

A Multivariate Technique to Develop Hybrid Water Quality Index of the Bengawan Solo River, Indonesia

Evellin Dewi Lusiana^{1,2*}, Mohammad Mahmudi^{1,2}, Sarah Mega Hutahaean¹, Arief Darmawan¹, Nanik Retno Buwono^{1,2}, Sulastri Arsad^{1,2}, Muhammad Musa^{1,2}

¹ Department of Aquatic Resources Management, Faculty of Fisheries and Marine Science, Universitas Brawijaya, Veteran Street, Malang 65145, Indonesia

² AquaRES Research Group, Faculty of Fisheries and Marine Science, Universitas Brawijaya, Veteran street, Malang 65145, Indonesia

* Corresponding author's e-mail: evellinlusiana@ub.ac.id

ABSTRACT

Surface water resource, such as river, is constantly contaminated by domestic and industrial pollutants. In order to properly manage the water resource, a composite index for water quality assessment, such as water quality index (WQI), has been designed to monitor and evaluate the properties of surface water. However, this index is quite subjective in terms of determination of relative weights. A principal component analysis (PCA) can be used to reduce the dimension and subjectivity of water quality variables. The purpose of this study was to implement the use of hybrid PCA and WQI methods to assess and monitor the water quality of the Bengawan Solo River, which is located in Java Island, Indonesia. The result suggested that COD, BOD, TSS, TDS, nitrate, nitrite, and ammonia were the main factors that determine water quality of the Bengawan Solo River. Furthermore, it was revealed that most samples from the river showed water quality status as slightly polluted. In addition to this, the seasonal variation of the PCWI values indicated a significant increase of water pollution in the Bengawan Solo River per year.

Keywords: principal component analysis, WQI, variable selection, water quality assessment.

INTRODUCTION

Water resources are heavily utilised natural resources. They are constantly contaminated and in short supply all over the world [Boretti and Rosa, 2019]. As a result, it is critical to pay close attention to improve as well as maintain their availability and quality. Hence, reliable approaches to evaluate water resources are required for long-term development and human health safety [Dunca, 2018]. Compared to groundwater, the surface water resources are most susceptible to domestic and industrial pollution [Walker et al., 2019]. Their continuous deterioration causes a serious harm to the environments that thrive therein. As a result, a thorough and attentive strategy to monitor and measure surface water is required to maintain its safety, since it affects public health [Massoud, 2012].

In order to properly manage the water resources, the information with regard to water quality is critical [Tripathi and Singal, 2019]. The standards for water quality have been created on an international and regional basis to successfully monitor and evaluate the water resources. Nevertheless, they provide judgement based on specific factors and do not present a full view of the issue [Kannel et al., 2007; Rosemond et al., 2008]. As a result, a composite index for water quality assessment, such as water quality index (WQI), has been designed to monitor and evaluate the properties of surface water [Banda and Kumarasamy, 2020].

WQI is a method that summarizes numerous water quality metrics into a single number. The index enables the interpretation of water quality in a single number [Horton, 1965; Brown et al., 1970]. A benefit of WQI classification is its ability to measure the heterogeneity of the system. It

is based on simple additive weighting (SAW) approach that combines independent criteria the relative relevance of which is reflected by subjective weights [Praus, 2019]. Water quality varies greatly on a geographical and temporal basis, and hence, frequent monitoring produces a complicated and vast dataset that includes a large number of factors, many of which are difficult to interpret. Different multivariate statistical approaches such as Principal Component Analysis (PCA) allow better understanding the results and reduce subjectivity in the process [Kazi et al., 2009; Esdras et al., 2017].

PCA is a multivariate statistical method that determines the interrelation between initial variables and converts them into independent principal factors [Jolliffe and Cadima, 2016]. It removed the relationship between the water quality variables and the effort of variable selection, thus considerably facilitating the analysis. When the water quality is represented as an index, the correlation of factors might lead to inappropriate classifications. Moreover, PCA has already been used to analyse the water quality data and minimise the number of variables in the system without giving up any information [Mahapatra et al., 2012].

The purpose of this study was to implement the use of hybrid PCA and WQI methods to assess and monitor the water quality of the Bengawan Solo River. This river is one of the longest rivers on the island of Java (600 km), Indonesia, which crosses two provinces, namely Central Java and East Java. The upstream and downstream

sections of the river do not physically satisfy the standards for clean water, as evidenced by an unpleasant odour, yellow-black river water, and a large amount of rubbish on the riverbanks [As-tuti, 2015]. Another issue is that the town uses the water from the Bengawan Solo River for fishing, agriculture, industry, home activities, and other purposes [Dani et al., 2015]. Therefore, it is essential to assess the overall water quality status of the river by using hybrid method which has never been reported in previous studies.

MATERIALS AND METHODS

Study area

The research was conducted by collecting secondary monthly data from the Bengawan Solo River Basin Center (BBWS) between 2016 and 2020. The study involved 12 water quality parameters such as temperature (°C), dissolved oxygen (DO, mg/L), total dissolved solid (TDS, mg/L), total suspended solid (TSS, mg/L), nitrate (mg/L), nitrite (mg/L), ammonia (mg/L), phosphate (mg/L), chemical oxygen demand (COD, mg/L), biological oxygen demand (BOD, mg/L), and total coliform (MPN/100 mL). There were 7 sampling stations that are described in Figure 1 and Table 1.

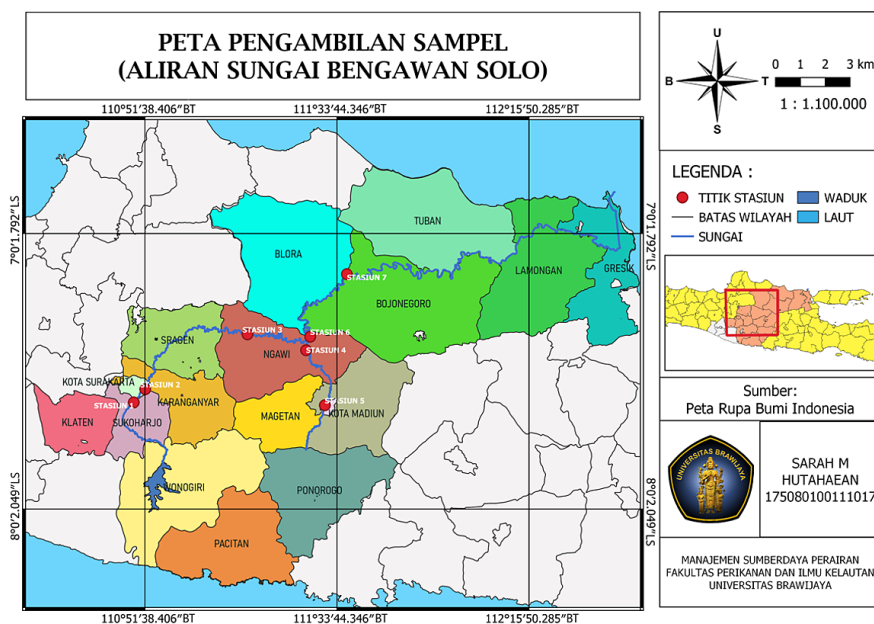


Figure 1. Map of sampling station for Bengawan Solo river monitoring

Table 1. The location of station used for collection of water samples

Station number	Location
1	Bacem Bridge Area of Surakarta City
2	Jurug Bridge Area, Sukoharjo Regency
3	Kajangan Area, Ngawi Regency
4	Ketonggo Bridge, Ngawi Regency
5	A. Yani Area, Madiun City
6	Napel Area, Ngawi Regency
7	Cepu Bridge, Bojonegoro Regency

Data analysis

Water Quality Index (WQI)

WQI measures water quality in terms of an index number that represents the overall water quality for any intended purpose. It is described as a rating that reflects the aggregate effect of several water quality criteria that were taken into account while calculating the WQI [Ewaid and Abed, 2017]. The formula of WQI is defined as follows

$$Q_i = \frac{(M_i - I_i)}{(S_i - I_i)} \times 100 \quad (1)$$

$$WQI = \frac{\sum_{i=1}^p W_i Q_i}{\sum_{i=1}^p W_i} \quad (2)$$

where: Q_i = rating for i -th parameter;
 W_i = unit weight for i -th parameter;
 M_i = measured i -th parameter;
 I_i = ideal value for i -th parameter; and
 S_i = standard value for i -th parameter. The weight values in this study were taken from various studies [Alobaidy et al., 2010; Ahmed et al., 2019].

Principal Component Analysis (PCA)

The benefit of PCA is that it decreases the number of variables to a smaller number of factors that can then be used to sort variables and clusters of observations with similar features based on these factors. It also eliminates explanatory factors that are not relevant. The sample size of data must be large enough for correlations to converge into mutually exclusive variables when doing the PCA [Hutcheson and Sofroniou, 1999]. Therefore, a sampling adequacy test that uses the

KMO test is required. Additionally, by using the Bartlett test, the homogeneity of datasets need to be tested [Mohd Matore et al., 2019]. Every principal component for a complete set of data is a linear function of the parameter variables and may be expressed as follows:

$$PC_i = e_1 X_1 + e_2 X_2 + \dots + e_n X_p \quad (3)$$

where: e_i = i -th loading factor;
 X_i = i -th explanatory factor;
 $i = 1, 2, \dots, p$; and
 p = number of explanatory variables.

The general procedure of PCA involves construction of a correlation or covariance matrix of the dataset, dimension reduction that is based on the explained variance of principal components (>80%), and factor rotation for clearer interpretation of individual variables [Rencher and Christensen, 2012].

Principal Component-Water Quality Index (PCWI)

A hybrid method that is obtained by integrating PCA and WQI is called PCWI. The main idea of PCWI involves replacing the weight unit of WQI with eigenvalue of principal component that is resulted from PCA. Therefore, the determination of weight unit is data-driven or more objective than the classic WQI. The formula of PCWI is expressed as follows [Praus, 2019]:

$$PCWI = \frac{\sum_{i=1}^n \lambda_i PC_i}{\sum_{i=1}^n \lambda_i} \quad (4)$$

where: λ_i = eigenvalue of i -th component and
 PC_i = the i -th principal component.

Data transformation

In this study, standard normal (Z-score) transformation was used for each variable, since the research variables have different units and magnitudes. Moreover, it is also utilised to compare the result between WQI and PCWI. The formula of Z-score transformation is denoted as follows [Lusiana and Mahmudi, 2020]:

$$Z_{ij} = \frac{X_{ij} - \bar{X}_i}{s_i} \quad (1)$$

where: Z_{ij} = Z-score of i -th parameter and j -th observation;

X_{ij} = value of i -th parameter and j -th observation;

\bar{X}_i = mean of the i -th parameter;

s_i = standard deviation of the i -th parameter; $i = 1, 2, \dots, p$;

$j = 1, 2, \dots, n$;

p = number of parameters/explanatory variable; and

n = number of observations.

Applicability of PCWI

The application of PCWI to assess the water quality of the Bengawan Solo river was carried out by performing the classification that was proposed by [Praus, 2019] and box-whisker plot, along with one-way ANOVA and Tukey test [Midway et al., 2020; Musa et al., 2020].

RESULTS AND DISCUSSIONS

Descriptive summary of variables

The following Table 2 shows the descriptive summary of the water quality parameter that was observed from the Bengawan Solo River. The standard value that was used in this study was based on the regulation of Indonesia Ministry of Environment, 2001 [Ministry of Environment, 2001]. The result indicated that the minimum value of each parameter met the standard value, except for DO. Meanwhile, both maximum and mean values of nitrite, ammonia, BOD, and total coliform exceeded the standard value.

Correlation between variables

Table 3 revealed that BOD, COD, and DO have significant correlation coefficient (bold). These parameters were basically the oxygen uptake measures that further indicated the intersect details of the water quality characteristics and relevance of PCA [Khelif and Boudoukha, 2018]. Moreover, nitrate, ammonia, total phosphate, TSS, and TDS also exhibited significant correlations that represented organic matter characteristics of the waters.

Results of principal component analysis on water quality parameters of the Bengawan Solo River

The Kaiser-Meyer-Olkin (KMO) and Bartlett sphericity tests were used to assess the suitability of PCA. These tests were used to ensure that the sample was adequate [Rencher and Christensen, 2012] and that each variable was independent [Jolliffe and Cadima, 2016]. KMO test statistics was equal to 0.57 (> 0.50) and Bartlett's p -value was 2.2×10^{-16} (p -value < 0.05) that suggests that the data is appropriate for PCA.

Figure 2 shows the percentages of explained variance of each main component. The scree plot helped in selecting major components and comprehending the underlying data structure. It was suggested that the cumulative proportion of variance that represents the standard rule for dimension reduction in PCA may explain at least 80% of the variation [Rencher and Christensen, 2012]. As a result, the first eight main components were kept to account for 84.51% of the dataset's variability.

Table 2. Summary of water quality parameters

Variable	Min	Max	Mean	Std. dev	Standard value
Temperature	23.12	32	27.291	1.604	<i>deviation ± 3</i>
pH	0	8.89	7.374	0.787	6–9
TDS	35	840	225.663	114.555	< 1000
TSS	1	1130	83.11	146.286	< 100
DO	0	9.74	5.833	1.839	> 3
Nitrate	0.001	13.44	2.725	2.472	< 20
Nitrite	0	3.8	0.254	0.421	< 0.06
Ammonia	0	7.026	0.273	0.568	< 0.50
Phosphate	0	2.66	0.269	0.346	< 1
COD	4.3	441.7	21.319	28.439	< 40
BOD	0.3	103	6.789	6.37	< 6
Total coliform	0	2400000	553496	2630259	< 10000

Table 3. Correlation matrix of water quality variables

Variable	Temperature	pH	TDS	TSS	DO	Nitrate	Nitrite	Ammonia	Phosphate	COD	BOD	Total coliform
Temperature	1.000	-0.126	0.210	-0.353	0.076	0.022	-0.086	-0.209	0.042	-0.028	-0.127	0.033
pH		1.000	-0.001	-0.008	0.119	-0.231	0.152	0.130	-0.168	-0.042	0.106	-0.101
TDS			1.000	-0.225	-0.024	0.106	-0.006	0.202	0.096	0.315	0.134	0.104
TSS				1.000	-0.060	0.108	-0.059	-0.033	-0.011	0.013	-0.021	-0.058
DO					1.000	-0.128	-0.093	-0.127	-0.051	-0.281	-0.283	0.054
Nitrate						1.000	0.067	0.070	-0.049	0.233	0.017	0.072
Nitrite							1.000	0.089	-0.081	0.037	0.115	0.068
Ammonia								1.000	0.078	0.313	0.164	-0.062
Phosphate									1.000	0.169	0.096	-0.016
COD										1.000	0.814	0.078
BOD											1.000	-0.019
Total coliform												1.000

PCA often includes principal component interpretation that is required to comprehend the data structure. Table 4 summarises the rotated component loadings that can be used to explain relationships between raw variables. COD and BOD were major sources of saturation in the first principal component (PC1). All of these variables describe oxygen uptake in water. Temperature, TDS, and TSS had an effect on the second main component (PC2). Meanwhile, pH, nitrate, nitrite, and ammonia had the greatest impact on PC3, PC4, and PC5. PC6 and PC7 were also affected by phosphate and total coliform. Finally, DO (PC8) played a role in the less important PCs.

The BOD and COD levels in rivers have traditionally been used to measure the extent of organic contamination that is caused by an excess of organic materials [Tangahu et al., 2019]. COD

measures the quantity of oxygen that is absorbed by organic pollutants, whereas BOD measures the amount of oxygen that is required by aerobic biological organisms to decompose organic waste [Khotimah et al., 2021]. The two measures are utilised as primary criterion to assess aquatic ecosystem resources because they indicate the quantity of oxygen-depleting organic pollutants in the water body. These parameters have been widely recognised as major indicators for river pollution [Lee and Nikraz, 2015].

Erosion of surrounding surface soil and stream banks, abrasion of streambed, and accumulation of dissolved organic matter or chemical deposition of inorganic particles within water column are all processes that create TSS in streams [Hudson-Edwards, 2003]. Meanwhile, organic TDS in streams is produced by organic compounds that

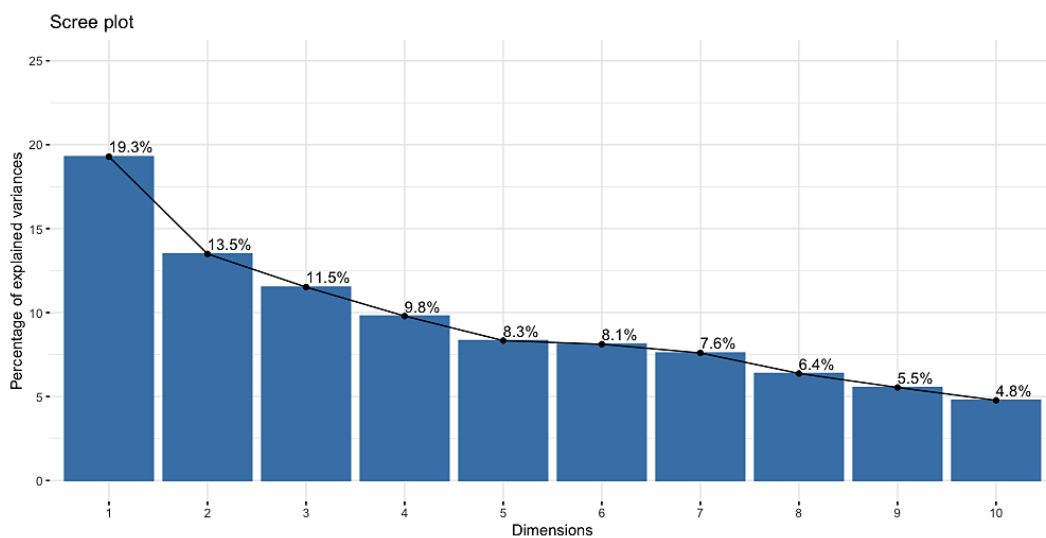


Figure 2. Scree plot of PCA

Table 4. Rotated loading factor of PCA

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Temperature	0.0858	-0.5992	0.0814	-0.0149	0.2717	-0.1837	-0.1078	0.1706
pH	0.0058	0.3011	0.6179	0.0140	-0.0326	0.0443	-0.1973	0.1451
TDS	-0.2699	-0.4121	0.2028	0.0483	-0.3659	-0.0906	-0.1435	-0.0133
TSS	0.0122	0.4778	-0.4136	-0.0290	-0.1680	0.1453	-0.1920	0.2245
DO	0.3147	-0.1054	0.2327	-0.0233	-0.3923	0.3121	-0.2547	0.5484
Nitrate	-0.1830	-0.0934	-0.4334	0.4542	-0.2087	-0.3354	-0.1754	0.3460
Nitrite	-0.1135	0.1559	0.2561	0.5327	0.1979	-0.0705	0.5728	0.3851
Ammonia	-0.3197	0.1395	0.2085	-0.0337	-0.5897	-0.2812	0.1865	-0.2782
Phosphate	-0.1510	-0.1733	-0.1732	-0.5245	-0.1742	0.2075	0.5794	0.3494
COD	-0.5996	-0.0486	-0.0327	-0.0506	0.0779	0.1517	-0.2201	0.1176
BOD	-0.5372	0.0961	0.1043	-0.1131	0.3450	0.2263	-0.1967	0.1188
Coliform	-0.0337	-0.2167	-0.0801	0.4629	-0.1528	0.7237	0.1036	-0.3210

are formed during biological growth (e.g., plant roots and microorganisms) and breakdown of biological materials within the river or stream bank. TSS raises turbidity of a water body that reduces penetration of light and, as a result, inhibits the photosynthetic rate of aquatic plants, possibly leading to oxygen deprivation [Bilotta and Brazier, 2008]. Furthermore, TSS can cause fish deaths by blocking their gills [Bilotta and Brazier, 2008]. In sluggish stream sections, suspended silt may settle to the streambed, thereby suffocating fish eggs or any other benthos. TDS is a total measurement of the quantity of soluble components in a river or lake. The precise ions and levels that contribute to TDS may have ecotoxicological consequences [Weber-Scannell and Duffy, 2007]. In addition to this, TDS may pollute groundwater through aquifer recharge and violate drinking water requirements. Increases in TSS or TDS in the stream may indicate a human effect that may subsequently be examined by assessing further water quality measures and component quantities [Butler and Ford, 2018].

On the other hand, increased nitrate levels have a negative impact on the aquatic ecosystems since they cause eutrophication, toxic algal blooms, and hypoxia [Mahmudi et al., 2020]. Furthermore, excessive nitrate levels in drinking water can raise the risk of illnesses and health impacts such as methaemoglobinaemia, diabetes, spontaneous abortion, thyroid disease, and stomach cancer [Ward et al., 2018]. As a result, it is critical to determine the nitrate sources to develop effective nitrogen contamination mitigation techniques and ensure long-term viability of water resources [Bastani and Harter, 2019]. For instance, the river water quality in China is in serious jeopardy and high nitrate content in the

water is a major contributor to poor water quality [Xue et al., 2016]. Even the presence of nitrite, a different type of nitrogen, has the potential to cause cancer. As a result, nitrogen pollution is a serious environmental issue that should be taken seriously [Xu et al., 2014].

Comparison of PCWI and WQI

The PCWI results were validated by comparing them to the conventional approach of WQI. WQI relative weights were established based on their relevance to an aquatic system, which implies that they are subjective. The values were adopted from some previous studies that dealt with river water quality evaluation [Alobaidy et al., 2010; Ahmed et al., 2019]. A correlation analysis between the Z-scores of PCWI and WQI was performed. The result is shown in Figure 3.

As shown in Figure 3, WQI and PCWI have strong ($R > 0.90$) and significant ($p < 0.001$) coefficient correlation. This implies a good consensus of both indices in water quality assessment. It is worth noting that, unlike WQI, PCWI operates with objective weights that are calculated for specific water components [Praus, 2019; Tripathi and Singal, 2019].

Application of PCWI for water quality assessment

The PCWI values of the Bengawan Solo river were evaluated by classifying them in analogy to Shewhart control chart [MacGregor and Kourti, 1995; Praus, 2019]. The classification result for this study is shown in Table 5. It is visible that more than half of the PCWI values that indicate the water quality status of the Bengawan Solo

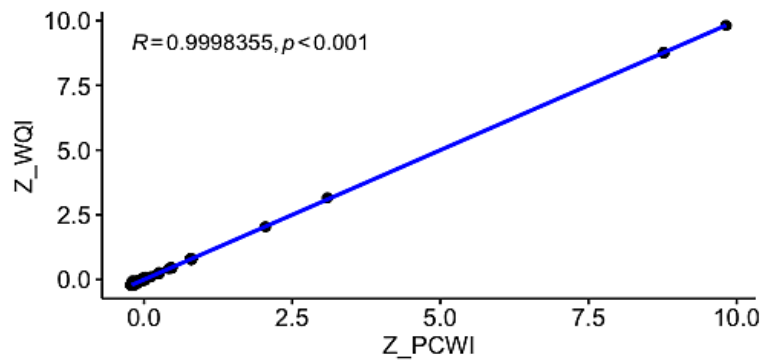


Figure 3. Correlation plot between WQI and PCWI

River is classified as slightly polluted. However, 30 samples suggested that the water quality of the river was fairly polluted, 1 sample was heavily polluted, and 5 samples were unsuitable.

Furthermore, PCWI was also employed to analyse the seasonal variation of water quality of the Bengawan Solo River over various years, as presented in Figure 4. It can be seen that the highest index was found in 2020. Moreover, the result of ANOVA indicated that there was significant difference of PCWIs from 2016 to 2020 ($F = 11.9, p < 0.0001$). Furthermore, Tukey test revealed that PCWI in 2020 was statistically significant compared to the previous years, which was denoted by its letter notation. This result suggested that the level of water pollution in the Bengawan Solo River has significantly increased from year-to-year.

The Bengawan Solo River is the largest river on Java Island that carries water from a drainage area (DPS) that spans $\pm 16,100 \text{ km}^2$. The water from the Bengawan Solo river is used for agriculture, pumping, and meeting the water demands of households, which are now met by the local water company supply, community wells, or directly from rivers [Sri Wahyu Kusumastuti et al., 2021]. The volume of wastewater that is generated in urban areas has grown significantly due to fast population expansion, economic activity, industrial output, and changes in water use behaviour [Setyaningrum and Agustina, 2020]. As discussed before, the parameters that become the most dominant characteristics in the Bengawan Solo River are COD, BOD, TSS, TDS, nitrate, nitrite, and ammonia. Anthropogenic activities around the river stream contributed to organic pollutant load that was discharged to the Bengawan Solo River. In addition to this, the high level of ammonia comes from urine and faeces due to traditional dense settlement that is adjacent to the river and factory

waste that is carried from upstream to downstream of the Bengawan Solo River. Furthermore, the concentration of BOD and COD in the Bengawan Solo River shows domestic waste as the dominant pollutant, which implies it is difficult to degrade. Ever since the industries were established in the Sukoharjo-Karanganyar region, the water quality of the Bengawan Solo river has been known to be deteriorate [Sri Wahyu Kusumastuti et al., 2021].

CONCLUSIONS

Freshwater pollution is a major concern worldwide, as it brings negative impact on water sustainability as well as public health. Compared to groundwater, surface water resources, such as rivers, are most susceptible to domestic and industrial pollution. Therefore, water resources monitoring and assessment is crucial for water management purpose. A novel hybrid technique that combines PCA and WQI (or PCWI) suggested COD, BOD, TSS, TDS, nitrate, nitrite, and ammonia as the main factors that determine the water quality of the Bengawan Solo River. Furthermore, it was also revealed that most samples of the river showed the water quality status as slightly polluted. The seasonal variation of PCWI values indicated significant increase of water pollution in the Bengawan Solo River per year.

Table 5. Classification of PCWI values of Bengawan Solo River

Classification	Class	Range	N observed
Good	I	-3 to -2	0
Slightly polluted	II	-2 to 0	329
Fairly polluted	III	0 to 2	30
Heavily polluted	IV	2 to 3	1
Unsuitable	V	>3	5

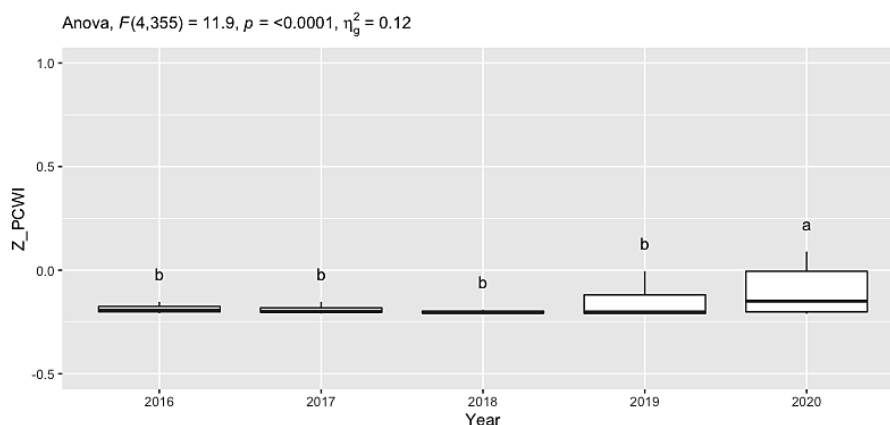


Figure 4. PCWI box-whisker plot of Bengawan Solo River

Acknowledgements

This work was supported by Hibah Pengantuan Kelompok Kajian Universitas Brawijaya 2021 with contract number 1204.9/UN10.C10/PN/2021.

REFERENCES

- Ahmed U., Mumtaz R., Anwar H., Shah A.A., Irfan R. 2019. Efficient Water Quality Prediction Using Supervised. *Water*, 11(1), 1–14.
- Alobaidy A.H.M.J., Abid H.S., Maulood B.K. 2010. Application of water quality index for assessment of Dokan Lake ecosystem, Kurdistan Region, Iraq. *Journal of Water Resource and Protection*, 2(9), 792–798.
- Astuti B.C. 2015. Kualitas air sumur Desa Bantaran Sungai Bengawan Solo berdasarkan aspek kemasayarakatan dan standar menteri kesehatan. *Jurnal Matematika, Sains, dan Teknologi*, 16(1), 18–25.
- Banda T.D., Kumarasamy M.V. 2020. Development of water quality indices (WQIs): A review. *Polish Journal of Environmental Studies*, 29(3), 2011–2021.
- Bastani M., Harter T. 2019. Source area management practices as remediation tool to address groundwater nitrate pollution in drinking supply wells. *Journal of Contaminant Hydrology*, 226, 103521.
- Bilotta G.S., Brazier R.E. 2008. Understanding the influence of suspended solids on water quality and aquatic biota. *Water Research*, 42(12), 2849–2861.
- Boretti A., Rosa L. 2019. Reassessing the projections of the World Water Development Report. *npj Clean Water*, 2(1), 15.
- Brown R.M., McClelland N.I., Deininger R.A., Tozzer R. 1970. Water quality index-do we dare? *Water Sewage Works*, 117(10), 339–343.
- Butler B.A., Ford R.G. 2018. Evaluating relationships between total dissolved solids (TDS) and total suspended solids (TSS) in a mining-influenced watershed. *Mine water and the environment*, 37(1), 18–30.
- Dani T., Suripin, Sudarno. 2015. Analisis daya tampung beban cemar di DAS Bengawan Solo segmen Kota Surakarta dan Kabupaten Karanganyar dengan model Qual2Kw. *Jurnal Ilmu Lingkungan*, 13(2), 92–102.
- Dunca A.M., 2018. Water Pollution and Water Quality Assessment of Major Transboundary Rivers from Banat (Romania). *Journal of Chemistry*, 2018, 9073763.
- Esdras J., Zandagba B., Adandedji F.M., Lokonon B.E., Chabi A., Dan O., Mama D. 2017. Application Use of Water Quality Index (WQI) and Multivariate Analysis for Nokoué Lake Water Quality Assessment. *American Journal of Environmental Science and Engineering*, 1(4), 117–127.
- Ewaid S.H., Abed S.A. 2017. Water quality index for Al-Gharraf River, southern Iraq. *Egyptian Journal of Aquatic Research*, 43(2), 117–122.
- Horton R.K. 1965. An index-number system for rating water quality. *Journal of Water Pollution Control Federation*, 37(3), 300–306.
- Hudson-Edwards K.A. 2003. Sources, mineralogy, chemistry and fate of heavy metal-bearing particles in mining-affected river systems. *Mineralogical Magazine*, 67(2), 205–217.
- Hutcheson G., Sofroniou N. 1999. *The multivariate social scientist: introductory statistics using generalized linear models*. SAGE Publications, London.
- Jolliffe I.T., Cadima J. 2016. Principal component analysis: a review and recent developments. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 374(2065), 20150202.
- Kannel P.R., Lee S., Lee Y.S., Kanel S.R., Khan S.P. 2007. Application of water quality indices and dissolved oxygen as indicators for river water classification and urban impact assessment. *Environmental Monitoring and Assessment*, 132(1), 93–110.

19. Kazi T.G., Arain M.B., Jamali M.K., Jalbani N., Afridi H.I., Sarfraz R.A., Baig J.A., Shah A.Q. 2009. Assessment of water quality of polluted lake using multivariate statistical techniques: A case study. *Ecotoxicology and Environmental Safety*, 72(2), 301–309.
20. Khelif S., Boudoukha A. 2018. Multivariate statistical characterization of groundwater quality in Fesdis, East of Algeria. *Journal of Water and Land Development*, 37(1), 65–74.
21. Khotimah C.K., Handayani D.S., Sudarno, Syaf-rudin, Hadiwidodo M., Wardhana I.W. 2021. The efficiency of biological oxygen demand removal in domestic wastewater treatment using multi soil layering. *IOP Conference Series: Earth and Environmental Science*, 802(1).
22. Lee A.H., Nikraz H. 2015. BOD: COD Ratio as an Indicator for River Pollution. *International Proceedings of Chemical, Biological and Environmental Engineering*, 88(1), 89–94.
23. Lusiana E.D., Mahmudi M. 2020. *Teori dan Praktik Analisis Data Univariat dengan PAST*. UB Press, Malang.
24. MacGregor J.F., Kourti T. 1995. Statistical process control of multivariate processes. *Control Engineering Practice*, 3(3), 403–414.
25. Mahapatra S.S., Sahu M., Patel R.K., Panda B.N. 2012. Prediction of Water Quality Using Principal Component Analysis. *Water Quality, Exposure and Health*, 4(2), 93–104.
26. Mahmudi M., Serihollo L.G., Herawati E.Y., Lusiana E.D., Buwono N.R. 2020. A count model approach on the occurrences of harmful algal blooms (HABs) in Ambon Bay. *Egyptian Journal of Aquatic Research*, 46(4), 347–353.
27. Massoud M.A. 2012. Assessment of water quality along a recreational section of the Damour River in Lebanon using the water quality index. *Environmental Monitoring and Assessment*, 184(7), 4151–4160.
28. Midway S., Robertson M., Flinn S., Kaller M. 2020. Comparing multiple comparisons: practical guidance for choosing the best multiple comparisons test. *PeerJ.*, 8, 1–26.
29. Ministry of Environment. 2001. *Peraturan Pemerintah Republik Indonesia Tentang Pengelolaan Kualitas Air Dan Pengendalian Pencemaran Air*. Indonesia.
30. Mohd Matore M.E.E., Khairani A.Z., Adnan R. 2019. Exploratory factor analysis (EFA) for adversity quotient (AQ) instrument among youth. *Journal of Critical Reviews*, 6(6), 234–242.
31. Musa M., Lusiana E.D., Buwono N.R., Arsad S., Mahmudi M. 2020. The effectiveness of silvofishery system in water treatment in intensive whiteleg shrimp (*Litopenaeus vannamei*) ponds, probolinggo district, East Java, Indonesia. *Biodiversitas*, 21(10), 4695–4701.
32. Praus P. 2019. Principal component weighted index for wastewater quality monitoring. *Water (Switzerland)*, 11(11).
33. Rencher A.C., Christensen W.F. 2012. *Methods of multivariate analysis*. Wiley, New Jersey.
34. Rosemond S. de, Duro D.C., Dubé M. 2008. Comparative analysis of regional water quality in Canada using the Water Quality Index. *Environmental Monitoring and Assessment*, 156(1), 223.
35. Setyaningrum D., Agustina L. 2020. Analisis Kualitas Air Di Daerah Aliran Sungai Bengawan Solo Wilayah Kabupaten Bojonegoro. *Samakia: Jurnal Ilmu Perikanan*, 11(1), 1–9.
36. Sri Wahyu Kusumastuti A., Moh Bisri B., Solichin C., Tri Budi P D. 2021. Water Quality Monitoring and Evaluation in the Bengawan Solo River Region. *IOP Conference Series: Earth and Environmental Science*, 641(1).
37. Tangahu B.V., Ningsih D.A., Kurniawan S.B., Imron M.F. 2019. Study of BOD and COD removal in batik wastewater using *Scirpus grossus* and *Iris pseudacorus* with intermittent exposure system. *Journal of Ecological Engineering*, 20(5), 130–134.
38. Tripathi M., Singal S.K. 2019. Use of Principal Component Analysis for parameter selection for development of a novel Water Quality Index: A case study of river Ganga India. *Ecological Indicators*, 96(May 2018), 430–436.
39. Walker D.B., Baumgartner D.J., Gerba C.P., Fitzsimmons K. 2019. Chapter 16 – Surface Water Pollution. In: Brusseau, M.L., Pepper, I.L., Gerba, C.P.B.T.-E. and P.S. (Third E. (Eds.)). *Academic Press*, 261–292.
40. Ward M.H., Jones R.R., Brender J.D., de Kok T.M., Weyer P.J., Nolan B.T., Villanueva C.M., van Breda S.G. 2018. Drinking water nitrate and human health: An updated review. *International Journal of Environmental Research And Public Health*, 15(7), 1557.
41. Weber-Scannell P.K., Duffy L.K. 2007. Effects of total dissolved solids on aquatic organisms: A review of literature and recommendation for salmonid species. *American Journal of Environmental Sciences*, 3(1), 1–6.
42. Xu Z., Zhang X., Xie J., Yuan G., Tang X., Sun X., Yu G. 2014. Total Nitrogen Concentrations in Surface Water of Typical Agro- and Forest Ecosystems in China, 2004–2009. *PLOS ONE*, 9(3), e92850.
43. Xue Y., Song J., Zhang Y., Kong F., Wen M., Zhang G. 2016. Nitrate Pollution and Preliminary Source Identification of Surface Water in a Semi-Arid River Basin, Using Isotopic and Hydrochemical Approaches. *Water*.