Water Science and Engineering xxxx, xxx(xxx): xxx

HOSTED BY
ELSEVIER

Available online at www.sciencedirect.com

Water Science and Engineering

journal homepage: wse.hhu.edu.cn

Superior decomposition of xenobiotic RB5 dye using three-dimensional electrochemical treatment: Response surface methodology modelling, artificial intelligence, and machine learning-based optimisation approaches

Voravich Ganthavee*, Antoine P. Trzcinski

School of Agriculture and Environmental Science, University of Southern Queensland, Toowoomba, QLD 4350, Australia

Received 28 November 2023; accepted 13 May 2024
Available online ■ ■ ■

Abstract

The highly efficient electrochemical treatment technology for dye-polluted wastewater is one of hot research topics in industrial wastewater treatment. This study reported a three-dimensional electrochemical treatment process integrating graphite intercalation compound (GIC) adsorption, direct anodic oxidation, and ·OH oxidation for decolourising Reactive Black 5 (RB5) from aqueous solutions. The electrochemical process was optimised using the novel progressive central composite design−response surface methodology (CCD−NPRSM), hybrid artificial neural network−extreme gradient boosting (hybrid ANN−XGBoost), and classification and regression trees (CART). CCD−NPRSM and hybrid ANN−XGBoost were employed to minimise errors in evaluating the electrochemical process involving three manipulated operational parameters: current density, electrolysis (treatment) time, and initial dye concentration. The optimised decolourisation efficiencies were 99.30%, 96.63%, and 99.14% for CCD−NPRSM, hybrid ANN−XGBoost, and CART, respectively, compared to the 98.46% RB5 removal rate observed experimentally under optimum conditions: approximately 20 mA/cm^2 of current density, 20 min of electrolysis time, and 65 mg/L of RB5. The optimised mineralisation efficiencies ranged between 89% and 92% for different models based on total organic carbon (TOC). Experimental studies confirmed that the predictive efficiency of optimised models ranked in the descending order of hybrid ANN−XGBoost, CCD−NPRSM, and CART. Model validation using analysis of variance (ANOVA) revealed that hybrid ANN−XGBoost had a mean squared error (MSE) and a coefficient of determination (R^2) of approximately 0.014 and 0.998, respectively, for the RB5 removal efficiency, outperforming CCD−NPRSM with MSE and R^2 of 0.518 and 0.998, respectively. Overall, the hybrid ANN−XGBoost approach is the most feasible technique for assessing the electrochemical treatment efficiency in RB5 dye wastewater decolourisation.

Keywords: Three-dimensional electrochemical treatment; Dye-polluted wastewater; Artificial intelligence; Machine learning; Optimisation; Analysis of variance; Error function analysis

1. Introduction

Dye-contaminated water can prevent the penetration of sunlight into water and limit photosynthetic activity in the marine environment, thereby polluting the aquatic environment and threatening the lives of both organisms and humans. Notably, Reactive Black 5 (RB5) has been widely utilised by textile, printing, and leather industries due to its intense black colouration, remarkable solubility, and adhesive properties, making it one of the most suitable options for dyeing cotton and other cellulose fibres (Droguett et al., 2020; Feng et al., 2022). Its favourable dyeing properties meet most requirements of textile manufacturers. However, RB5's chemical stability contributes to significant ecological toxicity when industrial effluents containing it are discharged into the marine environment without proper control measures.

Electrochemical oxidation is emerging as an attractive alternative method for wastewater treatment to replace

* Corresponding author.
E-mail address: Voravich.Ganthavee@unisq.edu.au (Voravich Ganthavee).
Peer review under responsibility of Hohai University.

https://doi.org/10.1016/j.wse.2024.05.003

Please cite this article as: Ganthavee, V., Trzcinski, A.P., Superior decomposition of xenobiotic RB5 dye using three-dimensional electrochemical treatment: Response surface methodology modelling, artificial intelligence, and machine learning-based optimisation approaches, Water Science and Engineering, https://doi.org/10.1016/j.wse.2024.05.003

conventional processes owing to substantial quantities of toxic pollutants generated by various industrial processes, particularly in dye wastewater. In the electrochemical oxidation process, organic pollutants undergo removal via two mechanisms: (1) direct oxidation, wherein electron transfer occurs directly from organics to the electrode surface, or (2) indirect oxidation, wherein electron transfer from adsorbed organic species results in the generation of oxidising species that further oxidise pollutants (Ganiyu et al., 2021). Powerful oxidising agents such as hydroxyl radical (·OH), active chlorine species, and sulphate radical ($SO_4^{\bullet -}$) generated during the electrochemical oxidation process can degrade organic pollutants or even mineralise them completely into CO_2 and H_2O (Fu et al., 2023).

The three-dimensional (3D) electrooxidation technology has recently emerged as a powerful method for wastewater treatment. It involves using a 3D electrode reactor, wherein a third electrode, namely graphite intercalation compound (GIC), is incorporated into the reactor and positioned between the anode and cathode. In contrast, conventional two-dimensional (2D) electrochemical reactors lack a particle electrode apart from the anode and cathode. The advantages of the 3D electrochemical oxidation process stem from the electroactive surface area of the particle electrode, which enhances the reaction process, space−time yield, and current efficiency (Li et al., 2021). GIC exhibits superior electrocatalytic efficiency and regenerative capabilities, capable of restoring adsorptive capacity even after several adsorption−regeneration cycles, thereby leading to sustained catalytic oxidation performance (Trzcinski and Harada, 2023). However, thorough evaluation of the influence of operational variables on the electrochemical oxidation process is seldom undertaken. Variations in operational variables can influence process conditions in various ways, necessitating a comprehensive examination of their combined effects on the electrochemical system's overall responses. Moreover, the response surface methodology (RSM) serves as an optimisation tool offering substantial benefits in terms of cost reduction, including reduced energy consumption, enhanced value management, and conservation of valuable resources such as energy and materials (Dong et al., 2023).

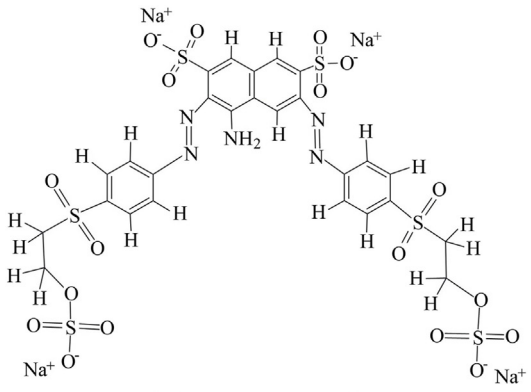
This study investigated the efficacy of 3D electrochemical treatment for removing RB5 xenobiotic dye from water, using an electrically conductive GIC. An RSM with a face-centred central composite design (CCD) was developed to construct a mathematical model for predicting dye and total organic carbon (TOC) removal efficiencies, current efficiency, electrical energy consumption for RB5 and TOC removal, and annual electricity cost. In addition, the correlation of these dependent variables with input parameters (including current density, electrolysis time, and initial dye concentration) was quantified. Various artificial intelligence and machine learning techniques were utilised to assess the predictive efficiency of response variables.

2. Materials and methods

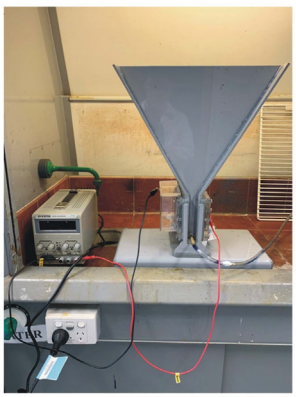
2.1. Materials and electrochemical reactor

All chemicals, including RB5 dye powder (empirical formula: $C_{26}H_{21}N_5Na_4O_{19}S_6$) with a molecular weight of 991.82 g/mol, were purchased from Sigma-Aldrich, Australia. HCl (32%; RCI Labscan) and NaCl (99.7%; Chem-Supply) were used as received. Three stock solutions with various initial dye concentrations (C_0) of 30 mg/L, 65 mg/L, and 100 mg/L were prepared by dissolving RB5 in distilled water. A schematic diagram illustrating the experimental setup for continuous adsorption and electrochemical regeneration is shown in Fig. 1.

The adsorbent employed in this study is an expandable GIC purchased from Sigma-Aldrich (P/N: 808121). At least 75% of the flakes possess sizes greater than 300 μm. GIC has no porous structure and exhibits a relatively low electroactive surface area of approximately 1 m^2/g (Hussain et al., 2016), with high conductivity (0.8 S/cm) (Asghar et al., 2014). The 3D electrochemical reactor used in this experiment was designed to remove RB5 from an aqueous solution (Fig. 1). The design of the 3D electrochemical reactor adhered to the standards employed in Trzcinski and Harada (2023). RB5 concentrations were measured using an ultraviolet−visible spectrophotometer (DR6000, HACH Co.) at the maximum



(a) Chemical structure of RB5



(b) Image of sequential batch electrochemical reactor

Fig. 1. Chemical structure of RB5 and image of sequential batch electrochemical reactor.

Please cite this article as: Ganthavee, V., Trzcinski, A.P., Superior decomposition of xenobiotic RB5 dye using three-dimensional electrochemical treatment: Response surface methodology modelling, artificial intelligence, and machine learning-based optimisation approaches, Water Science and Engineering, https://doi.org/10.1016/j.wse.2024.05.003

absorbance wavelength of 596 nm. To quantify the mineralisation efficiency of RB5-contaminated water, TOC measurement was conducted using a TOC analyzer (Shimadzu VCHS/CSN, Japan).

2.2. Experimental design, modelling, and optimisation

2.2.1. CCD−RSM procedure

An RSM-based face-centred CCD was conducted using Minitab software to configure, model, and optimise the operational parameters affecting a response with minimal experimental runs (Asgari et al., 2020). CCD stands out as one of the most well-established techniques within RSM for determining the correlation between operational parameters and experimental responses, in terms of linear, interactive, and partial or full quadratic effects (Pavlović et al., 2014). The selection of operational parameters was meticulous, aiming to maximise the performance of the electrochemical system within a reasonable experimental domain to facilitate optimisation and yield meaningful outcomes. For instance, dye concentrations in actual textile wastewater typically fall between 10 mg/L and 200 mg/L (Gahr et al., 1994; Laing, 1991). Conversely, recommended applied current densities and electrolysis times typically range from 10 mA/cm^2 to 30 mA/cm^2 and from 10 min to 30 min, respectively (Chen et al., 2018). Extremely high current density can induce undesirable side reactions due to the rapid formation of intermediate breakdown products from organic pollutants, potentially compromising overall treatment efficiency. Therefore, this study investigated the influence of three key operational parameters (current density, electrolysis (treatment) time, and initial dye concentration) on the performance of the 3D electrochemical system.

2.2.2. ANN procedure

In addition to the novel progressive central composite design−response surface methodology (CCD−NPRSM), the artificial neural network (ANN) method was also employed for modelling and predicting responses affected by operational parameters (Fig. 2). The number of neurons within the hidden layer was investigated within a range of 1−20 to determine the optimum number of neurons with minimum mean squared error (MSE) while striving for a high coefficient of determination (R^2) for each response variable. Further analyses incorporated experimental data stipulated in the data matrix in Table A.1 in Appendix A into the ANN model, with 70% of the data allocated for training and 30% for validation, randomly classified in three categories containing input parameters. The curve fitting of the ANN model relies on operational parameter values, and variations in input variables significantly affect the degree of fitness. To mitigate computational issues, all input variable values and experimental efficiencies were normalised into a Gaussian distribution within the range of 0.1−0.9 using Eq. (1) (Asgari et al., 2020):

$$y_i = 0.1 + \frac{0.8(x_i - x_{\min})}{x_{\max} - x_{\min}} \quad (1)$$

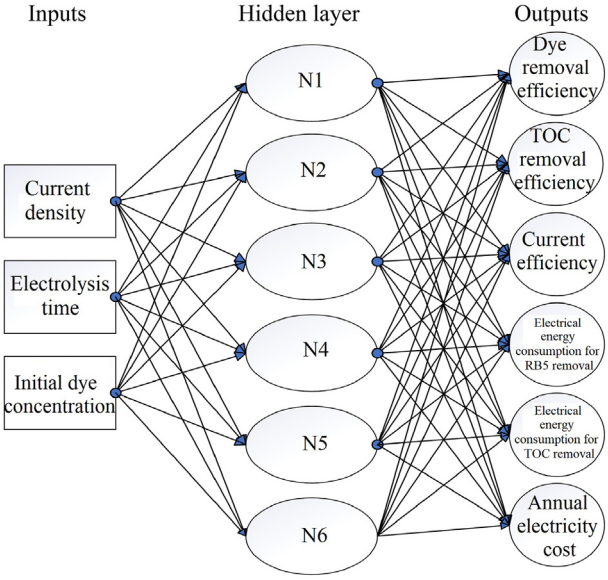


Fig. 2. ANN network with topology.

where x_i is the ith input variable; y_i is the normalized value of x_i; and $x_{\max}$ and $x_{\min}$ are the maximum and minimum values of x_i, respectively. The ANN analysis and modelling were performed using MATLAB R2023a. The performance of ANN models in curving fitting was evaluated using statistical error function analyses such as MSE, R^2, adjusted coefficient of determination (R^2_{adj}), root mean squared error (RMSE), and mean absolute percentage error (MAPE).

2.2.3. Optimisation procedure

2.2.3.1. CCD−NPRSM optimisation. To predict responses, CCD−NPRSM optimisation was performed using an empirical second-order polynomial equation nested within a higher-order polynomial equation as a transfer function to establish multilevel nested models. The optimisation procedure was based on maximum dye and TOC removal efficiencies, maximum current efficiency, minimum electrical energy consumption for RB5 and TOC removal, and minimum annual electricity cost. In the CCD−NPRSM approach, optimisation follows the desired function derived from statistical software. To evaluate fitness and prediction accuracy, composite desirability was used to define the objective function based on the weighted geometric mean of individual desirabilities for response variables to determine the optimal conditions (Askari et al., 2017). The weighted geometric mean of individual desirabilities (D) is expressed as

$$D = \left(\prod_{i=1}^{n} d_i^{w_i} \right)^{1/w} \quad (2)$$

where d_i is the individual desirability for the ith response, w_i is the importance of the ith response, w is the element weight, and n is the number of responses. If each response holds

Please cite this article as: Ganthavee, V., Trzcinski, A.P., Superior decomposition of xenobiotic RB5 dye using three-dimensional electrochemical treatment: Response surface methodology modelling, artificial intelligence, and machine learning-based optimisation approaches, Water Science and Engineering, https://doi.org/10.1016/j.wse.2024.05.003

critical importance or significance, the composite desirability (D_c) can be expressed as

$$D_c = \prod_{i=1}^{n} d_i^{1/n} \tag{3}$$

2.2.3.2. XGBoost-based optimisation. Extreme gradient boosting (XGBoost) is an ensemble method with weaker models instead of more robust models such as ANN. Nonetheless, XGBoost can be used to reinforce ANN optimisation. Similar to classification and regression trees (CART) models, XGBoost comprises regression trees. It employs a second-order Taylor expansion of the loss function, integrating a regular term to find the optimal solution to balance the decline in the loss function, manage model complexity, and mitigate overfitting issues (Wang et al., 2022). The estimated output of the model for any given sample is obtained by summing leaves assigned to each sample corresponding to each regression tree (Ching et al., 2022):

$$\hat{y} = \sum_{k=1}^{K_b} f_k(x_i) \tag{4}$$

where $\hat{y}$ is the predicted value, f_k is the kth boosted function, and K_b is the number of boosted functions. Regression trees are added to the ensemble, such as f_t (the boosted function of variables for t iterations), yielding a new regression tree to minimise learning objectives. Unlike a single model with a pre-defined structure, which can be optimised in Euclidean space (Ching et al., 2022), XGBoost can be integrated with ANN to create a hybrid model, thereby reducing errors and enhancing prediction efficiency.

2.2.3.3. CCD−NPRSM. The second-order CCD−NPRSM may offer satisfactory curve fitting but can produce significantly lower MSE and RMSE when using highly non-linear and complex mathematical functions. Through model transformation, a higher order (6th order) polynomial mathematical function is used as a nested transfer function to modulate curve fitting, reduce MSE and RMSE, and improve the correlation coefficient (Zheng et al., 2022). Therefore, CCD−NPRSM emerges as another research focus to generate multilevel nested models. When integrated with a feedback control loop, improvements in RB5 and TOC removal efficiencies can be achieved (Fig. 3). The proposed CCD−NPRSM entails a step-by-step procedure for a non-linear regression algorithm:

$$Y_i^*(x_k) = f\left[K(x_i,x_j)G(x_i,x_j)\right] \quad x_1 \in [10,30], \; x_2 \in [10,30], x_3 \in [30,100] \tag{5}$$

$$K(x_i,x_j)G(x_i,x_j) = Y_i(x_i,x_j) = \beta_0 + \sum_{i=1}^{k}(\beta_i x_i) + \sum_{i=1}^{k}\left(\beta_{ii}x_i^2\right) + \sum_{i=1}^{k}\sum_{j=i+1}^{k}\left(\beta_{ij}x_i x_j\right) + \gamma \tag{6}$$

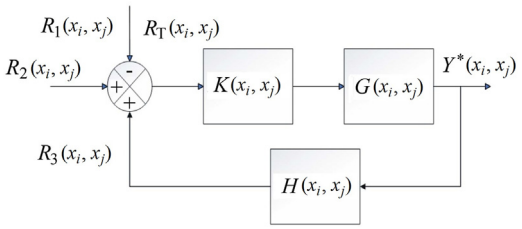


Fig. 3. Derivation of transfer function from block diagram (with R_T denoting sum of input functions and Y^* representing output function).

$$\frac{Y_i^*(x_k)}{R_i(x_i,x_j)} = \frac{K(x_i,x_j)G(x_i,x_j)}{1 + K(x_i,x_j)G(x_i,x_j)H(x_i,x_j)} \tag{7}$$

where Y_i^* is the ith output function, f is the composite function, K the transfer function of the first system, G is the transfer function of the second system, Y_i is ith the predicted response in the form of a polynomial function, β_0 is the intercept or regression coefficient, β_i is the linear coefficient, β_{ii} is the quadratic coefficient, β_{ij} is the interaction coefficient, γ is the experimental or residual error, R_i is the Laplace transform of the ith input function, and H is the closed-loop transfer function.

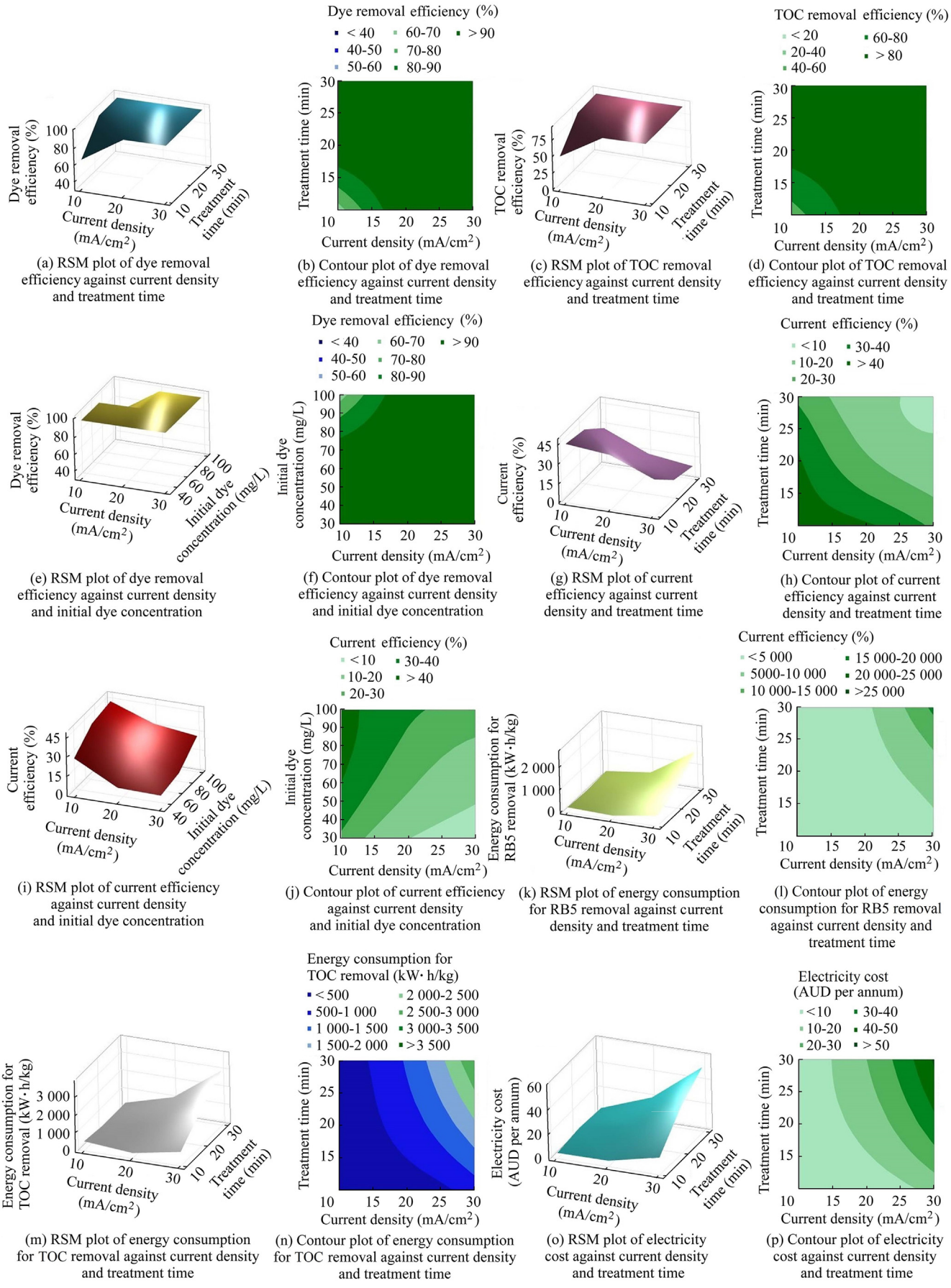
2.2.3.4. Predictive analytics by CART machine learning optimisation. CART machine learning optimisation stands out as one of the best-in-class approaches, not only fitting more accurate models when combined with experimental data but also for handling larger datasets with more variables, messy or missing data, outliers, and non-linear relationships. With the power of the original CART, it offers visualisations of predicted values and interactive effects to achieve optimal prediction accuracy (Okagbue et al., 2021). Initially employed in the bootstrap aggregation method to tackle complex non-linear problems, it delivers the most accurate model obtained from the proprietary predictive analytics of CART.

3. Results and discussion

3.1. Optimisation study using 3D response surface plots

In the experimental study, the effects of operational parameters, such as current density, electrolysis time, and initial dye concentration, electrical energy consumption for RB5 and TOC removal, current efficiency, and annual electricity cost were investigated using RSM optimisation via CCD. The primary aim was to determine the optimal current density, electrolysis time, and initial dye concentration to achieve maximum dye and TOC removal efficiencies, minimize electrical energy consumption, enhance current efficiency, and manage annual electricity cost. Batch experimental runs were conducted according to the CCD design of experiments to three-dimensionally visualise the effects of independent variables on targeted responses by optimising results within the experimental conditions. A general finding revealed that an

Please cite this article as: Ganthavee, V., Trzcinski, A.P., Superior decomposition of xenobiotic RB5 dye using three-dimensional electrochemical treatment: Response surface methodology modelling, artificial intelligence, and machine learning-based optimisation approaches, Water Science and Engineering, https://doi.org/10.1016/j.wse.2024.05.003



(a) RSM plot of dye removal efficiency against current density and treatment time

(b) Contour plot of dye removal efficiency against current density and treatment time

(c) RSM plot of TOC removal efficiency against current density and treatment time

(d) Contour plot of TOC removal efficiency against current density and treatment time

(e) RSM plot of dye removal efficiency against current density and initial dye concentration

(f) Contour plot of dye removal efficiency against current density and initial dye concentration

(g) RSM plot of current efficiency against current density and treatment time

(h) Contour plot of current efficiency against current density and treatment time

(i) RSM plot of current efficiency against current density and initial dye concentration

(j) Contour plot of current efficiency against current density and initial dye concentration

(k) RSM plot of energy consumption for RB5 removal against current density and treatment time

(l) Contour plot of energy consumption for RB5 removal against current density and treatment time

(m) RSM plot of energy consumption for TOC removal against current density and treatment time

(n) Contour plot of energy consumption for TOC removal against current density and treatment time

(o) RSM plot of electricity cost against current density and treatment time

(p) Contour plot of electricity cost against current density and treatment time

Fig. 4. 3D RSM optimisation plots of interaction effects of current density, electrolysis (treatment) time, and initial dye concentration.

Please cite this article as: Ganthavee, V., Trzcinski, A.P., Superior decomposition of xenobiotic RB5 dye using three-dimensional electrochemical treatment: Response surface methodology modelling, artificial intelligence, and machine learning-based optimisation approaches, Water Science and Engineering, https://doi.org/10.1016/j.wse.2024.05.003

increase in current density led to increases in dye and TOC removal efficiencies in the batch runs.

Fig. 4(a) shows the 3D response surface for dye removal efficiency as a function of current density and treatment time. To achieve over 95% decolourisation efficiency, Fig. 4(b) indicates a necessity for a current density exceeding 18 mA/cm^2 and an electrolysis time exceeding 10 min (dark-shaded green regions). Among the examined operational parameters, current density proved the most critical variable. Its effect on targeted responses, such as dye and TOC removal efficiencies, was predominant, particularly with the presence of electro-generated oxidising species like hydroxyl radicals and active chlorine species due to high current density employed for oxidising dye contaminants in water. Moreover, Fig. 4(d) shows that the TOC removal efficiency surpassed 90% when current density exceeded 18 mA/cm^2 and initial dye concentration was approximately 36 mg/L (dark-shaded green regions). However, beyond 18 mA/cm^2, an increase in current density for a similar initial dye concentration resulted in the TOC removal efficiency exceeding 90% (Fig. 4(c)). As shown in Fig. 4(i), an increase in current density from 20 mA/cm^2 to 30 mA/cm^2 led to a decrease in current efficiency below 10%. This result indicated that higher current density induced side reactions. In addition, the effect of operational parameters on electrochemical redox reactions is depicted in Section A.2 in Appendix A. The reaction mechanisms of RB5 and its intermediates in the electrochemical system are detailed in Fig. A.1 (Feng et al., 2016) and Section A.3 in Appendix A.

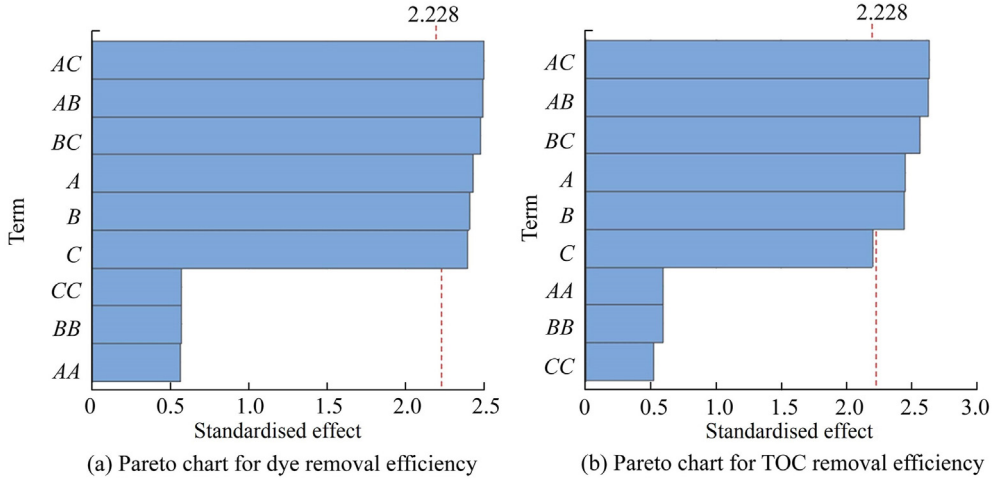
Fig. 4(h) shows that when current density exceeded 25 mA/cm^2 with an electrolysis time greater than 25 min, the current efficiency of the electrochemical reactor decreased below 10%. This result indicated that a higher current density led to side reactions due to the accumulation of intermediate transformation byproducts, offsetting the current efficiency of the electrochemical system. Consequently, the degradation efficiency of RB5 pollutants in water might be adversely affected as some of the current generated from the anodic oxidation process was lost through side reactions. Fig. 4(h) also demonstrates that at a low current density of 10 mA/cm^2 and an electrolysis time shorter than 15 min, current efficiency increased beyond 40%, indicating a more efficient untilisation of current to generate radical species. Despite this, dye and TOC removal efficiencies remained lower than 80%, highlighting an inverse relationship between current efficiency and dye and TOC removal efficiencies. Furthermore, although increasing the initial dye concentration from 50 mg/L to 100 mg/L at a low current density of 10 mA/cm^2 boosted current efficiency beyond 40%, it might compromise dye and TOC removal efficiencies. This indicated that current density exerted the most substantial impact on dye and TOC removal efficiencies. A prolonged electrolysis duration significantly enhanced electrolytic efficiency, thereby maximising dye and TOC removal efficiencies. As shown in Fig. 4(b), the interactive effect of current density and electrolysis time on dye removal efficiency was evident from the elliptical or saddle pattern of the contour plot. Similarly, Fig. 4(d) highlights the interactive effect of current density and electrolysis time on TOC removal efficiency, with the contour plot exhibiting a similar elliptical or saddle pattern, signifying its notable impact. Notably, the TOC removal efficiency was lower than the dye removal efficiency due to the presence of the residual fragments of dye molecules in the aqueous solution.

As shown in Fig. 4(f), the elliptical or saddle pattern of the contour plot indicated a significant interactive effect of current density and initial dye concentration on dye removal efficiency. Similarly, Fig. 4(k) and (l) indicates a significant interactive effect of current density and electrolysis time on electrical energy consumption for RB5 removal. Furthermore, Fig. 4(m) shows that the interactive effect of current density and treatment time on electrical energy consumption for TOC removal was more pronounced, as indicated by the degree of curvature. The elliptical or saddle pattern of the contour plot (Fig. 4(n)) indicates the greater significance of this interactive effect on electrical energy consumption for TOC removal compared to the interactive effect on electrical energy consumption for RB5 removal, implying a higher electrical energy requirement to oxidise or electrolyse TOC to transform all RB5 dye molecules into inert and non-toxic end products such as CO_2 and H_2O. However, Fig. 4(o) shows that the interaction between current density and treatment time significantly affected annual electricity cost if current density surpassed 20 mA/cm^2. No significant curvature is present in the response surface plot at a lower current density less than 20 mA/cm^2. The mild circular pattern in the contour plot (Fig. 4(p)), particularly below a current density of 20 mA/cm^2, indicated that the interactive effect of current density and treatment time may not significantly affect annual electricity cost, implying that the electrochemical treatment process for RB5 removal might not contribute to significant electrical energy consumption when current density and treatment time were below 20 mA/cm^2 and 15 min, respectively. This finding underscored the cost-effectiveness of the 3D electrochemical treatment process. Detailed mathematical expressions or functions of various targeted responses are provided in Appendix A.

3.2. ANOVA analysis

The significance and validity of the generated CCD−NPRSM models were assessed through analysis of variance (ANOVA), as presented in Table A.2 (Zhang et al., 2013) and Fig. A.2 in Appendix A. The quadratic model yielded F-values of 4.42 and 4.57, with corresponding p-values of 0.015 and 0.013, indicating statistical significance of the models for dye and TOC removal efficiencies. Moreover, the F-values of current density (5.88 and 6.01) and their associated p-values (0.036 and 0.034) indicated that current density significantly affected dye and TOC removal efficiencies. The magnitude of the responses and the absolute values of standardised effects delineate the most significant to the least significant effects, providing a reference line to gauge statistical significance. According to the ANOVA results, high F-values and low P-values for dye removal efficiency (Y_1)

Please cite this article as: Ganthavee, V., Trzcinski, A.P., Superior decomposition of xenobiotic RB5 dye using three-dimensional electrochemical treatment: Response surface methodology modelling, artificial intelligence, and machine learning-based optimisation approaches, Water Science and Engineering, https://doi.org/10.1016/j.wse.2024.05.003



(a) Pareto chart for dye removal efficiency (b) Pareto chart for TOC removal efficiency

Fig. 5. Pareto charts representing significant effect of three different factors on dye and TOC removal efficiencies.

($p < 0.05$ and $F = 4.42$) and TOC removal efficiency (Y_2) ($p < 0.05$ and $F = 4.57$) indicated the considerable significance of the hybrid CCD−NPRSM models. The experimental results demonstrated the accurate fit of the models with the theoretical models governing the relationship between independent variables and responses.

In the Pareto chart shown in Fig. 5(a), the bars representing current density (factor A) and treatment time (factor B) surpassed the reference line with an absolute value of 2.228. This signified a statistically significant effect of current density and electrolysis time on dye removal efficiency, with a significance level below 0.05. Moreover, synergistic effects like combined current density and treatment time (AB), combined current density and initial dye concentration (AC, with factor C denoting the initial dye concentration), and combined treatment time and initial dye concentration (BC) also demonstrated statistically significant effects on dye removal efficiency. Although ANOVA analysis indicated a slight difference with a p-value of 0.015 for initial dye concentration, the Pareto chart reveals that the bar did not cross the reference line of 2.228 (Fig. 5(b)), less than the absolute value of standardised effects. This indicated that the magnitude of the effect of initial dye concentration on TOC removal efficiency might not be substantially significant.

The one-way ANOVA revealed F-values of 5.88 and 5.77 for current density and treatment time, respectively, in contrast to an F-value of 5.72 for initial dye concentration in Y_1. This suggested that the variances of current density and treatment time were significantly different from the mean of initial dye concentration. In addition, current density and treatment time also exerted a more significant effect on dye removal efficiency than initial dye concentration. As shown in Table A.2 in Appendix A, the lack-of-fit F-values surpassed the p-values for all parameters, underscoring the statistical significance of the models for both dye and TOC removal efficiencies.

3.3. Predictive accuracy of developed models

The TOC removal efficiency provides a comprehensive insight into the overall mineralisation efficiency of dye pollutants, It serves as a key indicator of the extent to which the toxicity of dye pollutants can be converted into inert and non-toxic CO_2 and water, thus facilitating the complete abatement of RB5 from contaminated water. To assess the optimisation and predictive capabilities of CCD−NPRSM and hybrid ANN−XGBoost models in terms of TOC removal efficiency, 20 experimental runs were conducted at specified levels of operational variables (Table 1). The comparison between experimental and predicted TOC removal efficiencies revealed that all three models could accurately predict values close to the experimental data.

The statistical significance of the three models was evaluated using error function analysis, including MSE and R^2, to gauge the fitness of the models with experimental data. As

Table 1
Comparison between optimised and non-optimised experimental results.

Model or experimental result	I (mA/cm^2)	T_E (min)	C_0 (mg/L)	E_{dye} (%)	E_{TOC} (%)	E_C (%)	C_{dye} (kW·h/kg)	C_{TOC} (kW·h/kg)	C_E (AUD per annum)	D_c
CCD−NPRSM optimisation	20	20	65	99.303 7	89.759 3	22.542 2	2 540.74	715.38	14.020 8	0.805 0
Hybrid ANN−XGBoost optimisation	20	20	65	99.626 2	90.471 1	23.075 4	2 934.30	734.26	14.530 6	0.794 9
CART optimisation	20	20	65	99.137 0	89.680 0	26.055 8	4 996.63	1 043.92	18.237 0	0.774 3
Experimental result for validation	20	20	65	98.462 2	89.178 3	23.344 9	2 336.55	611.48	12.053 3	

Note: I is the current density, T_E is the electrolysis time, C_0 is the initial dye concentration, E_{dye} is the dye removal efficiency, E_{TOC} is the TOC removal efficiency, E_C is the current efficiency, C_{dye} is the electrical energy consumption for RB5 removal, C_{TOC} is the electrical energy consumption for TOC removal, C_E is the electricity cost, and D_c is the composite desirability.

Please cite this article as: Ganthavee, V., Trzcinski, A.P., Superior decomposition of xenobiotic RB5 dye using three-dimensional electrochemical treatment: Response surface methodology modelling, artificial intelligence, and machine learning-based optimisation approaches, Water Science and Engineering, https://doi.org/10.1016/j.wse.2024.05.003

Table 2
Performance of developed models.

Model	Response	R^2	R^2_{adj}	MSE	RMSE	MAPE
CCD−NPRSM	E_{dye}	0.998	0.998	0.518	0.720	0.495
	E_{TOC}	0.997	0.997	1.010	1.005	0.925
	E_c	0.995	0.995	0.858	0.926	3.605
	C_{dye} (ANN−NPRSM)	0.945	0.942	5.536×10^6 kW•h/kg	2.353×10^3 kW•h/kg	0.001
	C_{TOC} (ANN−NPRSM)	0.981	0.980	7.625×10^3 kW•h/kg	87.321 kW•h/kg	9.210×10^{-5}
	C_E	0.982	0.981	3.818 AUD per annum	1.954 AUD per annum	18.529
Hybrid ANN−XGBoost	E_{dye}	0.998	0.997	0.014	0.120	6.340×10^{-5}
	E_{TOC}	0.998	0.997	0.007	0.081	7.120×10^{-5}
	E_c	0.998	0.997	6.630×10^{-5}	0.008	2.390×10^{-4}
	C_{dye} (ANN−NPRSM)	0.998	0.997	1.035 kW•h/kg	1.018 kW•h/kg	1.400×10^{-6}
	C_{TOC} (ANN−NPRSM)	0.998	0.997	1.158 kW•h/kg	1.076 kW•h/kg	6.530×10^{-6}
	C_E	0.998	0.997	0.010 AUD per annum	0.099 AUD per annum	3.720×10^{-4}
CART	E_{dye}	0.991	0.990	2.004	1.416	1.011
	E_{TOC}	0.987	0.986	4.290	2.071	2.013
	E_c	0.132	0.083	155.407	12.466	60.936
	C_{dye} (ANN−NPRSM)	0.116	0.067	2.797×10^7 kW•h/kg	5.288×10^3 kW•h/kg	110.290
	C_{TOC} (ANN−NPRSM)	0.253	0.212	5.580×10^5 kW•h/kg	746.997 kW•h/kg	74.044
	C_E	0.268	0.228	159.419 AUD per annum	12.626 AUD per annum	99.800

Note: E_{dye} is the dye removal efficiency, E_{TOC} is the TOC removal efficiency, E_C is the current efficiency, C_{dye} is the electrical energy consumption for RB5 removal, C_{TOC} is the electrical energy consumption for TOC removal, and C_E is the electricity cost.

shown in Table 2, CCD−NPRSM and hybrid ANN−XGBoost models achieved R^2 values of 0.998 and 0.998 for dye removal efficiency, respectively. In contrast, the CART model obtained an R^2 value of 0.991 with an MSE value significantly greater than those of CCD−NPRSM and hybrid ANN−XGBoost models. The hybrid ANN−XGBoost model outperformed CCD−NPRSM and CART models in terms of overall prediction efficiency (Table 2). In contrast, the CART algorithm yielded significantly higher MSE, RMSE, and MAPE values than CCD−NPRSM for all response variables, highlighting its susceptibility to high variances across samples and instability in managing noise and data changes. With disadvantages of overfitting, high variances, and great biases, the CART model tends to make the decision tree structure increasingly unstable when predicting certain response variables or anomalies with high fluctuations. To address these issues, hybrid ANN−XGBoost performs parallel tree boosting and offers unequal accuracy in predictions using advanced multiple hyperparameter tuning techniques to optimise loss functions. These features make the hybrid ANN−XGBoost model suitable for managing large datasets with high residual or bias errors, with advantages of capturing complex patterns in combined datasets containing multiple response variables.

The CCD−NPRSM model performed much better than expectations, achieving R^2 and MSE values of 0.997 and 1.010 for TOC removal efficiency. Error function analysis revealed that regarding electrical energy consumption for TOC removal, hybrid ANN-XGBoost yielded R^2 and MSE values of 0.998 and 1.158 kW•h/kg, respectively, significantly outperforming the CART model with R^2 and MSE values of 0.253 and 5.580×10^5 kW•h/kg. Furthermore, CCD−NPRSM demonstrated superior management of high residual errors or variances compared to the CART algorithm. Notably, hybrid ANN−XGBoost outperformed both CCD−NPRSM and CART in terms of R^2 and MSE for majority of response variables. The hybrid intelligence of combined ANN and XGBoost minimises training time and avoids undesirable convergence to local optimal solutions, efficiently handling complex operational parameters for more accurate predictions through global optimisation. Therefore, the observed accuracy of predicted responses confirmed the feasibility of hybrid ANN−XGBoost optimisation in modelling RB5 aqueous systems. Except for certain response variables such as electrical energy consumption for TOC removal, the combination of ANN and CCD−NPRSM optimisation significantly reduced MSE and RMSE to 1.158 kW•h/kg and 1.076 kW•h/kg, respectively. CART exhibited limited prediction and optimisation capabilities, due to certain nature of datasets and its susceptibility to noises and overfitting.

In conclusion, each modelling method offers distinct advantages tailored to specific wastewater treatment processes. Although the CCD−NPRSM approach can reveal the interactive effects of operational variables and their impact on responses via higher-order polynomial mathematical functions, hybrid ANN−XGBoost exhibits superior optimisation capabilities compared to both CCD−NPRSM and CART. Furthermore, hybrid ANN−XGBoost operates as a black-box model, relying primarily on data availability for accurate analysis, thus bypassing the need for intricate experimental designs. Conversely, CART serves as a robust analytical tool to mitigate some of RSM's limitations in predictive modelling, accommodating categorical and continuous data, and managing missing values or data clustering through non-parametric methods without inherent assumptions. However, CART proves less suitable for handling large datasets with extreme variances.

Please cite this article as: Ganthavee, V., Trzcinski, A.P., Superior decomposition of xenobiotic RB5 dye using three-dimensional electrochemical treatment: Response surface methodology modelling, artificial intelligence, and machine learning-based optimisation approaches, Water Science and Engineering, https://doi.org/10.1016/j.wse.2024.05.003

Moreover, Table 2 shows that both CCD−NPRSM and hybrid ANN−XGBoost models exhibited significantly superior overall optimisation and predictive capabilities compared to CART. The current efficiency and electrical energy consumption for RB5 and TOC removal, optimised by hybrid ANN−XGBoost and CCD−NPRSM, were significantly lower than those predicted by the CART model, aligning more closely with experimental data. Ultimately, the hybrid ANN−XGBoost and CCD−NPRSM algorithms can aid water authorities, environmental regulatory bodies, and water resources engineers in achieving exceptional results through hybrid modelling processes. Furthermore, they can be utilised to predict relationships among variables and optimise responses in scaled-up processes within real wastewater treatment systems.

3.4. Optimisation efficiency of electrochemical process

The primary purpose of this experimental study was to enhance the electrochemical process by optimising the operational parameters to boost dye and TOC removal efficiencies while minimising electrical energy consumption and reducing annual electricity cost, all without compromising the treatment efficiency. The optimised results, with a composite desirability of 0.805 0 (Table 1), underscored the precision of the results. Specifically, optimised data in Table 1 reveal that achieving a dye removal efficiency of 99% or higher for CCD−NPRSM necessitated a minimum current density of approximately 20 mA/cm^2 and an electrolysis time of 20 min for treating 65-mg/L RB5. Under optimised conditions using the hybrid ANN−XGBoost algorithm, merely 2 934.30 kW·h/kg of RB5 sufficed to attain a dye removal efficiency exceeding 99% within a 20-min electrolysis timeframe, indicating the remarkable energy efficiency of the electrochemical process. In contrast, employing the CART algorithm consumed 1 043.92 kW·h/kg of electrical energy for TOC removal to achieve over 99% and 89% removal efficiencies for dye and TOC. Comparative research has demonstrated similar results, achieving a 91.6% RB5 removal rate using 0.4 A of applied current over 50 min of electrolysis (Feng et al., 2022). In addition, their research showed that treating 0.5 L of the RB5 solution at an initial dye concentration of 4 mg/L required a maximum electrical energy consumption of 4.89 kW·h/m^3. Table 3 compares electrical energy consumption between experimental results and the literature. In addition, Table 1 shows that the TOC removal efficiency significantly surpassed non-optimised experimental results, highlighting the advantages of employing hybrid ANN−XGBoost and CCD−NPRSM optimisation techniques. Experimental results indicated that employing a current density of approximately 20 mA/cm^2 for dye solution treatment yielded notably higher current efficiency than the optimised result, suggesting mitigation of side reactions and intermediate oxidation byproduct formation, which enhanced potential current utilisation efficiency. However, undesirably high electrical energy consumption due to voltage fluctuations was observed under non-optimised conditions, leading to additional energy wastage, albeit with lower electricity costs attributed to moderately high current efficiency.

4. Conclusions

This study extensively investigated the electrochemical degradation of RB5 xenobiotic dye in a simulated dye solution using a 3D electrochemical process with GIC particle electrodes and a graphite anode. The effects of operational parameters on dye and TOC removal efficiencies, current efficiency, electrical energy consumption for RB5 and TOC removal, and electricity cost were optimised using CCD−NPRSM, hybrid ANN−XGBoost, CART algorithms, along with approximating functions, until satisfactory convergence of solutions was achieved to maximise fitness in the modelling. Key optimisation results showed that TOC mineralisation efficiencies of 89.76%, 90.47% and 89.68% were achieved using CCD−NPRSM, hybrid ANN−XGBoost, and CART optimisation techniques, respectively, compared to the non-optimised experimental result of 89.18%. Although CART optimisation accurately predicted observed RB5 and TOC removal efficiencies, errors for other response variables were significantly higher than those of CCD−NPRSM and hybrid ANN−XGBoost. In contrast, the

Table 3
Comparison of electrical energy consumption between experimental results and secondary sources from literature.

Adsorbent	Anode	Cathode	C_0	Reactor	C_{EE}	Reference
Granular activated carbon (3DER-GAC) particle electrodes	Ti/SnO_2-Sb/β-PbO_2	Ti substrate	50 mg/L (2,4-dichlorophenol)	Fluidised 3D electrochemical reactor	810 kW·h/kg for TOC using 3DER-GAC); 1 570 kW·h/kg for TOC using electrochemical oxidation (EO)	Samarghandi et al. (2021)
Polymer-based spherical activated carbon (AC)	Titanium coated with RuO_2-IrO_2-TiO_2	Stainless steel	10 mg/L (diclofenac (DCF) or sulfamethoxazole (SMX))	3D biofilm electrode reactor (3D-BERs)	38.5 kW·h/kg for DCF; 20 kW·h/kg for SMX	Soares et al. (2022)
Mn-Co/GAC particle electrode	Ti/RuO_2 electrodes	Ti/RuO_2 electrodes	150 mg/L (amoxicillin (AMX))	3D electrochemical reactor	73 kW·h/kg for AMX	Ma et al. (2022)
Granular activated carbon particle electrode	Ti/RuO_2-IrO_2	Titanium plate	1 000 mg/L (Rhodamine B)	3D electrochemical reactor	6.22 kW·h/kg for chemical oxygen demand (COD)	Ji et al. (2018)

Note: C_0 is the initial dye concentration, and C_{EE} is the electrical energy consumption.

Please cite this article as: Ganthavee, V., Trzcinski, A.P., Superior decomposition of xenobiotic RB5 dye using three-dimensional electrochemical treatment: Response surface methodology modelling, artificial intelligence, and machine learning-based optimisation approaches, Water Science and Engineering, https://doi.org/10.1016/j.wse.2024.05.003

predictive efficiency of hybrid ANN−XGBoost exceeded expectation to other optimisation methods. The overall findings confirmed the techno−economic viability, engineering feasibility, and environmental suitability of the 3D electrochemical process when optimised by either hybrid ANN−XGBoost or hybrid CCD−NPRSM.

Declaration of competing interest

The authors declare no conflicts of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.wse.2024.05.003.

References

Asgari, G., Shabanloo, A., Salari, M., Eslami, F., 2020. Sonophotocatalytic treatment of AB113 dye and real textile wastewater using ZnO/persulfate: Modeling by response surface methodology and artificial neural network. Environ. Res. 184, 109367. https://doi.org/10.1016/j.envres.2020.109367.

Asghar, H.M.A., Hussain, S.N., Sattar, H., Brown, N.W., Roberts, E.P.L., 2014. Electrochemically synthesized GIC-based adsorbents for water treatment through adsorption and electrochemical regeneration. J. Ind. Eng. Chem. 20(4), 2200−2207. https://doi.org/10.1016/j.jiec.2013.09.051.

Askari, H., Ghaedi, M., Dashtian, K., Azghandi, M.H.A., 2017. Rapid and high-capacity ultrasonic assisted adsorption of ternary toxic anionic dyes onto MOF-5-activated carbon: Artificial neural networks, partial least squares, desirability function and isotherm and kinetic study. Ultrason. Sonochem. 37, 71−82. https://doi.org/10.1016/j.ultsonch.2016.10.029.

Chen, L., Lei, C., Li, Z., Yang, B., Zhang, X., Lei, L., 2018. Electrochemical activation of sulfate by BDD anode in basic medium for efficient removal of organic pollutants. Chemosphere 210, 516−523. https://doi.org/10.1016/j.chemosphere.2018.07.043.

Ching, P.M.L., Zou, X., Wu, D., So, R.H.Y., Chen, G.H., 2022. Development of a wide-range soft sensor for predicting wastewater BOD5 using an eXtreme gradient boosting (XGBoost) machine. Environ. Res. 210, 112953. https://doi.org/10.1016/j.envres.2022.112953.

Dong, C., Zhou, N., Zhang, J., Lai, W., Xu, J., Chen, J., Yu, R., Che, Y., 2023. Optimized preparation of gangue waste-based geopolymer adsorbent based on improved response surface methodology for Cd(II) removal from wastewater. Environ. Res. 221, 115246. https://doi.org/10.1016/j.envres.2023.115246.

Droguett, T., Mora-Gómez, J., García-Gabaldón, M., Ortega, E., Mestre, S., Cifuentes, G., Pérez-Herranz, V., 2020. Electrochemical degradation of reactive black 5 using two-different reactor configuration. Sci. Rep. 10, 4482. https://doi.org/10.1038/s41598-020-61501-5.

Feng, L., Liu, J., Guo, Z., Pan, T., Wu, J., Li, X., Liu, B., Zheng, H., 2022. Reactive black 5 dyeing wastewater treatment by electrolysis-Ce(IV) electrochemical oxidation technology: Influencing factors, synergy and enhancement mechanisms. Sep. Purif. Technol. 285, 120314. https://doi.org/10.1016/j.seppur.2021.120314.

Feng, Y., Yang, L., Liu, J., Logan, B.E., 2016. Electrochemical technologies for wastewater treatment and resource reclamation. Environ. Sci. Water Res. Technol. 2(5), 8−831. https://doi.org/10.1039/c5ew00289c.

Fu, R., Zhang, P.S., Jiang, Y.X., Sun, L., Sun, X.H., 2023. Wastewater treatment by anodic oxidation in electrochemical advanced oxidation process: Advance in mechanism, direct and indirect oxidation detection methods. Chemosphere 311, 136993. https://doi.org/10.1016/j.chemosphere.2022.136993.

Gahr, F., Hermanutz, F., Oppermann, W., 1994. Ozonation − An important technique to comply with new German laws for textile wastewater treatment. Water Sci. Technol. 30(3), 255−263. https://doi.org/10.2166/wst.1994.0115.

Ganiyu, S.O., Martínez-Huitle, C.A., Oturan, M.A., 2021. Electrochemical advanced oxidation processes for wastewater treatment: Advances in formation and detection of reactive species and mechanisms. Curr. Opin. Electrochem. 27, 100678. https://doi.org/10.1016/j.coelec.2020.100678.

Hussain, S.N., Trzcinski, A.P., Asghar, H.M.A., Sattar, H., Brown, N.W., Roberts, E.P.L., 2016. Disinfection performance of adsorption using graphite adsorbent coupled with electrochemical regeneration for various microorganisms present in water. J. Ind. Eng. Chem. 44, 216−225. https://doi.org/10.1016/j.jiec.2016.09.009.

Ji, J., Liu, Y., Yang, X., Xu, J., Li, X., 2018. Multiple response optimization for high efficiency energy saving treatment of rhodamine B wastewater in a three-dimensional electrochemical reactor. J. Environ. Manag. 218, 300−308. https://doi.org/10.1016/j.jenvman.2018.04.071.

Laing, I.G., 1991. The impact of effluent regulations on the dyeing industry. Rev. Prog. Coloration Relat. Top. 21(1), 56−71. https://doi.org/10.1111/j.1478-4408.1991.tb00081.x.

Li, H., Yang, H., Cheng, J., Hu, C., Yang, Z., Wu, C., 2021. Three-dimensional particle electrode system treatment of organic wastewater: A general review based on patents. J. Clean. Prod. 308, 127324. https://doi.org/10.1016/j.jclepro.2021.127324.

Ma, J., Gao, M., Liu, Q., Wang, Q., 2022. High efficiency three-dimensional electrochemical treatment of amoxicillin wastewater using Mn−Co/GAC particle electrodes and optimization of operating condition. Environ. Res. 209, 112728. https://doi.org/10.1016/j.envres.2022.112728.

Okagbue, H.I., Oguntunde, P.E., Obasi, E.C.M., Akhmetshin, E.M., 2021. Trends and usage pattern of SPSS and Minitab software in scientific research. J. Phys. Conf. 1734(1), 12017. https://doi.org/10.1088/1742-6596/1734/1/012017.

Pavlović, M.D., Buntić, A.V., Mihajlovski, K.R., Šiler-Marinković, S.S., Antonović, D.G., Radovanović, Ž., Dimitrijević-Branković, S.I., 2014. Rapid cationic dye adsorption on polyphenol-extracted coffee grounds—A response surface methodology approach. J. Taiwan Inst. Chem. Eng. 45(4), 1691−1699. https://doi.org/10.1016/j.jtice.2013.12.018.

Samarghandi, M.R., Dargahi, A., Rahmani, A., Shabanloo, A., Ansari, A., Nematollahi, D., 2021. Application of a fluidized three-dimensional electrochemical reactor with Ti/SnO_2−Sb/β-PbO_2 anode and granular activated carbon particles for degradation and mineralization of 2,4-dichlorophenol: Process optimization and degradation pathway. Chemosphere 279, 130640. https://doi.org/10.1016/j.chemosphere.2021.130640.

Soares, C., Correia-Sá, L., Paíga, P., Barbosa, C., Remor, P., Freitas, O.M., Moreira, M.M., Nouws, H.P.A., Correia, M., Ghanbari, A., et al., 2022. Removal of diclofenac and sulfamethoxazole from aqueous solutions and wastewaters using a three-dimensional electrochemical process. J. Environ. Chem. Eng. 10(5), 108419. https://doi.org/10.1016/j.jece.2022.108419.

Trzcinski, A.P., Harada, K., 2023. Adsorption of PFOS onto graphite intercalated compound and analysis of degradation by-products during electrochemical oxidation. Chemosphere 323, 138268. https://doi.org/10.1016/j.chemosphere.2023.138268.

Wang, M., Li, X., Lei, M., Duan, L., Chen, H., 2022. Human health risk identification of petrochemical sites based on extreme gradient boosting. Ecotoxicol. Environ. Saf. 233, 113332. https://doi.org/10.1016/j.ecoenv.2022.113332.

Zhang, C., Jiang, Y., Li, Y., Hu, Z., Zhou, L., Zhou, M., 2013. Three-dimensional electrochemical process for wastewater treatment: A general review. Chem. Eng. J. 228, 455−467. https://doi.org/10.1016/j.cej.2013.05.033.

Zheng, G., Ariffin, M.K.A.B.M., Ahmad, S.A.B., Aziz, N.B.A., Xu, W., 2022. A novel progressive response surface method for high-order polynomial metamodel. J. Phys. Conf. 2224(1), 12045. https://doi.org/10.1088/1742-6596/2224/1/012045.

Please cite this article as: Ganthavee, V., Trzcinski, A.P., Superior decomposition of xenobiotic RB5 dye using three-dimensional electrochemical treatment: Response surface methodology modelling, artificial intelligence, and machine learning-based optimisation approaches, Water Science and Engineering, https://doi.org/10.1016/j.wse.2024.05.003