A Mamdani Fuzzy-C4.5 Hybrid Model for River Water Quality Assessment: A Case Study of the Beringin River

Fianti[1*], Risma Dwi Rahmawati[2], Sunarno[3], Suharto Linuwih[4], Siti Wahyuni[5], Fifin Dewi Ratnasari[6]
[1,2,3,4,5,6]*Program Studi Fisika, Universitas Negeri Semarang, Indonesia*
*corr-author: fianti@mail.unnes.ac.id

Abstract - Urban and industrial activities along the Beringin River in Semarang City cause water pollution in the river. Therefore, assessing water quality is necessary an effort to preserve the environment, safeguard human well-being, and provide early warnings of potential decline. This study used a Mamdani fuzzy logic for development water quality assessment. Input variables consisted of TSS, TDS, Ph, DO, BOD, COD, Nitrate, and Nitrite with water quality as the output. System development was begun by determination of fuzzy set domains for fuzzification and defuzzification. Fuzzy rules were formulated using IF-THEN relationships integrated with C4.5 Algorithm, an algorithm was used to select the most relevant attributes and pruning redundant branches to simplify complex decision-making system, in accordance to the health and environmental standards. The validation of the water quality assessment system was carried out by comparing the Pollution Index (PI) calculation data of 27 data points over a five-year period. The validation results showed accuracy system of 88.9%, with the water quality was categorized as moderately polluted. These findings indicated that the water quality assessment system have high accuracy and can be specifically applied in the Beringin River.

Keywords: River water quality, decision-making system, mamdani fuzzy logic, C4.5 algorithm.

I. INTRODUCTION

River water quality continues to decline due to industrial and urban activity along the riverbanks, while river flow dynamics have contributed to the spread of pollution [1]. In major cities, water pollution caused by domestic waste, urban development, and industrial expansion has significantly reduced river water quality. This condition is further exacerbated by inadequate river carrying capacity and limited wastewater treatment facilities [2-4]. Previous studies have stated that poor land use management accompanied by inadequate draining system has become a major factor contributing to river pollution [5]. These conditions highlight the urgent need to develop a water quality assessment system as an effort to mitigate river pollution [6]. The Beringin River, which extends for 15.5 km with a catchment area of 32.8 km², is located in an area with high industrial and domestic activities, making this river highly susceptible to pollution. [7]. Such conditions indicate the necessity of developing a system capable to assessing water quality. One approach is through the application of a fuzzy logic-based system. Fuzzy logic is selected due to its advantages in providing more adaptive modeling against sudden environmental changes. Furthermore, fuzzy logic has been recognized as capable of addressing complex problems by representing the concept of partial truth, which differs from Boolean logic that tends to be rigid by representing all situation in binary terms (1/0, true/false) [8-11]. Fuzzy logic encompasses various methods, with two popular ones being the Mamdanni and Sugeno methods. These methods exhibit significant differences, where the Mamdani method provides reasoning that more closely resembles human intuition, while Sugeno method excels in optimization and computational efficiency [12-15]. In its development, fuzzy logic can be integrated with Decision Trees (DT) as an effort to simplify datasets classification processes and support the construction of fuzzy rules [16-19]. One of the most popular algorithms for developing DT is the C4.5 algorithm, which enhances interpretability and efficiency in data mining [20-22].

One of the key efforts in assessing and controlling river water pollution involves measuring and analyzing water quality using the Pollution Index (PI), which evaluates water quality relative to regulatory standards [23, 24]. The PI method has the advantage of being able to utilize periodic water quality data through sequential measurement over a defined time span, which allows for a more efficient assessment of water quality. This approach helps reduce coasts as well as research duration [25]. Equation 1 presents the formula used in calculating PI, while the water quality is determined based on Table I.

$$PI_j = \sqrt{\frac{\left(C_i/L_{ij}\right)_M^{\ 2} + \left(C_i/L_{ij}\right)_R^{\ 2}}{2}} \qquad (1)$$

Description:

PI_j = Pollution Index for designated use j

C_i = Measured concentration of parameter i

L_{ij} = Permissible concentration of parameter i for designated j

$(C_i/L_{ij})_M$ = Maximum index of (C_i/L_{ij})

$(C_i/L_{ij})_R$ = Average index of (C_i/L_{ij})

Advances in the development of water quality assessment systems have been widely achieved by utilizing computational intelligence, including DT [26,27], support vector machines [28,29], and neural networks [30]. Previous studies optimized fuzzy logic by integrating genetic algorithms to improve system accuracy [31]. Other studies focused on integrating water quality indices with GIS-based spatial analysis to provide broader environmental insights [32]. Despite the significant contributions of earlier research, several challenges remain. First, the complexity of fuzzy rule bases often leads to inefficiency. Second, system testing has frequently relied on generalized datasets rather than being tailored to local river conditions. Third, outputs are not always calibrated against established standards such as the PI). These issues form the rationale for conducting this study.

This study integrates the C4.5 algorithm into Mamdani fuzzy logic to reduce rule base complexity and thereby improve efficiency. Furthermore, the framework is specifically applied to the Beringin River to generate localized insights into pollution dynamics. In addition, validation using PI ensures that the system's outputs remain consistent with applicable regulations. The success of previous studies in applying fuzzy logic to analyze water quality [33-35] provides concentrate that fuzzy logic systems are both feasible and necessary to be developed for assessing river water quality, particularly for evaluating the pollution level of the Beringin River. This effort supports environmental preservation, strengthens water resource management, and safeguards the health of communities living along the riverbanks.

II. METHOD

A. Datasets

This study utilizes sample data from the Beringin River, located in Ngaliyan District, Semarang City, for the years 2017, 2018, 2019, 2020, and 2021. The data obtained consisted of watr quality parameters, including concentrations of Total Suspended Solids (TSS), Total Dissolved Solids (TDS), pH, Dissolved Oxygen (DO), Biochemical Oxygen Demand (BOD), Chemical Oxygen Demand (COD), Nitrate, and Nitrite. Fig.1 was showed the sampling locations map along the Beringin River. Furthermore, eight water quality parameters serve as input variables with water quality classification as the system's output (Table II). The data were obtained from the Environmental Agency of Semarang City.

Data collection was conducted to obtain information related to the research object, which was subsequently processed according to fuzzy system requirements to serve as input and test data. Interviews were also carried out with laboratory technicians at the Environmental Agency of Semarang City to gather relevant insights. The fuzzy system was designed using the Mamdani fuzzy logic method implemented in MATLAB using a combination of the minimum function for the fuzzy rule implication method and the maximum function for the aggregation of rule outputs. The developed fuzzy system was then optimized to ensure sufficient decision-making accuracy by comparing assessment results generated through fuzzy logic with those obtained using the IP method. The stages of developing fuzzy system are illustrated in Fig. 2.

TABLE I
PI CLASSIFICATION

No	Score	Status
1	PI ≤ 1	Good
2	1< PI ≤ 5	Lightly polluted
3	5 < PI ≤ 10	Moderately polluted
4	PI >10	Heavily polluted

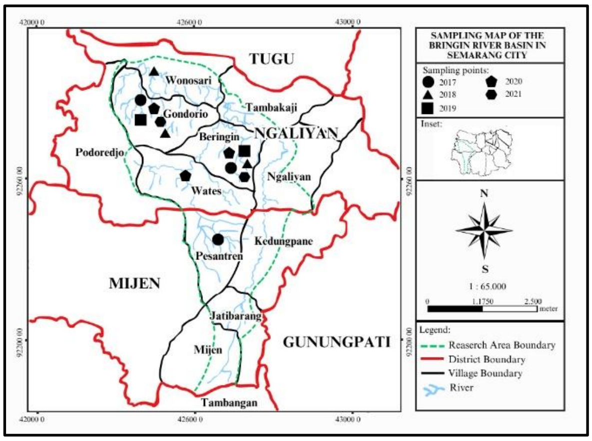

Fig. 1 Map of the beringin river watershed

TABLE II
INPUT AND OUTPUT VARIABLE

No	Variable	Unit	Status
1.	TSS	mg/L	Input
2.	TDS	mg/L	Input
3.	pH		Input
4.	DO	mg/L	Input
5.	BOD	mg/L	Input
6.	COD	mg/L	Input
7.	Nitrate	mg/L	Input
8.	Nitrite	mg/L	Input
9.	Water Quality		Output

The initial stage involved the data input procedure, in which a mechanism was constructed to accommodate crisp values derived from the water quality parameters of the Beringin River. The subsequent stage prepared the fuzzification procedure, aimed at converting the crisp input values into fuzzy sets and calculating their degrees of membership, thereby enabling the inputs to be classified into appropriate fuzzy sets. The domains of each fuzzy set for the respective water parameters were formulated based on several reference standards. Following this, fuzzy rules were developed and compiled into a rule base. At this stage, linguistic rules were constructed to describe the relationship between input parameters and the expected system output using IF–THEN implications. To optimize the rule base by reducing the number of fuzzy rules and minimizing the time required to determine appropriate rules, the C4.5 algorithm was integrated into the rule base using WEKA Software. The subsequent stage encompassed the inference procedure, in which reasoning was established using the predefined fuzzy inputs and fuzzy rules to generate corresponding fuzzy outputs. Finally, the defuzzification procedure was conducted to transform these fuzzy outputs into crisp values, based on the specified membership functions.

B. *Fuzzy rule generation and pruning using the C4.5 algorithm*

The process of rule generation was conducted using WEKA Software with the C4.5 algorithm [36]. The procedure began with data preprocessing, including cleaning and preparing dataset as well as defining the target attribute. Subsequently, information gain was calculated based on entropy to select the most appropriate attribute as the root node. The DT was then constructed recursively, with each node representing an attribute and each branch indicating a specific condition. Fuzzy rules were extracted from the tree paths extending from the root to the leaf nodes, thereby producing a set of if–then rules. To reduce tree complexity, prevent overfitting, and improve generalization, error-based pruning was applied, in which the error rate of a subtree was calculated and the subtree was removed if its error exceeded that of the parent node [37]. The final output was a simplified, efficient, and accurate rule base for evaluating water quality. Furthermore, the accuracy of the fuzzy system was determined by comparing the system's output with the calculated of PI.

C. *Water Quality Parameter Standards*

Water quality standards were employed as a reference for defining the domains of fuzzy sets during the fuzzification process (where water parameters served as fuzzy inputs) and the defuzzification process (where water quality served as the fuzzy output). Several regulatory standards were used to establish the reference values for river water quality parameters [38-43]. Conflicts among regulatory threshold values were resolved through fuzzy aggregation and expert validation. The boundaries of the membership functions were not solely adopted from regulatory standards but were also justified using empirical data analysis and domain-specific literature [44-46].

III. RESULT AND DISCUSSION

A. *Datasets*

Descriptive statistics were applied to analyze the datasets obtained. Organic parameters such as BOD, COD, and DO were exhibited considerable variation with right-skewed distributions, which established them as the primary indicators in formulating the rule base. TSS also showed high variability and was closely related to the physical condition of the water quality. Meanwhile, TDS and nitrite displayed highly abnormal distribution, making it more appropriate as supporting rather that determining indicator.

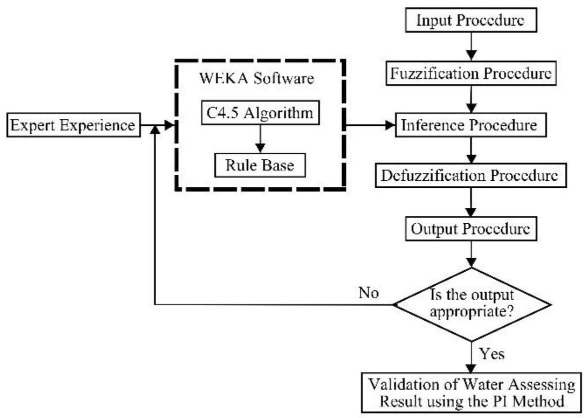

Fig. 2 Development stages of the fuzzy system

B. Fuzzy Sets for Input and Output Variables

By considering several water quality standards, the domains of fuzzy sets used in the fuzzification process were defined, as well as those used in the defuzzification process. The analysis of these regulatory standards resulted in the formulation of fuzzy set domains. During the fuzzification procedure, crisp input values were transformed into fuzzy values, and the degree of membership for each input was determined. This allowed the inputs to be classified into the appropriate fuzzy sets.

C. Rule Base

The rule base consisted of a collection of fuzzy rules incorporated into the fuzzy system. The fuzzy rules were constructed by enumerating all possible combinations of water quality parameters against the corresponding water quality classifications, based on the previously defined fuzzy sets and the regulatory standards for water quality according to its intended use. This process resulted in a total of 6.561 fuzzy rules, each formulated using the implication structure: IF fuzzy input set THEN fuzzy output set.

The C4.5 algorithm was applied to the rule base to generate a DT using the WEKA Software, which identified and eliminated redundant fuzzy rules. This optimization reduced memory usage and improve inference efficiency within the Mamdani fuzzy logic system. Consequently, the fuzzy rules were reduced to 541.

D. Inference and Defuzzification

The inference process in the system was carried out using the MIN-MAX implication function within the Mamdani fuzzy logic method. Data processing was conducted using MATLAB tools. The defuzzification process involved extracting fuzzy outputs obtained from rule evaluation and mapping them into an output membership function. To derive the final water quality output, the centroid method was applied. The centroid method also known as the center of area (COA) produced crisp values from fuzzy numbers by calculating the center point of the fuzzy region. According to [47] the COA method was considered superior because it did not disregard error values in its calculations, making control using COA more reliable, realistic, and smoother. Another study also reported that the development of a water quality assessment system for the Gajahwong watershed using Mamdani fuzzy logic with the COA method yielded favorable results, even when applied to a larger set of parameters (11 in total) [33].

The defuzzification results provided annual river water quality classifications based on parameter data, as illustrated in Fig. 4. In 2017, the average water quality derived from data at locations 1 to 7 was 3.58, categorized as lightly polluted. The data indicated that from 2017 to 2021, the water quality of Beringin River tended to increase, although it remained within the same fuzzy set interval, namely lightly polluted. This finding differed from previous studies, which reported that river water quality in Central, Java was categorized as moderately to heavily polluted [48]. The discrepancy was attributed to the limited fuzzy rule employed in the previous study, consisting of only 11 rules combined with the very broad scope of regional analysis.

E. Validation of Water Quality Assesing Results

Fuzzy system testing was performed to evaluated the performance of the developed Beringin River water quality assessment system. The fuzzy system's outputs were compared with those obtained using the PI. This evaluation provided the percentage accuracy of the developed system. Validation data were derived from the Water Quality Report issued by the Environmental Agency of Semarang City for the Beringin River from 2017 to 2021. A total of 27 data points were analyzed in the testing phase. The test results are presented in Table III, which shows a comparison between the classification outcomes based on fuzzy logic and those derived from the PI method.

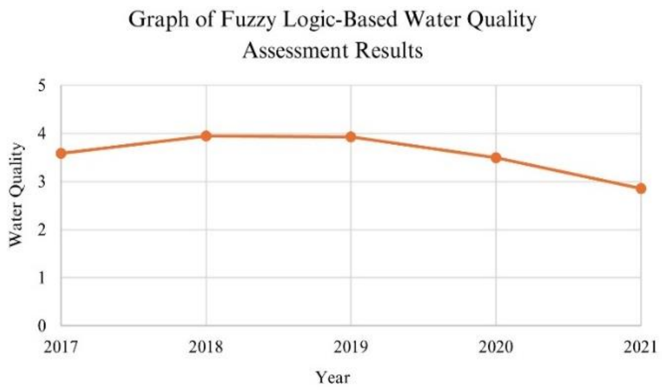

Fig. 3 Annual water quality status graph

TABLE III
COMPARATIVE RESULTS OF FUZZY LOGIC OUTPUT VERSUS POLLUTION INDEX METHOD

Year	Data Point	Input Parameter								Water Quality Determination		Result
		TSS	TDS	pH	DO	BOD	COD	Nitrate	Nitrite	Fuzzy Logic Output	PI output	
2017	1	21	280	6	6.19	15	63	0.311	0.03	Lightly polluted	Lightly polluted	Matching
2017	2	68	281	7	6.44	10	32	1.906	0.017	Lightly polluted	Lightly polluted	Matching
2017	3	133	650	6	5.1	35	63	10.143	0.250	Moderately polluted	Moderately polluted	Matching
2017	4	8	322	6	7.5	13	40	1.013	0.005	Lightly polluted	Lightly polluted	Matching
2017	5	12	243	7	7.2	10	45	1.662	2.006	Moderately polluted	Moderately polluted	Matching
2017	6	15	239	7	7.87	20	64	1.665	0.019	Lightly polluted	Lightly polluted	Matching
2017	7	10	449	7	6.9	26	64	1.662	0.011	Lightly polluted	Lightly polluted	Matching
2018	8	25	254	8	3.54	21	42	0.701	0.655	Lightly polluted	Lightly polluted	Matching
…	…	…	…	…	…	…	…	…	…	…	…	…
2021	27	88	264	8	2.79	10	11	2.207	0.143	Lightly polluted	Lightly polluted	Matching

Description:
- Parameter 1: TSS (mg/L)
- Parameter 2: TDS (mg/L)
- Parameter 3: DO (mg/L)
- Parameter 4: pH
- Parameter 5: BOD (mg/L)
- Parameter 6: COD (mg/L)
- Parameter 7: Nitrate (mg/L)
- Parameter 8: Nitrite (mg/L)

After obtaining the water quality classification results using both fuzzy system and the PI method, the percentage of agreement between the two approaches was calculated. A total of 27 data points were analyzed by using the confusion matrix (Table IV). The limitations of the dataset in this study posed challenges to statistical validation, as small amount of sample tended to be unstable and highly sensitive to minor variations in the data, thereby reducing the reliability of the fuzzy system. In addition, the risk of bias increased due to the lack of data diversity, which caused the fuzzy system to learn overly specific patterns ad diminished its ability to generalize to boarder populations. Other challenge that emerged was the reduced stability of fuzzy rules and DT, which became less robust and consequently undermined the reliability of the fuzzy system [49].

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+FP+TN+FN}} = \frac{21+3}{27} = 0.889; (88.9\%)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} = \frac{21}{21} = 1$$

TABLE IV
CONFUSION MATRIX

		Predict class			
		Good	Lightly Polluted	Moderately Polluted	Heavily Polluted
True class	**Good**	0	0	0	0
	Lightly Polluted	0	21	3	0
	Moderately Polluted	0	0	3	0
	Heavily Polluted	0	0	0	0

Recall/ sensitivity/ True Positive Rate (TPR) = $\frac{TP}{(TP+FN)} = \frac{21}{(21+3)} = 0.875$; (87.5%)

Specificity/ True Negative Rate (TNR) = $\frac{TN}{(TN+FP)} = \frac{3}{(3+0)} = 1$

Based on the confusion matrix, it was observed that the class distribution was not balanced. Evaluating of the good and heavily polluted water quality classes was not conducted due to data limitations, which rendered metrics such as recall and precision meaningless for these categories. Overall, the fuzzy system demonstrated good performance for the dominant class over the past five years but was not trained to detect the rarely occurring classes, leading to overfitting. The imbalance in distribution within the confusion matrix could not fully represent the overall capability of the fuzzy system. Therefore, this fuzzy system was considered less reliable.

Based on the validation results, the fuzzy system produced an accuracy of 88.9%. This result reflects a high level of success in classifying water quality and demonstrates the effectiveness of the Mamdani fuzzy logic system developed in previous research for assessing the water quality of the Winongo River using the Mamdani fuzzy logic approach, the system achieved an accuracy rate of 93.333% [24]. This result demonstrates the effectiveness of the fuzzy logic approach in classifying water quality conditions. Other studies have also demonstrated the effectiveness of fuzzy logic methods in producing precise outputs, serving as a reliable decision support tool for water quality assessment [50]. Fuzzy-logic-based methods was demonstrated to be appropriate to address uncertainty and subjectivity in drinking water quality assessment; they are an effective method for managing complicated, uncertain water systems and predicting drinking water quality [51]. Although the fuzzy system occasionally failed to recognize certain positive cases (recall < 1), the overall high accuracy (88.9%) indicates that the Mamdani fuzzy logic method is suitable for assessing river water quality, particularly in the Beringin River. This finding is further supported by the results of the linear regression analysis between PI and fuzzy system assessment (Fig. 5).

Based on Fig.5, the regression analysis yielded the equation Y=1.1156 X-0.4441 , with a coefficient of determination R^2=0.6471. The slope of 1.1156 indicates that for every one-unit increase in fuzzy system assessment, the performance index (PI) is expected to rise by approximately 1.12 units. The negative intercept value (-0.4441) suggests that when the fuzzy system assessment is zero, the PI approaches a value close to zero. The R^2 value of 0.6471 demonstrates that 64.7% of the variation in PI can be explained by fuzzy system assessment, while the remaining 35.3% is attributable to other factors not incorporated into the fuzzy system assessment. Although fuzzy system assessment plays a substantial role in predicting PI, additional variables outside the current framework also contribute meaningfully to performance outcomes. This result could be a suggestion to stake holder of environmental agency in improving river water quality in Semarang City.

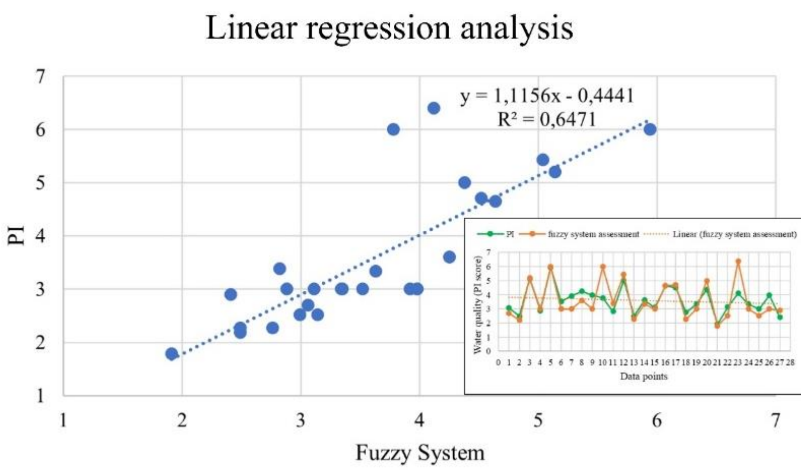

Fig. 4 Linear regression analysis between fuzzy system assessment and PI

IV. CONCLUSION

The Hybrid Mamdani Fuzzy-C4.5 Algorithm assessing the water quality level was successfully developed and applied to the Beringin River in Semarang City. The construction of fuzzy sets was based on established water quality standards, with eight water quality parameters (TSS, TDS, pH, DO, BOD, COD, Nitrate, and Nitrite) serving as input variables and one output variable representing overall water quality. The fuzzy system was supported by a rule base comprising 541 fuzzy rules generated through the integration of the C4.5 algorithm using the WEKA software. Based on Mamdani fuzzy inference, the water quality status of the Beringin River was classified as lightly polluted. Validation using five consecutive years of data through the PI method indicated a high accuracy rate of 88.9%. However, due to the imbalaced class distribution, the fuzzy system assessment was considered less reliable. The limitation of the dataset, consisting of only 27 data points, restricted the generalizability of the findings and therefore require further validation with more diverse dataset. The imbalance of the dataset also affected the fuzzy system, making it prone to overfitting toward the dominant class. In addition, the outputs had not been fully compared with established standards such as the Water Quality Index (WQI), which weakened the integration. Future research could be improved by employing a larger and more diverse dataset that represented the full range of parameters used in the system. In this way, the fuzzy system would have been more robust and reliable. The integrity of the fuzzy system could be enhancing by comparing against established standards such as the Water Quality Index (WQI). Furthermore, integration with Adaptive Neuro-Fuzzy Inference System (ANFIS) could be applied to enhance predictive accuracy, as well as integration with Geographic Information System (GIS)-based spatial analysis to provide more comprehensive environmental insights.

ACKNOWLEDGEMENT

The authors are thankful to the Faculty of Mathematics and Science of Universitas Negeri Semarang of research fund with DPA number: DPA 023.17.2.690645/2024.04/2024 due to an agreement of faculty research 2024 with number: 196.28.3/UN37/PPK.04/2024

REFERENCES

[1] N. F. Salsabila, M. Raharjo, and T. Joko, “Indeks Pencemaran Air Sungai dan Persebaran Penyakit yang Ditularkan Air (Waterborne Diseases): Suatu Kajian Sistematis,” *Environ. Occup. Health Saf. J.*, vol. 4, no. 1, pp. 24–34, Jul. 2023, doi: 10.24853/eohjs.4.1.24-34.

[2] J. J. Kim and J. H. Ryu, “Decision-making of LID-BMPs for Adaptive Water Management at the Boise River Watershed in a Changing Global Environment,” *Water (Switzerland)*, vol. 12, no. 9, pp. 1–17, Sep. 2020, doi: 10.3390/w12092436.

[3] Y. Wang, Y. Liu, J. Huang, T. Wu, and J. Huang, “Analysis and Prevention of Urban River Pollution,” in *J. Phys.: Conf. Ser.*, Institute of Physics Publishing, Jun. 2020. doi: 10.1088/1742-6596/1549/2/022056.

[4] B. Yohannes, S. W. Utamo, and H. Agustina, “Kajian Kualitas Air Sungai dan Upaya Pengendalian Pencemaran Air (Studi di Sungai Krukut, Jakarta Selatan),” *Indones. J. Environ. Educ. Manag.*, vol. 4, no. 2, pp. 2549–5798, Jul. 2019, doi: 10.21009/jgg.042.05.

[5] D. Roosmini, M. A. Septiono, N. E. Putri, H. M. Shabrina, I. R. S. Salami, and H. D. Ariesyady, “River Water Pollution Condition in Upper Part of Brantas River and Bengawan Solo River,” in *IOP Conf. Ser.: Earth Environ. Sci.*, Institute of Physics Publishing, Jan. 2018, pp. 1–7. doi: 10.1088/1755-1315/106/1/012059.

[6] Y. Carriazo-Regino, R. Baena-Navarro, F. Torres-Hoyos, J. Vergara-Villadiego, and S. Roa-Prada, “IoT-Based Drinking Water Quality Measurement: Systematic Literature Review,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 28, no. 1, pp. 405–418, Oct. 2022, doi: 10.11591/ijeecs.v28.i1.pp405-418.

[7] D. Anwariani, “Pengaruh Air Limbah Domestik Terhadap Kualitas Sungai,” *OSF*, pp. 1–6, 2019, doi: 10.31227/osf.io/8nxsj.

[8] A. Setiawan, B. Yanto, and K. Yasdomi, *Logika Fuzzy dengan Matlab (Contoh Kasus Penelitian Penyakit Bayi dengan Fuzzy Tsukamoto)*. Denpasar: Jayapangus Press Books, 2018.

[9] L. A. Zadeh, “Fuzzy Sets,” *Inf. Control*, vol. 8, no. 3, pp. 338–353, 1965, doi: 10.1016/S0019-9958(65)90241-X.

[10] M. A. R. Pohan, J. Utama, and B. Herdiana, “Novel Motion Planning Strategy with Fuzzy Logic for Improving Safety in Autonomous Vehicles in Response to Risky Road User Behaviors,” *ASEAN J. Sci. Eng. (AJSE)*, vol. 4, no. 3, pp. 471–484, Dec. 2024, doi: 10.17509/ajse.v4i3.75343.

[11] S. Syofian, A. Setiawan, M. Fathan, and R. Siregar, “CO Detection in High-Mileage Vehicle Cabins and Traffic Density Analysis Using Fuzzy Logic,” *JUITA J. Inform.*, vol. 13, pp. 229–234, Aug. 2025, doi: 10.30595/juita.v13i2.26854.

[12] J. Dombi and E. Tóth-Laufer, “Reducing the Computational Requirements in the Mamdani-type

Fuzzy Control," *Acta Polytech. Hung.*, vol. 17, no. 3, pp. 25–41, 2020, [Online]. Available: https://www.epa.oszk.hu/02400/02461/00098/pdf/

[13] M. N. Sinaga, N. A. Hasibuan, and A. M. H. Sihite, "Sistem Pakar Diagnosa Kifosis menerapkan Metode Fuzzy Mamdani," *KOMIK: Konf. Nas. Teknol. Inf. Komput.*, vol. 4, no. 1, Oct. 2020, doi: 10.30865/komik.v4i1.2716.

[14] F. A. Sihombing, "Kajian Fuzzy Metode Mamdani dan Fuzzy Metode Sugeno serta Implementasinya," *Innov. J. Soc. Sci. Res.*, vol. 4, no. 4, pp. 4940–4955, Jul. 2024, doi: 10.31004/innovative.v4i4.

[15] M. A. Yulianto, "Implementasi FIS Sugeno pada Algoritma C4.5 Berbasis Particle Swarm Optimization (PSO) untuk Prediksi Prestasi Siswa," *J. Artif. Intell. Innov. Appl.*, vol. 1, no. 1, pp. 12–22, Feb. 2020, [Online]. Available: https://openjournal.unpam.ac.id/index.php/JOAIIA/article/view/4272

[16] S. Lotfi, M. Ghasemzadeh, M. Mohsenzadeh, and M. Mirzarezaee, "Scalable Decision Tree Based on Fuzzy Partitioning and an Incremental Approach," *Int. J. Electr. Comput. Eng.*, vol. 12, no. 4, pp. 4228–4234, Aug. 2022, doi: 10.11591/ijece.v12i4.pp4228-4234.

[17] A. S. Eesa, Z. Orman, and A. M. A. Brifcani, "a Novel Feature-Selection Approach Based on the Cuttlefish Optimization Algorithm for Intrusion Detection Systems," *Expert Syst. Appl.*, vol. 42, no. 5, pp. 2670–2679, Apr. 2015, doi: 10.1016/j.eswa.2014.11.009.

[18] B. T. Jijo and A. M. Abdulazeez, "Classification Based on Decision Tree Algorithm for Machine Learning," *J. Appl. Sci. Technol. Trends*, vol. 2, no. 01, pp. 20–28, Mar. 2021, doi: 10.38094/jastt20165.

[19] Y. Dani and M. A. Ginting, "Comparison of Iris dataset classification with Gaussian naïve Bayes and decision tree algorithms," *Int. J. Electr. Comput. Eng.*, vol. 14, no. 2, pp. 1959–1968, Apr. 2024, doi: 10.11591/ijece.v14i2.pp1959-1968.

[20] G. L. Pritalia, "Penerapan Algoritma C4.5 untuk Penentuan Ketersediaan Barang E-commerce," *Indones. J. Inf. Syst.*, vol. 1, no. 1, pp. 47–56, Aug. 2018, doi: 10.24002/ijis.v1i1.1727.

[21] M. A. Abdillah, A. Setyanto, and Sudarmawan, "Implementasi Decision Tree Algoritma C4.5 Untuk Memprediksi Kesuksesan Pendidikan Karakter," *J. Tek. Informasi*, vol. 15, no. 2, pp. 59–69, Jul. 2020, doi: 10.35842/jtir.v15i2.349.

[22] J. R. Quinlan, *C4.5: Programs for Machine Learning*. in Morgan Kaufmann series in machine learning. Morgan Kaufmann, 1993.

[23] Menteri Negara Lingkungan Hidup, *Keputusan Menteri Negara Lingkungan Hidup Nomor: 115 Tahun 2003 tentang Pedoman Penentuan Status Mutu Air*. Jakarta, Indonesia, 2003. [Online]. Available: https://www.regulasip.id/ebooks/10016-keputusan-menteri-negara-lingkungan-hidup-nomor-:-115-tahun-2003-tentang-pedoman-penentuan-status-mutu-air-32e636b9

[24] Y. Anggraini and E. Wardhani, "Studi Mutu Air Sungai Cibaligo Kota Cimahi Provinsi Jawa Barat dengan Metode Indeks Pencemar," *Serambi Eng.*, vol. 6, no. 1, pp. 1478–1487, Jan. 2021, doi: 10.32672/jse.v6i1.2589.

[25] S. Rahayu and H. Pristianto, "Studi Penentuan Status Mutu Dan Kualitas Air Sungai Klawili Km.12 Kota Sorong," *J. Tek. Sipil*, vol. 5, no. 1, pp. 35–41, Mar. 2019, doi: 10.33506/rb.v5i1.744.

[26] L. J. Shangguan and Z. Sun, "Enhancing data quality and real-time sharing performance in water informatics through decision tree mining algorithm," *Meas. Sensors*, vol. 31, Feb. 2024, doi: 10.1016/j.measen.2023.101024.

[27] H. Tahraoui, A. Amrane, A. E. Belhadj, and J. Zhang, "Modeling the organic matter of water using the decision tree coupled with bootstrap aggregated and least-squares boosting," *Environ. Technol. Innov.*, vol. 27, Aug. 2022, doi: 10.1016/j.eti.2022.102419.

[28] L. Wang, Y. Huang, J. Ren, H. Zhou, L. Chen, and H. Zhou, "Double-transfer learning-based least squares wavelet support vector machine for water quality classification," *Desalin. Water Treat.*, vol. 320, Oct. 2024, doi: 10.1016/j.dwt.2024.100714.

[29] Z. Jamshidzadeh, M. Ehteram, and H. Shabanian, "Bidirectional Long Short-Term Memory (BILSTM) - Support Vector Machine: A new machine learning model for predicting water quality parameters," *Ain Shams Eng. J.*, vol. 15, no. 3, Mar. 2024, doi: 10.1016/j.asej.2023.102510.

[30] M. A. Kumar, N. Srinivas, P. Ramya, N. Ahlawat, J. Sharma, and F. Vinod, "Monitoring the quality of ground water in pipelines using deep neural network model," *Groundw. Sustain. Dev.*, vol. 24, Feb. 2024, doi: 10.1016/j.gsd.2023.101073.

[31] A. Mazouzi, N. Hadroug, W. Alayed, A. Hafaifa, A. Iratni, and A. Kouzou, "Comprehensive optimization of fuzzy logic-based energy management system for fuel-cell hybrid electric vehicle using genetic algorithm," *Int. J. Hydrogen Energy*, vol. 81, pp. 889–905, Sep. 2024, doi: 10.1016/j.ijhydene.2024.07.237.

[32] M. T. Islam *et al.*, "Regional irrigation water quality index for the Old Brahmaputra River, Bangladesh: A multivariate and GIS-based spatiotemporal assessment," *Results Eng.*, vol. 24, Dec. 2024, doi: 10.1016/j.rineng.2024.103667.

[33] N. A. Dewanti and A. M. Abadi, "Fuzzy Logic Application as a tool for Classifying Water Quality Status in Gajahwong River, Yogyakarta, Indonesia," in *IOP Conf. Ser.: Mater. Sci. Eng.*, Institute of Physics Publishing, Jul. 2019, pp. 1–8. doi: 10.1088/1757-899X/546/3/032005.

[34] N. C. Putra, Jayanta, and Y. Widiastiwi, "Penerapan Logika Fuzzy untuk Mendeteksi Kualitas Air Higiene

Sanitasi menggunakan Metode Sugeno (Studi Kasus: Air Tanah Kota Bekasi)," in *Sem. Nas. Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)*, Aug. 2020, pp. 693–706. [Online]. Available: https://conference.upnvj.ac.id/index.php/senamika/issue/archive

[35] S. Oktavia, S. Musdalifah, and D. Lusiyanti, "Implementasi Sistem Inferensi Fuzzy Pada Pengambilan Keputusan Penentuan Kualitas Air PDAM," *J. Ilm. Mat. Terapan*, vol. 17, no. 1, pp. 118–128, Jun. 2020, doi: 10.22487/2540766x.2020.v17.i1.15187.

[36] A. V. Pahuriray, J. D. Basanta, J. C. T. Arroyo, and A. J. P. Delima, "Flexible Learning Experience Analyzer (FLExA): Sentiment Analysis of College Students through Machine Learning Algorithms with Comparative Analysis using WEKA," *Int. J. Emerg. Technol. Adv. Eng.*, vol. 12, no. 12, pp. 1–15, Dec. 2022, doi: 10.46338/ijetae1222_01.

[37] J. Li, X. Xu, N. Li, H. Pan, and M. Li, "Research on a Classification Model for Online Course Quality Based on C4.5 Algorithm and Naive Bayes," in *Adv. Transdiscipl. Eng.*, IOS Press BV, Aug. 2025, pp. 490–499. doi: 10.3233/ATDE250634.

[38] Presiden Republik Indonesia, *Peraturan Pemerintah No 22 Tahun 2021 tentang Baku Mutu Air yang Disesuaikan untuk Kegunaannya*. Jakarta, Indonesia, 2021. [Online]. Available: https://peraturan.bpk.go.id/Details/161852/pp-no-22-tahun-2021

[39] Menteri Kesehatan Republik Indonesia, *Peraturan Menteri Kesehatan Republik Indonesia Nomor 32 Tahun 2017 Khusus tentang Parameter Kimia dalam Standar Baku Mutu Kesehatan Lingkungan untuk Media Air*. Jakarta, Indonesia, 2017. [Online]. Available: https://peraturan.bpk.go.id/Details/112092/permenkes-no-32-tahun-2017

[40] Menteri Lingkungan Hidup dan Kehutanan Republik Indonesia, *Peraturan Menteri Lingkungan Hidup dan Kehutanan Republik Indonesia Nomor: P.68/Menlhk/Setjen/Kum.1/8/2016 tentang Baku Mutu Air Limbah Domestik*. Jakarta, Indonesia, 2016. [Online]. Available: https://peraturan.go.id/id/permen-lhk-no-p-68-menlhk-setjen-2016-tahun-2016

[41] Pemerintah Daerah Kalimantan Timur, *Peraturan Daerah Provinsi Kalimantan Timur Nomor 2 Tahun 2011 tentang Pengelolaan Kualitas Air dan Pengendalian Pencemaran Air*. Samarinda, Indonesia, 2011. [Online]. Available: https://peraturan.bpk.go.id/Details/32978/perda-prov-kalimantan-timur-no-2-tahun-2011

[42] Menteri Kesehatan, *Peraturan Menteri Kesehatan Republik Indonesia Nomor 416/Menkes/Per/IX/1990*. Jakarta, Indonesia, 1990. [Online]. Available: https://labkesda.kapuaskab.go.id/public/deploy/pdf/1762311251_1c7efecaf84074f3e5ef.pdf

[43] Menteri Kesehatan Republik Indonesia, *Permenkes Nomor 492/Menkes/Per/IV/2010 tentang Persyaratan Kualitas Air Minum*. Jakarta, Indonesia, 2010. [Online]. Available: https://jdih.pom.go.id/download/rule/350/492/2010/Persyaratan%20Kualitas%20Air%20Minum

[44] M. M. H. Masum, M. B. Islam, A. Hoque, and M. A. Hossen, "Impacts of firefighting effluents from a sugar refinery fire on river water quality in Bangladesh," *NPJ Clean Water*, vol. 4, Dec. 2025, doi: 10.1016/j.clwat.2025.100150.

[45] N. Huda, T. Ahmed, M. H. Masum, N. Faruque, and M. S. Islam, "Assessment of surface water quality using advanced statistical techniques around an urban landfill: A multi-parameter analysis," *City Environ. Interact.*, vol. 28, Dec. 2025, doi: 10.1016/j.cacint.2025.100237.

[46] N. Khatri, S. Tyagi, D. Rawtani, and M. Tharmavaram, "Assessment of river water quality through application of indices: a case study River Sabarmati, Gujarat, India," *Sustain. Water Resour. Manag.*, vol. 6, no. 6, Dec. 2020, doi: 10.1007/s40899-020-00459-8.

[47] Fianti and S. Hartati, "Verifikasi Unjuk Kerja Pengendali Berlogika Samar pada Proses Pemanasan Ruang," *J. Kaunia*, vol. IV, no. 1, pp. 57–69, Apr. 2008, [Online]. Available: https://digilib.uin-suka.ac.id/id/eprint/7921/

[48] M. Widyastiti, "Penerapan Metode Fuzzy Mamdani untuk Menentukan Tingkat Kualitas Air Sungai Lintas Provinsi di Pulau Jawa," *Ekologia*, vol. 18, no. 1, pp. 17–24, Apr. 2018, [Online]. Available: https://journal.unpak.ac.id/index.php/ekologia/issue/view/150/showToc

[49] J. Manurung, B. Sinaga, P. M. Hasugian, Logaraj, and S. Ramen, "Analisis Algoritma C4.5 dan Fuzzy Sugeno untuk Optimasi Rule Base Fuzzy," *J. Sist. Inf. Ilmu Komput. Prima*, vol. 5, no. 2, pp. 166–171, Feb. 2022, doi: 10.34012/jurnalsisteminformasidanilmukomputer.v5i2.2488.

[50] H. Fakhrurroja, E. T. Nuryatno, A. Munandar, M. Fahmi, and N. A. Mahardiono, "Water quality assessment monitoring system using fuzzy logic and the internet of things," *J. mechatron. electr. power veh. technol.*, vol. 14, no. 2, pp. 198–207, Dec. 2023, doi: 10.14203/j.mev.2023.v14.198-207.

[51] Y. Barzegar, I. Gorelova, F. Bellini, and F. D'Ascenzo, "Drinking Water Quality Assessment Using a Fuzzy Inference System Method: A Case Study of Rome (Italy)," *Int. J. Environ. Res. Public Health*, vol. 20, no. 15, Aug. 2023, doi: 10.3390/ijerph20156522.