

Analisis Klaster K-Means Terhadap Kualitas Air Sungai di Kota Surakarta, Jawa Tengah

(K-Means Cluster Analysis on River Water Quality in Surakarta City, Central Java)

Yuli Handayani^{1*}, Syafrudin Syafrudin², Suherman Suherman³

¹Program Magister Ilmu Lingkungan, Program Pascasarjana, Universitas Diponegoro

²Jurusan Teknik Lingkungan, Fakultas Teknik, Universitas Diponegoro

³Departemen Teknik Kimia, Fakultas Teknik, Universitas Diponegoro

E-mail: handayani.yuli23@gmail.com

ABSTRAK

Penelitian ini bertujuan untuk membagi Sungai Pepe Atas Hulu, Sungai Pepe Bawah Hulu, Sungai Anyar Hilir, Sungai Jenes Hilir, Sungai Kedung Jumbleng Hilir, Sungai Anyar Tengah, Sungai Pepe Bawah Hilir, Sungai Anyar Hulu, Sungai Gajah Putih Hulu, Sungai Brojo Hilir, Sungai Premulung Tengah, Sungai Premulung Hulu, dan Sungai Premulung Hilir menjadi beberapa klaster berdasarkan tipologi dan tingkat pencemaran yang hampir sama. Data diperoleh dari hasil pemantauan kualitas air yang dilakukan oleh Dinas Lingkungan Hidup Kota Surakarta pada tahun 2020. Terdapat 12 parameter yang digunakan yaitu TSS, TDS, pH, BOD, COD, DO, P, NO₃-N, NH₃, NO₂-N, *Total Coliform*, dan *Fecal Coliform*. Dengan algoritma *k-means* diperoleh dua klaster dengan parameter pembeda *Total Coliform* dan *Fecal Coliform*. Dengan kandungan 3.790.050 MPN/100ml *Total Coliform* dan 604.400 MPN/100ml *Fecal Coliform*, Sungai Jenes Hilir mempunyai kualitas air yang lebih buruk dibandingkan sungai lainnya dalam hal kandungan pencemar *Total Coliform* dan *Fecal Coliform*.

Kata kunci: Analisis klaster, *k-means*, kualitas air sungai, Surakarta

ABSTRACT

This study aims to divide the Upstream Pepe Atas River, Upstream Pepe Bawah River, Downstream Anyar River, Downstream Jenes River, Downstream Kedung Jumbleng River, Middle Anyar River, Downstream Pepe Bawah River, Upstream Anyar River, Upstream White Elephant River, Downstream Brojo River, Middle Premulung River, Upstream Premulung River, and Downstream Premulung River into several clusters based on almost the same typology and pollution level. The data was obtained from the results of water quality monitoring conducted by the Surakarta City Environmental Agency in 2020. There are 12 parameters used, namely TSS, TDS, pH, BOD, COD, DO, P, NO₃-N, NH₃, NO₂-N, Total Coliform, and Fecal Coliform. With the k-means algorithm, two clusters were obtained with differentiating parameters Total Coliform and Fecal Coliform. With a content of 3,790,050 MPN/100ml Total Coliform and 604,400 MPN/100ml Fecal Coliform, the Jenes Hilir River has poorer water quality than other rivers in terms of pollutant content of Total Coliform and Fecal Coliform.

Keywords: Cluster analysis, *k-means*, river water quality, Surakarta

PENDAHULUAN

Sejak awal peradaban manusia, sungai digunakan sebagai tempat pembuangan limbah domestik, industri, dan limbah pertanian. Pada dasarnya sungai-sungai yang memiliki area aliran air yang luas dapat melakukan proses pemurnian diri. Proses pemurnian sendiri meliputi pengenceran dan dispersi, aerasi, oksidasi, reduksi dan dengan peran banyak mikroorganisme akan membuat zat-zat beracun menjadi tidak beracun (Garg, 2018). Di Indonesia, air sungai sangat diperlukan untuk pemenuhan kebutuhan rumah tangga sebagai penyuplai air minum, untuk irigasi di bidang pertanian, mendukung usaha perikanan dan pariwisata, bahkan digunakan sebagai sarana transportasi. Tingginya laju urbanisasi dan cepatnya pertumbuhan industri menjadikan pencemaran air sebagai masalah yang mengkhawatirkan. Pada tahun 2019 sebanyak 67,84% rumah tangga di daerah perkotaan pembuangan air limbah domestiknya berakhir di got/selokan/sungai yang mengakibatkan air sungai tercemar (BPS, 2020). Sehingga penting untuk dilakukan penelitian mengenai kualitas air sungai terutama di daerah perkotaan.

Telah dilakukan banyak penelitian tentang pengelompokan kualitas air sungai menggunakan analisis klaster dengan metode hirarki, di antaranya pengelompokan kualitas air Sungai Varuna di Varanasi, India (Dabgerwal & Tripathi, 2016), air Sungai Markanda, India (Garg, 2018), air Sungai Luan, Cina (Zhao & Cui, 2009), air Sungai Ganga, India (Dimri et al., 2021), dan air Sungai Muda, Malaysia (Azhar et al., 2015). Begitu pula penelitian pengelompokan kualitas air sungai dengan metode *k-means*, antara lain pengelompokan air

sungai di Yogyakarta (Sukmawati, 2019) dan air Sungai Kreung Manee, Aceh Utara (Irfannur & Khairan, 2021), air Sungai Klang, Malaysia (Sharif et al., 2015), dan air Sungai Haihe, Cina (Zou et al., 2015). Beberapa tahun terakhir, analisis tentang kualitas air sungai di Kota Surakarta, Provinsi Jawa Tengah telah dilakukan oleh para peneliti di antaranya Landhiani (2016), Yuliasuti (2011), Rezagama et al. (2019), dan Pitasari (2016) namun tidak diteliti tentang pengelompokan sungai berdasarkan parameter-parameter kualitas air sungai.

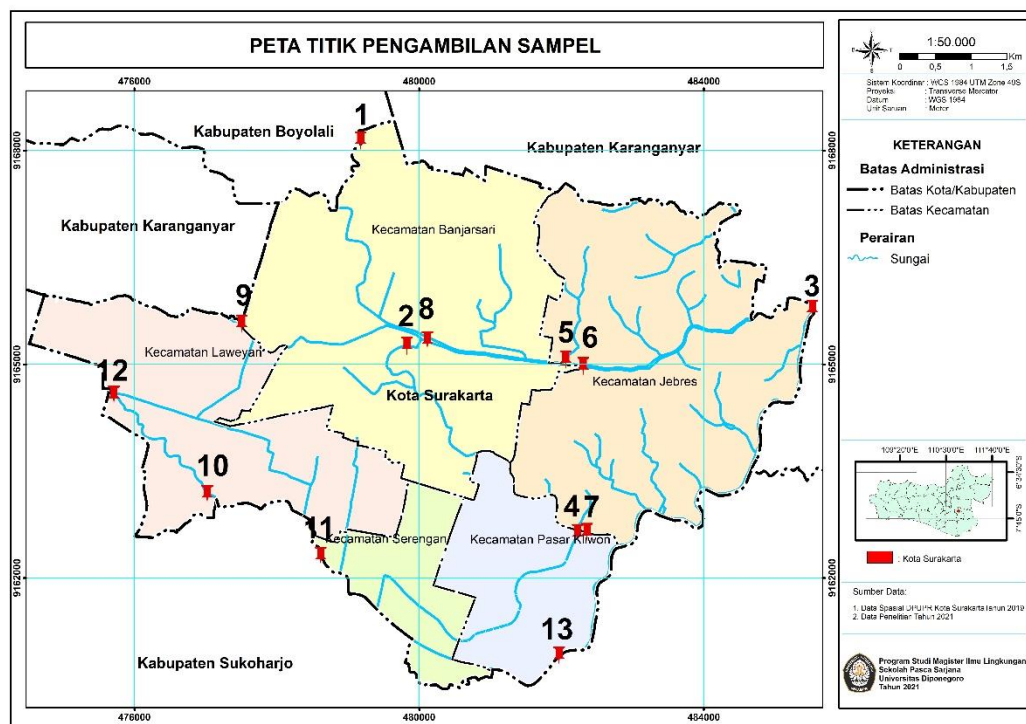
Setiap tahun Dinas Lingkungan Hidup Kota Surakarta melakukan pemantauan kualitas air sungai terutama sungai-sungai yang diperkirakan rawan tercemar limbah domestik dan industri (DLH Kota Surakarta, 2019). Dari data hasil pemantauan yang dilakukan pada tahun 2020, peneliti ingin membagi lokasi-lokasi sampel menjadi beberapa kelompok/klaster berdasarkan tipologi dan tingkat pencemaran yang hampir sama sehingga diharapkan dapat meningkatkan kewaspadaan warga dalam memanfaatkan air sungai, dan bermanfaat bagi Pemerintah Kota Surakarta dalam menentukan kebijakan yang tepat dalam peningkatan kualitas sungai.

METODE

Lokasi dan Variabel Penelitian

Surakarta adalah sebuah kota yang terletak antara $110^{\circ} 45' 15''$ dan $110^{\circ} 45' 35''$ Bujur Timur dan antara $7^{\circ} 36'$ dan $7^{\circ} 56'$ Lintang Selatan di Provinsi Jawa Tengah. Berdasarkan sensus penduduk tahun 2020, Badan Pusat Statistik mencatat penduduk Kota Surakarta sebanyak 522,36 ribu jiwa menghuni wilayah Kota Surakarta seluas 44,04 kilometer persegi. Rata-rata suhu udara pada tahun 2020 berkisar 25°C sampai $28,2^{\circ}\text{C}$, rata-rata kelembaban udara berkisar 71% sampai 88%, dan banyaknya curah hujan berkisar 2 mm sampai 275,3 mm (BPS Kota Surakarta, 2021).

Penelitian ini menggunakan data sekunder berupa hasil pemantauan uji kualitas air sungai di 13 titik lokasi pengambilan sampel yang dilakukan oleh Dinas Lingkungan Hidup Kota Surakarta pada tahun 2020 seperti yang dapat dilihat pada Gambar 1. Lokasi pengambilan sampel tersebut yaitu (1) Sungai Pepe Atas Hulu, (2) Sungai Pepe Bawah Hulu, (3) Sungai Anyar Hilir, (4) Sungai Jenes Hilir, (5) Sungai Kedung Jumbleng Hilir, (6) Sungai Anyar Tengah, (7) Sungai Pepe Bawah Hilir, (8) Sungai Anyar Hulu, (9) Sungai Gajah Putih Hulu, (10) Sungai Brojo Hilir, (11) Sungai Premulung Tengah, (12) Sungai Premulung Hulu, dan (13) Sungai Premulung Hilir.



Gambar 1. Lokasi Pengambilan Sampel

Variabel yang diamati pada penelitian ini adalah 12 parameter kualitas air sungai yang terdiri dari *Total Suspended Solid (TSS)*, *Total Dissolve Solid (TDS)*, *Power of Hydrogen (pH)*, *Biochemical Oxygen Demand*

(BOD), *Chemical Oxygen Demand* (COD), *Dissolved Oxygen* (DO), *Total Fosfat* (P), Nitrat (NO₃-N), Amonia (NH₃), Nitrit (NO₂-N), *Total Coliform*, dan *Fecal Coliform* (Tabel 1).

Tabel 1. Parameter kualitas air sungai.

Parameter	Pengertian	Hubungan dengan kualitas air
(1)	(2)	(3)
1. TSS	Tingkat kekeruhan air yang dapat menghalangi cahaya masuk ke air sehingga mengganggu proses fotosintesis dan respirasi biota air (DLH Kota Surakarta, 2019)	Semakin tinggi jumlah TSS maka kualitas air semakin buruk
2. TDS	Partikel-partikel terlarut yang tidak dapat disaring dengan kertas berpori 0,45 µm (Hadidjah, 2019)	Semakin tinggi jumlah TDS maka kualitas air semakin buruk
3. pH	Derajat keasaman air, berpengaruh pada aktivitas biologis dan kimia dalam air (DLH Kota Surakarta, 2019)	Berada pada rentang 6-9
4. BOD	Jumlah oksigen yang dibutuhkan untuk dekomposisi biologis bahan organik di bawah kondisi aerobik pada suhu standar dan waktu inkubasi (Droste & Gehr, 2019)	Semakin tinggi jumlah BOD maka kualitas air semakin buruk
5. COD	Jumlah oksigen yang dibutuhkan untuk menstabilkan bahan organik (Droste dan Gehr, 2019)	Semakin tinggi jumlah COD maka kualitas air semakin buruk
6. DO	Konsentrasi oksigen terlarut yang sangat penting bagi kelangsungan hidup organisme di air (Droste & Gehr, 2019)	Semakin tinggi jumlah DO maka kualitas air semakin baik
7. P	Secara natural, fosfat tersaring dalam air tanah selanjutnya diterima oleh sistem perairan (DLH Kota Surakarta, 2019)	Semakin tinggi jumlah P maka kualitas air semakin buruk
8. NO ₃ -N	Berasal dari proses oksidasi sempurna senyawa nitrogen dan ammonia di perairan (Trilaksono et al., 2014)	Semakin tinggi jumlah NO ₃ -N maka kualitas air semakin buruk
9. NH ₃	Limbah yang berasal dari aktivitas metabolisme manusia maupun biota air (Sawyer dalam Trilaksono et al. (2014))	Semakin tinggi jumlah NH ₃ maka kualitas air semakin buruk
10. NO ₂ -N	Ditemukan dalam jumlah yang lebih sedikit daripada NO ₃ -N, karena dalam perairan alami segera dioksidasi menjadi NO ₃ -N (Trilaksono et al., 2014).	Semakin tinggi jumlah NO ₂ -N maka kualitas air semakin buruk
11. <i>Total Coliform</i>	Parameter bakteri yang menunjukkan aman atau tidak air untuk dikonsumsi karena dapat menyebabkan penyakit diare (DLH Kota Surakarta, 2019)	Semakin tinggi jumlah <i>Total Coliform</i> maka kualitas air semakin buruk
12. <i>Fecal Coliform</i>	Parameter bakteri yang lebih akurat daripada <i>Total Coliform</i> terhadap indikasi pencemaran dari kotoran manusia atau hewan (DLH Kota Surakarta, 2019)	Semakin tinggi jumlah <i>Fecal Coliform</i> maka kualitas air semakin buruk

Sumber: data diolah dari (DLH Kota Surakarta, 2019), (Droste & Gehr, 2019), (Lampiran VI Peraturan Pemerintah Republik Indonesia Nomor 22 Tahun 2021 Tentang Penyelenggaraan Perlindungan dan Pengelolaan Lingkungan Hidup, 2021), (Hadidjah, 2019), (Trilaksono et al., 2014)

Algoritma K-Means

Pengklasteran *k-means* digunakan untuk mengekstrak kluster dari *dataset* yang telah dioptimasi dengan pemilihan karakteristik khusus (Javadi et al., 2017). Objek-objek yang memiliki karakteristik mirip dikelompokkan menjadi satu kluster yang sama, sedangkan objek-objek dengan karakteristik berbeda dikelompokkan menjadi kluster yang lainnya (Rahmadina et al., 2020). Menurut Faisal dalam Suryono & Triscowati (2020), pengklasteran dengan metode *k-means* bertujuan membagi data dalam beberapa kelompok menggunakan sistem partisi untuk mendapatkan nilai variasi paling kecil dalam satu kluster dan mendapatkan nilai variasi terbesar antar kluster. Setiap kluster akan mempunyai titik pusat yang disebut dengan *centroid*. Algoritma pengklasteran *k-means* seperti yang dijelaskan oleh Singh et al. (2013), Jatnika et al. (2020), dan Suryono & Triscowati (2020) adalah sebagai berikut:

- a. Siapkan *dataset* yang akan diklasterkan
- b. Tentukan jumlah kluster (*k*) yang ingin dibuat
- c. Tentukan *centroid* sebagai titik pusat kluster. Pada iterasi pertama *centroid* dipilih secara acak
- d. Hitung jarak terdekat setiap data ke *centroid* menggunakan persamaan *Euclidean Distance*, yaitu:

$$de = \sqrt{\sum_{i=1}^n (x_i - c_i)^2} \dots\dots\dots(1)$$

dimana:

- x_i = data pada variabel ke *i*
- c_i = *centroid* pada variabel ke *i*
- n = banyak parameter

- e. Tentukan kelompok berdasarkan jarak terpendek ke *centroid* dari setiap kluster

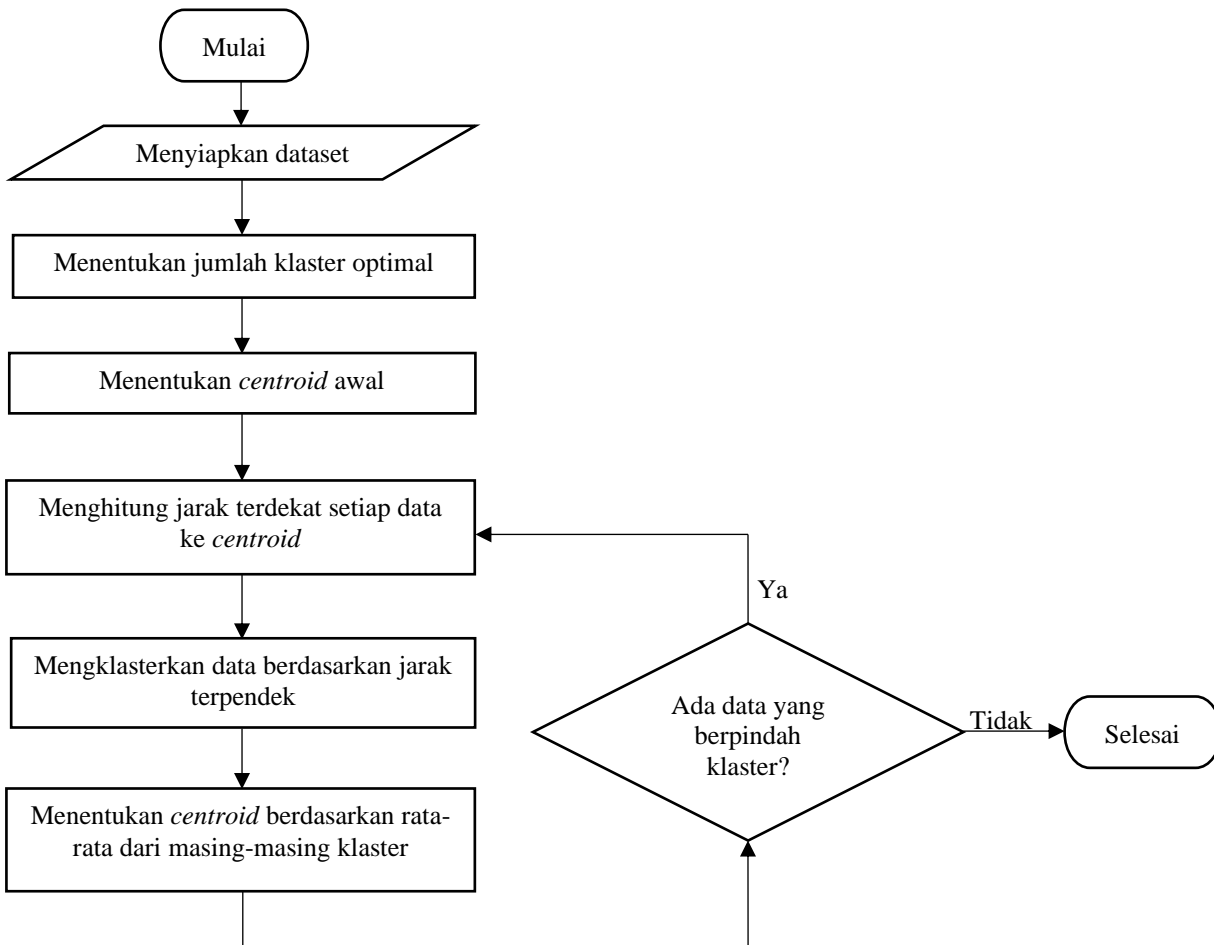
- f. Pada iterasi kedua hingga iterasi terakhir, *centroid* ditentukan berdasarkan rata-rata dari masing-masing kluster menggunakan persamaan seperti di bawah:
- $$C_i = \frac{1}{N_k} \sum_{j=1}^{N_k} X_{ji} \dots\dots\dots (2)$$
- dimana:
- C_i = *centroid* baru pada variabel ke i
 - N_k = banyak anggota kluster ke k
 - X_{ji} = data ke j variabel ke i untuk kluster yang dimaksud
- g. Ulangi proses dari langkah e sampai kelompok data pada iterasi terakhir mempunyai hasil yang sama dengan kumpulan data pada iterasi sebelumnya
- h. Jika anggota kluster sudah tidak mengalami perpindahan kluster lagi, maka algoritma pengklasteran selesai.

Metode Silhouette

Pada algoritma pengklasteran dengan metode *k-means* di atas, disebutkan bahwa jumlah kluster (k) harus ditentukan terlebih dahulu sebelumnya. Maka dari itu sebaiknya dipilih jumlah k yang optimal sehingga tercapai tujuan pengklasteran. Salah satu cara untuk mendapatkan jumlah k optimal adalah dengan metode *silhouette* di samping metode *elbow* (Rahmadina et al., 2020). Metode *silhouette* menggunakan *silhouette scores*. Semakin tinggi *silhouette scores*, maka semakin baik kluster yang terbentuk (Saputra et al., 2020).

Flow Chart Analisis Pengklasteran K-Means

Analisis dalam penelitian ini dibantu oleh perangkat lunak *Orange* dan *SPSS*. Langkah-langkah pengolahan data dan analisis pengklasteran dengan algoritma *k-means* disajikan dalam Gambar 2.



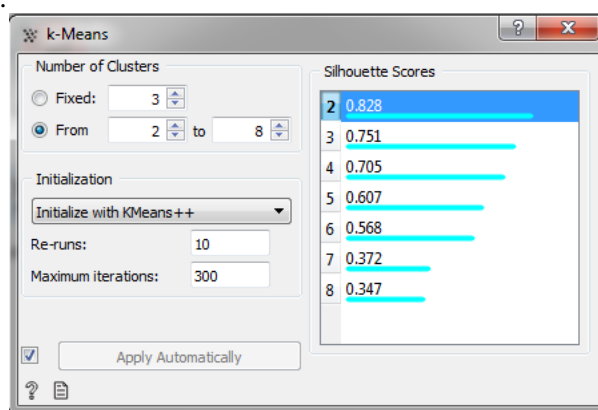
Gambar 2. Flow Chart Analisis Kluster K-Means

HASIL DAN PEMBAHASAN

Uji kualitas air sungai di Kota Surakarta dilakukan dalam tiga kali periode pengukuran selama tahun 2020. Periode pertama pada 13 Februari 2020, periode kedua pada 15 Juli 2020, dan periode ketiga pada 30 September sampai 14 Oktober 2020. Hasil rata-rata pengukuran selama tiga periode tersebut kemudian diolah dengan perangkat lunak *Orange* dan *SPSS*. Sebanyak 12 parameter kualitas air sungai yang dianalisis dalam penelitian ini ditampilkan dalam Tabel 1. Data kandungan BOD yang tertulis < 2 mg/L dalam penelitian ini dianggap bernilai 2 mg/L supaya dapat diolah menggunakan perangkat lunak tersebut.

Penentuan asumsi hubungan parameter dengan kualitas air di Tabel 1 kolom (3) berdasarkan pada baku mutu air sungai yang tertuang dalam Lampiran VI Peraturan Pemerintah Republik Indonesia Nomor 22 Tahun 2021 Tentang Penyelenggaraan Perlindungan dan Pengelolaan Lingkungan Hidup. Data kualitas air sungai Kota Surakarta yang terdiri dari 13 sungai berisi informasi 12 parameter tersebut selanjutnya dilakukan pengklasteran untuk mendapatkan informasi lokasi-lokasi dengan tingkat kualitas air signifikan berbeda dibandingkan lokasi lainnya.

Dengan bantuan perangkat lunak *Orange*, algoritma *k-means* diterapkan dengan *silhouette scores*. Dapat dilihat pada Gambar 3 bahwa *silhouette scores* tertinggi pada kluster sebanyak dua. Maka rekomendasi terbaik adalah dua kluster dengan hasil seperti pada Tabel 2. Kluster 1 beranggotakan 12 sungai yaitu Sungai Pepe Atas Hulu, Sungai Pepe Bawah Hulu, Sungai Anyar Hilir, Sungai Kedung Jumbleng Hilir, Sungai Anyar Tengah, Sungai Pepe Bawah Hilir, Sungai Anyar Hulu, Sungai Gajah Putih Hulu, Sungai Brojo Hilir, Sungai Premulung Tengah, Sungai Premulung Hulu, Sungai Premulung Hilir. Sedangkan hanya Sungai Jenes Hilir yang termasuk dalam kluster 2.



Gambar 3. Tampilan *Silhouette Score* pada *Output Orange*

Tabel 2. Penempatan anggota kluster berdasarkan *silhouette*.

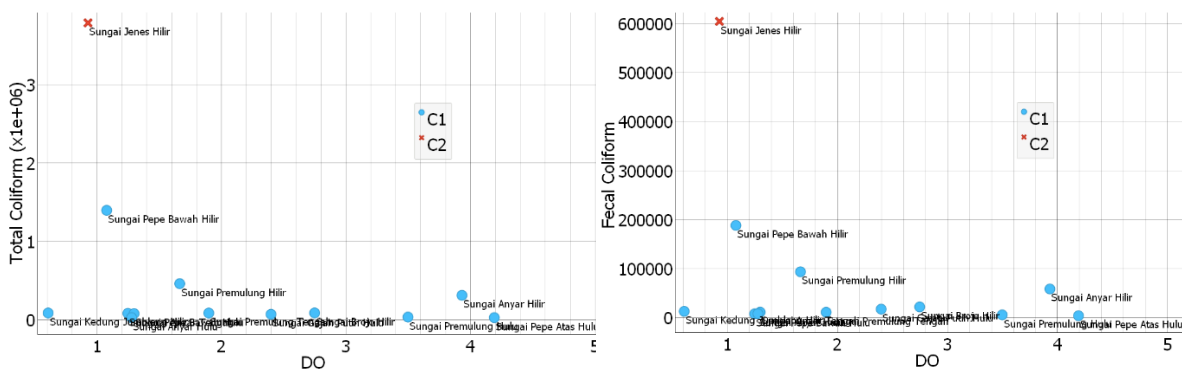
	Sungai	Kluster	<i>Silhouette</i>
1	Sungai Pepe Atas Hulu	1	0,740387
2	Sungai Pepe Bawah Hulu	1	0,741475
3	Sungai Anyar Hilir	1	0,734874
4	Sungai Jenes Hilir	2	0,5
5	Sungai Kedung Jumbleng Hilir	1	0,741481
6	Sungai Anyar Tengah	1	0,741497
7	Sungai Pepe Bawah Hilir	1	0,640127
8	Sungai Anyar Hulu	1	0,740398
9	Sungai Gajah Putih Hulu	1	0,741375
10	Sungai Brojo Hilir	1	0,74138
11	Sungai Premulung Tengah	1	0,741489
12	Sungai Premulung Hulu	1	0,740609
13	Sungai Premulung Hilir	1	0,728282

Selanjutnya dibantu perangkat lunak *SPSS* untuk mendapatkan informasi parameter mana yang signifikan membedakan kluster 1 dan kluster 2. Tabel 3 menunjukkan bahwa parameter yang signifikan membedakan antara kluster 1 dan kluster 2 adalah *Total Coliform* dan *Fecal Coliform* karena nilai Sig<0,05. Kemudian yang menarik untuk dicari lebih lanjut adalah, khususnya kluster 2 (Sungai Jenes Hilir), apakah kandungan *Total*

Coliform dan *Fecal Coliform* signifikan lebih rendah atau signifikan lebih tinggi daripada kluster 1. Untuk mengetahui hal tersebut, maka dibuat contoh scatter plot antara DO di sumbu-x terhadap *Total Coliform* dan *Fecal Coliform* di sumbu-y (Gambar 4). Nampak bahwa kandungan *Total Coliform* dan *Fecal Coliform* di C2 signifikan lebih tinggi daripada C1. Sungai Jenes Hilir mengandung 3.790.050 MPN/100ml *Total Coliform* dan 604.400 MPN/100ml *Fecal Coliform*. Sebagaimana telah dijelaskan dalam Tabel 1, semakin tinggi jumlah *Total Coliform* dan *Fecal Coliform* maka kualitas air semakin buruk, sehingga dapat disimpulkan bahwa Sungai Jenes Hilir mempunyai kualitas air yang lebih buruk dibandingkan sungai lainnya dalam hal kandungan pencemar *Total Coliform* dan *Fecal Coliform*.

Tabel 3. Nilai signifikansi masing-masing parameter

Parameter	TSS	TDS	pH	BOD	COD	DO	P	NO ₃ -N	NH ₃	NO ₂ -N	Total Coliform	Fecal Coliform
Sig.	0,71	0,80	0,54	0,78	0,94	0,35	0,27	0,84	0,94	0,54	0,00	0,00

Gambar 4. Scatter Plot antara DO terhadap *Total Coliform* dan *Fecal Coliform*

KESIMPULAN

Sungai Jenes Hilir mempunyai kualitas air lebih buruk daripada Sungai Pepe Atas Hulu, Sungai Pepe Bawah Hulu, Sungai Anyar Hilir, Sungai Kedung Jumbleng Hilir, Sungai Anyar Tengah, Sungai Pepe Bawah Hilir, Sungai Anyar Hulu, Sungai Gajah Putih Hulu, Sungai Brojo Hilir, Sungai Premulung Tengah, Sungai Premulung Hulu, Sungai Premulung Hilir berdasarkan pencemar *Total Coliform* dan *Fecal Coliform*. Informasi tersebut dapat memberikan masukan kepada Pemerintah Daerah Kota Surakarta tentang kerawanan penyakit diare masyarakat yang bertempat tinggal di sekitar Sungai Jenes Hilir. Selain itu Pemerintah Daerah Kota Surakarta dapat menentukan kebijakan yang tepat mengenai penataan ruang dan wilayah sekitar Sungai Jenes Hilir dalam upaya peningkatan kualitas air sungai. Penelitian selanjutnya disarankan penghitungan jarak terdekat setiap data ke *centroid* menggunakan metode selain *Euclidean Distance* yaitu *Minkowski Distance* dan *Manhattan Distance*.

DAFTAR PUSTAKA

- Azhar, S. C., Aris, A. Z., Yusoff, M. K., Ramli, M. F., & Juahir, H. (2015). Classification of River Water Quality Using Multivariate Analysis. *Procedia Environmental Sciences*, 30, 79–84. <https://doi.org/10.1016/j.proenv.2015.10.014>
- BPS. (2020). *Statistik Lingkungan Hidup Indonesia 2020*. Badan Pusat Statistik.
- BPS Kota Surakarta. (2021). *Kota Surakarta Dalam Angka 2021*. Badan Pusat Statistik Kota Surakarta.
- Dabgerwal, D. K., & Tripathi, S. K. (2016). Assessment of surface water quality using hierarchical cluster analysis. *International Journal of Environment*, 5(1), 32–44. <https://doi.org/10.3126/ije.v5i1.14563>
- Dimri, D., Daverey, A., Kumar, A., & Sharma, A. (2021). Monitoring water quality of River Ganga using multivariate techniques and WQI (Water Quality Index) in Western Himalayan region of Uttarakhand, India. *Environmental Nanotechnology, Monitoring and Management*, 15(July 2020), 100375. <https://doi.org/10.1016/j.enmm.2020.100375>
- DLH Kota Surakarta. (2019). *Dokumen Informasi Kinerja Pengelolaan Lingkungan Hidup Daerah Kota Surakarta Tahun 2019*. Dinas Lingkungan Hidup Kota Surakarta.

- Droste, R., & Gehr, R. (2019). *Theory and Practice of Water and Wastewater Treatment* (2nd ed.). John Wiley & Sons, Inc.
- Garg, V. (2018). Assessment of Water Quality using Principal Component Analysis and Cluster Analysis for River Markanda, India. *International Journal for Research in Applied Science and Engineering Technology*, 6(4), 3240–3246. <https://doi.org/10.22214/ijraset.2018.4539>
- Hadidjah, K. (2019). Analisa Tingkat Pencemaran Air pada Sumur Gali Rt 15, Desa Bangun Rejo, Kecamatan Tenggarong Seberang. *Buletin Loupe*, 15(2), 31–36.
- Irfannur, I., & Khairan, K. (2021). Analisis Parameter Fisika Kimia Kualitas Perairan di Sungai Krueg Mane Aceh Utara. *Arwana: Jurnal Ilmiah Program Studi Perairan*, 3(1), 16–23. <https://doi.org/10.51179/jipsbp.v3i1.450>
- Jatnika, H., Huda, M., Amelia, R. R., Manuhutu, M. A., Windarto, A. P., Sumantrie, P., & Waluyo, A. (2020). Analysis of Data Mining in the Group of Water Pollution Areas using the K-Means Method in Indonesia. *Journal of Physics: Conference Series*, 1783, 1–7.
- Javadi, S., Hashemy, S. ., Mohammadi, K., Howard, K. W. ., & Neshat, A. (2017). Classification of Aquifer Vulnerability Using K-means Cluster Analysis. *Journal of Hydrology*, 549, 27–37.
- Lampiran VI Peraturan Pemerintah Republik Indonesia Nomor 22 Tahun 2021 Tentang Penyelenggaraan Perlindungan dan Pengelolaan Lingkungan Hidup, (2021).
- Landhiani, J. N. (2016). *Analisis Kualitas Air Sungai Pepe dan Sungai Anyar yang Melintas di Kota Surakarta*. Universitas Sebelas Maret.
- Pitasari, I. S. (2016). *Analisis Kualitas Air dan Kandungan Logam Berat Kromium (Cr) Pada Air, Sedimen, dan Daging Ikan Nila (Oreochromis niloticus Linn.) di Sungai Premulung Kota Surakarta*. Universitas Sebelas Maret.
- Rahmadina, R. P., Arien, Y., & Sutikno. (2020). *Analisis Cluster Virus Corona (COVID-19) di Indonesia pada 2 Maret 2020 – 12 April 2020 dengan Metode K-Means Clustering*.
- Rezagama, A., Sarminingsih, A., Rahmadani, A. Y., & Aini, A. N. (2019). Pemodelan Peningkatan Kualitas Air Sungai melalui Variasi Debit Suplesi. *TEKNIK*, 40(2), 106–114.
- Saputra, D. M., Saputra, D., & Oswari, L. D. (2020). Effect of Distance Metrics in Determining K-Value in K-Means Clustering Using Elbow and Silhouette Method. *Sriwijaya International Conference on Information Technology and Its Applications (SICONIAN 2019)*, 341–346.
- Sharif, S. M., Kusin, F. M., Asha'ari, Z. H., & Aris, A. Z. (2015). Characterization of Water Quality Conditions in the Klang River Basin, Malaysia Using Self Organizing Map and K-means Algorithm. *Procedia Environmental Sciences*, 30, 73–78.
- Singh, A., Yadav, A., & Rana, A. (2013). K-means with Three different Distance Metrics. *International Journal of Computer Applications*, 67(10), 13–17.
- Sukmawati, D. T. (2019). *Penentuan Kategori Kualitas Air Sungai dengan Menggunakan Metode K-Means*. Universitas Islam Majapahit.
- Suryono, H., & Triscowati, D. W. (2020). Pengelompokan Provinsi di Indonesia Berdasarkan Resiko Covid-19 dan Ketahanan Pangan. *Seminar Nasional Official Statistics 2020: Statistics in New Normal: A Challenge of Big Data and Official Statistics*, 116–123.
- Trilaksono, G., Sudarno, & Handayani, D. S. (2014). Studi Penentuan Daya Tampung Beban Pencemaran Air Sungai Melalui Pendekatan Software QUAL2E dan Metode Neraca Massa (Studi Kasus: Sungai Garang, Jawa Tengah). *Jurnal Teknik Lingkungan*, 3(1), 1–12.
- Yuliasuti, E. (2011). *Kajian Kualitas Air Sungai Ngringo Karanganyar dalam Upaya Pengendalian Pencemaran Air*. Universitas Diponegoro.
- Zhao, Z. W., & Cui, F. Y. (2009). Multivariate statistical analysis for the surface water quality of the Luan River, China. *Journal of Zhejiang University: Science A*, 10(1), 142–148. <https://doi.org/10.1631/jzus.A0850094>
- Zou, H., Zou, Z., & Wang, X. (2015). An Enhanced K-Means Algorithm for Water Quality Analysis of The Haihe River in China. *International Journal of Environmental Research and Public Health*, 12(11), 14400–14413. <https://doi.org/10.3390/ijerph121114400>