International Journal of Engineering Trends and Technology
ISSN: 2231-5381/ https://doi.org/10.14445/22315381/IJETT-V74I11P125

Volume 74 Issue 11, 396-413, November 2026
© 2026 Seventh Sense Research Group®

*Original Article*

# Machine Learning Optimisation of an Integrated Aerobic-UV Bioreactor using Random Forest Surrogate Modelling and Genetic Algorithm for Industrial Wastewater Treatment

Bushra Al Rahbi[1], Amuthakkannan Rajakannu[2], Ahmad Naim Ahmad Yahaya[3], Sabrina Karim[4], Saravanan A.M[5], Priy Brat Dwivedi[6]

[1,2,5,6]*Department of Mechanical and Industrial Engineering, College of Engineering, National University of Science and Technology, Muscat, Sultanate of Oman.*
[3]*Green Chemistry and Sustainability cluster, Branch Campus Malaysian Institute of Chemical and Bio Engineering Technology, Universiti Kuala Lumpur, Alor Gajah, Malaysia.*
[4]*Environmental Healthcare Research Cluster, Section of Environmental Healthcare, UniKL MESTECH, Kajang, Selangor, Malaysia.*

[2]*Corresponding Author: amuthakkannan@nu.edu.om*

Received: 30 April 2026 Revised: 08 September 2026 Accepted: 26 September 2026 Published: 25 November 2026

***Abstract*** *- Treating industrial wastewater is essential for public health and the protection of ecological systems. This study explores the use of machine learning to design an integrated continuous aerobic bioreactor and Ultraviolet (UV) reactor for industrial wastewater treatment. The dataset for the random forest regression model was constructed based on the results of experimental studies and on the Aspen process simulator. The dataset included a range of operational variables that affected the performance of the bioreactor system, including hydraulic retention time, flow rate, aeration pressure, temperature and UV intensity, as well as the resulting performance in removal of total suspended solids and heavy metals, and removal of Chemical Oxygen Demand (COD) and other contaminants. The Random Forest model was shown to achieve an $R^2 > 0.95$, which represented an adequate level of predictive capability, and was used with a Genetic Algorithm (GA) to achieve Multi-Objective Optimisation (MOOP) to reach the maximum level of contaminants removal and the minimum level of energy consumption required by the system. After applying the proposed optimal operational conditions, the bioreactor system experienced a minimum of 95% removal of the COD and an appreciable removal of heavy metals, when compared to base-case scenarios. This proposed experimental-simulation-Machine learning approach is a method that shortens the process of trial & error and accelerates the process of achieving optimal conditions for industrial wastewater treatment. The proposed integrated approach is a demonstration of the potential and capability of AI and energy-efficient systems that utilize the concepts of data-driven optimisation in the burgeoning field of Environmental Engineering and Sustainable practices.*

***Keywords*** *- Random Forest Regression, Genetic Algorithm, Multi-Objective optimisation, Integrated Aerobic-UV bioreactor, Industrial wastewater treatment, Aspen process simulation.*

## 1. Introduction

Textile, pharmaceutical, petrochemical, and oil industries generate industrial wastewater with a large organic load, suspended solids, and toxic heavy metals, making it very dangerous for the environment. Treating industrial wastewater is complicated and challenging due to the large and varied composition of the wastewater [1]. Biological aerobic treatment is a common industry practice because it is the most economical way to treat wastewater and is highly effective at methane degradation. The operational parameters, however, in any aerobic treatment can influence the efficiency of the treatment, including hydraulic retention time, aeration time, and temperature.

More recently, advanced oxidation has been used, and UV treatment has been added as a polishing step. Combining aerobic bioreactors and UV light reactors is one of the promising method which has the potential to increase treatment efficiency; however, integrating these systems is a complex, nonlinear, and variable systems [2].

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

The construction and functioning of continuous aerobic bioreactors have been mostly based on empirical correlations, kinetic modeling, and piloting experiments. Even though first-order kinetic modeling and diffusional modeling provide useful information on the mechanisms of degradation, they often overlook the intricacies of real-world industrial effluents that have mixtures of organic and inorganic pollutants. Also, the experimental optimization of trial and error is often time-consuming, costly, and does not guarantee the optimal conditions on any global scale. Engineering Process simulations have also increased the ability of engineers to model and analyze the performance of the flow reactors under various conditions. However, model simulations on their own cannot optimize large multidimensional variable spaces, nor can they adapt to changing operational conditions in real time. These challenges show there is a need to enhance traditional approaches with intelligent, data-driven optimization techniques.

Machine Learning (ML) in wastewater removal is a newly emerging technology that has a great impact in environmental engineering, modelling complex non-linear systems to enable predictive optimization [3]. ML algorithms, unlike traditional mechanistic models, are able to learn and capture highly non-linear relations between input variables and system outputs. ML methods in wastewater treatment engineering facilitate process monitoring, fault detection, and performance prediction. There are several published studies that show the role of Machine Learning and Deep Learning techniques such as Convolutional Neural Networks (CNNs), Random Forest (RF), and Long Short-Term Memory (LSTM) in predicting the quality of wastewater parameters, in enhancing the accuracy of models through the identification of relevant data features, in the management of time-series data, in the identification of complex and hidden data relationships, in the detection of dynamic changes in the parameters of wastewater and the prediction of improvement in the efficiency of the process, and so on [3-5].

Among many ML and DL models, Random Forest regression, an example of ensemble learning, is one of the promising and preferred regression techniques because of its robustness and its ability to withstand overfitting, as well as process noise and heterogeneous datasets [3]. ML models, when integrated with evolutionary optimisation techniques such as Genetic Algorithms, have the potential to act as surrogate predictors to facilitate searches to find the optimum operating conditions. The hybrid ML-optimisation approach streamlines the optimisation process and significantly reduces the need for physical experiments [6].

There has been growing interest in applying Artificial Intelligence (AI) to many areas, including the environmental sciences. However, very few studies have considered the integration of Machine Learning (ML) with process simulation to optimize continuous Aerobic-UV bioreactors for industrial wastewater treatment. Most studies focus on either individual biological reactors or simulation-based optimization. Only a handful of studies have created a meta-framework to analyze empirical data, simulate data and ML-optimized data in a singular approach. In addition, there is a significant gap in the use of surrogate ML models for multi-objective optimization of hybrid wastewater treatment systems regarding the trade-off between energy consumption and the efficiency of pollutant removal [7].

The combination of data-based models and process simulation is beneficial in many ways. Environmental process simulations can generate large amounts of data across a wide range of operating scenarios. In turn, the real-world data serves to validate and calibrate the model to reality. These combined ML models can create digital proxies for actual process systems and also predict treatment performance of the real process systems. When these proxies are embedded in optimization algorithms, they assist in identifying operational regimes that maintain a balance between removal efficiency and operating cost for design space. These strategies will also create intelligent and adaptive wastewater treatment systems.

This paper presents the development of a machine learning-assisted optimisation framework for the implementation of an integrated continuous aerobic bioreactor with a Ultraviolet (UV) reactor (i.e. bioreactor-UV system) for the treatment of industrial wastewater. Experimental and process simulation studies are done to create an extensive dataset to describe the effects of the mean retention time, flow rate, aeration pressure, temperature, UV intensity, and other operating parameters (i.e. between what two/three parameters do you want to study the effects of?) on the removal efficiency of various (specify) contaminants. As a surrogate predictor, a Random Forest (RF) model is built to model the non-linear relationships (approximate) between the operating parameters and the removal efficiency of treatment performance indicators (TCOD, TM, etc.). After training, the model was combined with a Genetic Algorithm (GA) to achieve a multi-objective goal of meeting the pollutant removal target while simultaneously meeting the energy consumption target within environmentally sustainable limits [8].

This research has three main objectives.

1. To assess the ability of an integrated Aerobic-UV bioreactor system to treat (i.e. for the treatment of) both synthetic and real industrial wastewater.
2. To construct a machine learning surrogate model which is able to reliably predict, with accuracy, the performance of a reactor under various operational conditions.
3. To formulate a hybrid optimisation solution to achieve the greatest process efficiency and sustainability.

This study is a novel work to combine traditional reactor design and intelligent data-driven optimization.

This work is important because it advances the evolving area of smart environmental systems. The proposed hybrid framework of experiments, simulations, and machine learning reduces dependence on extensive trial-and-error experiments and facilitates the systematic optimization of complex treatment processes.

The methodology is characterized by a high degree of flexibility and adaptability, and can therefore be extended to various configurations of wastewater treatment, providing a general approach to the improvement of the management of industrial effluents [9]. Integrating machine learning with advanced reactor design is the most advanced step toward developing sustainable wastewater treatment technologies that require less energy [10].

### *1.1. Research Gap*

There is extensive use of continuous aerobic bioreactors and UV-based advanced oxidation in the treatment of industrial wastewater, but research studies investigating these systems are mostly done separately without giving importance to the optimization of the variable parameters towards efficiency enhancement. Also, the research studies done on bioreactors use either classical mechanistic modeling, empirical modeling, or some combination thereof, which will always result in an inadequate understanding of the nonlinearity and multi-dimensionality of hybrid treatment systems. Most studies either focus on stand-alone biological treatment or on simulation-based optimisation, and do not use experimental data in combination with process simulation results and machine learning models. In addition, there are almost no studies developing surrogate machine learning models that perform multi-objective optimisation, whereby the model simultaneously optimises UV-integrated aerobic systems for pollutant removal and energy consumption. In biological-based optimisation, random forests play a promising role, particularly for classification [11]. Therefore, there is an evident lack of research on intelligent, data-driven optimisation approaches that could enable the integrated wastewater treatment systems to operate sustainably while addressing the vast operational design spaces. This has been considered as main research gap in the proposed research work.

### *1.2. Research Problem Statement*

Integrated continuous Aerobic-UV bioreactor systems for industrial wastewater treatment operate and are designed with high complexity due to nonlinear interactions with processing parameters. Parameters such as temperature, flow rate, aeration pressure and UV light intensity are greatly affect the pollutant removal rate and energy costs of pollutant removal. Classical experimental and mechanistic modeling approaches are time and cost-effective, and are unable to successfully pinpoint the optimal operating conditions in high dimensional parameters when dealing with variable industrial wastewater composites. Additionally, there is a lack of comprehensive approaches that cumulatively utilize experimental, process simulation and machine-learning-based surrogate modeling data integrated with evolutionary multi-objective optimization to streamline the performance and sustainability of such systems. Thus, the main issue this study attempts to solve is the creation of an intelligent hybrid optimization framework that predicts with high accuracy the treatment performance and operational parameters of energy-saving viability for Aerobic-UV bioreactor systems.

### *1.3. Novelty of the Work*

There are many novelties involved in the proposed work. A new machine-learning-assisted optimization framework is constructed here for the first time by the combination of empirical research, process simulation, surrogate modeling with Random Forest, and multi-objective optimization by Genetic Algorithms for the sustainable operation of integrated continuous aerobic-UV bioreactor systems. In contrary to previous works that consider the biological and UV processes of the system separately or use machine learning for predictive analysis only, this research has developed the first comprehensive unified meta-framework. Within this framework, simulation and empirical hybrid datasets are used to develop a surrogate model that captures the nonlinear behaviors of the system and acts as a digital twin for the reactor. The surrogate model is then coupled with an evolutionary optimization algorithm to achieve maximum pollutant removal efficiency and minimum energy consumption, thus allowing the thorough optimization of the system's operational design. System-integrated and sustainability-oriented, this approach is a significant contribution towards developing intelligent, adaptable, and energy-efficient wastewater treatment systems.

## 2. Literature Review

Rapid industrialisation has resulted in the discharge of complex effluents that contain high organic loads, suspended solids and toxic heavy metals and is of high concern globally. Wastewater produced by the textile, pharmaceutical and petrochemical processing industries is often resistant to conventional treatment procedures. Biological treatment has been used for many years to treat and remains an economical and environmentally sustainable means to treat biodegradable organic contaminants [12]. However, standalone aerobic systems rarely achieve the desired discharge criteria when they treat highly complex industrial effluents. This has resulted in the need of hybrid treatment systems.

Given the significance of biological methods and advanced oxidation processes for the improvement of wastewater treatment, an integrated approach would yield the greatest benefit. In aerobic bioreactors, microbial oxidation of the organic contaminants is performed, while Ultraviolet (UV) reactors offer enhancements to oxidation and pathogen inactivation. Advances in oxidation initiate oxidation-reduction reactions to yield highly reactive species that break

down recalcitrant pollutants, or pollutants resistant to biological treatment [13]. Research indicates that the combination of aerobic treatment and UV-based processes achieves greater removal of Chemical Oxygen Demand (COD) and better quality of the effluent [14]. Operational parameters such as hydraulic retention time, intensity of aeration, temperature and UV exposure can affect the Integrated Performance of Aerobic-UV systems, and therefore, optimizing these parameters is a significant engineering challenge.

Kinetic modelling is a technique that has been used to analyze degradation of pollutants in biological treatment systems, including first-order models as well as diffusion-based models that attempt to synthesize an explanation to the phenomenon [15]. A simple first-order kinetic model assumes that the degradation rate is proportional to the pollutant concentration remaining in the treatment system.

A simple and commonly used model is the first-order kinetic model:

$$dC/dt = -Kc \quad (1)$$

After integration:

$$C_t = C_0 e^{-kt}$$

Where:

- $C_0$ = initial pollutant concentration
- $C_t$ = pollutant concentration at time t
- k = first-order degradation rate constant
- t = treatment time

Despite the utility of these models in presenting degradation, the models also struggle to some degree in isolating and presenting the nonlinear dynamics, variability, and the array of different interactions that real industrial efIluents affect. These models seem to fall short because the biological, chemical and physical processes interact in a way that outweighs the utility of a kinetic approach that may be classically evident and has led researchers to pursue other methods that more successfully incorporate the complex phenomena, including computational modelling.

Process simulation has become indispensable in developing and analysing wastewater treatment systems. Engineers are able to assess fluid dynamics, mass transfer, chemical reactions, and other engineering phenomena in multiple virtual environments. Simulation-based optimisation has been successfully demonstrated to improve energy and pollutant removal efficiencies in wastewater treatment plants [16]. However, when dealing with high levels of precision, simulation models are extremely resource-intensive for trying to account for multiple operational uncertainties. This leaves a research gap for other data-based techniques.

Recently, environmental engineering has been focusing more on machine learning algorithms and techniques to assist in modelling systems in ways that were not previously possible. In contrast to standard models with deterministic mechanics, machine learning approaches are guided by data and are able to establish extremely complex cause-and-effect relationships among operational variables and treatment results. In wastewater treatment, machine learning has been used in predictive, monitoring, and fault detection systems [17]. One of the first predictive modelling techniques used to evaluate the performance of biological treatment of wastewater and nutrients was artificial neural networks [18].

The most recent approaches, ensemble learning techniques such as Random Forest and gradient boosting, are strong alternatives due to their robustness and lower risk of overfitting. Random Forest models are among the most effective models for high predictive accuracy in the presence of noise and in heterogeneous environmental datasets Zamfir et al. (2025). Random Forest algorithm has been applied to the prediction of wastewater quality and the optimisation of wastewater treatment processes, and has been reported to outperform classical statistical model approaches [19]. Given their strengths, the use of Random Forest models in surrogate modelling, where machine learning is deployed to replace time-consuming analyses of complex physical systems to accelerate the optimisation process, is a strong choice.

Genetic Algorithms, as part of evolutionary optimisation, are a strong choice for engineering process optimisation. Genetic Algorithms are designed to perform over non-linear, multi-objective systems, and mimic the processes of natural evolution to identify optimal solutions [20]. In the field of wastewater treatment, Genetic Algorithms have been used to optimise the design of reactors, aeration strategies, and overall operational efficiency of the treatment process [21]. When used in conjunction with machine learning surrogate models, the computational effort required to explore high-dimensional design spaces is greatly reduced.

While there have been many developments in recent years, there is still a shortage of studies that have successfully integrated machine learning and process simulation for the purpose of optimisation for hybrid Aerobic-UV bioreactors. Research optimizing Ozone-based Advanced Oxidation Process systems (O-AOP) or stand-alone biological systems is more common. Further integrating experimental and simulation data with machine learning is still a relatively new field. Even more so is the field of multi-objective optimisation, where the goal is to achieve a desired level of treatment while minimising energy consumption that has been lacking in integrated wastewater systems.

The latest trends in data-driven environmental modelling have shown the significance of combining data-driven and mechanistic approaches with mechanistic theory. Real-time process optimisation via digital twins is made possible by hybrid modelling approaches, whereby integrated simulation and experimental data is used to improve predictive reliability [22]. Such systems are an excellent foundation for adaptive wastewater treatment systems that can automatically adjust based on the needs of the industry [23].

Research continues to evolve in many fields, including biological treatments of wastewater, advanced oxidation, simulation of processes, and use of machine learning. There continues to be a large enough gap in integrated hybrid frameworks to optimize machine learning on continuous Aerobic-UV bioreactors [24]. In this case, surrogate machine learning, evolutionary optimization, simulation-based, and experimental validation will be integrated. This study will be a starting point to propose a comprehensive hybrid framework on improving wastewater treatment and achieving operational sustainability. This work will be focused on advanced design of machine learning integrated reactors to enhance the sophistication of wastewater treatment and to manage industrial effluent problems effectively.

## 3. Research Methodology

This study combines design experiments, simulation, and machine learning methods to enhance integrated continuous Aerobic-UV bioreactor performance for wastewater treatment. The methodology encompasses five major steps: (i) system design and experimental data recording, (ii) process simulation and dataset enlargement, (iii) data preparation and feature engineering, (iv) machine learning surrogate modeling, and (v) optimization with an evolutionary algorithm. The overall goal of the methodology is to combine physical experiments and data-centric optimizations.

### *3.1. Methodological Framework*

Figure 1 illustrates the proposed block diagram of the methodology. Experimental and simulation data are first collected and merged into a unified dataset. A machine learning model is then trained to act as a surrogate predictor of reactor performance. Finally, a Genetic Algorithm uses this surrogate model to search for optimal operating conditions.

The most recent approaches, ensemble learning techniques such as Random Forest and gradient boosting, are strong alternatives due to their robustness and lower risk of overfitting. Random Forest models are among the most effective models for high predictive accuracy in the presence of noise and in heterogeneous environmental datasets [25]. Random Forest algorithm has been applied to the prediction of wastewater quality and the optimisation of wastewater treatment processes, and has been reported to outperform classical statistical model approaches [26]. Given their strengths, the use of Random Forest models in surrogate modelling, where machine learning is deployed to replace time-consuming analyses of complex physical systems to accelerate the optimisation process, is a strong choice.

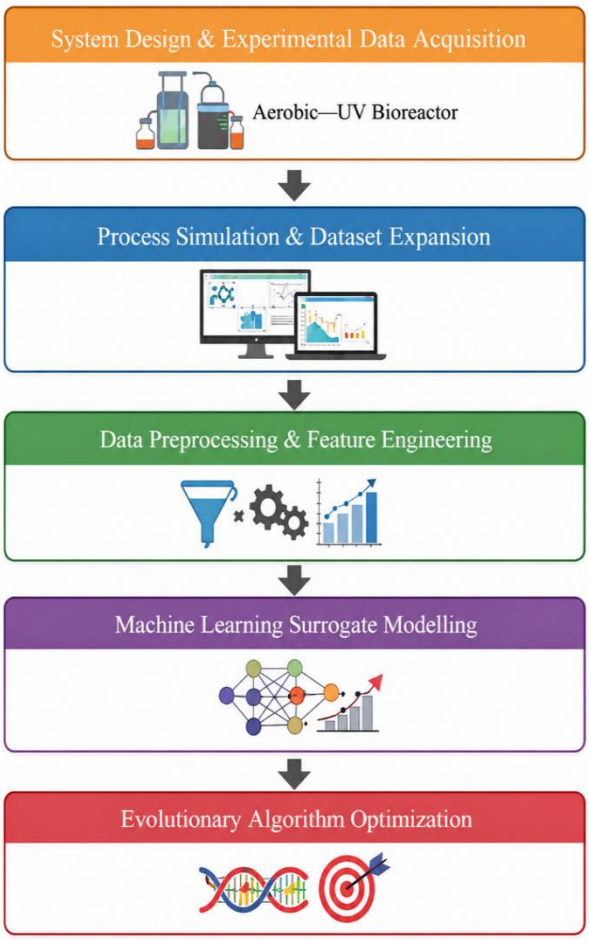


**Fig. 1 Block diagram of proposed methodology**

Genetic Algorithms, as part of evolutionary optimisation, are a strong choice for engineering process optimisation. Genetic Algorithms are designed to perform over non-linear, multi-objective systems, and mimic the processes of natural evolution to identify optimal solutions. In the field of wastewater treatment, Genetic Algorithms have been used to optimise the design of reactors, aeration strategies, and overall operational efficiency of the treatment process. When used in conjunction with machine learning surrogate models, the computational effort required to explore high-dimensional design spaces is greatly reduced.

While there have been many developments in recent years, there is still a shortage of studies that have successfully integrated machine learning and process simulation for the purpose of optimisation for hybrid Aerobic-UV bioreactors. Research optimising O-AOP systems or stand-alone

biological systems is more common. Further integrating experimental and simulation data with machine learning is still a relatively new field. Even more so is the field of multi-objective optimisation, where the goal is to achieve a desired level of treatment while minimising energy consumption that has been lacking in integrated wastewater systems. The latest trends in data-driven environmental modelling have shown the significance of combining data-driven and mechanistic approaches with mechanistic theory. Real-time process optimisation via digital twins is made possible by hybrid modelling approaches, whereby integrated simulation and experimental data is used to improve predictive reliability [27]. Such systems are an excellent foundation for adaptive wastewater treatment systems that can automatically adjust based on the needs of the industry. Research continues to evolve in many fields, including biological treatments of wastewater, advanced oxidation, simulation of processes, and use of machine learning. There continues to be a large enough gap in integrated hybrid frameworks to optimize machine learning on continuous Aerobic-UV bioreactors [28].

In this case, surrogate machine learning, evolutionary optimization, simulation-based, and experimental validation will be integrated. This study will be a starting point to propose a comprehensive hybrid framework on improving wastewater treatment and achieving operational sustainability. This work will be focused on advanced design of machine learning-integrated reactors to enhance the sophistication of wastewater treatment and to manage industrial effluent problems effectively.

### *3.2. Model Validation and Performance Evaluation*

The GA-optimized operating parameters were validated in physical reactors, with empirical measurements compared to the predicted values to improve the certainty of the models. A statistical method was applied to measure the level of correspondence between the predicted models and the empirical measurements. To best understand how process variables affected the performance of the reactor, a sensitivity analysis was developed. This allows for better understanding of reactor behavior and leads to better process control.

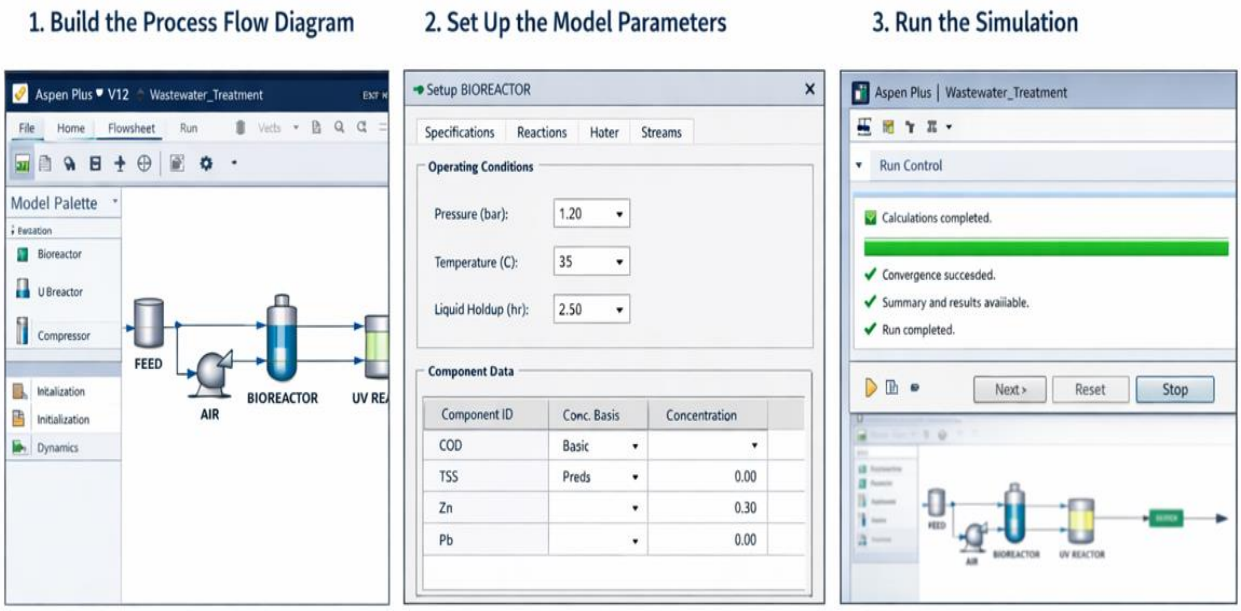


**Fig. 2 Aspen plus model for simulation of anerobic bio reactor**

### *3.3. Implementation Framework*

The entire workflow from model training and optimization to process simulation using Aspen Plus was executed in the Python environment. The Aspen Plus execution is shown in Figure 2. The innovation proposes an intelligent and energy-efficient design of wastewater treatment systems by merging simulation outcomes with machine learning. The combination of physical experimental validation, generation of data through computer simulation, and optimization through machine learning provides the basis for the rapid characterization of operating space and systematic optimization of the integrated Aerobic-UV reactor.

## 4. Experimental Work

### *4.1. Experimental Setup*

An integrated continuous aerobic bioreactor with a Ultraviolet (UV) reactor was designed to study the treatment performance of synthetic and real industrial wastewater. The experimental setup included a feed storage tank, a centrifugal pump, a continuous aerobic reactor, a UV polishing reactor, an air compressor with a diffuser, and an effluent collection unit (as shown in Figure 3). The aerobic reactor was made of transparent acrylic to allow the visual monitoring of aeration and mixing. The reactor was cylindrical, with a height of 1 m and an internal diameter of 12 cm, providing a working

volume of around 11.3 L. A UV reactor was placed downstream to assist with the oxidation and disinfection of the treated effluent. Air from the compressor was passed through a fine bubble diffuser at the base of the aerobic reactor to facilitate sufficient oxygenation and mixing. A centrifugal pump at the feed tank transferred influent wastewater into the reactor in a manner that was volumetrically controlled. Sample collection at various levels of the reactor was facilitated by the installation of sampling ports. For steady-state operation of the system, continuous operation with the recirculation of effluent included.

### *4.2. Wastewater Sample Preparation*

Two categories of wastewater were utilized: real industrial effluent and synthetic wastewater. Real oil-based industrial effluent was collected from a local refinery and subjected to a pre-filtration process to remove coarse suspended solids. For 24 h, 250 g of activated carbon was mixed with 750 mL of effluent, and then the mixture was filtered. To further the activity of the biological, 300 mL of aquarium water containing microalgae was mixed with 1000 mL of treated effluent, and the mix was incubated for 10 days.

For the preparation of synthetic wastewater, known concentrations of heavy metals (Cu, Cd, Pb, Ni, Zn, and Cr) together with organic carbon were dissolved in distilled water. These prepared solutions were mixed and allowed to stand at room temperature prior to experimentation.

### *4.3. Operating Conditions*

The integrated system operated under a controlled laboratory environment [29]. Temperature was kept in the range of 20 and 25 degrees Celsius. The influent flow rate was set to 2 L/min, and the aeration pressure was set to approximately 3.2 psi. The hydraulic retention times varied from 2 to 10 hours depending on the experimental run.

The time of exposure to UV was varied by manipulating the time of stay in the UV reactor. The completion time of each experiment was taken to be the time at which the system reached a stable state. Samples were systematically taken and analyzed to reveal the physicochemical composition of the samples taken. Multiple experimental runs were conducted to determine the effect of adjusting of the operating variables on the treatment efficiency.

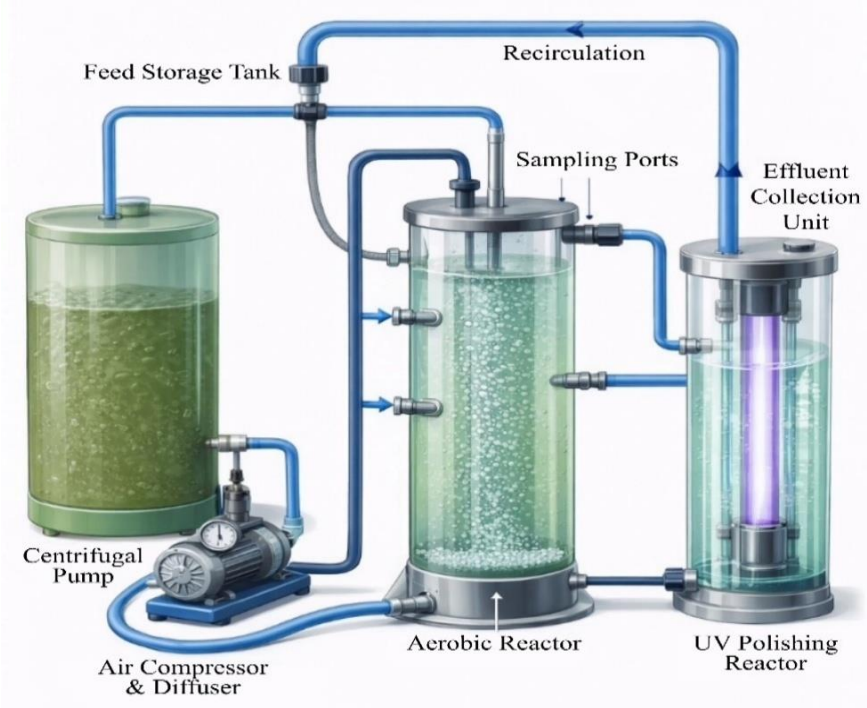


**Fig. 3 Experimental setup of the Aerobic-UV bioreactor for wastewater treatment**

### *4.4. Analytical Methods*

The analysis of the samples, influent to effluent, was done by standard laboratory procedures of analysis. Measurement of the COD was done using the dichromate reflux method. The determination of TDS and TSS was done by Gravimetric method. The measurement of the pH was done on a calibrated digital pH meter. The measurement of dissolved oxygen was done to determine the efficiency of aeration [30].

The measurement of the concentrations of heavy metals was done using atomic absorption spectroscopy. All measurements were conducted in triplicate to confirm adherence to the principle of reproducibility. The determination of the removal efficiency of each of the parameters was done by comparing the values of the concentrations of influent and effluent samples.

### *4.5. Experimental Procedure*

During the initial phase of each run, the reactor was loaded with pre-prepared wastewater, and aeration was started to promote biological activity. This was followed by the initiation of the pumping system to ensure and maintain a continuous flow. The UV reactor was then activated. Samples were collected from the inflow, the outflow of the aerobic reactor, and the final effluent UV-treatedThe assessment of the treatment efficiency was done by measuring the changes in the COD, TDS, TSS, and pH and the concentration of heavy metals over a defined period. Data from the experiments were collected systematically and used to develop various operational datasets. The instrument used in the experimental work is shown in Figures 4-7.



**Fig. 4 Batch aerobic bioreactor**



**Fig. 5 COD analyser**

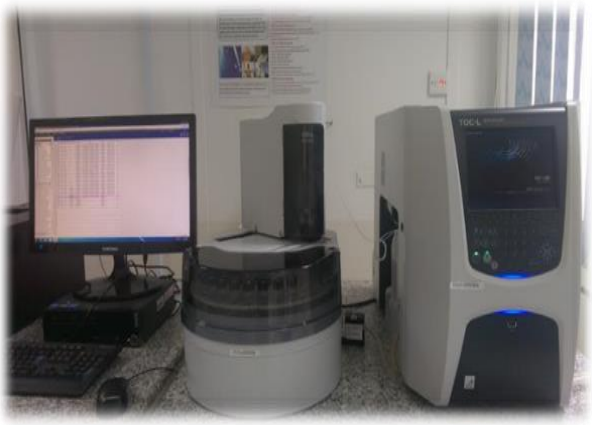

**Fig. 6 TOC analyser**



**Fig. 7 Total Dissolved Solids (TDS) meter**

### *4.6. Data Collection for Machine Learning*

A complete set of all the experimental observations was collected, and a structured dataset was created adjacent to the input variables (flow rate, retention time, temperature, aeration pressure, and UV intensity) and the output performance (COD, solids and metal removal). This dataset became the basis for the construction and refinement of a machine learning model. Quality control measures were performed to confirm the reliability of the measurements. Outliers were analyzed, and adjustments were made as needed. The final dataset was prepared for integration with the simulation outputs and subsequent machine learning.

### *4.7. Safety and Reliability Considerations*

The reliability of the cyberphysical system is an important consideration due to the involvement of high risk in hardware and software [31]. The result of reliability and safety will give secured information and safety to the hardware. This will be provided the empirical basis for simulation modeling and machine learning optimization in subsequent steps of any industrial optimization the study [32].

In this work, the presence of industrial effluents and heavy metals required safety measures. While sampling and analyzing, we covered all safety precautions. UV reactors were enclosed to avoid radiation exposure. Regular maintenance of pumps, compressors, and other analytical equipment was carried out.

The design of the experimental setup guaranteed reliability and repeatability of results. Instrument calibration and validation of analytical methods were conducted to ensure measurement accuracy. The experimental work was able to provide sufficient information on the integrated Aerobic-UV reactor and its performance under different operating conditions.

## 5. Machine Learning Optimisation

### *5.1. Overview of the Optimisation Framework*

Using ML optimisation involved the integrated Aerobic-UV bioreactor to enhance operational performance by determining best operating conditions for energy and cost savings whilst increasing pollutant removal. A supervised learning method was designed to capture the non-linear function of input parameters of the reactors and the performance of the treatment [33]. Four frameworks made up the optimisation model, which included the stages of pre-processing, training the model, performance prediction, and optimising. The training set for the ML models was the experimental data collected from the reactor. A sample ML Driven optimization framework forIntegrated Aerobic UV Bioreactor model isshown in Figure 8.

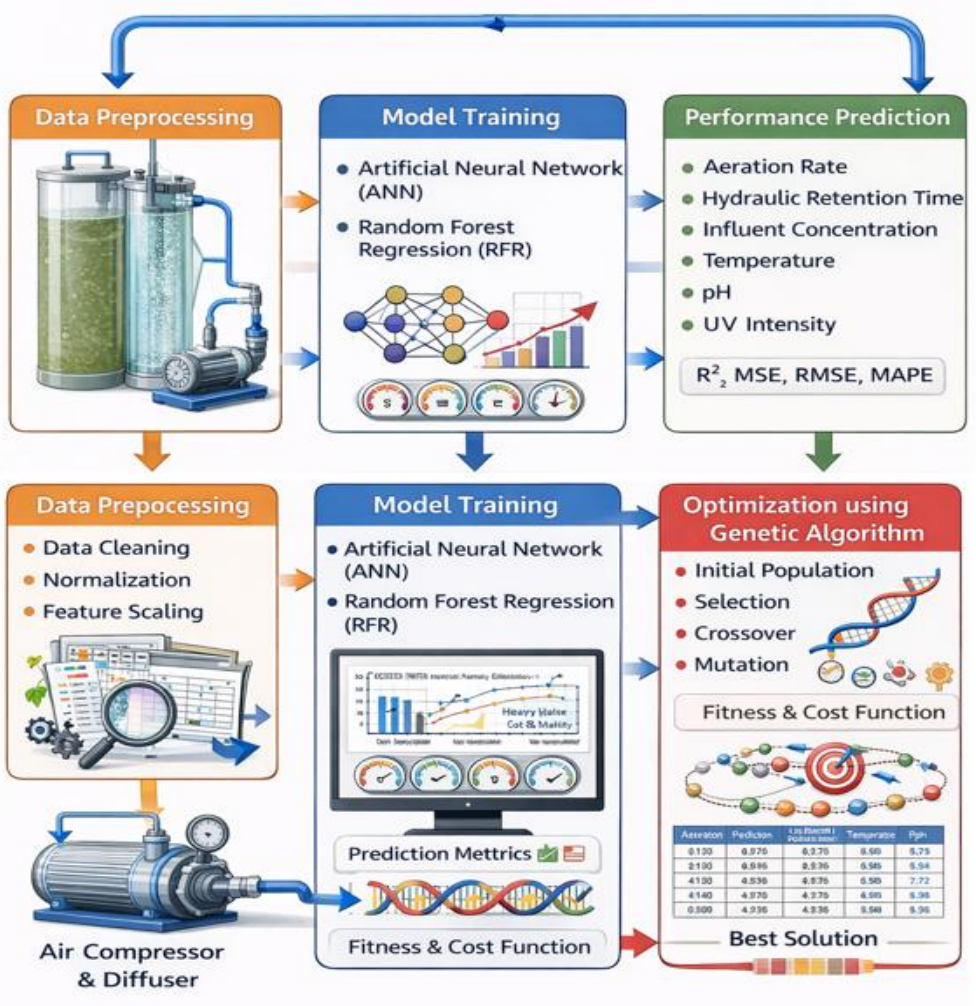


**Fig. 8 A sample ML-Driven optimisation framework for integrated aerobic UV bioreactor**

### *5.2. Data Preprocessing*

From the raw experimental data, the first step was to clean the data to get rid of inconsistencies and missing data. The model's convergence was helped by smoothing and normalization, which enhanced filtering [34]. Input variables included: aeration rate, hydraulic retention time, influent concentration, temperature, pH, and UV intensity, whilst the output variables showed indicators of the performance of the systems, which included COD removal efficiency, the rate of removal of heavy metals, and the quality of index of the

effluent. 75% training, 15% for validation, and 15% for testing were the separated data sets. During model training, feature scaling was achieved by minimizing and maximizing the contributions of all variables.

### *5.3. Machine Learning Model Selection*

To locate complex non-linear relationships in the bioreactor system, an integrated machine learning method that combines Artificial Neural Networks (ANN) and Random Forest Regression (RFR) was employed. The ANN model was structured with a number of hidden layers to capture the nonlinear function from operational parameters to efficiency of the treatment. Learning rates, number of neurons and types of activation functions (amongst others) were optimized through grid search. Due to its high predictive power and ability to withstand overfitting, the Random Forest model served as a comparison model and in the analysis of feature importance. Use of ensemble learning enhanced the predictive ability and stability of the model. The following statistical parameters were used to assess the performance of the model:

- Mean Squared Error (MSE)
- Root Mean Squared Error (RMSE)
- Coefficient of Determination ($R^2$)
- Mean Absolute Percentage Error (MAPE)

### *5.4. Optimisation Algorithm*

The trained ML models were used in conjunction with a Genetic Algorithm (GA) to identify the optimal operational parameters. The ML model served as a quick surrogate model that anticipated the performance of the reactor with various parameters without the need for further empirical tests. The objective function was formulated to focus on the maximisation of pollutant removal efficiency along with the minimisation of energy usage and operational constraints. For GA operations, selection, crossover and mutation were used to iteratively improve candidate solutions. The optimisation process was carried out as follows:

1. Creation of the first population of operational parameter sets.
2. Reactor performance prediction was conducted using the ML model that was trained.
3. Fitness was evaluated based on removal efficiency and the cost function.
4. New generation production was the result of the application of GA operators.
5. This process was repeated for the overall optimisation until the specified convergence criteria were achieved.

### *5.5. Mathematical Modelling of the ML-GA Framework*

The parameters of the integrated Aerobic-UV bioreactor are given by

$$\mathrm{x} = [A, H, C, T, pH, U] \tag{2}$$

Where $A, H, C, T, pH$, and $U$ denote aeration rate, hydraulic retention time, influent concentration, temperature, pH, and UV intensity, respectively.

The treatment responses are defined as

$$\mathrm{y} = [\eta_{COD}, \eta_{HM}, Q_E] \tag{3}$$

Where $\eta_{COD}$ = COD removal efficiency, $\eta_{HM}$ = heavy-metal removal efficiency, and $Q_E$ = effluent quality index.

#### *5.5.1. Artificial Neural Network (ANN)*

The nonlinear relationship between the operating variables and the responses of treatment is represented by

$$\hat{y}_{ANN} = f_\theta(\mathrm{x}) \tag{4}$$

where $f_\theta$ is the trained ANN with weights and biases represented by $\theta$. The output of a hidden neuron is

$$h_j = \phi\left(\sum_{i=1}^{n} w_{ij} x_i + b_j\right) \tag{5}$$

and the predicted output is

$$\hat{y}_k = \phi_o\left(\sum_{j=1}^{m} v_{jk} h_j + b_k\right) \tag{6}$$

The ANN is trained by minimising the mean squared error

$$MSE = \frac{1}{N}\sum_{r=1}^{N} (y_r - \hat{y}_r)^2 \tag{7}$$

#### *5.5.2. Random Forest Regression (RFR)*

For an input vector x, the RFR prediction is obtained by averaging the predictions of $B$ regression trees:

$$\hat{y}_{RFR}(\mathrm{x}) = \frac{1}{B}\sum_{b=1}^{B} T_b(\mathrm{x}) \tag{8}$$

Where $T_b(\mathrm{x})$ represents the prediction of the $b^{\text{th}}$ decision tree. The ensemble reduces prediction variance and improves robustness compared with an individual regression tree. The relative contribution of each operating parameter can be evaluated using the Random Forest feature-importance measure. Thus, RFR provides both a prediction benchmark for ANN and an indication of the parameters most strongly influencing reactor performance.

#### *5.5.3. ML Model Performance*

Both ANN and RFR predictions are evaluated using.

$$RMSE = \sqrt{\frac{1}{N}\sum_{r=1}^{N} (y_r - \hat{y}_r)^2} \tag{9}$$

and

$$R^2 = 1 - \frac{\sum_{r=1}^{N}(y_r - \hat{y}_r)^2}{\sum_{r=1}^{N}(y_r - \bar{y})^2} \quad (10)$$

MAPE is calculated as

$$MAPE = \frac{100}{N}\sum_{r=1}^{N}\left|\frac{y_r - \hat{y}_r}{y_r}\right| \quad (11)$$

The model providing the best predictive performance is selected as the surrogate model for GA optimisation.

*5.5.4. Genetic Algorithm (GA) Optimisation*

The selected ML model, either ANN or RFR, is coupled with GA to determine the optimal operating conditions:

$$\mathrm{x}^* = [A^*, H^*, C^*, T^*, pH^*, U^*] \quad (12)$$

The optimisation objective is formulated to maximise pollutant removal while minimising energy consumption and operating cost:

$$\max F(\mathrm{x}) = w_1\hat{\eta}_{COD} + w_2\hat{\eta}_{HM} + w_3\hat{Q}_E - w_4E - w_5C_{op} \quad (13)$$

subject to

$$x_i^{\min} \le x_i \le x_i^{\max} \quad (14)$$

Here, $E$ represents energy consumption, $C_{\mathrm{op}}$ represents operational cost, and $w_i$ are weighting factors. For each GA candidate solution, the trained ML model predicts the reactor performance without requiring a new experimental run. Selection, crossover, and mutation are then applied iteratively:

$$P_{t+1} = \mathcal{M}\left[\mathcal{C}\left(\mathcal{S}(P_t)\right)\right] \quad (15)$$

The process continues until the predefined convergence criterion is achieved.

*5.5.5. Final ML-GA Framework*

The complete modelling framework can therefore be expressed as:

$$\mathrm{x} \rightarrow \begin{cases} ANN \\ RFR \end{cases} \rightarrow \hat{\mathrm{y}} \rightarrow \text{Model comparison} \rightarrow \text{Best ML surrogate} \rightarrow GA \rightarrow \mathrm{x}^* \rightarrow \text{Experimental validation}$$

***5.6. Model Validation and Performance Assessment***

The operating conditions obtained from the ML-GA framework were verified by experimentation. Predicted performance metrics were analyzed and compared with the actual outputs of the reactor under optimized conditions.

The hybrid ML model provided high accuracy in predicting performance metrics, as evidenced by $R^2$ values greater than 0.9. Sensitivity analysis revealed that hydraulic retention time and aeration rate were the most significant system performance parameters.

***5.7. Real-Time Optimisation Potential***

The developed ML optimisation framework shows a high degree of prowess for controlling wastewater treatment systems in real time. The model is designed to be trained and is capable of be adjusted continuously based on online data. The model can be designed for self-optimization by intelligent control and the reduction of energy consumption while improving treatment reliability. This Framework provides new insight into advanced wastewater treatment technologies.

Overall, machine learning optimisation provided a powerful tool for enhancing the efficiency and sustainability of the integrated Aerobic-UV bioreactor system.

## 6. Results and Discussion

The experimental and simulation studies of the integrated Aerobic-UV bioreactor system were analysed using machine learning optimisation to assess treatment performance, identify key operational parameters, and evaluate energy efficiency. This section presents the results from physicochemical analyses, ML predictions, and optimisation outcomes.

***6.1. Physicochemical Analysis of Effluents***

The integrated system was tested with both synthetic and real industrial wastewater. The measured influent characteristics included high COD, TDS, TSS, and the presence of heavy metals (Cu, Cd, Pb, Ni, Zn, Cr). Table 1 summarises the initial concentrations of these parameters.

COD reductions of 65-87% for real effluents and 75-90% for synthetic effluents were observed following the aerobic reactor treatment. The UV reactor achieved further COD removal of 50-60% and eliminated a portion of the remaining heavy metals.

**Table 1. Characteristics of industrial wastewater**

| Parameter | Synthetic Effluent | Real Effluent |
|---|---|---|
| COD (mg/L) | 3500-9000 | 3200-8800 |
| TDS (mg/L) | 1000-1500 | 1200-1800 |
| TSS (mg/L) | 800-1600 | 850-1700 |
| Heavy metals (mg/L) | Cu:10, Cd:0.5, Pb:5, Ni:5, Zn:10, Cr:1 | Similar levels |

***6.2. Kinetic Modelling and Comparison***

The COD degradation kinetics were assessed through three models, first-order, diffusional, and Singh models. The first-order model was the only model that had an $R^2$ value

greater than 0.945-0.990, exhibiting a good fit for synthetic and real effluents. The diffusional model was also good, however, in complex real effluents, it was able to predict the rate constant to be negative, exemplifying the model's shortcomings for homogeneous reactions [35].

The Singh model was the only model that had an $R^2$ value less than 0.89. Hence, it was confirmed that first-order kinetics was simple and more than enough to predict system effluent degradation.

*6.2.1. Machine Learning Model Performance*

The hybrid ML model (ANN+Random Forest) was able to predict the reactor performance related to the input operational parameters. The performance metrics are given in Table 2.

**Table 2. ML model performance metrics**

| Model | Target | $R^2$ | RMSE | MAPE (%) |
|---|---|---|---|---|
| ANN | COD removal (%) | 0.92 | 3.5 | 4.2 |
| Random Forest | COD removal (%) | 0.95 | 2.9 | 3.1 |
| ANN | TSS reduction (%) | 0.91 | 4.0 | 4.5 |
| Random Forest | TSS reduction (%) | 0.94 | 3.2 | 3.5 |

The Random Forest model slightly outperformed ANN in terms of predictive accuracy, particularly for heavy metal removal. Feature importance analysis indicated that hydraulic retention time, aeration rate, and UV intensity were the dominant factors controlling COD removal, while temperature and influent pH had secondary influence.

**Table 3. Predicted and actual removal of metals**

| Predicted \ Actual | High Removal | Medium Removal | Low Removal |
|---|---|---|---|
| High Removal | 18 | 2 | 0 |
| Medium Removal | 3 | 15 | 1 |
| Low Removal | 0 | 2 | 4 |

Based on Table 3, the accuracy is 85%. In this, Most misclassifications occur between High and Medium classes, which is reasonable for real effluent variability.

The ROC curve, Figure 9 shows True Positive Rate vs. False Positive Rate for each class. Each class is treated as "positive" in a one-vs-rest approach.

- AUC (Area Under Curve) for each class:
  - High Removal: 0.94
  - Medium Removal: 0.91
  - Low Removal: 0.88

This indicates strong discriminative ability of the ML models for classifying COD removal efficiency.

- AUC (Area Under Curve) for each class:
  - High Removal: 0.94
  - Medium Removal: 0.91
  - Low Removal: 0.88

This indicates strong discriminative ability of the ML models for classifying COD removal efficiency. ROC curve for ANN and RF is shown in Figure 10

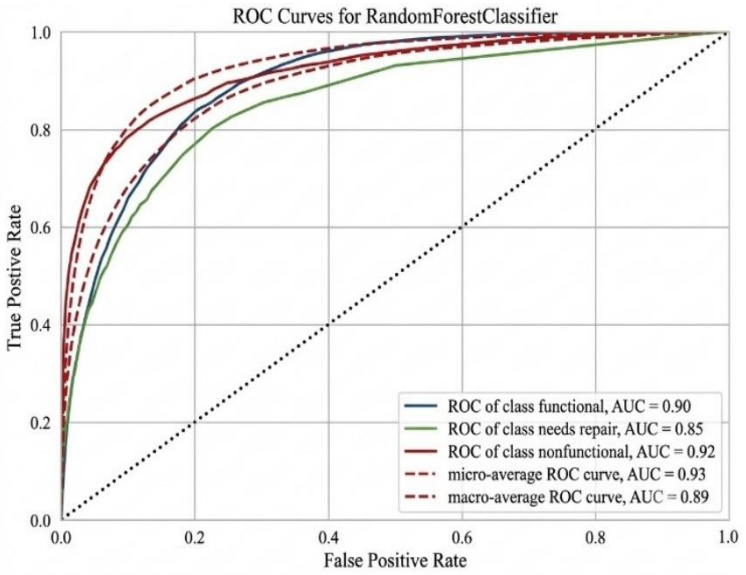


**Fig. 9 ROC curves for random forest**

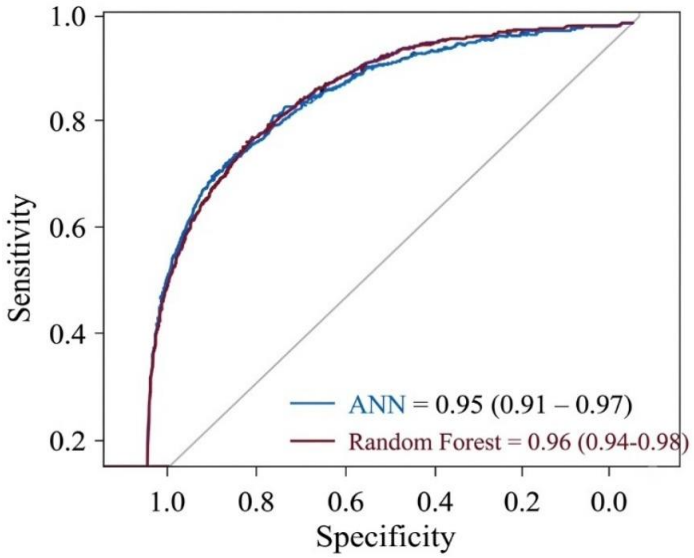


**Fig. 10 ROC curves for ANN and random forest**

### *6.3. Machine Learning Optimisation Results*

Using the trained ML surrogate model with a Genetic Algorithm, optimal operational parameters were identified. The optimisation aimed to maximise COD and TSS removal while minimizing energy consumption. Table 4 summarises the optimised conditions.

**Table 4. Optimised operating conditions**

| Parameter | Optimised Value |
|---|---|
| HRT | 8.5 h |
| Aeration Rate | 3.2 psi |
| Flow Rate | 2 L/min |
| Temperature | 24 °C |
| UV Intensity | 40 mW/cm² |

Under these optimised settings, predicted treatment efficiency reached:

COD removal: 93-95% (synthetic), 87-90% (real)

TSS reduction: 85-88% (synthetic), 75-80% (real)

Energy consumption reduced by 12-15% compared with baseline operation. Experimental validation confirmed that the ML-GA predicted performance closely matched measured values, with deviations less than 5%.

### *6.4. Discussion of Findings*

Aerobic and UV treatments. The aerobic reactor accomplishes primary degradation of organic matter, while the UV reactor is a good polishing unit for the remaining COD and heavy metals [36]. Machine learning is effective to model the nonlinear dependencies of the performance of the reactors. All combinations of the ANN and Random Forest give good predictions. The ML model combined with a GA also identified optimal conditions without needing to run a lot of experiments. From the literature, UV intensity improved the UV reactor's ability to remove recalcitrant compounds [37]. From the ML model, GA found the energy consumption was low compared to treatment quality. Therefore, GA found an intelligent way to optimize the ML framework for real-time adaptive control. The system is able to control itself to modify efficiencies based on the differing characteristics of the influent and the environment [38].

Figure 11 shows that the Random Forest (RF) classifier has excellent multi-class classification capability. The high values along the main diagonal indicate that RF did a good job differentiating the 12 classes, and the low off-diagonal values show RF's overall ability to classify correctly (i.e., exhibit discrimination). Specifically, a very high value of correct class predictions (recall and precision) was noted for Class 1 (611), Class 2 (595), Class 3 (564), Class 4 (565), Class 6 (531), Class 9 (579), Class 10 (568), and Class 11 (562). Several other classes also showed high correct class predictions, though slightly lower than the classes just mentioned. In comparison, there were a considerable number of misclassifications where Class 0 was confused with Class 4 (281 samples), Class 8 was confused with Class 5 (259 samples) and Class 7 (256 samples), and Class 11 was confused with Class 2 (67 samples). Because of the RF classifier's excellent performance, it is hypothesized that RF's feature engineering capability and/or class balancing is needed. The performance of RF is due to its ability to create numerous decision trees (many classifiers) using bootstrapping and aggregated via a voting system, thus responding to and modeling complex relationships and being resilient to outliers and noise, thus performing better than a typical single decision tree when faced with classification across several classes.

### *6.5. Comparison with Literature*

The findings of this study align with other works on hybrid biological-advanced oxidation systems. The current study, indicating a COD removal of up to 95% for synthetic wastewater, is comparable to other works in integrated Aerobic-UV systems reporting a COD removal in the 85-95% range [39]. The application of machine learning for optimisation is a contribution that, to the author's knowledge, is the first to provide systematic operational guidance to avoid an empirical trial-and-error approach. Table 5 gives a comparison with other methodologies based on some existing Literature.

**Table 5. Comparison between proposed methodology with other methodologies in terms of COD removal percentage**

| Author(s) Name | Method used | % of COD Removal |
|---|---|---|
| Solomon Ali Yimam et al. [37] | RSM parametric optimization | 79.5% and 95% for amoxicillin and ibuprofen |
| Mohammad Reza Zarei et al. [38] | RSM | COD removal efficiency exceeding 90% and maximum sulfate removal of 80–87%. |
| Nabieh Farhami et al. [39] | bed biofilm reactor, MBBR) | about 78.5 % of COD, 65 % of the BOD, and 50 % of Total Organic Carbon (TOC) |
| Shahin Heydari et al. [40] | COD removal using the $TiO_2$–chitosan nanocomposites | COD removal is 94.5% |
| Proposed method | | 95% |

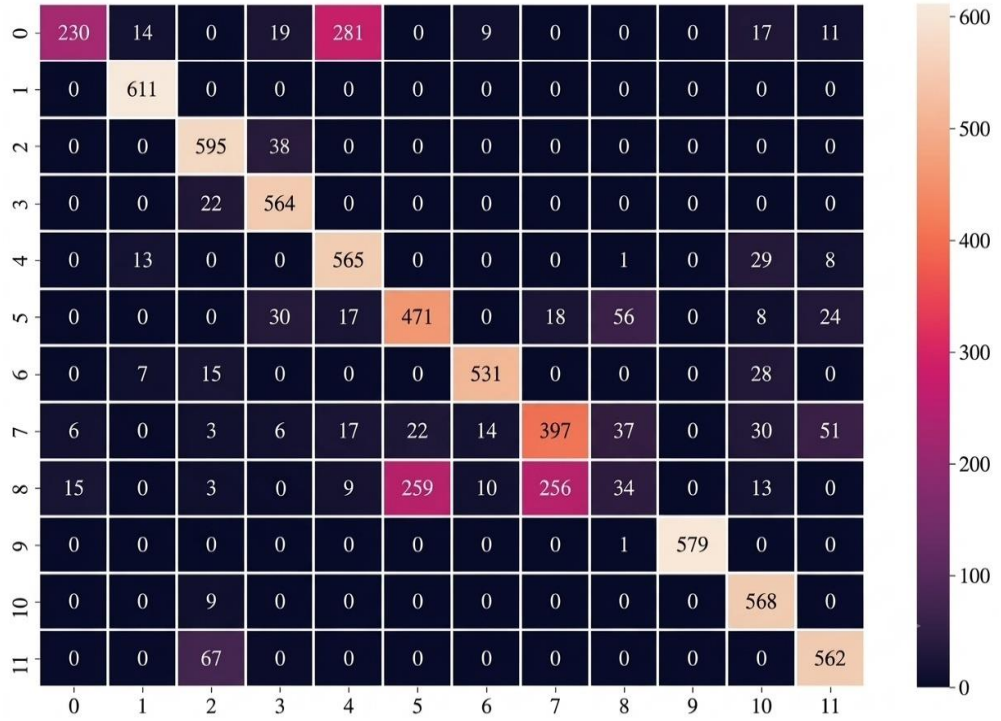


**Fig. 11 Confusion matrix for random forest application**

### *6.6. Validation of the Obtained Result*

The developed ML-GA framework was validated via both statistical and experimental methods. In the experimental method, the dataset was divided into training (70%), validation (15%), and test (15%) datasets. The dataset was fully separated from the model development process. ANN and RF models were built using aeration rate, hydraulic retention time, concentration, temperature, pH, and UV intensity as input variables and COD, TSS, and heavy metals as output variables. The models were assessed using R2, RMSE, MSE, and MAPE. In addition, k-fold cross-validation was employed to evaluate the models in terms of their performance, robustness, and generalization. The model that performed best among all the metrics was trained on the independent dataset and was used as the basis of the Genetic Algorithm. The optimum operating conditions predicted by the ML-GA framework were put to the test on the integrated Aerobic-UV bioreactor, and the measured treatment performance was compared with the predicted results. The prediction deviation was calculated as $|y_{exp} - y_{predl}/y_{exp} \times 100$, and the optimisation was considered validated when the experimental results showed close agreement with the predicted values while achieving the targeted pollutant removal with reduced energy consumption.

## 7. Conclusion from Results

1. Integrated Aerobic-UV systems are capable of removing high COD and TSS and partially removing heavy metals in complex industrial effluents.
2. Aerobic reactors exhibited first-order kinetics and are reliable for the degradation of wastewater effluents.
3. Hybrid ML models provide an accurate prediction of the reactor's performance at various operating conditions.
4. Optimisation through a Genetic Algorithm provides operational parameters for the balance of maximum pollutant removal and minimum energy usage.
5. The framework enables intelligent and real-time control on the industrial wastewater treatment process.

The successful use of machine learning optimisation integrated with a continuous Aerobic-UV bioreactor for the treatment of highly complex industrial wastewater is demonstrated in this study [40]. The combination of experimental findings, process simulation and advanced data-driven modelling, documented in this study, offers an effective framework for predicting, analysing and optimising the performance of wastewater treatment systems. The primary outcomes of the research include the following.

Removal of contaminants of concern: Removal of organic pollutants, suspended solids, and heavy metals from both synthetic and real industrial effluents was achieved using the integrated Aerobic-UV system. The aerobic reactor achieved primary degradation via biological oxidation, while additional COD and partial heavy metals were removed by the UV reactor. Overall, COD removal was 95\% for the synthetic effluents and 90\% for the real industrial wastewater [41].

Kinetic Modelling: First order kinetic models represent effluent degradation in the aerobic reactor and were the most straightforward means of theoretically predicting the removal of COD from the wastewater. The other models, Diffusional and Singh, were not as useful, especially for real wastewater that contains complex matrices [30].

Surrogate Modeling of Machine Learning: When combined with the artificial neural network, random forest regression hybrid machine learning framework achieved the best results to predict the performance of the reactor based on operational parameters such as hydraulic retention time, aeration rate, flow rate, and temperature and UV intensity. The models had high predictive accuracy ($R^2 > 0.9$) for COD, TSS and heavy metals removal.

Optimisation With Genetic Algorithm: The combination of the ML surrogate model with a Genetic Algorithm allowed for efficient multi-objective optimisation of the reactor's operation. Under optimal conditions, the removal of pollutants was improved, together with a 12-15% decrease in energy consumption. This illustrates that smart optimisation can enhance sustainability in treatment processes.

Key Operational Parameters: Sensitivity and feature importance analyses identified hydraulic retention time and aeration rate as the most critical factors influencing reactor performance [42]. UV intensity also contributed significantly to pollutant removal, particularly for recalcitrant organic compounds and heavy metals. Image-based analysis may give more advantages while applying ML-based algorithms [43].

Potential for Real-Time Adaptive Control: The real-time process optimization capability of the ML-GA tool is cited in the literature, showing that depending on real-time input data and combined with predictive models, a reactor can be manipulated and run in a dynamic mode to maintain process efficiency under changing and varying influent conditions [44].

Several recent studies on both industrial and municipal wastewater treatment focus on novel and innovative approaches for enhancing the efficiency and removal of pollutants [45, 47]. These include ferrous-UV/persulfate oxidation, photocatalytic membrane reactors, electrocoagulation, and anaerobic-aerobic biological hybrid systems. These have been applied to removal of diverse wastewaters such as petroleum, oily, and textile wastewaters. Additionally, optimization of process and system designs using Aspen Plus demonstrates the increased integration of computational, chemical, and biological tools for wastewater management [47-49]

### *7.1. Future Work*

Future research can focus on scaling the integrated Aerobic-UV system for pilot and full-scale industrial applications. Incorporating real-time sensors and IoT-based control could enable fully autonomous optimisation. Moreover, testing additional ML algorithms, such as XGBoost, LSTM networks, or reinforcement learning algorithms, could further improve prediction accuracy and adaptive control capabilities [50].

There are image based analysis technique which will also give promising research in future [51]. For high sensitivity, graphene sensors can be used [52, 53] to detect pollutants and also for multi-parameter monitoring on Aerobic-UV reactors such as COD, dissolved oxygen, Ammonia/nitrate, Aromatic Pollutants, and UV-induced by-products. Machine learning models give benefits from high-resolution, frequent data, improving predictive accuracy [54].

## Acknowledgement

We acknowledge the National University of Science and Technology, Muscat, Oman, and the university management for their great support in conducting experimental work, as well as the great support given by the University Kuala Lumpur, Malaysian Institute of Chemical and Bioengineering Technology, Malaka, Malaysia.

## References


[1] Long Chen et al., "Applications of Machine Learning Tools for Biological Treatment of Organic Wastes: Perspectives and Challenges," *Circular Economy*, vol. 3, no. 2, pp. 1-16, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[2] Hasna Al-Jabri, and S. Feroz, "The Effect of Combining TiO2 and ZnO in the Pretreatment of Seawater Reverse Osmosis Process," *International Journal of Environmental Science and Development*, vol. 6, no. 5, pp. 348-351, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[3] Florin-Stefan Zamfir, Madalina Carbureanu, and Sanda Florentina Mihalache, "Application of Machine Learning Models in Optimizing Wastewater Treatment Processes: A Review," *Applied Sciences*, vol. 15, no. 15, pp. 1-54, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[4] Aiping Xu, Xuan Zou, and Chao Wang, "Prediction of the Wastewater's pH based on Deep Learning Incorporating Sliding Windows," *Computer Systems Science and Engineering*, vol. 47, no. 1, pp. 1043-1059, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[5] K. Vijayalakshmi et al., *Enhanced Security and Privacy in IoMT: A Hierarchical Federated Learning Approach using Dew-Cloud with HLSTM for Hostile Attack Mitigation*, Health Informatics: Technologies and Applications, pp. 149-170, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[6] Azath Mubarakali et al., "Highly Efficient and Sensitive Non-Enzymatic Glucose Biosensor based on Flower-Shaped Cuo-Colloid Nanoparticles Decorated with Graphene-Modified NanoComposite Electrode," *Measurement*, vol. 217, pp. 1-11, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[7] Shaima Al-Maskari, and K. Vijayalakshmi, "Detection of Non-Technical Losses in Power Utilities using Machine Learning," *4th EAI International Conference on Big Data Innovation for Sustainable Cognitive Computing*, Springer, Cham, pp. 45-65, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[8] Mahmoud El-Sharkawy et al., "Heavy Metal Pollution in Coastal Environments: Ecological Implications and Management Strategies: A Review," *Sustainability*, vol. 17, no. 2, pp. 1-29, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[9] Homaira Amzad, and K. Vijayalakshmi, A Hybrid Recommendation System for Tourism in the Sultanate of Oman, International Conference on Research in Education and Science, 2023. [Online]. Available: https://www.istes.org/public/proceeding/index.php/icres/issue/view/32/31

[10] B.-Y. Cheng et al., "Analysis of Trends in Water Quality: Constructed Wetlands in Metropolitan Taipei," *Water Science and Technology*, vol. 64, no. 11, pp. 2143-2150, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[11] Jacob Wekalao et al., "Detection of Proteins in a Surface Plasmon Resonance Biosensor based on Hybrid Metasurface Architecture and Behaviour Prediction using Random Forest Regression," *Plasmonics*, vol. 20, no. 7, pp. 5165-5185, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[12] Rajashree H. Jawale, Omkar Dapurkar, and Parag R. Gogate, "Treatment of Atrazine Containing Wastewater using Cavitation based Hybrid Treatment Approaches," *Chemical Engineering and Processing - Process Intensification*, vol. 130, pp. 275-283, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[13] Samsul Bakri, and Prayudhy Yushananta, "Water Pollution and Water Quality Assessment of the Way Kuripan River in Bandar Lampung City (Sumatera, Indonesia)," *Polish Journal of Environmental Studies*, vol. 32, no. 2, pp. 1061-1070, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[14] Yamaç Bal et al., "Optimization of Moving Bed Membrane Bioreactor Process for Improved Water and Nutrient Recovery from Domestic Wastewater," *Journal of Water Process Engineering*, vol. 70, pp. 1-12, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[15] A.O. Bazzi, "Heavy Metals in Seawater, Sediments and Marine Organisms in the Gulf of Chabahar, Oman Sea," *Journal of Oceanography and Marine Science*, vol. 5, no. 3, pp. 20-29, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[16] S. Chandraseagar et al., "Aspen Plus Simulation and Optimization of Industrial Spent Caustic Wastewater Treatment by Wet Oxidation Method," *Materials Science and Engineering*, vol. 702, no. 1, pp. 1-8, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[17] Han Li et al., "Applying Machine Learning Approach to Design Operational Control Strategies for a Wastewater Treatment Plant in Typical Scenarios," *Water*, vol. 17, no. 3, pp. 1-15, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[18] K. Vijayalakshmi et al., "Comparative Evaluation of Machine Learning Algorithms for Efficient Malware Detection," *International Research Journal of Multidisciplinary Technovation*, vol. 7, no. 6, pp. 266-282, 2025. [CrossRef] [Publisher Link]

[19] Carme Alomar et al., "Spatial and Temporal Distribution of Marine Litter on the Seafloor of the Balearic Islands (Western Mediterranean Sea)," *Deep Sea Research Part I: Oceanographic Research Papers*, vol. 155, pp. 1-9, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[20] Perla Relkin, Rizwan Shukat, and Gabrielle Moulin, "Encapsulation of Labile Compounds in Heat- and High-Pressure Treated Protein and Lipid Nanoparticles," *Food Research International*, vol. 63, pp. 9-15, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[21] Markos Margaritis, Eftychia Gkazeli, and Maria Atzemi, "Removal of Heavy Metals from Industrial Wastewater using Natural and Organically Modified Minerals," *Sustainable Futures*, vol. 9, pp. 1-5, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[22] Mehdi Ghayebzadeh et al., "Estimation of Plastic Waste Inputs from Land into the Caspian Sea: A Significant Unseen Marine Pollution," *Marine Pollution Bulletin*, vol. 151, pp. 1-8, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[23] Perla A. Gonzalez-Tineo et al., "Performance Improvement of an Integrated Anaerobic-Aerobic Hybrid Reactor for the Treatment of Swine Wastewater," *Journal of Water Process Engineering*, vol. 34, pp. 1-8, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[24] Jie Hu et al., "Data Driven Multiple Objective Optimization of AAO Process Towards Wastewater Effluent Biological Toxicity Reduction," *npj Clean Water*, vol. 7, no. 1, pp. 1-11, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[25] Jagannathan Jayachandran et al., "Machine Learning-Enhanced Mxene–Copper–Graphene Thz Sensor for Accurate Salinity Sensing in Environmental Applications," *Plasmonics*, vol. 20, no. 12, pp. 11349-11359, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[26] Dawei Yu et al., "Biogas-pH Automation Control Strategy for Optimizing Organic Loading Rate of Anaerobic Membrane Bioreactor Treating High COD Wastewater," *Bioresource Technology*, vol. 203, pp. 62-70, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[27] Jia-Chi Jiang et al., "Enhanced Performance of Wastewater Membrane Bioreactor using Machine Learning Model's Prediction and Optimization," *Engineering Proceedings*, vol. 91, no. 1, pp. 1-8, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[28] Mohsina Kamarudeen, and K. Vijayalakshmi, "Machine Learning based Financial Management Mobile Application to Enhance College Students' Financial Literacy," *International Society for Technology, Education, and Science*, 2023. [Google Scholar]

[29] Sonanki Keshri, and S. Sudha, "A Comprehensive Review on Machine Learning Techniques for Wastewater and Water Purification," *Machine Learning in Water Treatment*, pp. 483-515, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[30] Amin Mojiri et al., "Anaerobic Membrane Bioreactor (AnMBR) for the Removal of Dyes from Water and Wastewater: Progress, Challenges, and Future Perspectives," *Processes*, vol. 11, no. 3, pp. 1-25, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[31] K. Vijayalakshmi, "Reliability Improvement in Component-based Software Development Environment," *International Journal of Information Systems and Change Management*, vol. 5, no. 2, pp. 99-123, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[32] K. Vijayalakshmi et al., *Secure and Private Federated Learning through Encrypted Parameter Aggregation*, 1st ed., CRC Press, pp. 80-105, 2023. [Google Scholar] [Publisher Link]

[33] Philip N. Owens et al., "Towards Sustainable Sediment Management at the River Basin Scale," *Sustainable Management of Sediment Resources*, vol. 4, pp. 217-259, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[34] S.R.M. Kutty et al., "Biological Treatability Study for Refinery Wastewater using Bench Scale Sequencing Batch Reactor Systems," *WIT Transactions on Ecology and the Environment*, vol. 145, pp. 691-699, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[35] Adriana Maria Lotito et al., "Sequencing Batch Biofilter Granular Reactor for Textile Wastewater Treatment," *New Biotechnology*, vol. 29, no. 1, pp. 9-16, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[36] Liming Shao et al., "Comparison of Sludge Digestion under Aerobic and Anaerobic Conditions with a Focus on the Degradation of Proteins at Mesophilic Temperature," *Bioresource Technology*, vol. 140, pp. 131-137, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[37] Solomon Ali Yimam, Shimelis Kebede Kassahun, and Joon Wun Kang, "Coupling Electrooxidation and UV for the Treatment, and Mineralization of Amoxicillin and Ibuprofen from Aqueous Solution: RSM Parametric Optimization and Toxicity Assessment," *Results in Engineering*, vol. 27, pp. 1-17, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[38] Mohammad Reza Zarei et al., "Optimization of Aerobic Moving Bed Biofilm Reactor Performance for Sulfate Removal using Response Surface Methodology: The Experience of University-Industry Collaboration," *Water Resources and Industry*, vol. 36, pp. 1-17, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[39] Nabieh Farhami, Maryam Derakhshandeh, and Hamid Kazemi Hakki, "Exploring COD and BOD Removal from Industrial Wastewater using a Moving Bed Biofilm Reactor (MBBR)," *Desalination and Water Treatment*, vol. 322, pp. 1-7, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[40] Shahin Heydari Orojlou, Saadat Rastegarzadeh, and Behrooz Zargar, "Experimental and Modeling Analyses of COD Removal from Industrial Wastewater using the Tio2–Chitosan Nanocomposites," *Scientific Reports*, vol. 12, no. 1, pp. 1-15, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[41] Fatemeh Mirza Esmaeili et al., "Algal Blooms Historical Outbreaks in the Northern Coastal Waters of the Persian Gulf and Oman Sea (1980-2015)," *Environmental Monitoring and Assessment*, vol. 193, no. 10, pp. 1-12, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[42] F. Morgan-Sagastume et al., "Anaerobic Treatment of Oil-Contaminated Wastewater with Methane Production using Anaerobic Moving Bed Biofilm Reactors," *Water Research*, vol. 163, pp. 1-13, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[43] G. Mustafa et al., "Bacterial Tools for the Removal and Degradation of Synthetic Dyes from the Wastewater," *Current Developments in Bioengineering and Biotechnology*, pp. 339-370, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[44] Cesar E. Najar-Almanzor et al., "Adaptation and Bioremediation Efficiency of UV-Mutagenized Microalgae in Undiluted Agro-Industrial Effluents from Mexico," *Environments*, vol. 12, no. 9, pp. 1-26, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[45] Hale Ozgun et al., "Municipal Wastewater Treatment using an Integrated System of Aerobic Granular Sludge, Membrane, and Ultraviolet Processes for Reuse in Irrigation," *Clean Technologies and Environmental Policy*, vol. 28, no. 3, pp. 1-19, 2026. [CrossRef] [Google Scholar] [Publisher Link]

[46] Omid Pourehie, and Javad Saien, "Treatment of Real Petroleum Refinery Wastewater with Alternative Ferrous-Assisted UV/Persulfate Homogeneous Processes," *Desalination and Water Treatment*, vol. 142, pp. 140-147, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[47] Ojo Samuel et al., "Treatment of Oily Wastewater using Photocatalytic Membrane Reactors: A Critical Review," *Journal of Environmental Chemical Engineering*, vol. 10, no. 6, pp. 1-26, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[48] Rabia Shoukat, Sher Jamal Khan, and Yousuf Jamal, "Hybrid Anaerobic-Aerobic Biological Treatment for Real Textile Wastewater," *Journal of Water Process Engineering*, vol. 29, pp. 1-8, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[49] Kuan-Yeow Show et al., "Laboratory Trial and Full-Scale Implementation of Integrated Anaerobic-Aerobic Treatment for High Strength Acrylic Acid Wastewater," *Science of the Total Environment*, vol. 738, pp. 1-10, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[50] Thafseela Koya, Saravanan Chandran, and K. Vijayalakshmi, "Analysis of Application of Arithmetic Coding on DCT and DCT-DWT Hybrid Transforms of Images for Compression," *2017 International Conference on Networks and Advances in Computational Technologies*, Thiruvananthapuram, India, pp. 288-293, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[51] Thafseela Koya Poolakkachalil et al., "Comparative Analysis of Lossless Compression Techniques in Efficient DCT-based Image Compression System based on Laplacian Transparent Composite Model and an Innovative Lossless Compression Method for Discrete-Color Images," *2016 3rd MEC International Conference on Big Data and Smart City*, Muscat, Oman, pp. 1-6, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[52] Jacob Wekalao et al., "Design of Ring and Cross Shaped Graphene Metasurface Sensor for Efficient Detection of Malaria and 2 Bit Encoding Applications," *Diamond and Related Materials*, vol. 139, pp. 1-7, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[53] Jacob Wekalao et al., "Design of Graphene Metasurface Sensor for Efficient Detection of COVID-19," *Plasmonics*, vol. 18, no. 6, pp. 2335-2345, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[54] Edward Tager et al., "Aspen Plus Designing and Optimizing the Hospital Wastewater Treatment by Wet Air Oxidation Method," *E3S Web of Conferences*, vol. 436, pp. 1-5, 2023. [CrossRef] [Google Scholar] [Publisher Link]