Jurnal Teknik Informatika (JUTIF) Vol. 6, No. 4, Agustus 2025, Page. 2223-2238
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4847

Prediction of Turbidity Removal Time in Electrocoagulation Wastewater Using Random Forest, XGBoost, and Others: A Data-Driven Information System Approach

Sinung Suakanto*[1], Tan Lian See[2], Zatul Alwani Shaffiei[3], Taufiq Maulana Firdaus[4], Muharman Lubis[5], Anggera Bayuwindra[6]

[1,4,5]Information System, Telkom University, Indonesia
[2,3]Malaysia - Japan International Institute of Technology, Universiti Teknologi Malaysia, Malaysia
[6]School of Electrical Engineering and Informatics, Bandung of Institute of Technology, Indonesia

Email: [1]sinung@telkomuniversity.ac.id

Received : Jun 9, 2025; Revised : Jun 30, 2025; Accepted : Aug 16, 2025; Published : Aug 19, 2025

Abstract

Electrocoagulation is an effective and environmentally friendly technology for treating wastewater by removing contaminants such as turbidity, heavy metals, and organic compounds. Accurately predicting turbidity removal time is essential for optimizing treatment performance and operational efficiency. However, this is challenging due to complex, nonlinear relationships between multiple parameters including current, voltage, electrode configuration, conductivity, and turbidity removal rate. This study aims to develop a predictive framework by comparing six supervised regression models, namely Linear Regression, Polynomial Regression, Random Forest, Support Vector Regression (SVR), XGBoost, and Long Short-Term Memory (LSTM), using key electrocoagulation parameters. After extensive data preprocessing, a dataset of 281 samples was used for training and validation. Among them, Random Forest achieved the best performance ($R^2 = 0.876$, RMSE = 601.15). A data-driven information system is proposed to integrate these predictive capabilities for real-time monitoring and control. By improving turbidity prediction accuracy, the system enables the sustainable utilization of water as a valuable asset, even in its wastewater form. The approach enhances decision-making by providing intelligent feedback for process optimization. This research contributes to the advancement of intelligent, sustainable wastewater treatment systems by integrating machine learning prediction models with practical process control applications in informatics.

Keywords : *Electrocoagulation Turbidity Removal, Random Forest, Wastewater Treatment, Water Resource Asset.*

1. INTRODUCTION

Water is a fundamental asset in industrial production, playing a vital role in operations such as washing, cooling, transporting raw materials, and serving as a solvent or processing medium, and is therefore considered a supporting asset that directly supports manufacturing efficiency and output [1], [2]. Because this process depends on water to support the operational activities, the availability and quality of water directly impacts its operational efficiency and sustainability. However, the use of water in industrial processes inevitably generates wastewater, which contains a variety of pollutants that must be removed before the water can be reused or safely discharged [3]. Once water has been utilized and mixed with various process materials, it is often regarded as a non-productive asset that must be discarded; however, its disposal must adhere to strict environmental regulations and compliance standards to prevent ecological harm [4].

Unfortunately, the disposal process is often overlooked as merely a waste elimination step [5]. Untreated or inadequately treated wastewater can damage ecosystems, pollute water sources, and pose

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 4, Agustus 2025, Page. 2223-2238
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4847

health risks. Proper treatment is essential not only for operations but also as an environmental obligation. Discharges must meet international water and wastewater management regulations, which, like local laws, require environmental impact assessments before projects that could harm ecosystems.

Among the most common pollutants found in industrial wastewater are Chemical Oxygen Demand (COD), which indicates the presence of organic compounds; Total Phosphorus (TP), often originating from detergents or biological waste; Nitrite (NO_2^-), a product of nitrogen transformations; Ammonia Nitrogen (NH_3-N), a byproduct of protein and urea decomposition; benzene, toluene, ethyl benzene, and xylenes, commonly known as BTEX, along with polyaromatic hydrocarbons (PAHs); alkylphenols (APs) [6]; and Total Suspended Solids (TSS) [7], which include fine particulate matter [8], [9]. These contaminants are particularly prevalent in wastewater generated by food and beverage processing industries, livestock and slaughterhouse operations, industrial activities, textile, and even domestic wastewater treatment plants [10], [11], [12].Without effective treatment, these pollutants can contribute to eutrophication, oxygen depletion, and long-term aquatic degradation [13].

One of the promising technologies for addressing such complex wastewater compositions is electrocoagulation (EC) [14]. Electrocoagulation involves the in-situ generation of coagulant species via the dissolution of sacrificial electrodes (typically aluminum or iron) under an electric current. Electrocoagulation has emerged as a promising technology for the treatment of wastewater due to its efficiency in removing contaminants such as turbidity, heavy metals, and organic matter [15] and even oil or diesel from drilling fluids wastewater [16], [17]. These coagulants destabilize and aggregate suspended, colloidal, and dissolved contaminants, enabling their removal through sedimentation or flotation. Compared to conventional chemical coagulation and other conventional methods, EC destabilizes, and aggregates suspended, colloidal, and dissolved contaminants for removal via sedimentation or flotation, uses fewer chemicals, requires less space, is easier to operate, and removes a broad range of pollutants. While future integration with other separation techniques is anticipated [18], [19], [20], focusing on standalone EC remains essential for assessing operational efficiency [21].

Despite its advantages, electrocoagulation performance is influenced by a multitude of factors, including electrode type, current density, pH, inter-electrode spacing, and processing time [22]. Among these, Purification time or turbidity removal time, defined as the duration required for effective contaminant separation, is a key metric that influences both treatment efficiency and energy consumption [23] [24]. Traditionally, the determination of clarification time relies on empirical approaches or repetitive laboratory trials, which are resource-intensive and impractical for dynamic real-time applications.

In industrial settings, wastewater volume and composition often vary with production cycles, time of day, or operational loads. Using a fixed EC configuration, such as constant voltage, current, or electrode count, can reduce efficiency. High-power setups may waste energy or over-treat during low-load periods, while under high-load conditions, they may fail to achieve adequate purification level [25]. This mismatch between operational configuration and wastewater characteristics presents a clear decision-making challenge for process engineers.

Electrocoagulation (EC) has increasingly gained attention as an effective and environmentally friendly technique for water and wastewater purification. Recent research has largely focused on treatment performance and the underlying mechanisms of EC, as well as the integration of EC with other processes to address emerging pollutants [26]. However, there remains a limited exploration of how machine learning can be leveraged to support the operational management of EC systems. There are only limited studies that have applied machine learning for water turbidity removal using EC or similar methods (see Table 1). Several parameters influencing the EC process include current/voltage, electrolyte concentration, inter-electrode distance, electrolysis time and electrode combinations (Fe–Fe and Al–Al); the best results are achieved by optimizing the combination of these factor [27], [28]. Most prior works,

Vol. 6, No. 4, Agustus 2025, Page. 2223-2238
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4847

such as those by Jery et al [29] and Khan et al) [30], focused on predicting pollutant removal efficiency using ML but did not address temporal aspects like purification time.

To address this, there is a growing need for the development of data-driven decision support systems (DSS) that can intelligently recommend EC process settings based on real-time or predicted wastewater profiles. In this context, machine learning (ML) offers a robust framework for building predictive models that can estimate turbidity removal time under various input conditions. The turbidity removal time can be assessed by tracking the reduction in turbidity levels, typically from a highly turbid state (above 95%) to a significantly clearer state (below 2%), indicating that the water has reached an acceptable level of cleanliness. Such models can be embedded into a DSS to guide operators in selecting optimal EC parameters—such as voltage, current, or electrode count—ensuring that treatment is adaptive, cost-effective, and energy-efficient. By aligning the treatment configuration with the actual characteristics of incoming wastewater, this approach supports more sustainable and responsive process control.

This study aims to develops and evaluates six machine learning models including Linear Regression, Polynomial Regression, Random Forest, Support Vector Regression, XGBoost, and LSTM to predict turbidity removal time using key electrocoagulation parameters. Unlike previous studies that focused on pollutant removal efficiency such as COD, nitrogen, or phosphate, this study uniquely targets turbidity removal time as a predictive variable. By shifting the focus toward the time dimension of turbidity reduction, this work offers new insights into process optimization and dynamic control of electrocoagulation. This novel contribution enhances real-time operational decision-making and supports the development of data-driven information systems for sustainable wastewater treatment. A summary of the state of the art, highlighting this study's position relative to previous work, is provided in Table 1.

Table 1. Research State of The Art

Author	Technology	Predicted Variable	Input Variable	Model	Evaluation Result
Jery et al (2023) [29]	Electrocoagulation (EC)	COD removal (%)	Current density, pH, COD conc., electrode area, NaCl conc., time	Artificial Neural Network	MAE = 1.12%, R^2 = 0.99
Khan et al (2024) [30]	Sequencing Batch Reactor (SBR)	Nutrient removal efficiency (SND, total nitrogen)	Wastewater concentration, HRT, mixing ratios	CatBoost (Machine Learning)	SND efficiency 69%, total nitrogen removal 66%, nutrient removal 88–98%, COD removal 93%
Zakoor et al (2023) [31]	EC combined with MBR	Nitrate (NO_3^-) removal, Phosphate (PO_4^{3-}) removal	Temperature, pH, DO, initial concentrations of NO_3^- and PO_4^{3-}	Artificial Neural Network (ANN)	Removal efficiency: 98.1% (PO_4^{3-}); ANN model accuracy: 98.1% (PO_4^{3-})
This research	Electrocoagulation (EC)	Turbidity removal time	Current, voltage, number of electrodes, electrode spacing, conductivity, turbidity removal rate	Various Machine Learning	Discussed at next section

Vol. 6, No. 4, Agustus 2025, Page. 2223-2238
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4847

2. METHOD

Electrocoagulation is an advanced wastewater treatment process that utilizes electrical current to remove contaminants from water. In this process, metal electrodes commonly aluminum are submerged in wastewater and connected to a power supply. When electric current passes through the electrodes, metal ions are released into the solution, which then react with pollutants such as suspended solids, colloidal particles, and dissolved substances. These reactions cause the pollutants to coagulate into larger aggregates, or flocs, which can be easily separated from the water by sedimentation or flotation. Electrocoagulation is considered an effective, environmentally friendly technique due to its ability to treat a wide range of pollutants without adding chemical coagulants, and it offers advantages like reduced sludge production and simplified operation.

2.1. System Architecture

The system architecture developed in this study is illustrated in Figure 1. The setup includes an electrocoagulation (EC) treatment box, which receives input in the form of wastewater. Within this box, several electrodes are installed with specific spacing configurations to facilitate the electrocoagulation process. A turbidity sensor is integrated into the system to measure the turbidity level of the wastewater throughout the treatment. The measurement results are stored and transmitted to a cloud server, enabling remote data storage and monitoring.

Several experimental trials were conducted using wastewater samples with high turbidity levels (above 95%) and treated until low turbidity levels (below 2%) were achieved. Each trial involved varying key process parameters, including current, voltage, number of electrodes, and the distance between electrodes. Turbidity levels were recorded at fixed time intervals (e.g., every 1 or 5 minutes), and all measurement data were compiled to form a structured dataset for model training and evaluation**.**

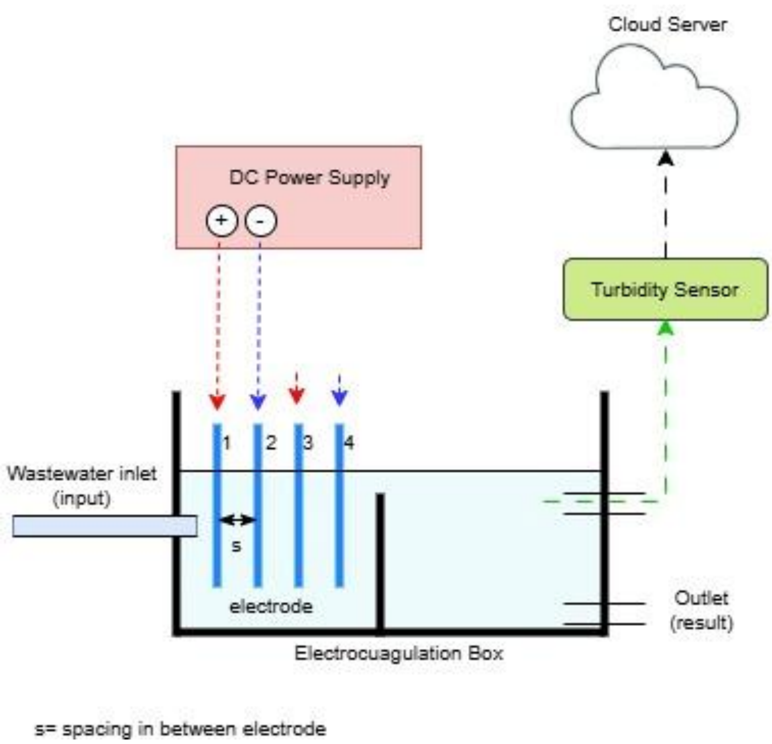

Figure 1. Electrocoagulation Architecture

An example of the physical implementation of the system can be seen in Figure 2.

Vol. 6, No. 4, Agustus 2025, Page. 2223-2238
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4847

Figure 2. Electrocoagulation System

2.2. Research Methodology

The dataset used in this study was obtained through direct measurement using a self-constructed experimental setup, as shown in Figure 2. The measurement parameters used as input variables for the model include current, voltage, number of electrodes, spacing between electrodes, and conductivity. The monitored output parameters consist of time logging and turbidity level. Based on these measurements, turbidity removal (%) is calculated as the difference between the initial turbidity level and the turbidity level at a given time point. Furthermore, the turbidity removal time is determined as the duration required for the wastewater to transition from a high turbidity state (cloudy condition) to a low turbidity state (clear condition). This time-based performance metric serves as the target variable in the predictive modeling process. Examples of turbid (high turbidity level) and clear water (low turbidity level) conditions can be seen in Figure 3.

Figure 3. An Example of Wastewater before and after treatment

Vol. 6, No. 4, Agustus 2025, Page. 2223-2238
P-ISSN: 2723-3863
https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4847

The research methodology framework is illustrated in Figure 4, presenting a clear overview of the key stages in data processing, model development, and evaluation. The workflow consists of two main phases: training and testing. In the **training phase**, the dataset is first subjected to data preprocessing to clean and prepare the data for analysis. Following this, various machine learning algorithms are applied to train predictive models. Hyperparameter tuning is performed to optimize model performance and ensure the best possible fit to the training data. The outcome of this phase is a finalized trained model ready for evaluation. During the **testing phase**, the trained model is applied to a separate, previously unseen dataset (test data) to generate predictions. These predictions are then compared with the actual observed values (ground truth) in the test data to assess model accuracy and reliability.

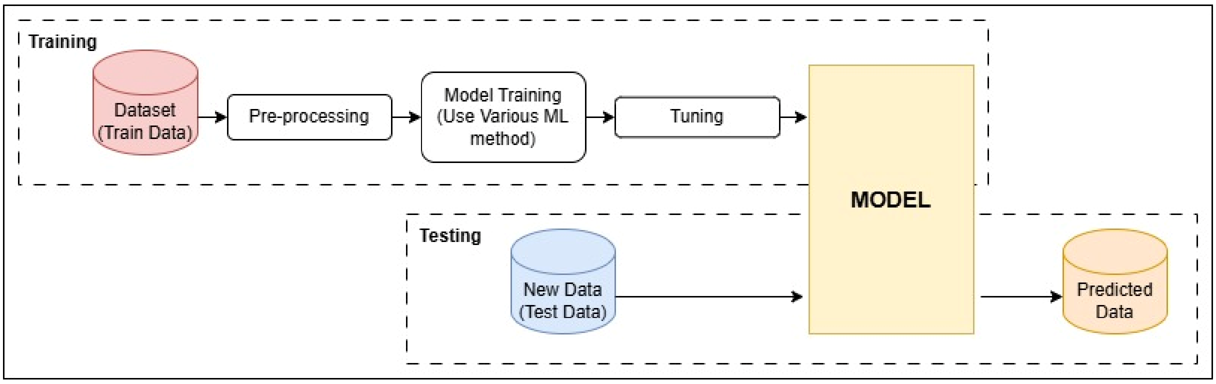

Figure 4. Research Method

During the preprocessing stage, the raw dataset was first loaded, and relevant features were selected, including electrocoagulation process parameters and turbidity removal percentage. Null values in the collected data were identified and handled using deletion. The input features were then standardized using a StandardScaler to normalize the data distribution and ensure that all variables contribute equally to the model training.

The input variables in this study consist of six features derived from electrocoagulation process parameters, which include: Average Current (A), Voltage (V), Number of Electrodes, Electrode Spacing (mm), Conductivity, and Turbidity Removal (%). These features are used to predict the output variable, which is the required reaction time (in seconds) for the electrocoagulation process.

A total of six regression models were implemented and compared: Linear Regression, Polynomial Regression (degree=2), Random Forest, Support Vector Regression (SVR), XGBoost, and Long Short-Term Memory (LSTM) neural network. The dataset was standardized and split into training and testing sets to ensure fair model evaluation. Predictions from all models were also tested on new input scenarios to assess their performance in real-world cases.

The selection of these six models was based on their diverse learning characteristics: Linear and Polynomial Regression represent traditional statistical approaches; Random Forest and XGBoost offer tree-based ensemble learning known for robustness and interpretability; SVR provides strong generalization for small- to medium-sized datasets; and LSTM captures sequential patterns useful for time-dependent observations such as turbidity reduction over time.

Model performance is evaluated using standard regression metrics, including R^2 (coefficient of determination), Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Pearson correlation coefficient. The mathematical formulations for R^2, RMSE, MAE and Pearson correlation are provided in Equations (1), (2), (3) and (4), respectively, where y_i^{true} denotes the actual value, y_i^{pred} is the predicted value, $\bar{y}_i^{true}$ is the mean of actual values, $\bar{y}_i^{pred}$ is the mean of predicted values, and n is the total number of observations.

Vol. 6, No. 4, Agustus 2025, Page. 2223-2238
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4847

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i^{true} - y_i^{pred})^2}{\sum_{i=1}^{n}(y_i^{true} - \overline{y}_i^{pred})^2} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i^{true} - y_i^{pred})^2} \quad (2)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i^{true} - y_i^{pred}| \quad (3)$$

$$r = \frac{\sum_{i=1}^{n}(y_i^{pred} - \bar{y}_i^{pred})(y_i^{true} - \bar{y}_i^{true})}{\sqrt{\sum_{i=1}^{n}(y_i^{pred} - \bar{y}_i^{pred})^2} . \sqrt{\sum_{i=1}^{n}(y_i^{true} - \bar{y}_i^{true})^2}} \quad (4)$$

3. RESULT

The findings of this study are organized into three parts: data exploration, model performance comparison, and analysis of the best-performing model. The initial analysis focuses on understanding the dataset structure, parameter distribution, and their relationship with turbidity removal time

3.1. Dataset

The initial dataset contained 49,856 entries, which was reduced to 281 after data cleaning and preprocessing. Preprocessing included handling missing values, deleting null values, detecting anomalies, encoding categorical variables into numerical codes, and standardizing input features using StandardScaler. After preprocessing, data aggregation was conducted by grouping and averaging relevant measurements to obtain representative values for each experiment. This significant reduction was due to differences in sampling intervals: while most sensor data were recorded every minute, turbidity measurements were taken only every five minutes. To maintain consistency and ensure complete feature representation for each data point, only rows with full measurements—including turbidity—were retained. These steps ensured the dataset was clean, consistent, and suitable for model training. An example of the processed dataset is shown in Table 2.

Table 2. Example of Dataset

Time (sec)	Average current (A)	Volt (V)	No. of electrode	Spacing electrode (mm)	pH	Conduct-ivity	Initial Turbidity	Current Turbidity	Turbidity removal (%)
0	1.251	10	6	20	7.143	367.469	75.39	75.39	0
296	1.251	10	6	20	7.275	362.781	75.39	61.38	18.58337
588	1.251	10	6	20	7.406	355.094	75.39	53.53	28.99589
878	1.251	10	6	20	7.477	346.75	75.39	50.31	33.26701
1177	1.251	10	6	20	7.566	337.656	75.39	49.59	34.22205
1467	1.251	10	6	20	7.665	332.312	75.39	45.23	40.00531
1758	1.251	10	6	20	7.758	323.781	75.39	32.95	56.29394
2049	1.251	10	6	20	7.857	315.438	75.39	26.98	64.21276
2340	1.251	10	6	20	7.944	310	75.39	24.76	67.15745
2616	1.251	10	6	20	8.015	298.188	75.39	15.27	79.74532
2879	1.251	10	6	20	8.082	293.312	75.39	11.01	85.39594
3171	1.251	10	6	20	8.144	288.438	75.39	8.88	88.22125
3459	1.251	10	6	20	8.204	284.219	75.39	2.45	96.75023
3753	1.251	10	6	20	8.253	277.656	75.39	1.65	97.81138

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 4, Agustus 2025, Page. 2223-2238
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4847

3.2. Data Exploration

In this section, data exploration is conducted to understand the relationships among the input variables for the electrocoagulation process. The analysis begins with a correlation matrix to examine interrelationships between features such as Average Current, Voltage, Number of Electrodes, Electrode Spacing, Conductivity, and Turbidity Removal. The full correlation matrix is presented in Figure 5.

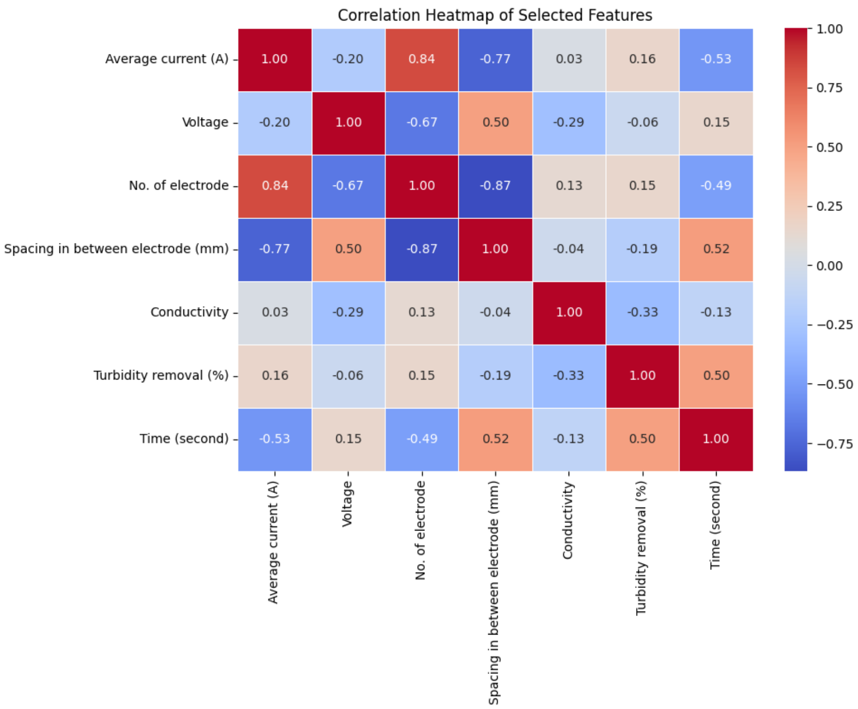

Figure 5. Correlation Matrix of Electrocoagulation Features

When aiming to predict "Time (second)", the correlation analysis provides valuable insights into the relationships between this target variable and the input features. "Turbidity removal (%)" exhibits the strongest positive correlation with "Time (second)" (0.50), suggesting that longer treatment durations tend to result in greater turbidity removal. Additionally, "Spacing in between electrode (mm)" shows a moderate positive correlation (0.52), indicating that larger electrode spacing may generally require longer operational times. On the other hand, "No. of electrode" has a moderate negative correlation (-0.49), implying that systems with more electrodes might achieve the desired outcome in a shorter time. Similarly, "Average current (A)" is moderately negatively correlated with time (-0.53), suggesting that higher current may accelerate the process. "Voltage" and "Conductivity" show weaker negative correlations (-0.15 and -0.13, respectively), indicating they may have less direct influence on the prediction of time. Overall, these patterns highlight the key features that could be most influential in modeling and predicting the required operational time in the electrocoagulation process.

The distribution of the input variables is presented in Figure 6, which provides a comprehensive overview of all variables. We can observe a variety of distribution shapes across the features. The provided histograms display the frequency distribution of each selected variable. "Average current (A)" and "No. of electrode" exhibit bimodal distributions, with peaks at lower and higher values, suggesting two distinct operating regimes or settings. "Voltage" also shows a bimodal distribution, heavily concentrated at 10V and 25V. "Spacing in between electrode (mm)" is highly skewed to the left, with a large number of observations at smaller spacing and a smaller peak around 90mm. "Conductivity" appears to have a relatively normal distribution but is slightly skewed to the left, with most values clustered around 350 µS/cm. "Turbidity removal (%)" has a somewhat uniform distribution, with frequencies spread across various percentages, indicating a wide range of removal efficiencies.

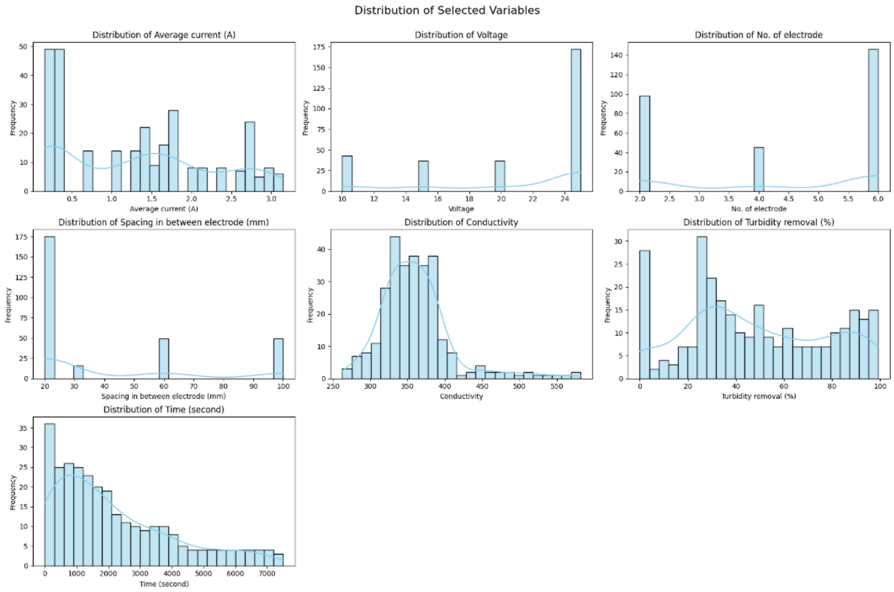

Figure 6. Distribution Histograms of Key Electrocoagulation Features

3.3. Model Development and Evaluation

The dataset was split into 80% for training and 20% for testing. Several machine learning algorithms were applied to predict the electrocoagulation reaction time based on six input features. The performance of each model is summarized in Table 3.

Table 3. Model Accuracy

Model	R2	MSE	MAE	RMSE	Correlation
Linear Regression	0.665143	972856.2	739.7013	986.3347	0.818570709
Polynomial Regression	0.862191	400375.2	472.4686	632.7521	0.929968619
XGBoost	0.864684	393132.5	**438.1191**	627.0028	0.93132205
Random Forest	**0.875614**	**361377.4**	467.87	**601.1467**	**0.937571542**
SVM	-0.06361	3090081	1267.273	1757.863	0.772176785
LSTM	-0.64171	4769636	1561.186	2183.95	0.637716002

Table 3 presents the performance metrics of various machine learning models used to predict "Time (second)" in the electrocoagulation process. Among the models evaluated, the Random Forest model achieved the best performance, with the highest R^2 value of 0.8756, indicating strong predictive power. It also recorded the lowest RMSE (601.15) and a high correlation (0.9376) between predicted and actual values.

The XGBoost and Polynomial Regression models followed closely, with R^2 values of 0.8647 and 0.8622, respectively, and similarly low errors, suggesting that these models are also well-suited for capturing complex patterns in the data. In contrast, the Support Vector Machine (SVM) and LSTM models performed poorly, with negative R^2 values (indicating worse performance than simply predicting the mean), and significantly higher error metrics. This suggests that these models may not be appropriate for this particular prediction task, possibly due to insufficient data or poor model tuning.

Vol. 6, No. 4, Agustus 2025, Page. 2223-2238
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4847

4. DISCUSSIONS

This section discusses key findings, including model performance implications, strengths and limitations of the approaches, and their relevance to predicting electrocoagulation duration. It also highlights how these results can inform future research and practical applications in wastewater treatment optimization

4.1. Discussion

The Random Forest model outperformed other algorithms, achieving the highest R^2 value (0.876) and lowest error metrics (RMSE = 601.15). This superior performance can be attributed to Random Forest's ability to capture complex nonlinear relationships and interactions among features without overfitting. The model's robustness to noise and capacity to handle multicollinearity among input variables further contributed to its accuracy. These characteristics make Random Forest particularly suitable for electrocoagulation process data, where nonlinear effects and parameter interdependencies are common.

The XGBoost and Polynomial Regression (degree 2) models also delivered strong results, with R^2 values close to Random Forest and relatively low error rates. XGBoost's gradient boosting framework allows it to iteratively refine predictions, improving model generalization on unseen data. Polynomial Regression, although simpler, demonstrated that introducing nonlinear terms helped capture curvature in the relationship between input features and reaction time. However, its performance was slightly inferior to ensemble-based models, likely due to limited flexibility in modeling complex interactions.

Conversely, Support Vector Regression (SVR) and Long Short-Term Memory (LSTM) networks underperformed considerably, with negative R^2 values and high error metrics. The poor SVR results might stem from insufficient hyperparameter tuning or its limitations in handling noisy or highly variable datasets. LSTM's low performance could be due to the relatively small size of the final cleaned dataset (281 entries), which is inadequate for training deep learning models that require extensive data for learning temporal dependencies and complex patterns.

The correlation analysis provided valuable insights into feature relevance. The moderate positive correlation of electrode spacing and turbidity removal percentage with reaction time indicates their direct impact on process duration, while negative correlations of current and number of electrodes suggest these parameters can accelerate treatment. These findings align with established electrocoagulation theory, reinforcing the importance of process optimization based on operating parameters.

Compared to previous studies such as Jery et al [29], Khan et al [30], and Zakoor et al [31] which focused on predicting pollutant removal efficiencies like COD, nitrogen, or phosphate, this study provides a novel contribution by targeting turbidity removal time. While prior research achieved high R^2 values using ANN or gradient boosting for removal percentages, none addressed the time dimension in electrocoagulation treatment. This work fills that gap by offering a time-based prediction framework that enhances operational decision-making and real-time adaptability.

From a practical perspective, deploying the Random Forest model within a data-driven information system can enable real-time prediction of turbidity removal time, facilitating dynamic process control and optimization in wastewater treatment plants. From a computer science and informatics standpoint, this integration demonstrates the potential of machine learning for real-time decision support in cyber-physical systems, contributing to the advancement of intelligent environmental monitoring infrastructures.

However, several limitations should be noted. First, the dataset size after cleaning was relatively small, which may reduce the model's ability to generalize. Second, the current study focuses on a

Vol. 6, No. 4, Agustus 2025, Page. 2223-2238
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4847

specific electrocoagulation setup and wastewater profile; as such, the model's applicability to different wastewater types, flow conditions, or electrode configurations has not been validated. These factors may limit the scalability and robustness of the model across broader industrial settings Future work should explore expanding the dataset with more diverse operational conditions and incorporating additional relevant parameters such as temperature and pH. Moreover, integrating real-time sensor data and feedback loops could enhance model adaptability and accuracy in live systems.

4.2. Implication of the Study and Decision Support System Development

This study presents important practical and theoretical contributions. Practically, the ability to predict turbidity removal time with good accuracy enables wastewater treatment operators to optimize electrocoagulation processes, improving efficiency, reducing energy use, and lowering operational costs. The integration of machine learning models like Random Forest into a data-driven system provides valuable, real-time decision support for better process control and system reliability.

Importantly, the developed predictive model offers a foundational tool for basic prediction scenarios. Given known input variables such as voltage, current, and flow rate, the model can estimate the required purification time accurately. This predictive capacity allows operators to anticipate processing durations and better plan operational workflows. An example of a simple application implementing this prediction is illustrated in Figure 7, demonstrating how real-time input data can translate into actionable process time estimates.

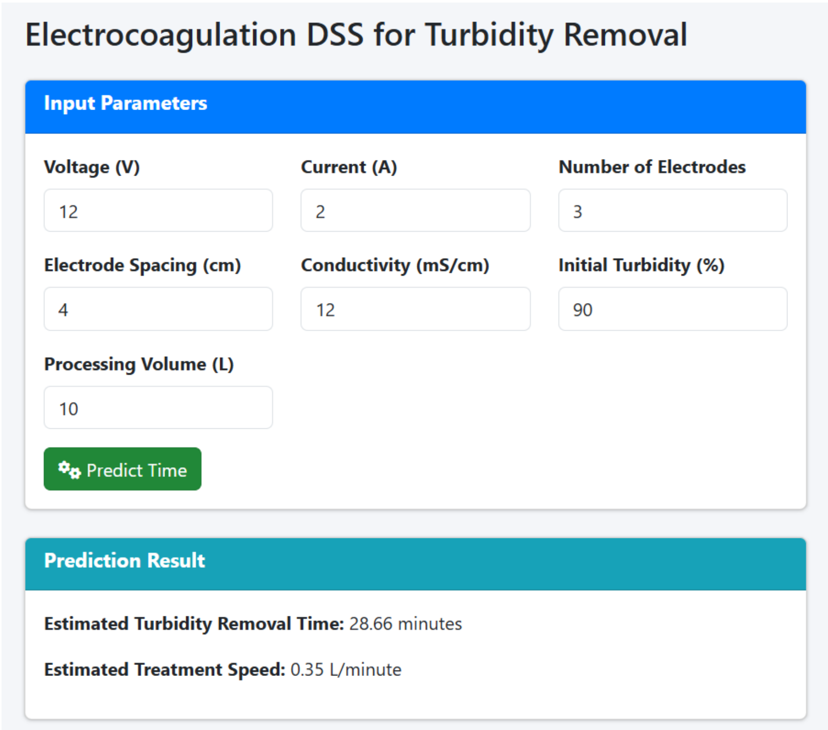

Figure 7. Basic Electrocoagulation DSS Application for predict Turbidity Removal Time

In a more advanced context, integrating the model into a live wastewater treatment control system could enable adaptive process management. By linking the system with sensors measuring real-time flow rate and water quality parameters, the electrocoagulation system can dynamically adjust voltage and current levels. This adaptive mechanism ensures that processing time aligns with the varying flow rates entering the system, thereby optimizing treatment speed without compromising quality. Such feedback-driven control can significantly improve energy efficiency and process stability, ultimately supporting the development of smarter, more sustainable wastewater treatment infrastructures. As a

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871

Vol. 6, No. 4, Agustus 2025, Page. 2223-2238
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4847

practical illustration, a simple Decision Support System (DSS) application can be used to adjust voltage, current, and other parameters in real time based on the dynamic flow rate entering the system, enabling responsive and efficient process control, as shown in Figure 8. It can be observed from the application example that energy efficiency is achieved by allowing current and voltage to vary dynamically rather than remain constant, adapting to the changes in incoming wastewater flow rate.

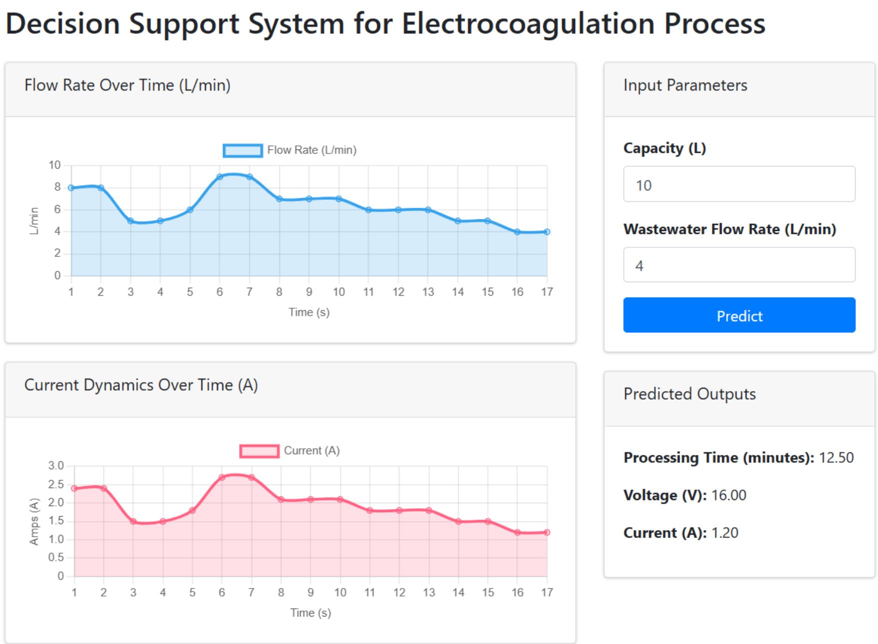

Figure 8. Adaptive System to Control Parameter for Electrocoagulation

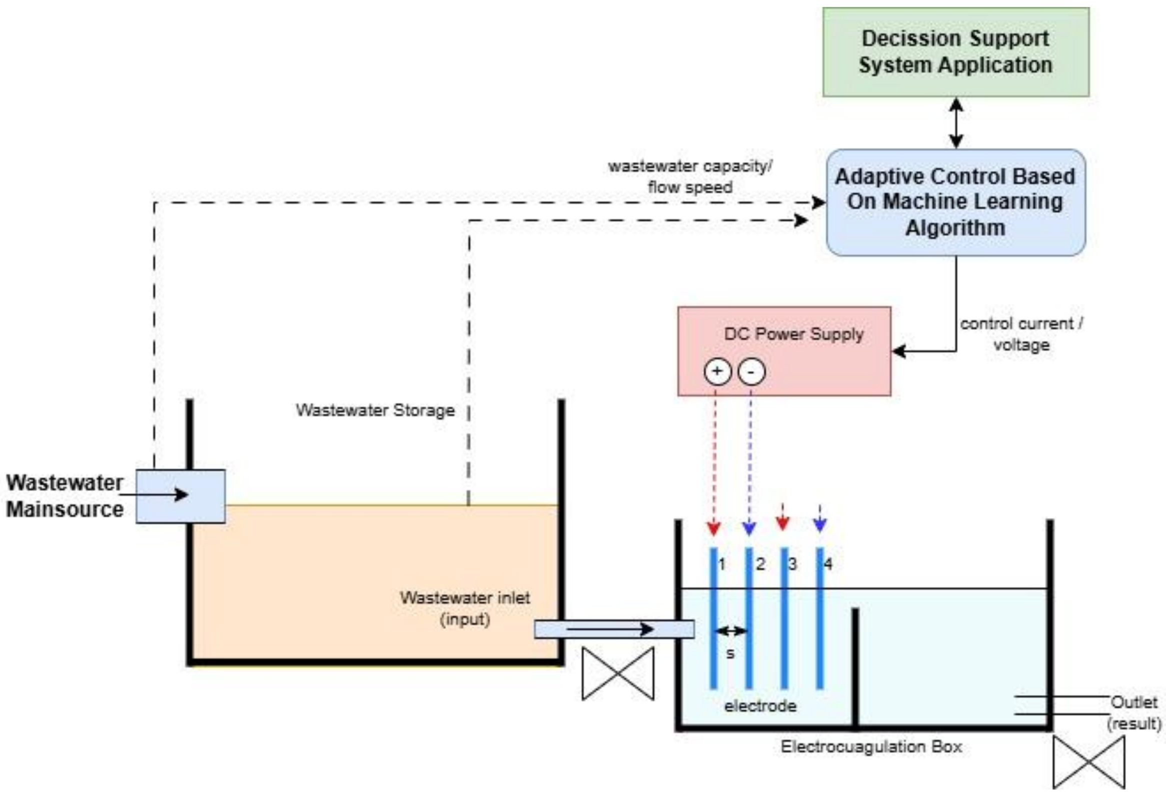

Figure 9. Adaptive Control System for Wastewater Treatment system

The proposed control system architecture for future development is illustrated in Figure 9. As shown in the figure, the system is capable of detecting or measuring the current wastewater capacity and flow speed. Based on these inputs, the system can predict the required processing time (e.g., x minutes) to completely treat the incoming wastewater. Leveraging machine learning models with adaptive control logic, the system can then determine the optimal voltage and current levels needed for the electrocoagulation process. As a result, the adaptive control block dynamically adjusts the electrical parameters in real-time, ensuring efficient operation that matches the current load and flow characteristics.

This research advances waste discharge asset management by showing how real process data can support predictive modeling. The approach is adaptable, scalable, and enables integration of electrocoagulation with machine learning-based control for more precise, energy-efficient, and sustainable treatment systems. Overall, the findings from this study not only demonstrate the feasibility of predicting turbidity removal time using machine learning, but also lay the groundwork for future adaptive wastewater treatment systems.

5. CONCLUSION

This study successfully developed and evaluated multiple machine learning models to predict the turbidity removal time in electrocoagulation wastewater treatment processes. Among the models tested, the Random Forest algorithm demonstrated the best performance, offering high accuracy and reliability in prediction. The research highlights the potential of data-driven approaches to optimize wastewater treatment by enabling precise process control and operational efficiency. This study demonstrates the feasibility of integrating ensemble learning models with intelligent decision systems in the context of environmental informatics, contributing to the advancement of predictive process control in wastewater management. By treating water not merely as waste but as a reusable asset, this research reinforces the strategic importance of water in sustainable industrial operations. The findings also provide a solid foundation for future work on automated control systems and the broader application of machine learning in environmental engineering. Future directions may include multi-objective optimization that balances energy efficiency and treatment speed, real-time integration with sensor data, and deployment in diverse wastewater treatment settings to assess scalability and generalizability. Overall, this study contributes valuable insights toward sustainable and intelligent wastewater treatment management.

CONFLICT OF INTEREST

The authors declares that there is no conflict of interest between the authors or with research object in this paper.

ACKNOWLEDGEMENT

We would like to sincerely thank to Universiti Teknologi Malaysia for funding this research through Digital-Infused Research Grant (PY/2024/00454) and Telkom University for their valuable funding and support that made this research possible.

REFERENCES

[1] S. Suakanto, A. Setiawan, and W. A. Nurtrisha, "Data Modeling for Productive Asset and Supporting Asset," in *2023 International Conference on Electrical Engineering and Informatics (ICEEI)*, 2023, pp. 1–6. doi: 10.1109/ICEEI59426.2023.10346719.

[2] Y. G. Asfaha, A. K. Tekile, and F. Zewge, "Hybrid process of electrocoagulation and electrooxidation system for wastewater treatment: A review," *Clean Eng Technol*, vol. 4, p. 100261, 2021, doi: https://doi.org/10.1016/j.clet.2021.100261.

Vol. 6, No. 4, Agustus 2025, Page. 2223-2238
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4847

[3] F. Y. AlJaberi *et al.*, "Can electrocoagulation technology be integrated with wastewater treatment systems to improve treatment efficiency?," *Environ Res*, vol. 214, p. 113890, 2022, doi: https://doi.org/10.1016/j.envres.2022.113890.
[4] A. Shahedi, A. K. Darban, F. Taghipour, and A. Jamshidi-Zanjani, "A review on industrial wastewater treatment via electrocoagulation processes," *Curr Opin Electrochem*, vol. 22, pp. 154–169, 2020, doi: https://doi.org/10.1016/j.coelec.2020.05.009.
[5] S. Davuluri, A. S. T S, N. P.V., and M. S. Kumar, "Industrial wastewater treatment by electrocoagulation process," *Sep Sci Technol*, vol. 55, pp. 1–33, 2019, doi: 10.1080/01496395.2019.1671866.
[6] H. I. Eldos, M. Khan, N. Zouari, S. Saeed, and M. A. Al-Ghouti, "Characterization and assessment of process water from oil and gas production: A case study of process wastewater in Qatar," *Case Studies in Chemical and Environmental Engineering*, vol. 6, p. 100210, 2022, doi: https://doi.org/10.1016/j.cscee.2022.100210.
[7] P. N. Alam, Yulianis, H. L. Pasya, R. Aditya, I. N. Aslam, and K. Pontas, "Acid mine wastewater treatment using electrocoagulation method," *Mater Today Proc*, vol. 63, pp. S434–S437, 2022, doi: https://doi.org/10.1016/j.matpr.2022.04.089.
[8] P. V. Ngobeni, M. Basitere, and A. Thole, "Treatment of poultry slaughterhouse wastewater using electrocoagulation: a review," *Water Pract Technol*, vol. 17, no. 1, pp. 38–59, Nov. 2021, doi: 10.2166/wpt.2021.108.
[9] J. Akansha, P. V Nidheesh, A. Gopinath, K. V Anupama, and M. Suresh Kumar, "Treatment of dairy industry wastewater by combined aerated electrocoagulation and phytoremediation process," *Chemosphere*, vol. 253, p. 126652, 2020, doi: https://doi.org/10.1016/j.chemosphere.2020.126652.
[10] L. Bilińska, K. Blus, M. Gmurek, and S. Ledakowicz, "Coupling of electrocoagulation and ozone treatment for textile wastewater reuse," *Chemical Engineering Journal*, vol. 358, pp. 992–1001, 2019, doi: https://doi.org/10.1016/j.cej.2018.10.093.
[11] W. Zhang, M. Zhang, J. Yao, and J. Long, "Industrial indigo dyeing wastewater purification: Effective COD removal with peroxi-AC electrocoagulation system," *Arabian Journal of Chemistry*, vol. 16, no. 4, p. 104607, 2023, doi: https://doi.org/10.1016/j.arabjc.2023.104607.
[12] N. M. Genawi, M. H. Ibrahim, M. H. El-Naas, and A. E. Alshaik, "Chromium Removal from Tannery Wastewater by Electrocoagulation: Optimization and Sludge Characterization," *Water (Basel)*, vol. 12, no. 5, 2020, doi: 10.3390/w12051374.
[13] F. E. Titchou, H. Zazou, H. Afanga, J. El Gaayda, R. A. Akbour, and M. Hamdani, "Removal of Persistent Organic Pollutants (POPs) from water and wastewater by adsorption and electrocoagulation process," *Groundw Sustain Dev*, vol. 13, p. 100575, 2021, doi: https://doi.org/10.1016/j.gsd.2021.100575.
[14] A. Tahreen, M. S. Jami, and F. Ali, "Role of electrocoagulation in wastewater treatment: A developmental review," *Journal of Water Process Engineering*, vol. 37, p. 101440, 2020, doi: https://doi.org/10.1016/j.jwpe.2020.101440.
[15] S. Boinpally, A. Kolla, J. Kainthola, R. Kodali, and J. Vemuri, "A state-of-the-art review of the electrocoagulation technology for wastewater treatment," *Water Cycle*, vol. 4, pp. 26–36, 2023, doi: https://doi.org/10.1016/j.watcyc.2023.01.001.
[16] C. An, G. Huang, Y. Yao, and S. Zhao, "Emerging usage of electrocoagulation technology for oil removal from wastewater: A review," *Science of The Total Environment*, vol. 579, pp. 537–556, 2017, doi: https://doi.org/10.1016/j.scitotenv.2016.11.062.
[17] A. M. H. Elnenay, E. Nassef, G. F. Malash, and M. H. A. Magid, "Treatment of drilling fluids wastewater by electrocoagulation," *Egyptian Journal of Petroleum*, vol. 26, no. 1, pp. 203–208, 2017, doi: https://doi.org/10.1016/j.ejpe.2016.03.005.
[18] A. Shokri and M. S. Fard, "A critical review in electrocoagulation technology applied for oil removal in industrial wastewater," *Chemosphere*, vol. 288, p. 132355, 2022, doi: https://doi.org/10.1016/j.chemosphere.2021.132355.
[19] K. E. ADOU, A. R. KOUAKOU, A. D. EHOUMAN, R. D. TYAGI, P. DROGUI, and K. ADOUBY, "Coupling anaerobic digestion process and electrocoagulation using iron and

P-ISSN: 2723-3863
https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4847

aluminium electrodes for slaughterhouse wastewater treatment," *Sci Afr*, vol. 16, p. e01238, 2022, doi: https://doi.org/10.1016/j.sciaf.2022.e01238.

[20] N. S. Graça and A. E. Rodrigues, "The Combined Implementation of Electrocoagulation and Adsorption Processes for the Treatment of Wastewaters," *Clean Technologies*, vol. 4, no. 4, pp. 1020–1053, 2022, doi: 10.3390/cleantechnol4040063.

[21] I. Kabdaşlı, I. Arslan-Alaton, T. Ölmez-Hancı, and O. Tünay, "Electrocoagulation applications for industrial wastewaters: a critical review," *Environmental Technology Reviews*, vol. 1, no. 1, pp. 2–45, 2012, doi: 10.1080/21622515.2012.715390.

[22] S. Boinpally, A. Kolla, J. Kainthola, R. Kodali, and J. Vemuri, "A state-of-the-art review of the electrocoagulation technology for wastewater treatment," 2023. doi: 10.1016/j.watcyc.2023.01.001.

[23] W. Zhang, M. Zhang, J. Yao, and J. Long, "Industrial indigo dyeing wastewater purification: Effective COD removal with peroxi-AC electrocoagulation system," *Arabian Journal of Chemistry*, vol. 16, no. 4, 2023, doi: 10.1016/j.arabjc.2023.104607.

[24] F. Y. AlJaberi *et al.*, "Can electrocoagulation technology be integrated with wastewater treatment systems to improve treatment efficiency?," *Environ Res*, vol. 214, 2022, doi: 10.1016/j.envres.2022.113890.

[25] P. P. Das, M. Sharma, and M. K. Purkait, "Recent progress on electrocoagulation process for wastewater treatment: A review," *Sep Purif Technol*, vol. 292, p. 121058, 2022, doi: https://doi.org/10.1016/j.seppur.2022.121058.

[26] Y. Mao, Y. Zhao, and S. Cotterill, "Examining Current and Future Applications of Electrocoagulation in Wastewater Treatment," *Water (Basel)*, vol. 15, no. 8, 2023, doi: 10.3390/w15081455.

[27] M. Ebba, P. Asaithambi, and E. Alemayehu, "Investigation on operating parameters and cost using an electrocoagulation process for wastewater treatment," *Appl Water Sci*, vol. 11, no. 11, p. 175, 2021, doi: 10.1007/s13201-021-01517-y.

[28] M. Shen *et al.*, "Efficient removal of microplastics from wastewater by an electrocoagulation process," *Chemical Engineering Journal*, vol. 428, p. 131161, 2022, doi: https://doi.org/10.1016/j.cej.2021.131161.

[29] A. El Jery, H. M. Salman, N. Al-Ansari, S. S. Sammen, M. A. J. Maktoof, and H. A Z Al-Bonsrulah, "Optimization of oil industry wastewater treatment system and proposing empirical correlations for chemical oxygen demand removal using electrocoagulation and predicting the system's performance by artificial neural network.," *PeerJ*, vol. 11, p. e15852, 2023, doi: 10.7717/peerj.15852.

[30] N. A. Khan, S. Singh, P. C. Ramamurthy, and I. H. Aljundi, "Exploring nutrient removal mechanisms in column-type SBR with simultaneous nitrification and denitrification," *J Environ Manage*, vol. 349, p. 119485, 2024, doi: https://doi.org/10.1016/j.jenvman.2023.119485.

[31] Zakoor, Yousif, Gzar, Hatem, and Zakur, Yahya, "Municipal wastewater treatment and recycle by an electrocoagulation process and a submerged membrane bioreactor system," *E3S Web Conf.*, vol. 448, p. 3038, 2023, doi: 10.1051/e3sconf/202344803038.

Vol. 6, No. 4, Agustus 2025, Page. 2223-2238
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.4847