PLTA. Hal ini dikarenakan curah hujan adalah sumber air dan juga sumber tenaga bagi sebuah PLTA. Tinggi muka air waduk yang kurang dari normal tidak akan dapat memutar turbin generator listrik. Akibatnya secara otomatis energi listrik juga tidak dapat dihasilkan. Estimasi tinggi muka air waduk dapat diketahui dengan mengetahui tinggi curah hujan yang jatuh pada daerah tangkapan air (*catchment area*) waduk Saguling.

Saguling adalah PLTA yang terletak pada Daerah Aliran Sungai (DAS) Citarum dengan kapasitas terpasang (700 MW) dan energi yang dibangkitkan adalah paling besar (2156 GWH/tahun) di antara dua PLTA lain yang terletak pada DAS yang sama yaitu PLTA Cirata dan Jatiluhur (Supriatin,

2004). Hal ini dikarenakan PLTA Saguling terletak pada hulu sungai Citarum. Bagian hulu sungai cenderung memiliki debit air yang lebih besar daripada bagian hilir sungai. PLTA Saguling berfungsi untuk menghasilkan listrik dan meningkatkan kehandalan sistem kelistrikan di pulau Jawa dan pulau Bali. Gardu Induk Tegangan Ekstra Tinggi (GITET) Saguling diinterkoneksi dengan sistem kelistrikan Jawa-Bali melalui Saluran Udara Tegangan Ekstra Tinggi (SUTET) sebesar 500 mega volt (MV).

Saat ini sedikitnya terdapat tiga metode untuk menduga data curah hujan yang tidak lengkap (kosong). Ketiga metode itu adalah rata-rata aritmatika, *polygon Thiessen*, dan metode *isohet*. Dalam rangka melengkapi metode pengisian data curah hujan yang sudah ada sebelumnya (seperti tersebut pada pernyataan di atas), maka akan diperkenalkan metode *cluster* untuk mengisi data curah hujan yang tidak lengkap (kosong). Masalah dalam penelitian ini adalah belum diketahuinya efektivitas metode *cluster* untuk melengkapi data curah hujan yang kosong. Tujuan dari penelitian ini adalah menganalisis potensi dari metode *cluster* untuk melengkapi data curah hujan yang kosong. Sasaran dari penelitian ini adalah diperolehnya efektivitas metode *cluster* untuk melengkapi data curah hujan yang kosong (tidak lengkap). Kontribusi penelitian ini pada pengembangan sains atmosfer adalah menambah sebuah metode untuk mengisi data curah hujan yang kosong.

2 TINJAUAN PUSTAKA

Asdak (2002) dalam bukunya menyatakan bahwa terdapat tiga metode untuk mengisi data curah hujan yang kosong yaitu rata-rata aritmatika, *polygon Thiessen*, dan metode isohet. Menurut Dunne dan Leopold (1978) dan Wanielista (1990) metode rata-rata aritmatika dibedakan menjadi dua. Pertama, jika besarnya perbedaan antara curah hujan rata-rata tahunan dari masing-masing ketiga stasiun penakar hujan (P_a , P_b , P_c) tersebut dengan curah hujan rata-rata tahunan (P_x) dari alat penakar hujan yang akan diduga adalah kurang dari 10%, maka metode rata-rata aritmatika yang digunakan menggunakan persamaan berikut ini:

$$P_x = \frac{P_a + P_b + P_c}{3} \quad (2-1)$$

Kedua, jika besarnya perbedaan antara curah hujan rata-rata tahunan dari masing-masing stasiun penakar hujan (P_a dan P_b) tersebut dengan P_x adalah lebih dari 10%, maka metode rata-rata aritmatika yang digunakan menggunakan metode perbandingan

normal dengan persamaan seperti berikut ini:

$$P_x = \frac{1}{2} \left(\frac{N_x}{N_a} \right) P_a + \left(\frac{N_x}{N_b} \right) P_b \quad (2-2)$$

N_x , N_a , dan N_b adalah curah hujan normal jangka panjang dari tiga stasiun penakar hujan.

Pada metode *polygon Thiessen* setiap bagian wilayah (daerah) tangkapan hujan diwakili secara proporsional oleh satu alat penakar hujan. Pada metode *isohet*, hal pertama yang harus dilakukan adalah membuat garis-garis *isohet* (garis yang menghubungkan wilayah/daerah dengan curah hujan yang sama). Berdasarkan garis *isohet* akan diketahui wilayah cakupan hujan (daerah antara dua garis *isohet*) dengan besar curah hujan tertentu.

3 DATA DAN METODE

Data yang digunakan adalah data curah hujan harian tahun 2010 dari 11 stasiun klimatologi. Tabel 3-1 menyajikan letak geografi (koordinat geografi) lokasi penelitian dari 11 stasiun klimatologi dan elevasinya (ketinggian dari permukaan laut).

Tabel 3-1: LETAK GEOGRAFI LOKASI PENELITIAN

No.	Stasiun Klimatologi	Koordinat Geografi	Elevasi (m dpl)
1.	Cicalengka	6°59'LS, 107°03'BT	700
2.	Paseh	7°06'LS, 107°03'BT	876
3.	Chinchona	7°01'LS, 107°02'BT	1356
4.	Ciparay	7°01'LS, 107°02'BT	689
5.	Ujung Berung	6°56'LS, 107°02'BT	750
6.	Bandung	6°53'LS, 107°02'BT	782
7.	Cililin	6°56'LS, 107°01'BT	662
8.	Montaya	7°01' LS, 107°01'BT	1089
9.	Sukawana	6°47'LS, 107°02'BT	653
10.	Saguling	6°54'LS, 107°01'BT	658
11.	Cisondari	7°05'LS, 107°01'BT	836

Tabel 3-2: KRITERIA KOEFISIEN KORELASI

Koefisien Korelasi (r)	Analisis
$0 \leq r < 0,2$	Hubungan sangat kecil/sangat lemah
$0,2 \leq r < 0,4$	Hubungan kecil/lemah
$0,4 \leq r < 0,7$	Hubungan moderat/ sedang
$0,7 \leq r < 0,9$	Hubungan erat/kuat
$0,9 \leq r < 1$	Hubungan sangat erat /sangat kuat

Ke 11 stasiun klimatologi tersebut akan dikelompokkan data curah hujan hariannya dengan metode *cluster*. Analisis *cluster* adalah analisis data multivariat yang digunakan untuk mengelompokkan obyek berdasarkan kemiripan sifat obyek (Richard A. Johnson dalam Tantular, 2013). Analisis yang digunakan pada penelitian ini bahwa hasil *cluster* yang berupa dua stasiun klimatologi sebagai satu kelompok yang memiliki kemiripan data curah hujan harian sehingga ke dua stasiun klimatologi satu sama lain dapat digunakan untuk menduga data curah hujan yang tidak lengkap.

Validasi dari metode *cluster* untuk aplikasi mengisi data curah hujan harian yang kosong dilakukan dengan analisis koefisien korelasi data curah hujan harian antara dua stasiun klimatologi yang termasuk dalam *cluster* yang sama. Evaluasi untuk mengetahui apakah aplikasi metode *cluster* untuk mengisi data curah hujan harian yang kosong dapat dipergunakan atau tidak, akan dilakukan dengan menghitung persentase jumlah *cluster* (pasangan stasiun klimatologi) menurut nilai koefisien korelasinya (r). Gulford dalam Tantular (2013) menafsirkan nilai koefisien korelasi dengan kriteria seperti tersaji pada Tabel 3-2.

4 HASIL DAN PEMBAHASAN

Hasil pengolahan data dengan metode *cluster* diperoleh 120 kelompok stasiun penakar hujan. Satu kelompok *cluster* terdiri dari dua stasiun penakar hujan. Tabel 4-1 menyajikan koefisien korelasi antara 2 buah stasiun

klimatologi yang dihasilkan oleh metode *cluster* pada penelitian ini.

Dari Tabel 4-1 dapat diketahui bahwa metode *cluster* yang diaplikasikan untuk mengisi data curah hujan harian yang tidak lengkap (*blank*) hanya memiliki ketepatan antara 90%-100% dengan persentase 0,83%, ketepatan antara 70% sampai dengan 89% dengan persentase 6,67%, ketepatan 40% sampai dengan 69% dengan persentase 35%. Artinya dua stasiun penakar hujan yang diuji koefisien korelasinya dengan besar r antara 40% sampai dengan 100% adalah 42,5%. Hal ini menunjukkan metode *cluster* untuk mengisi data curah hujan yang kosong memiliki ketepatan 42,5%.

Kecilnya persentase ketepatan metode *cluster* untuk mengisi data curah hujan harian yang *blank* dapat disebabkan oleh sifat variabilitas curah hujan di Indonesia yang tinggi menurut tempat (*spasial*) dan waktu (*temporal*). Pada lokasi yang berdekatan seringkali terjadi perbedaan waktu jatuhnya hujan, lama hujan, dan intensitas hujan. Perbedaan variabilitas curah hujan disebabkan oleh faktor pengendali cuaca dan iklim yang berbeda antara satu lokasi dengan lokasi lain. Trewartha (1995) menyebutkan bahwa pola iklim pada suatu kawasan ditentukan oleh faktor pengendali iklim. Faktor pengendali iklim mencakup parameter letak lintang geografi, distribusi daratan dengan laut, sistem angin dan tekanan, arus laut, sifat dan kondisi daratan yang menonjol, sifat dan kondisi lokal.

Tabel 4-1: KOEFISIEN KORELASI DARI KELOMPOK STASIUN PENAKAR HUJAN

Koefisien Korelasi (r)	Jumlah Kelompok (cluster)	Persentase
$r < 0$	14	11,67%
$0 < r < 0,2$	26	21,67%
$0,2 \leq r < 0,4$	29	24,17%
$0,4 \leq r < 0,7$	42	35%
$0,7 \leq r < 0,9$	8	6,67%
$0,9 \leq r < 1$	1	0,83%

Tabel 4-2: JARAK ANTARA DUA STASIUN PENAKAR HUJAN

Stasiun A	Stasiun B	Jarak (km)
Cicalengka	Ciparay	3,70
Cicalengka	Ujung Berung	5,56
Cicalengka	Cililin	5,56
Cicalengka	Montaya	3,70
Paseh	Cisondari	1,85
Ciparay	Cisondari	7,41
Ciparay	Montaya	0
Ujung Berung	Bandung	5,56
Ujung Berung	Cililin	0
Ujung Berung	Saguling	3,70
Bandung	Saguling	1,85
Cililin	Saguling	3,70
Montaya	Cisondari	7,41

Pada penelitian ini rendahnya ketepatan metode *cluster* untuk diaplikasikan mengisi data curah hujan harian yang kosong lebih banyak disebabkan oleh sifat dan kondisi lokal setempat. Jika dilihat pada Tabel 3-1 tentang elevasi, masing-masing stasiun klimatologi (stasiun penakar hujan) memiliki perbedaan letak ketinggian (elevasi) yang berbeda-beda dan bergunung-gunung (tidak datar). Elevasi penakar hujan terendah di Sukawana (653 m dpl) dan elevasi tertinggi di Chinchona (1356 m dpl).

Rendahnya ketepatan metode *cluster* untuk mengisi data curah hujan yang kosong juga disebabkan oleh jauhnya jarak antara stasiun penakar hujan (tidak sesuai standar WMO). *World Meteorological Organization* (WMO) dalam Asdak (2002) menyebutkan bahwa satu alat penakar hujan untuk daerah berupa pulau kecil dapat mencakup daerah dengan luas 25 km², untuk daerah bergunung-gunung (seperti lokasi pada penelitian ini) satu

penakar hujan mewakili daerah dengan luas 100-250 km², daerah dengan topografi datar, satu penakar hujan meliputi luas wilayah 600-900 km². Pernyataan dari WMO tersebut jika disetarakan dengan jarak horisontal, maka satu penakar hujan pada daerah dengan topografi gunung dan tidak rata akan meliputi wilayah dengan jarak horizontal maksimum sebesar 8,92 km.

Untuk mengubah (konversi) dari letak lintang geografi stasiun penakar hujan ke jarak digunakan persamaan kesetaraan bahwa setiap selisih 1 menit lintang setara dengan 1852 m, sehingga setiap selisih 1° lintang setara dengan jarak 111 km. Tabel 4-2 menyajikan jarak (hasil konversi dari letak lintang) antara dua stasiun penakar hujan yang memiliki jarak lebih kecil dari 8,92 km (sesuai ketentuan WMO).

Berdasarkan Tabel 4-2 dapat diketahui bahwa dari 55 pasang stasiun penakar hujan yang diketahui jaraknya, hanya 13 pasang stasiun penakar hujan (23,63%) yang memiliki jarak di bawah

8,92 km. Sebesar 76,37% pasang jarak stasiun penakar hujan memiliki jarak yang lebih besar dari 8,92 km. Hal ini yang menyebabkan mengapa metode *cluster* jika digunakan untuk 11 stasiun penakar hujan memiliki ketepatan yang rendah (42,5%). Berdasarkan Tabel 4-2 juga dapat diketahui bahwa jarak antara dua stasiun yang lebih kecil dari 8,92 km (standar WMO) juga memiliki koefisien korelasi yang tinggi. Daerah atau nama-nama stasiun penakar hujan yang memiliki koefisien korelasi tinggi seperti tercantum pada Tabel 4-2. Metode *cluster* untuk mengisi data curah hujan yang tidak lengkap cocok untuk lokasi-lokasi yang memiliki jarak antara 2 stasiun penakar hujan yang kurang dari 8,92 km.

Secara umum ketelitian hasil prakiraan besarnya curah hujan yang diperoleh dengan menggunakan metode prakiraan besarnya curah hujan akan meningkat dengan; pertama kerapatan alat penakar hujan, ke dua adalah lama waktu pengukuran, ketiga adalah luas daerah tangkapan air yang dikaji.

5 KESIMPULAN

Berdasarkan hasil analisis *cluster* tidak semua kelompok yang dihasilkan oleh analisis *cluster* memiliki koefisien korelasi yang tinggi. Hanya 35% dari 120 kelompok yang dihasilkan oleh analisis *cluster* memiliki koefisien korelasi 40% -70%, sekitar 7% dari 120 kelompok memiliki koefisien korelasi 70% - 90%, bahkan terdapat kelompok

kelompok *cluster* yang memiliki nilai koefisien korelasi negatif. Berdasarkan analisis koefisien korelasi ($r > 0,43$), metode *cluster* untuk mengisi data curah hujan yang tidak lengkap memiliki ketepatan 42,5%.

UCAPAN TERIMA KASIH

Terima kasih kami ucapkan kepada pihak PT. Indonesia Power UBP Saguling, khususnya Bapak Suyatno atas pemberian data curah hujan harian 11 stasiun klimatologi dari tahun 1986-2010.

DAFTAR RUJUKAN

- Asdak, C, 2002. *Hidrologi dan Pengelolaan DAS*, UGM Press, Yogyakarta.
- Dunne, T, L. B. Leopold, 1978. *Water in Environmental Planning*, W. H. Freeman and Company, New York.
- Supriatin, L. S, 2004. *Dampak Sedimentasi Pada Waduk Saguling*, Tesis Pasca Sarjana Ilmu Lingkungan Universitas Indonesia, Tidak dipublikasikan.
- Tantular, B, 2013. *Modul Pelatihan Pengolahan dan Analisis Data Atmosfer*, Jurusan Statistika, Universitas Padjadjaran, Bandung.
- Trewartha, G, T, 1995. *Pengantar Iklim*, UGM Press, Yogyakarta.
- Wanielista, M. P, 1990. *Hydrology and Water Quantity Control*, Jhon Wiley and Sons, New York.