SUPERIOR DECOMPOSITION OF XENOBIOTIC DYES AND PHARMACEUTICAL CONTAMINANTS IN WASTEWATER USING RESPONSE SURFACE METHODOLOGY, ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR OPTIMISATION OF A NOVEL THREE-DIMENSIONAL ELECTROCHEMICAL TECHNOLOGY

A Thesis submitted by

Voravich Ganthavee
(BE/BPharmSc and MSc QS)

For the award of

Doctor of Philosophy

2024

Rapid global urbanisation and industrialisation have led to the widespread production of emerging anthropogenic contaminants such as xenobiotic dyes and pharmaceutical pollutants discharged into our natural waters. These pollutants are recalcitrant to environmental degradation and often escape into water from industrial effluent systems. In this work, a novel graphite intercalation compound (GIC) particle electrode was used to investigate the adsorption of synthetic dye pollutant, Reactive Black 5 (RB5), using a three-dimensional electrochemical reactor to decompose the anthropogenic dye pollutant. Various adsorption kinetics and isotherm models were used to characterise the adsorption phenomena of GIC and determine the viability of the sorption process. When coupled with electrochemical oxidation technology, remarkably high dye removal efficiency can be achieved, and GIC can be electrochemically regenerated. Optimisation studies were conducted using response surface methodology and ANOVA analysis to provide insight into the significance of selectivity reversal from the salting effect of xenobiotic textile dye on GIC adsorbent. Non-linear models were simulated using the kinetic data in the order: Elovich > Bangham > Pseudo-second order > Pseudo-first order. The Redlich-Peterson isotherm was calculated to have a dye-loading capacity of 0.7316 mg/g by non-linear regression analysis. A range of error function analyses were used to evaluate the accuracy and precision of regression models. The best dye removal efficiency achieved using three-dimensional electrochemical treatment was approximately 93% using a current density of 45.14 mA/cm^2, whereas the highest total organic carbon (TOC) removal efficiency was 67%. Various advanced artificial intelligence (AI) and machine learning (ML) optimisation techniques were used to enhance the prediction efficiency of dye and total organic carbon (TOC) removal efficiencies. The AI/ML optimised decolourisation efficiencies were 99.30%, 96.63% and 99.14% using central composite design-novel progressive response surface methodology (CCD-NPRSM), hybrid artificial neural network-eXtreme boosting gradient (ANN-XGBoost) ensemble, and classification and regression trees (CART), respectively. The prediction efficiency of optimised models ranked in the descending order of hybrid ANN-XGBoost, CCD-NPRSM and CART. The ANOVA results revealed that hybrid ANN-XGBoost ensemble yielded a mean square error (MSE) and coefficient of determination (R^2) of 0.014 and 0.998, outperforming CCD-NPRSM and with MSE and R^2 of 0.518 and 0.998. The overall result showed that the hybrid ANN-XGBoost approach is the most feasible technique for improving the prediction efficiency of RB5 dye wastewater decolourisation.

I Voravich Ganthavee, declare that the PhD Thesis entitled *Superior decomposition of xenobiotic dyes and pharmaceutical contaminants in wastewater using response surface methodology, artificial intelligence and machine learning for optimisation of a novel three-dimensional electrochemical technology* is not more than 100,000 words in length, including quotes and exclusive of tables, figures, appendices, bibliography, references, and footnotes.

This thesis is the work of Voravich Ganthavee except where otherwise acknowledged, with most of the contribution to the papers presented as a Thesis by Publication undertaken by the student. The work is original and has not previously been submitted for any other award, except where acknowledged.

Date: 02 August 2024

Endorsed by:

Dr. Antoine P. Trzcinski
Principal Supervisor

Dr. Sreeni Chadalavada
Associate Supervisor

Student and supervisors' signatures of endorsement are held at the University.

The following detail outlines the agreed share of contribution for candidate and co-authors in the published and submitted papers presented in this thesis:

Paper 1:
Voravich Ganthavee and Antoine P. Trzcinski (2023). Removal of pharmaceutically active compounds from wastewater using adsorption coupled with electrochemical oxidation technology: A critical review. *Journal of Industrial Engineering and Chemistry*, 20-35. https://doi.org/10.1016/j.jiec.2023.06.003 (Published)

The overall contribution of **Voravich Ganthavee** was 80% to the Conceptualization, Visualization, Validation, Investigation, Formal analysis, Data curation, Writing – Original Draft and revising the final submission; **Antoine P. Trzcinski**: Supervision, Project administration and Resources. Overall, his contribution to this manuscript is 20%.

Paper 2:
Voravich Ganthavee and Antoine P. Trzcinski (2024). Artificial intelligence and machine learning for the optimization of pharmaceutical wastewater treatment systems: a review. *Environmental Chemistry Letters*. https://doi.org/10.1007/s10311-024-01748-w (Published)

The overall contribution of **Voravich Ganthavee** was 90% to the Conceptualization, Visualization, Validation, Investigation, Formal analysis, Data curation, Writing – Original Draft and Revising the Final Submission; **Antoine P. Trzcinski**: Supervision, Review and Editing, Project Administration and Resources. Overall, his contribution to this manuscript is 10%.

Paper 3:
Ganthavee, V., Trzcinski, A.P. (2024). Removal of reactive black 5 in water using adsorption and electrochemical oxidation technology: kinetics, isotherms and mechanisms. *Int. J. Environ. Sci. Technol*. https://doi.org/10.1007/s13762-024-05696-4 (Published)

The overall contribution of **Voravich Ganthavee** was 80% to the Conceptualization, Visualization, Validation, Investigation, Formal analysis, Data curation, Writing – Original Draft and Revising the Final Submission; **Antoine P. Trzcinski**: Supervision, Review and

Editing, Project Administration and Resources. Overall, his contribution to this manuscript is 20%.

Paper 4:
Voravich Ganthavee and Antoine P. Trzcinski (2024). 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 (Published)

The overall contribution of **Voravich Ganthavee** was 80% to the Conceptualization, Visualization, Validation, Investigation, Formal analysis, Data curation, Writing – Original Draft and Revising the Final Submission; **Antoine P. Trzcinski**: Supervision, Review and Editing, Project Administration and Resources. Overall, his contribution to this manuscript is 20%.

Paper 5:
Voravich Ganthavee, Merenghege M R Fernando and Antoine P. Trzcinski (2024). Monte Carlo Simulation, Artificial Intelligence and Machine Learning-based Modelling and Optimization of Three-dimensional Electrochemical Treatment of Xenobiotic Dye Wastewater. *Environmental Processes*. https://doi.org/10.1007/s40710-024-00719-1 (Published)

The overall contribution of **Voravich Ganthavee** was 80% to the Conceptualization, Visualization, Methodology, Validation, Investigation, Formal analysis, Data curation, Writing – Original Draft and Revising the Final Submission; **Merenghege M R Fernando**: Investigation, Data Curation, Visualization; **Antoine P. Trzcinski**: Supervision, Review and Editing, Project Administration and Resources. Overall, their contribution to this manuscript is 20%.

Paper 6:
Voravich Ganthavee and Antoine P. Trzcinski (2024). Computational modelling of Indigo Carmine adsorption onto bone char: Application of Monte Carlo simulation, Bayesian networks, artificial intelligence and machine learning-based optimisation approaches. *Neural Computing and Applications*. (Submitted)

The overall contribution of **Voravich Ganthavee** was 80% to the Conceptualization, Visualization, Methodology, Validation, Investigation, Formal analysis, Data curation, Writing – Original Draft and Revising the Final Submission; **Antoine P. Trzcinski**: Supervision, Review and Editing, Project Administration and Resources. Overall, his contribution to this manuscript is 20%.

First and foremost, I acknowledge the traditional owners of the land on which I have lived and studied. This research has been supported by the Australian Government Research Training Program Scholarship. I would like to express my sincere gratitude to University of Southern Queensland and Australian Government for granting me a scholarship to study PhD. My heartfelt gratitude goes to my principal and associate supervisors, Dr Antoine P. Trzcinski and Dr Sreeni Chadalavada, for their unwavering support for my PhD study and, for their patience, dedication, knowledge, and enthusiasm. I am deeply grateful to my principal supervisor, who responded to my queries regularly. I would also like to thank my associate supervisor for his insight and valuable guidance. Much of my work would not have been published without the valuable advice from my principal and associate supervisors. I would also like to express my gratitude to Dr Susette Eberhard, Mr Adrian Blokland, Mr Brian Lenske and Mr Terry Byrne for supporting my experimental work. Their assistance was crucial for my PhD success.

Pursuing a PhD is challenging and demands unwavering determination, emotional resilience, self-discovery and exploration, intellectual prowess and self-dedication. It is not just about intellectual pursuit of knowledge but about personal growth and discovery. It is about cultivating a resilient mind, undaunted by the challenges, undermining criticisms and discouragements from others. The support that I received from my academic and work supervisors was crucial for my success. PhD is more than a path of academic achievement; it is a personal journey of self-directed learning driven by intrinsic motivation, aspirations and a strong will to succeed. It is a strong determination that supersedes setbacks and obstacles along the way. The will to succeed pushes me to grow personally and professionally, developing new skills and expanding my capabilities. Most importantly, it is a transcendental journey towards self-discovery; building a positive mindset is critical to success. It is about maintaining optimism in the face of challenges while embarking on a transformative journey towards self-fulfilment.

When accustomed to doing the hard thing, life becomes easy.
When accustomed to doing the easy thing, life becomes difficult.
It is a trade-off.
Hardship teaches you strength and perseverance like a river cut through rock.
We transcend to be fluid in our intelligence, yielding towards the flow of water, and accepting the true transformational power of our inner selves.
Water grows the plant, undeterred by the coldness of water, turning the discomfort into opportunities.

DEDICATED TO THE DIVINE WISDOM OF WATER

AC	Activated Carbon
AD	Adsorbent Dosage
ANN	Artificial Neural Network
AI	Artificial Intelligence
ANOVA	Analysis of Variance
CART	Classification and Regression Trees
CCD-NPRSM	Central Composite Design-Novel Progressive Response Surface Methodology
CECs	Contaminants of Emerging Concern
EAOPs	Electrochemical Advanced Oxidation Processes
EDCs	Endocrine-Disrupting Compounds
GIC	Graphite Intercalation Compound
IT	Information Technology
ML	Machine Learning
MOF	Metal-Organic Framework
XGBoost	Extreme Gradient Boosting
RB5	Reactive Black 5
RF	Random Forest
RSM	Response Surface Methodology
SVM	Support Vector Machine
TOC	Total Organic Carbon
TOCRE	Total Organic Carbon Removal Efficiency
UV	Ultraviolet
WWTPs	Wastewater Treatment Plants

1.1 Background

Global industrialization and rapid population growth increase the large-scale production of goods and services. Unprecedented levels of xenobiotics and anthropogenic pollutants are discharged or escaped into the aquatic environment. Due to chemically stable organic compounds, textile and pharmaceutical residue wastewater is notoriously challenging to degrade in industrial effluents (Liu et al., 2021). Contaminants of emerging concern (CECs), such as non-regulated xenobiotic dyes and pharmaceutical contaminants, are released daily into surface water (Saravanan et al., 2022). Xenobiotic dyes are often produced by textile, paper and pulp, leather, tannery and paint and dye manufacturing industries (Bilal et al., 2022). Non-regulated pharmaceutical substances and personal care products are endocrine-disrupting compounds (EDCs), pesticides and disinfection byproducts that can cause adverse impacts on human health and ecology (Samal et al., 2022). Some CECs are partially metabolized by organisms, rapidly bioaccumulating toxic metabolites released into the environment (Bosio et al., 2021). CECs are highly recalcitrant to environmental degradation. However, the toxicity levels related to the xenobiotic dyes have not been adequately investigated. Industries' frequent use of these dyes may have severe public health consequences if not adequately regulated. Dye pollutants produce intense water colouration and cause carcinogenic, teratogenic and mutagenic effects in humans and animals (Tang et al., 2020).

Moreover, the transformation of the synthetic dyes after decomposition may lead to the formation of toxic intermediate byproducts due to their reactivities with substances present in the environment (Xu et al., 2023). Dye-contaminated wastewater causes intense colouration, which indicates that toxic pollution needs to be eliminated (Obayomi et al., 2023). Otherwise, it can affect the aesthetic quality of the environment. Dyes in industrial effluents can cause considerable ecotoxicity and pose a significant risk to human health (Alrobei et al., 2021). Synthetic dyes often have complex chemical structures, such as substituted aromatic and heterocyclic groups, making them highly recalcitrant to environmental degradation (Nidheesh et al., 2018). The incomplete breakdown of these dyes may produce aromatic amines, a common intermediate byproduct derivative that is both highly toxic and recalcitrant to environmental biodegradation (McYotto et al., 2021).

Moreover, these synthetic dyes are also highly resistant to removal by conventional WWTPs. Common biological treatments, such as anaerobic-aerobic, may not effectively eliminate toxic dyes due to inadequate removal efficiency (Feng et al., 2022). Some dyes are

highly toxic to microorganisms. For this reason, the conventional biological treatment method is not recommended for dye treatment. Therefore, there is an urgent need to find a more suitable wastewater treatment method, such as three-dimensional electrochemical oxidation technology, which involves combined adsorption and electrochemical oxidation technology, to remove dye pollutants more effectively than conventional WWTPs. In addition, the developing world is under pressure to utilise this integrated three-dimensional electrochemical technology due to its high pollutant removal efficiency, cost-effectiveness, and ease of operability. (Liu et al., 2022). A synergistic effect of effective carbon adsorption and electrochemical oxidation can enhance the treatment efficiency (Liu et al., 2022). More interestingly, low-cost, green adsorption technology synthesized from renewable sources with strong electrocatalytic performance can help regenerate the carbon particle electrode (Liu et al., 2022).

Most importantly, additional chemical reagents are not required to help reduce the severity of secondary pollution (Yuan et al., 2022). When used appropriately with a novel adsorbent material such as GIC, the adsorbent can remove pollutants effectively and regenerate to its total adsorptive capacity when subjected to electrochemical regeneration (Sun et al., 2013). Compared to conventional activated carbon (AC), which has low electrical conductivity but a greater surface area, the activated carbon-containing adsorbed toxic contaminants may have to be incinerated or disposed of in a landfill, resulting in secondary pollution (Ye et al., 2022). Hence, AC adsorbent is less desirable than GIC unless its usage satisfies a particular adsorption regime without electrochemical treatment.

Furthermore, the influence of operating conditions on the performance of electrochemical oxidation technology is highly complex. Advanced optimisation techniques must be developed to control the operational conditions of three-dimensional electrochemical technology and other wastewater treatment technologies (Gadekar & Ahammed, 2019). Response surface methodology, artificial neural network (ANN), support vector machine (SVM), eXtreme gradient boosting (XGBoost), random forest (RF), classification and regression trees (CART) and Bayesian inference network are some of the artificial intelligence, machine learning and statistical optimization techniques used to enhance the prediction efficiency of pollutant removal efficiency by optimising for a range of operational variables and to determine the causal relationships among the random variables. Ensemble forecasting is a modelling approach combining data sources and different models to reduce uncertainties and enhance prediction efficiency, accuracy, and reliability of AI and ML optimization technologies (Wu & Levinson, 2021). The AI and ML ensembles can be used to

predict water quality to achieve safety of the drinking water supply system (Park et al., 2022). To ease the complexity of wastewater treatment systems, various artificial intelligence and machine learning ensembles were used to improve the intelligent systems and manage complex dynamics of mathematical models to optimise the operational conditions of the wastewater treatment systems effectively. The advanced AI and ML algorithms can be applied to existing operational management systems of wastewater treatment systems, forming critical components of advanced computing and software technology for enhancing dye and pharmaceutical wastewater treatment systems. With ever-growing issues of antimicrobial resistance genes, viral diseases and toxic, xenobiotic textile wastewater, future trends are forecasted to rely on developing more advanced AI and ML optimisation techniques to improve the process conditions. Since the overall research was centred around the three-dimensional electrochemical treatment, specific emphasis was given to extending current electrochemical treatment techniques towards integration of advanced AI ensembles to optimise the process conditions. In addition, Figure 1 shows that the textile processing industry in Bangladesh is one of the major polluters due to its widespread use of natural resources. Reactive dyes are commonly used for dyeing purposes. The dye combination is highly chemical intensive and requires large water consumption. Figure 1 shows the estimated amount of water consumed, wastewater generation, and chemical consumption in the textile wet processing units. High water consumption leads to depletion and contamination of groundwater.

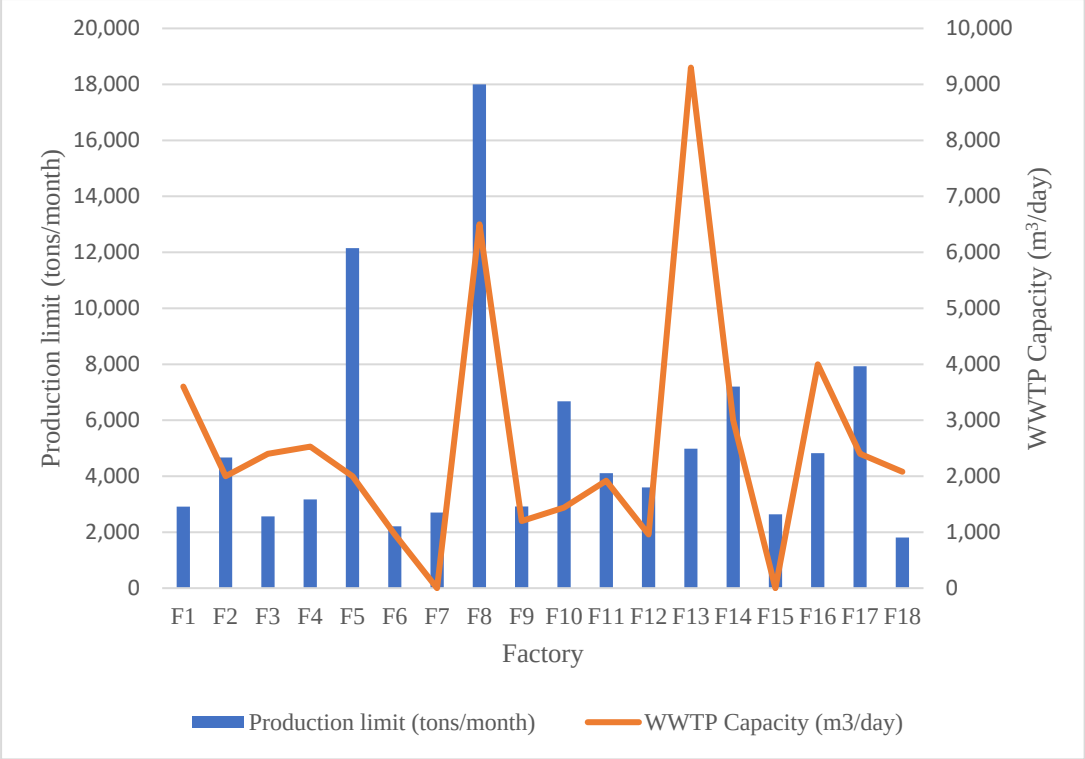


Figure 1. Water and chemical consumption of textile processing industry in Bangladesh (Uddin et al., 2023).

1.2 Significance of study

Xenobiotic dyes and pharmaceutical contaminants present in aquatic environment can pose a significant public health risk. These anthropogenic pollutants are increasingly present in the environment due to rapid urbanisation and industrialisation. Water pollution caused by textile and fashion industries often flies under the radar and is largely overlooked by environmental authorities. There are many treatment methods used to remove xenobiotic dyes and pharmaceutical pollutants from water and wastewater, such as conventional biological treatment (aerobic-anoxic-anaerobic), adsorption, advanced oxidation processes, Fenton's reagents, photocatalysis, membrane filtration, coagulation-flocculation, electrochemical treatment etc. Conventional activated sludge in biological treatment is not suitable due to excessive production of sludge, increasing costs associated with sludge management and transportation. In the context of advanced oxidation processes (AOPs), Fenton oxidation is an effective technology to degrade recalcitrant organic pollutants due to simplicity of process but conventional Fenton oxidation processes have some drawbacks, such as acidic pH condition, generation of iron sludge and high operational costs associated with high chemical input, low resource recovery process and energetically intensive (Bello et al., 2019).

On the other hand, coagulation-flocculation technique has exceptional performance in removing particles from wastewater effectively, causing small particles to clump together to

form flocs (Abujazar et al., 2022). The process requires the addition of chemical reagents to remove the pollutants by settlement. Using coagulants or flocculants is very costly; it also contributes to secondary pollution, and the process is energetically intensive due to limited energy recovery. Membrane filtration process, such as forward osmosis, membrane distillation and electrodialysis can advance nutrient recovery but often leads to membrane fouling, resulting in pressure drop across the membrane, and the process becomes more energetically intensive (Xie et al., 2016). The drawbacks of these conventional treatment techniquescan be addressed by using electrochemical oxidation technology, which is a more attractive alternative method for wastewater treatment to replace conventional processes owning to a substantial quantity of toxic organic pollutants removed from the industrial effluent systems, particularly xenobiotic dye wastewater. In addition, adsorption technology is a practical approach to remove contaminants from wastewater, particularly when the adsorbent is relatively inexpensive and easy to procure or manufacture. The adsorption process has minimal sludge generation, is less energetically intensive and does not require additional pre-treatment processes. The significance of a hybrid combination of wastewater treatment technologies is that they offer a significantly better performance in removing pollutants more effectively than using a single treatment method. Xenobiotic dyes and pharmaceutical contaminants are highly recalcitrant to environmental degradation and UV photolysis, leading to rapid accumulation in the aquatic environment. Hybrid wastewater treatment systems are more efficient at eliminating the toxic organic pollutants and its intermediate transformation byproducts from the breakdown of xenobiotic dyes and pharmaceutical contaminants more effectively than a single treatment method. When combining advanced AI and ML optimisation techniques to manage complex process parameters. Advanced computing and software technologies bolster the treatment efficiency of hybrid systems, facilitating data-driven analysis to minimise uncertainties and reducing fluctuations in effluent quality, costs and environmental risks (Zhang et al., 2023).

1.3 Research gaps

Our literature review identified the following research gaps:

1) The previous methods of optimisation techniques in past research are unreliable due to inherent uncertainties and error deviations in the estimated values, especially for pollutant removal efficiencies. Statistical optimisation techniques are not adequately accounted for due to the absence of a range of error function analyses applicable to adsorption kinetics and isotherms.

2) Advanced AI and ML ensembles have not been adequately covered in past research to enhance models' prediction efficiency, accuracy, and precision, especially for three-dimensional electrochemical technology. Most AI optimisation techniques, including RSM are applied to a simple adsorption or electrochemical oxidation process without a specific emphasis on using an ensemble to model a three-dimensional electrochemical oxidation technology.
3) Most RSM and AI/ML optimisation techniques were applied to a simple, conventional carbon-based adsorbent fabricated from an energetically intensive process. On the other hand, bone char adsorbent can be obtained from renewable and widely abundant agricultural sources. More interestingly, the pollutant removal efficiency of bone char can be subjected to AI and ML optimisations to improve the prediction efficiency, accuracy and precision of models.
4) The operational parameters used in the experiments were highly complex, requiring more novel RSM and/or AI and ML optimisation techniques. Past research has not accounted for any causal relationships between operational variables, including sensitivity analysis and Bayesian inference networks to account for the impact on the target variables.
5) Past research has not adequately accounted for the selectivity reversal, pH levels and salting effect of simulated, highly alkaline xenobiotic dye wastewater on the adsorptive capacity of GIC. In addition, the electrochemically regenerative performance of GIC was rarely emphasized in a three-dimensional electrochemical oxidation technology.

1.4 Research questions

There are seven main research questions in this study:

1) How can models' prediction efficiency, accuracy, and precision be enhanced to reduce uncertainties in the operational variables?
2) How to optimise the three-dimensional electrochemical oxidation technology and individual adsorption technology to enhance the pollutant removal rate?
3) How to electrochemically regenerate GIC adsorbent and improve the regeneration efficiency to recover its adsorptive capacity?
4) What are the optimal conditions to achieve the best pollutant removal efficiency using a three-dimensional electrochemical oxidation technology?

5) How to develop AI and ML ensembles, sensitivity analysis and other advanced statistical optimisation techniques to enhance the prediction efficiency, accuracy and precision of forecasting models?
6) How to optimise a number of experimental runs while accurately accounting for a range of process conditions to determine the most significant interactive effects of operational variables on targeted responses?
7) How can the thermodynamic conditions of xenobiotic dye wastewater be managed to improve the pollutant removal performance of bone char?

1.5 Research aims and objectives

The overall aim of this research project is to develop the best optimisation techniques to enhance the pollutant removal efficiency of a three-dimensional electrochemical reactor to achieve the most cost-effective manner to address the techno-economic aspects and energy efficiency of the wastewater treatment process for superior decomposition of xenobiotic dyes in aqueous solutions. The specific objectives are outlined below:

1) To optimise the electrical energy consumption of a three-dimensional electrochemical reactor using advanced AI and ML ensembles.
2) To optimise a range of operational variables to enhance targeted responses such as dye and TOC removal efficiencies to improve the mineralisation of toxic, xenobiotic dyes in wastewater.
3) To enhance the prediction efficiency, accuracy and precision of forecasting models using advanced RSM, AI and ML ensembles for a three-dimensional electrochemical oxidation technology and individual adsorption process.
4) To investigate the uncertainties and causal relationships between the operational parameters and conduct sensitivity analysis and Bayesian inference network analysis of the impact of interactive variables on the targeted responses.
5) To model the interactive effects of a range of operational variables on the targeted responses to enhance the electrochemical oxidation and adsorption efficiencies of a three-dimensional electrochemical reactor, GIC and bone char adsorbents.
6) To investigate the selectivity reversal, pH and salting effects of simulated, highly alkaline xenobiotic dye wastewater on adsorption.
7) To evaluate the adsorption kinetics and isotherm models using a range of error function analyses and/or statistical optimisation techniques.

8) To investigate the thermodynamic characteristics of xenobiotic dye wastewater on the pollutant removal efficiency of bone char adsorbent.
9) To investigate the regeneration efficiency of GIC adsorbent using a range of current densities, initial dye concentrations and electrolysis durations.

1.6 Organisation of the thesis

This PhD thesis consists of five chapters, and the organisation of the thesis is stipulated in Figure 1.

Chapter 1 provides the background information, purposes, and significance of the study, establishing a cohesive research narrative to guide the readers smoothly through the transition between ideas. It helps to resolve research gaps at the outset, minimise repetitions or tensions between the thesis components/chapters, and set a foundation to develop publication ideas that emerge from a comprehensive review of the whole thesis.

Chapter 2 represents the literature review for the study based on the contaminants of emerging concerns: **Paper 1** (Removal of pharmaceutically active compounds from wastewater using adsorption coupled with electrochemical oxidation technology: A critical review) presents the current three-dimensional electrochemical oxidation technology and its benefits in the mineralisation of xenobiotic pollutants in contaminated wastewater.

Chapter 3 represents **Paper 2** (Artificial intelligence and machine learning for the optimization of pharmaceutical wastewater treatment systems: a review) emphasized the benefits of artificial intelligence and machine learning optimisation of wastewater treatment plants to enhance the removal of specific anthropogenic pollutants such as pharmaceutically active compounds in wastewater. This review article sets a theme to justify using AI and ML optimisation techniques in wastewater treatment systems for various contaminants. These articles highlighted new insights and research gaps, setting a foundation for further development of this research project.

Chapter 4 represents **Paper 3** (Removal of reactive black 5 in water using adsorption and electrochemical oxidation technology: kinetics, isotherms and mechanisms), foundational experimentation with specific emphasis on using a combined adsorption and electrochemical oxidation technology, which is also known as a three-dimensional electrochemical oxidation technology, to remove xenobiotic dye contaminants from simulated textile wastewater synergistically. This research involved a critical investigation of the physicochemical properties of GIC adsorbent, using a range of adsorption kinetics and isotherms to evaluate the adsorption phenomena of the adsorbent. A range of error function analyses were

conducted to examine the uncertainty in the estimated values of adsorption kinetics and isotherms to achieve the accuracy and precision of calculations. The regeneration efficiency of GIC adsorbent was studied extensively. The impacts of selectivity reversal, salting effect and pH levels on GIC adsorption efficiency in a binary mixture of highly alkaline, simulated textile wastewater were investigated thoroughly. Electrochemical oxidation efficiency was evaluated using a range of current densities to enhance the mineralisation efficiency of xenobiotic dye pollutants.

Chapter 5 represents **Paper 4** (Superior decomposition of xenobiotic RB5 dye using three-dimensional electrochemical treatment: Response surface methodology modelling, artificial intelligence, and machine learning-based optimisation approaches), which strongly emphasized the significance of AI and ML algorithms or the roles of advanced software computing in the removal of xenobiotic dye pollutants from wastewater. Hybrid AI and ML optimisation techniques were applied to a three-dimensional electrochemical oxidation technology to manage a range of complex operational variables and achieve optimal conditions for enhancing pollutant removal efficiency. The benefits of using AI/ML ensembles include maximising the prediction efficiency of targeted variables, resulting in significant accuracy and precision of estimated variables.

Chapter 6 represents **Paper 5** (Monte Carlo Simulation, Artificial Intelligence and Machine Learning-based Modelling and Optimization of Three-dimensional Electrochemical Treatment of Xenobiotic Dye Wastewater), which investigated more advanced combination of AI and ML ensembles, such as integration of Monte Carlo simulations with artificial neural networks (ANN), support vector machine (SVM) and random forest (RF) algorithms generate various models for optimisation of three-dimensional electrochemical treatment of xenobiotic dye wastewater. Hybrid AI and ML optimisation techniques help to manage a range of complex operational variables by identifying the inherent system perturbations and estimation of uncertainties in predictive model platforms to achieve optimal conditions for enhancing the prediction efficiency of targeted variables. This resulted in better accuracy and precision of estimated variables to achieve optimal conditions for improving pollutant removal efficiency.

Chapter 7 presents **Paper 6** (Computational modelling of Indigo Carmine adsorption onto bone char: Application of Monte Carlo simulation, Bayesian networks, artificial intelligence and machine learning-based optimisation approaches), which involved using a green, renewable adsorbent material to remove xenobiotic dye from wastewater. The uniquely developed AI and ML ensembles, Monte Carlo simulations and Bayesian inference

network analysis were used to improve the predictive models of pollutant removal efficiency. This research used the unique optimisation techniques and computational fluid dynamics modelling to manage a range of operational variables by identifying the influential variables and its causal relationships, using sensitivity analysis to improve the adsorption technology. The levels of system perturbation and uncertainty were examined in detail to achieve optimal conditions of adsorption process.

Chapter 8 provides conclusions, future directions and final recommendations for this PhD research project.

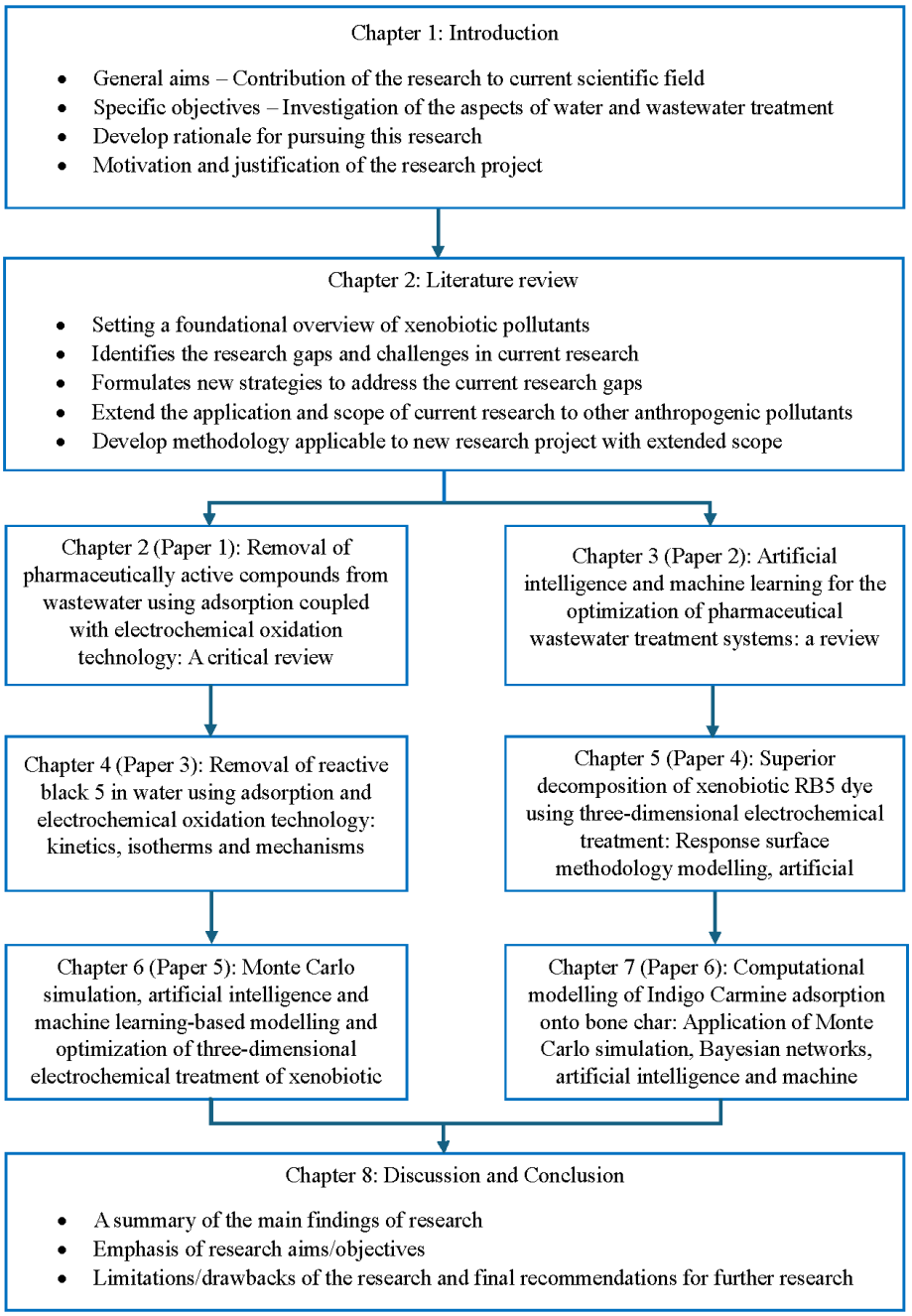


Figure 2. Organisation of the thesis.

This chapter is subdivided into two main sections corresponding to two review articles published during research. The first subsection provides a foundational overview of a

three-dimensional electrochemical oxidation technology to remove pharmaceutically active compounds from contaminated wastewater. The central emphasis is on using combined adsorption and electrochemical oxidation technology to remove a range of pharmaceutical contaminants by critically analysing a range of operating parameters to achieve optimal conditions for wastewater treatment. Understanding the effects of operating parameters on overall wastewater treatment efficiency helps to extend the scope of experimental studies towards more effective treatment of ubiquitous, xenobiotic dye pollutants. The second subsection emphasises the significance of using AI and ML ensemble optimisation techniques to enhance pollutant removal efficiency in existing wastewater treatment systems. This research helps to direct the focus on using a unique range of AI/ML ensembles to be included in future technical experimentation, albeit reactive dye pollutants will be the central focus of wastewater treatment. Overall, this chapter provides a comprehensive discussion, emphasizing the optimisation effects of operating parameters on targeted responses. More importantly, it sets a foundation for developing technical research, leading to final recommendations and future research directions.

CHAPTER 2: PAPER 1: REMOVAL OF PHARMACEUTICALLY ACTIVE COMPOUNDS FROM WASTEWATER USING ADSORPTION COUPLED WITH ELECTROCHEMICAL OXIDATION TECHNOLOGY: A CRITICAL REVIEW

2.1 Introduction

This review encompasses a thorough discussion and critical analysis of a specific wastewater treatment technology, i.e., a hybrid combination of adsorption and electrochemical oxidation technology. The inherent novelty of this hybrid wastewater treatment technology is the central theme that is largely emphasized throughout the research. This review was carefully crafted to explicitly analyse a range of suitable operating parameters to be adjusted to achieve the optimal conditions for electrochemical wastewater treatment. Pharmaceutical contaminants were considered as model pollutants rather than specific interest by the researchers. The potential application of this hybrid wastewater treatment technology was largely emphasized throughout the research, setting a foundation for future technical research to be conducted. Notwithstanding the chemical stability of pharmaceutical contaminants, a three-dimensional electrochemical oxidation technology involved a synergistic treatment of a range of pharmaceutical pollutants given that its mineralisation capability. A near complete decomposition of recalcitrant pharmaceutical contaminants in wastewater justified the technological capability of this hybrid wastewater treatment technology. Therefore, it is evidential that a three-dimensional electrochemical oxidation technology can be extended beyond its limits to treat other more ubiquitous contaminants such as xenobiotic dyes, albeit with strong certainty that the toxic dye pollutants can be degraded much more effectively than pharmaceutical contaminants. The effectiveness of using synergistic wastewater treatment techniques to eliminate pharmaceutical contaminants were largely emphasized throughout this review article. One of the main benefits include cost-effectiveness of adsorption technology used in the research. Others include the technoeconomic feasibility, practicality and environmental viability of the adsorption technology involving GIC adsorbent or a range of other adsorbent materials which can be obtained from agricultural or renewable sources. More significantly, electrochemical regeneration efficiency of nanocomposite adsorbent was critical to help recover the adsorptive capacity of adsorbent material, making it reusable for many cycles of adsorption and regeneration. On the other hand, the electrochemical oxidation technology is a more established electrochemical method for mineralising the pollutants but its electrooxidation efficiency can be enhanced using optimisation techniques to adjust operational variables such as current density, electrolysis time, pH level, adsorbent dosage, initial pharmaceutical concentration etc. The electrooxidation efficiency can also be enhanced using strong electrocatalytic anodic materials to improve mineralisation of pollutants. The electrocatalytic efficiency of anodes can be maximised using advanced electrode doping techniques and surface morphology tuning to help improve the advanced oxidation processes, resulting in rapid electro-generation

of powerful oxidants, such as hydroxyl and sulphate radicals, persulfate ions and active chlorine species to degrade pollutants effectively. More interestingly, the adsorption and electrochemical oxidation mechanisms for anodes were largely emphasized throughout the review article, highlighting the significance of improving electrocatalysis of electrochemical conversion or combustion of organic pollutants rather than just pharmaceutical contaminants. The differences between oxygen evolution reaction of non-active and active anodes and its effects on electrocatalytic activity were explained in great detail. The future directions and recommendations were provided at the end of the review article that emphasized the significance of renewable energy-driven electrochemical process. This renewable energy-driven electrochemical treatment technology can be applied to existing industrial wastewater treatment system where the adsorptive capacity of nanocomposite adsorbents can be continuously regenerated and electrooxidation efficiency of anodes can be enhanced to achieve cost-effectiveness of wastewater treatment process with robust pollutant degradation performance.

2.2 Links and implications

The central theme of this research is three-dimensional electrochemical oxidation technology. Although the research was initially conceived to apply a range of pollutants treated by three-dimensional electrochemical oxidation technology, the central focus was on the critical analysis of operating parameters applicable to a novel combination of adsorption and electrochemical oxidation processes. The technoeconomic feasibility, practicality and environmental viability of the three-dimensional electrochemical oxidation technology were largely emphasized throughout the review article rather than confining to a narrow focus on the realm of pharmaceutical wastewater treatment systems. There is currently a range of wastewater treatment technologies under consideration for treating pharmaceutical contaminants, but the current researchers have overlooked the potential benefits and broad application of three-dimensional electrochemical oxidation technology. There are numerous drawbacks associated with the current adsorption technology, such as poor reusability and recyclability of adsorbent materials.

Furthermore, the three-dimensional electrochemical oxidation technology offers an enhanced improvement in the mineralisation efficiency of organic pollutants. It allows the electrically conductive particle electrode to be regenerated, thereby recovering its adsorptive capacity to uptake pollutants continuously. The challenges associated with fabricating nanocomposite adsorbent material lie in finding electrically regenerative materials. The materials must be obtained from electrically conductive renewable sources to enhance the regeneration efficiency of adsorbents. The current thermal or chemical regeneration methods, such as incineration, solvent extraction, and landfill disposal of exhausted adsorbents, generate significant secondary pollution. The adsorbent fabrication technique leads to high energy consumption. Limited adsorptive capacity and regenerative properties of particle electrodes could downgrade its usage. Fine-tuning of nanoengineered novel metal-organic framework (MOF)-based adsorbent materials with defective structure can improve the adsorptive capacity and regeneration efficiency. The synergistic effects of adsorption and electrochemical oxidation technology can enhance pollutant removal and mineralisation efficiencies. When combined with renewable energy-driven electrochemical processes, energy efficiency can be achieved. However, there are challenges in adapting the existing electricity grid to power renewable energy-driven electrochemical advanced oxidation technologies.

CHAPTER 3: PAPER 2: ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR THE OPTIMIZATION OF PHARMACEUTICAL WASTEWATER TREATMENT SYSTEMS: A REVIEW

3.1 Introduction

This review article focuses explicitly on optimising pharmaceutical wastewater treatment systems, emphasising artificial intelligence and machine learning technologies. A conventional wastewater treatment system is highly complex due to numerous operational parameters that require effective control to improve water quality characteristics. This review article intends to bridge the gaps between AI/ML technologies and other critical advanced computing and information technology (IT) infrastructures such as blockchain technology, renewable energy, Big Data mining, cyber-physical systems, Internet of Things and automated smart grid power distribution networks. The combined advanced AI and IT computing techniques help to monitor fluctuations in contaminants in wastewater treatment plants, facilitating data analysis, diagnosing water quality and predicting process parameters. The AI/ML applications in pharmaceutical wastewater treatment systems are strongly highlighted throughout the review article, providing background information on numerous potential applications of AI technologies that can be extended beyond the realm of pharmaceutical wastewater treatment towards other treatment systems. Compared to pharmaceutical wastewater treatment, AI/ML technologies are equally applicable to textile dye wastewater treatment. This review article emphasises the usefulness of AI and advanced computing systems rather than just focusing on the process units of pharmaceutical wastewater systems.

Furthermore, the advancement of AI technologies leads to a remarkable transformation of conventional wastewater treatment plants towards zero waste generation, which is one of the most ideal pathways to achieve a circular economy, where value engineering and management processes can be added to the existing wastewater treatment systems to improve the outcomes. Similar to the toxicity of xenobiotic dyes, emerging pharmaceutical contaminants are equally culpable in causing adverse impacts on human health and environment. When pharmaceutically active compounds are discharged into the environment, the uptake of these compounds into human body can produce toxic metabolites, which have a wide range of side effects on non-target aquatic organisms, even at small concentrations. More critically, toxic drug metabolites can lead to multiple resistant strains or

antimicrobial-resistant genes and endocrine-disrupting intermediate compounds from the breakdown of parental compounds, causing significant carcinogenicity, teratogenicity and mutagenicity in humans and aquatic organisms.

Industrial effluent systems do not just contain pharmaceutically active compounds released into the natural waters; xenobiotic dyes, disinfection byproducts, personal care products, per- and poly-fluoroalkyl substances are among the toxic contaminants within the wastewater. Notwithstanding the narrow scope of pharmaceutical wastewater treatment systems, AI and ML optimisation technologies can be applied to other wastewater treatment contexts. For the readers' interest, the review article mainly focuses on using AI/ML optimisation techniques in pharmaceutical wastewater treatment systems to highlight its applicability. Numerous AI/ML algorithms can be applied to other process units within the pharmaceutical wastewater treatment systems to empower sustainable circularity, digital twin and intelligent data-driven operations, process control systems, and to support predictive platforms to achieve energy efficiency and minimise the spread of infectious diseases. AI/ML technologies can be used to predict and monitor wastewater quality characteristics such as chemical oxygen demand, biochemical oxygen demand, total suspended solids, total dissolved oxygen, total dissolved solids and many more. However, the cost of setting up complex computation infrastructure to facilitate AI systems is a financial impediment. AI systems require compatible hardware and software integrated into the wastewater management systems and other computational systems for proper functioning of AI. Deploying hardware and software systems into existing wastewater management systems results in large energy consumption due to the high demand for computational processing power. However, the complexity of AI infrastructure can be managed through optimisation and automation, but the debugging and troubleshooting of the process control systems can be a significant issue. Effective monitoring of process dynamic conditions requires advanced IT-powered technologies to facilitate data management of process control systems. Applying blockchain-related technologies helps facilitate sustainable wastewater and energy management systems. IT security vulnerabilities can be adverted when combined with the technological capabilities of the Internet of Things and advanced cyber-physical infrastructure. AI-powered process control systems help to minimise carbon footprint and remove barriers to resource recovery and energy management processes. Most interestingly, AI predictive platform improves models for measuring pharmaceutical wastewater quality and its constituents in complex process dynamic environment, minimising the operational cost and significantly improving the energy efficiency of wastewater treatment plants.

Environmental Chemistry Letters
https://doi.org/10.1007/s10311-024-01748-w

REVIEW ARTICLE

Artificial intelligence and machine learning for the optimization of pharmaceutical wastewater treatment systems: a review

Voravich Ganthavee[1] · Antoine Prandota Trzcinski[1]

Received: 14 November 2023 / Accepted: 2 May 2024

Abstract
The access to clean and drinkable water is becoming one of the major health issues because most natural waters are now polluted in the context of rapid industrialization and urbanization. Moreover, most pollutants such as antibiotics escape conventional wastewater treatments and are thus discharged in ecosystems, requiring advanced techniques for wastewater treatment. Here we review the use of artificial intelligence and machine learning to optimize pharmaceutical wastewater treatment systems, with focus on water quality, disinfection, renewable energy, biological treatment, blockchain technology, machine learning algorithms, big data, cyber-physical systems, and automated smart grid power distribution networks. Artificial intelligence allows for monitoring contaminants, facilitating data analysis, diagnosing water quality, easing autonomous decision-making, and predicting process parameters. We discuss advances in technical reliability, energy resources and wastewater management, cyber-resilience, security functionalities, and robust multidimensional performance of automated platform and distributed consortium, and stabilization of abnormal fluctuations in water quality parameters.

Keywords Algorithm · Cyber-security · Big data · Automation · Internet of things · Blockchain technology

Introduction

Rapid urbanization and population growth across the world have led to the widespread production of emerging contaminants, which puts significant pressure on wastewater treatment systems. Water scarcity drives our focus towards achieving maximum resource recovery. Zero waste generation is one of the ideal pathways towards achieving a circular economy, which brings remarkable transformation of wastewater treatment systems through commercialization by adding value management processes (Matheri et al. 2022). If left untreated and discharged from conventional wastewater systems, emerging pharmaceutical contaminants in aquatic or marine ecosystems can adversely impact human health and the environment (Osman et al. 2023; Priya et al. 2022).

When pharmaceutically active compounds are released into the environment through human metabolites, it can cause a wide range of side effects on non-target aquatic organisms even at minute concentrations (ng/L or µg/L). This leads to the development of multi-resistant strains and the formation of endocrine-disrupting chemicals from breakdown of intermediate by-products from parental compounds, causing significant carcinogenicity, mutagenicity, and teratogenicity in humans and aquatic organisms (Zhan et al. 2019). Among these pollutants, the types of contaminants that have become increasingly challenging to treat are pharmaceuticals and personal care products, disinfection by-products, and per- and poly-fluoroalkyl substances. In addition, wastewater treatment systems are highly complex and dependent on different environmental factors. Process parameters are optimized to tailor the control systems to improve the efficiency of wastewater treatment processes.

Although industrial and anthropogenic activities have introduced significant amounts of impurities and hazardous pollutants into our environment, several methods have been developed to minimize the effects of water pollution. These methods have its own merits in terms of the levels of water treatment quality and its varying effects on the environment. The treatment methods proposed by other researchers include coagulation–flocculation (Kooijman et al. 2020), membrane filtration (Ganiyu et al. 2015), ion

✉ Voravich Ganthavee
Voravich.Ganthavee@unisq.edu.au

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

Published online: 21 May 2024

exchange (Swanckaert et al. 2022), desalination (Shah et al. 2022), and biological treatment (Singh et al. 2023).

For biological treatment, the parameters used to characterize the levels of water treatment quality include biological oxygen demand and chemical oxygen demand. However, the conventional wastewater treatment used to purify or disinfect the wastewater is time-consuming and requires lengthy or arduous procedures (Safeer et al. 2022). To ease the complexity of wastewater treatment systems, artificial intelligence and machine learning algorithms are used to improve the intelligent systems and manage complex dynamics of mathematical models to effectively optimize the operational conditions of the wastewater treatment systems (Oruganti et al. 2023).

Currently, artificial intelligence and machine learning algorithms have been widely integrated into the existing operational management system of wastewater treatment plants, improving the water quality monitoring system (Chawishborwornworng et al. 2023), accuracy, and precision of model prediction (Serrano-Luján et al. 2022) and maximizing optimization efficiency of the process parameters (Zhang et al. 2023a). On the other hand, the theoretical or computational models developed for conventional wastewater treatment systems are overtly simplified based on the ideal assumptions rather than the real-world applicability of process models to make it practical for industrial purposes (Safeer et al. 2022).

Although empirical and statistical regression analyses are developed to predict the behaviour of process control systems, the complexity of real-world process dynamics and deviation in the non-linearity of regression models affect the accuracy of prediction (Özdoğan-Sarıkoç et al. 2023). Artificial intelligence can be incorporated into pharmaceutical wastewater treatment plants integrated with renewable energy technologies to forecast energy efficiency and offer advanced analytics for optimal energy management of pharmaceutical wastewater treatment systems.

Figure 1 depicts the critical components of advanced computing and software technology for improving pharmaceutical wastewater treatment systems. With the ever-growing issues of antimicrobial-resistant genes and viral diseases, future trends are forecasted to rely on developing more advanced artificial intelligence and machine learning algorithms to optimize the process conditions. This review is divided into five main topics encompassing the artificial intelligence applications in managing big data, strengthening cyber-physical systems, blockchain technology, and internet of things to improve the disinfection performance of pharmaceutical wastewater treatment systems.

Assessment of water quality

In the era of digital health and artificial intelligence, the challenges and perspectives for the future of electrochemical technologies, epidemiology, and interdisciplinary research can be bridged, unleashing the power of artificial intelligence and machine learning algorithms in diagnosing and treating infectious diseases (Tang and Cao 2023) and other antimicrobial-resistant genes developed from issues associated with water sanitation and environmental pollution, advancing both health informatics, precision medicine, and toxicogenomics related to improvement in water quality assessment of pharmaceutical wastewater effluent.

More interestingly, artificial intelligence and machine learning algorithms empower sustainable circularity, digital twin, and intelligent data-driven operations and process control systems, improving data mining, analysis, and

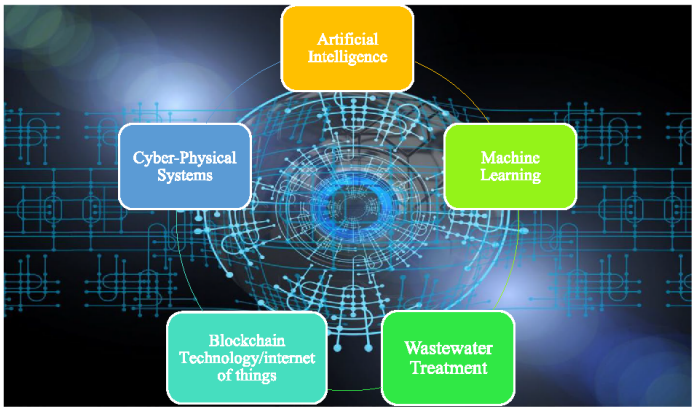


Fig. 1 Advanced computing and software technology allow to enhance technical reliability, cyber-resilience, energy resources management, and water quality in pharmaceutical wastewater treatment systems

prediction to support policymaking to achieve a circular economy and enhance energy efficiency, life cycle environmental and cost management technologies (Matheri et al. 2022; Osman et al. 2024). Artificial intelligence/machine learning algorithms can also be used to optimize complex process dynamics and non-linearity, using artificial neural network and adaptive neuro-fuzzy inference system and support vector machine software interfaces and other intelligent systems to assess water quality by predicting chemical oxygen demand, biochemical oxygen demand, total suspended solids, total dissolved solids concentrations in pharmaceutical wastewater (Safeer et al. 2022).

The artificial neural network was among the first machine learning algorithms developed based on perceptron (Park et al. 2022). An artificial neural network's model structure comprises three layers: input, hidden, and output. The hidden layer is a critical structure of an algorithm made up of nodes. Each node calculates the output variable for a series of steps using a nonlinear function called the activation function (Park et al. 2022). An increase in the number of hidden layers results in more complicated calculations due to additional predictions from input parameters. However, the problems associated with hidden layers are due to overfitting the training data and diminishing gradients during the optimization of the models (Jariwala et al. 2023).

To address this deficiency, a deep learning algorithm is used as an alternative function involving a rectified linear unit instead of a conventional sigmoidal function to minimize the problems associated with the vanishing effect of gradient. However, before the development of neuronal networks such as autocoders, feed-forward neural network, convolutional neural networks (Muniappan et al. 2023), recurrent neural network, and so on, there were various setbacks in artificial neural network architecture that needed to be explored.

The first significant issue associated with artificial neural network architecture is the non-existence of rules for defining neuronal network structures (Jariwala et al. 2023). The appropriate artificial neural network architecture design can be obtained through trial-and-error experience. This makes the process of developing artificial neural network architecture increasingly tedious.

Secondly, the artificial neural network architecture is hardware-dependent, which means the parallel processing power in computation becomes problematic because it is limited by the hardware properties (Jariwala et al. 2023). Hence, translating mathematical problems into numerical information leads to more issues related to artificial neural network architecture. This phenomenon involves unexplained network behaviour, constituting a probing solution and eventually leading to a fourth issue. The underlying issue associated with probing solutions is due to artificial neural network's justification and reliability, which may breach users' trust within the network.

When dealing financially with pharmaceutical companies, artificial intelligence and machine learning play a significant role in all aspects of drug discovery, wastewater treatment, and technological development processes. During wastewater treatment, the application of artificial intelligence can minimize the utilization of manpower and considerably reduce the expenditure on capital investment and maintenance costs related to treatment methods used.

On the other hand, the setbacks of wastewater treatment infrastructures can be attributed to the setting up of artificial intelligence infrastructures and computation technologies involving complex process control systems to improve the water quality at the output processes. There are several setbacks involved (Jariwala et al. 2023):

- The cost of setting up complex computation infrastructure to facilitate artificial intelligence systems becomes a financial impediment to small wastewater treatment industries and pharmaceutical firms. The requirement to install compatible hardware and software into the existing computational systems for proper functioning of artificial intelligence incurs significant financial expenditure.
- The speed of artificial intelligence algorithms affects the data processing power when it comes to accessing the data in real time to perform analysis and facilitate decision-making processes. Slow processing power and prolonged latency lead to undesirable consequences, resulting in delayed project timeline.
- Minimization of energy consumption is an important agenda when integrating compatible hardware with existing systems to deploy artificial intelligence technology. New integration systems with optimization modes to reduce power consumption would ease the financial burden on the business and wastewater treatment industry.
- The complexity of the artificial intelligence infrastructure can be managed through optimization and automation. Artificial intelligence technology can debug and troubleshoot any issues that arise rather than increasing computational complexity.
- Artificial intelligence systems require enormous computational energy to process and analyse data. Computational power grows immensely as the data grows, requiring algorithms to manage the voluminous data and minimize power consumption.
- Regular auditing and testing of machine learning models to improve the integrity of algorithms would help to streamline the deployment of artificial intelligence technology. This requires a diverse team of technical experts and personnel.

The optimization of analytical process conditions is significant. The characteristics and trace origins of water pollutants can be identified using unique artificial intelligence systems called the integrated long short-term memory network involving cross-correlation and association rules (Apriori) (Wang et al. 2019b).

Firstly, internet monitoring systems can acquire critical information about the pollutant sources entering the pharmaceutical wastewater treatment systems. The complex information on pollution incidents involving flow simulations, number of point sources at influent and effluent systems, and pollutant release processes can be interpreted using long short-term memory (Wang et al. 2019b). This method is computationally efficient because it deals with an artificial intelligence algorithm using a time-recursive neural network to predict critical events such as long intervals and delays of water pollutants on influent and effluent systems.

In addition, a convolutional long short-term memory provides a framework for sequencing learning problems using training data temporally to evaluate or predict the water quality pollutants in effluent systems (Wang et al. 2019b). However, there is currently a lack of robust mathematical expressions to correlate the measured parameters such as biochemical oxygen demand and chemical oxygen demand, total suspended solids, ammonia, organic nitrogen, and organic phosphorus content of wastewater, in which the data can only be obtained using online sensors. There are also uncertainties or perturbations in predicting biochemical oxygen demand and chemical oxygen demand values.

For this reason, integrating other artificial intelligence methods, such as gene expression programming and Monte Carlo simulation technique, can provide insights into estimating the levels of uncertainty or perturbations in wastewater process conditions. These techniques assess the sensitivity of target parameters and its influences on the variations in input parameters and scrutinize the interactions between the process parameters to evaluate the wastewater quality parameters (Aghdam et al. 2023). However, the online-based optimization technique has not been adequately applied to the bio-processing system due to the complexity of the biological behaviour.

Furthermore, the lack of data visualization techniques, low-quality industrial measurement systems, and understanding of underlying phenomena in wastewater treatment plants are ongoing issues. The data modelling approaches can be strengthened using artificial neural network, Gaussian process regression (Yao et al.), and polynomial chaos expansion to analyse the meta-models of wastewater treatment plants efficiently.

In addition, other artificial intelligence application tools such as expert systems (Wu et al. 2021), fuzzy logic (Mazhar et al. 2019), artificial neuro-fuzzy inference systems (Nam et al. 2023), support vector machine (Zhang et al. 2023b), knowledge-based systems (Liu et al. 2023), ruled-based systems (Victor et al. 2005), fuzzy logic control (Santín et al. 2018), pattern recognition (Gao et al. 2023), swarm intelligence (Negi et al. 2023), genetic algorithm (Aparna and Swarnalatha 2023), reinforcement learning (Wang et al. 2023a), hybrid systems (Tariq et al. 2021), and so on have gained its purposes in process control systems and prediction of water quality characteristics.

In addition, poor wastewater quality often leads to membrane fouling of filtration technologies used in the pharmaceutical wastewater treatment industry. Membrane fouling is a major obstacle hindering the widespread application of anaerobic membrane bioreactors to treat pharmaceutical wastewater (Niu et al. 2023).

Artificial intelligence algorithms and its modelling framework can predict membrane fouling phenomena in membrane filtration technologies using hyper-parameter optimization of artificial neural network and random forest to improve predictive capabilities (Niu et al. 2023; Yuan et al. 2023). In addition, artificial neural network and Bootstrap methods enhance the accuracy, robustness, and reliability of prediction tools to estimate the water quality indexes (Chawishborwornworng et al. 2023).

However, bootstrap programming adds a significant number of codes into the network, limiting the performance and processing speed of the software management system. On the other hand, the combination of artificial neural network and bootstrap algorithms improves the estimation of prediction error distributions, making it easier to analyse any faults or anomalies within the wastewater treatment systems (Mo et al. 2024).

In contrast, the main disadvantages of hybrid artificial intelligence models are complicated design constraints and uncertainties in predicted data arising from data clustering, making it challenging to discern exact data patterns to achieve optimal forecasting (Tikhamarine et al. 2020). For example, seasonal variation influences the wastewater streamflow and effluent quality; it is rather challenging to forecast the hydrological streamflow due to uncertainties in prediction (Ibrahim et al. 2022). In addition, Fig. 2 shows the artificial intelligence optimization framework applicable to various calculation tools for evaluating and predicting pharmaceutical wastewater treatment quality.

Overall, we observed that artificial intelligence applications in complex biological wastewater treatment systems are still developing, which could trigger severe and undesirable problems. Integrating artificial intelligence technologies may lead to system-wide compromise due to incompatibility with existing operational systems, a cascade of design errors, malfunctions, and possible cyber-attacks leading to other critical infrastructure failures, causing havoc in ecological systems and service availability to local communities. Hence, software and

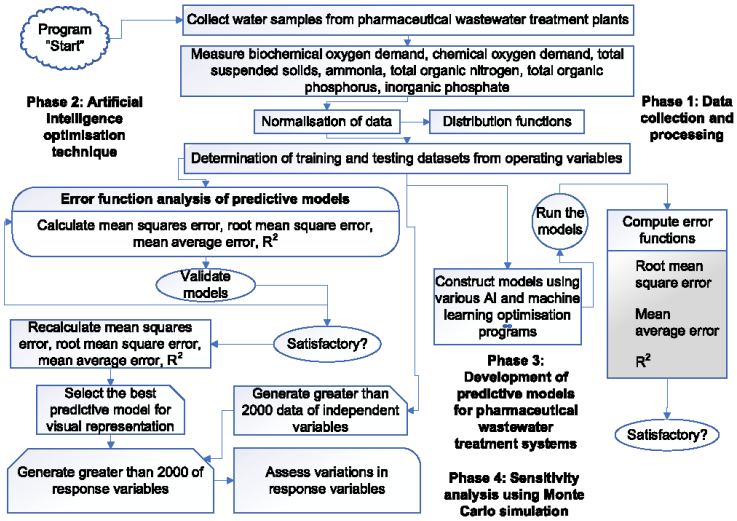


Fig. 2 Artificial intelligence research optimization framework for predicting pharmaceutical wastewater quality. Many data and model parameters can be arduous and challenging to manage, such as high-dimensional space and complex process control systems that require a powerful framework to assist computational resources in improving model performance, calibration, and optimization techniques. Error function analysis can determine the quality of predictive model performance to simulate process conditions and behaviour of pharmaceutical wastewater treatment systems

hardware functions of artificial intelligence technologies must account for the systemic risk and benefits of integrating advanced cyber-physical systems, data security infrastructure, blockchain technology, and the internet of things to achieve a robust system.

Disinfection

The concentration of disinfection by-products and severe acute respiratory syndrome coronavirus 2-related pharmaceuticals in wastewater effluents and surface water in aquatic environment impact the orchestration of coronavirus disease-19 pandemic. For example, a significant increase in concentrations of disinfection by-products such as trihalomethanes and haloacetic acids in hospital and pharmaceutical wastewater effluents and surface water ranging from 5.9 to 21.7 μg/L from wastewater discharge points increased ecotoxicities in aquatic environment (Zhang et al. 2022).

Wastewater-based epidemiology is one of the most effective surveillance tools for examining the sources of transmission of bacteria, microorganisms, and coronaviruses such as severe acute respiratory syndrome coronavirus 2 in wastewater. However, significant research gaps exist in addressing the difficulties and challenges in detecting, monitoring strategies, remediation, and disinfection methods of viruses in pharmaceutical and general wastewater (Bhattacharya et al. 2023).

More critically, there is a lack of regulatory framework and compliance related to the integration of artificial intelligence and machine learning technologies into existing pharmaceutical wastewater treatment systems, the uncertainties in technology efficiency for disinfection performance of wastewater treatment systems, and the economic viability of the wastewater treatment infrastructure, public or socio-economical resistance which may hinder practical implementation of artificial intelligence and blockchain-related technologies in wastewater treatment systems.

Moreover, these barriers can be minimized through collaborative efforts and systematic approaches from regulators, policymakers, engineers, and social scientists to translate innovative information technology into solutions to improve water sustainability in wastewater treatment domains (Robbins et al. 2022). However, method development and validation are the most significant challenges of implementing artificial intelligence and machine learning in complex process dynamics in pharmaceutical wastewater treatment systems.

Method validation is extremely critical to obtaining high-quality data (Corominas et al. 2018). A simple installation of sensors and cohesive maintenance efforts for optimizing process control systems do not guarantee adequate data quality, regardless of high computational processing power of information technology infrastructure. Dynamic wastewater processes are often characterized by constant changes at many different real-time scales, spanning from seconds to years in terms of dynamic pH changes, plant configuration, layout arrangement, and construction periods, which are also

critical conditions to how the processes change over time (Corominas et al. 2018).

Moreover, it is not practical to obtain voluminous, computationally expensive, and complex datasets, including real-time process dynamics over a meaningful period that requires uncompromised high-quality data (Ye et al. 2020). A large dataset of repositories requires continuous validation and comparison with predictive models for optimization, monitoring diagnostic purposes and updating control algorithms, which may require intensive labour and maintenance. There is a lack of standardized information technology protocols for selecting and implementing specific data analytic techniques equivalent to industry standards.

More advanced data management techniques are required to combine existing process systems with artificial intelligence, blockchain-related technologies, the internet of things, and cyber-physical systems. A plethora of methods have been developed or assessed. Still, challenges related to objective comparison between different industry artificial intelligence technologies, regulatory guidelines, validation limitations at full-scale systems, limited active and real-time data optimization, information sharing content, and quality affect the implementation of knowledge generation and artificial intelligence applications (Zhao et al. 2020).

The complexity of operational management systems increases with transmission routes of influent connecting to the pharmaceutical wastewater treatment plants, which involve several point sources discharged from hospitals, isolation centres, quarantine centres, and public places. The metropolitan and municipal wastewater plumbing systems are a significant pathway for spreading severe acute respiratory syndrome coronavirus 2. Severe acute respiratory syndrome coronavirus 2 is a member of a large family of viruses called coronaviruses that can infect people and some animals, causing mild to moderate respiratory illness.

The influent wastewater contains substantial viral loads of severe acute respiratory syndrome coronavirus 2 ribonucleic acid, a molecule present in most infected living organisms. Different treatment phases involving primary, secondary, and tertiary treatment methods are required to disinfect the pharmaceutical wastewater thoroughly. In primary physical treatment, the large, suspended solids in wastewater act as a physical barrier in removing viral particles (Bhattacharya et al. 2023).

Moreover, in the secondary treatment of wastewater treatment plants, diverse biological methods, including activated sludge process, membrane bioreactor, moving bed biofilm reactor, sequencing batch reactor, treatment ponds, and so on, are used to remove organic matter and large suspended solids from pharmaceutical wastewater. However, applying conventional activated sludge in large-scale hospitals and pharmaceutical wastewater treatments poses high energy consumption required for aeration, capital, and operational costs.

Although membrane bioreactors are progressively replacing the conventional activated sludge treatment systems and may achieve better treatment potential, the major drawback is due to membrane fouling, energy cost associated with aeration, and gradual reduction in membrane permeability resulting in the pressure fluctuation and greater energy consumption which lead to reduced performance at sizeable industrial-scale operations (Werkneh and Islam 2023).

In the tertiary treatment phase, various organics, turbidity, phosphorus, nitrogen, and other pathogenic microorganisms are removed using coagulation, advanced oxidation processes, filtration technologies, ultraviolet treatment, ozonation, chlorination, adsorbent materials such as titanium dioxide, carbon nanotubes, and other nanomaterials to inactivate viruses in the conventional wastewater treatment plants. It was reported that free residual chlorine species at a concentration of 0.5 mg/L required a contact time of 30 min at pH lower than 8, and 2.19 mg/L of chlorine dioxide is recommended for complete inactivation of severe acute respiratory syndrome coronavirus 2 in pharmaceutical wastewater (Bhattacharya et al. 2023).

More interestingly, artificial neural network can be applied to forecast the chlorination behaviour in the secondary pharmaceutical wastewater effluent containing ammonia, nitrate, and other pharmaceutical constituents. Disinfection of hospital wastewater results in changes in microbiome, resistome, and mobilome of wastewater and other bacterial communities and reduction in specific antibiotic resistance genes (Akhil et al. 2021; Rolbiecki et al. 2023).

An advanced control scheme can be developed to optimize the chlorination disinfection quality by integrating an artificial neural network model with fuzzy logic control to improve the chlorination process and minimize the cost of disinfection as well as maximizing disinfection efficiency while keeping the plant's budget within reach (Khawaga et al. 2019). However, the degrees of disinfection provided by direct chlorination were comparable to those attained by combining the conventional activated sludge process and chlorine treatment at conventional wastewater treatment plants (Azuma and Hayashi 2021).

In addition, the integrated photocatalytic-biological wastewater treatment systems are effective alternative processes for the removal of emerging pharmaceutical contaminants and pathogens, capable of achieving greater than 99% removal of chemical oxygen demand and nitrogen from the system with total disinfection of 10^6 colony-forming units/mL *E.coli* using hydroxyl radicals generated from photocatalysis (Ghosh et al. 2023). Moreover, colony-forming units estimate the number of active and viable microorganisms

in a sample. Artificial neural network and adaptive neuro-fuzzy inference system can be used to model the photocatalytic degradation process and mineralization efficiency of pharmaceutical and other organic pollutants while optimizing energy consumption and catalyst dosage for practical pharmaceutical wastewater treatment (Tabatabai-Yazdi et al. 2021).

Overall, we observed that artificial intelligence techniques can monitor complex variations in process conditions and accurately predict the performance of wastewater disinfection characteristics. However, extreme fluctuations in wastewater quality parameters during complex, full-scale disinfection processes, conventional biological wastewater treatment system, and predictive disinfection models may not handle intricate non-linearity issues, and immediate responses to remediate the disinfection level may not be effective. However, hybrid artificial intelligence technologies integrated with robust cyber-security infrastructures, blockchain technology, and distributed network design with the internet of things can facilitate autonomous wastewater treatment processes, reducing undesirable risks caused by incomplete disinfection processes.

Renewable energy

Current researchers rarely consider using renewable energy technologies for pharmaceutical wastewater treatment. Although conventional wastewater treatment plants are designed primarily to remove undissolved and dissolved wastewater, they are crucial in controlling water pollution and offering sanitary engineering. The additional energy generation potential of conventional wastewater treatment plants involves the utilization of digested sewage sludge for incineration, and electricity generation can provide a significant amount of energy and resource recovery (Zahmatkesh et al. 2022).

In addition, the settling properties of activated sludge having a sludge volume index greater than 150 mg/L could be susceptible to sludge bulking, which hinders the operation of the activated sludge process. This process may result in a mass proliferation of filamentous bacteria, impacting the techno-economic feasibility of the pharmaceutical wastewater treatment systems. For this reason, artificial neural network is very effective at simulating the nonlinear processes of sludge bulking, especially in various fluctuating environmental conditions (Deepnarain et al. 2020).

On the contrary, full industrial-scale pharmaceutical wastewater treatment systems comprise different physical, chemical, and biological processes that are highly complex and challenging to model using a linear method. Artificial neural network and multivariate statistics involving principal component analysis can model and extract valuable information by being adaptive and developing self-learning ability to discern the influence of process parameters (Guo et al. 2018; Verma and Suthar 2018). However, traditional principal component analysis is still limited by linear dimensionality reduction (Wang et al. 2019a).

On the other hand, nonlinear projection of principal component analysis can be determined using Gaussian process mapping, but the model lacks robustness and is susceptible to process noise (Wang et al. 2019a). When combined with another artificial intelligence technology, the artificial neural network model can optimize the process parameters for more accurate and robust results than regression-based mathematical models (Deepnarain et al. 2020).

The onsite nutrient recovery process of pharmaceutical wastewater treatment plants in which waste materials can be reused for other industrial purposes is crucial. The control components, such as energy distribution systems and metallurgical phosphorus recycling, can utilize activated sludge from wastewater treatment systems and transform it into energy and mineral products (Zahmatkesh et al. 2022). Artificial intelligence-powered renewable technologies can increase the sustainability of energy use at pharmaceutical wastewater treatment systems and reduce electricity costs or financial expenditure for energy supply.

Onsite renewable energy sources, such as solar, water, wind, and so on, can help minimize energy wastage and greenhouse gas emissions, saving economic costs immensely. Optimization via artificial intelligence can help to reduce the environmental impact of the combined energy systems using genetic algorithm to lessen the effect of carbon dioxide emission on the environment (Hai et al. 2022). However, integrating microgrids with renewable energy technologies and sharing with external grid networks are very challenging due to maintaining optimum power flows in the industry (Fan and Li 2023).

Overall, we observed that direct solar energy-assisted wastewater treatment with energy storage systems makes it convenient during day and night. Still, the installation and maintenance costs to achieve robust system efficiency affect effective renewable energy utilization. The complexity of adapting the existing electricity grid to a distributed energy network to utilize renewable energy resources in pharmaceutical wastewater treatment systems is still in its infancy.

Biological treatment

Artificial intelligence is one of the popular machine learning-based approaches in biological wastewater treatment simulations due to its high level of adaptability and learning strategies. Artificial neural network can be used to model

or predict biochemical oxygen demand and total suspended solids in removing activated sludge (Li et al. 2023b). It can also simulate total nitrogen, total phosphorus, and chemical oxygen demand for real-time dynamics of process conditions in wastewater treatment plants (Zaghloul and Achari 2022). However, the training datasets require high computational power and may not apply to small-scale industrial plants.

Unlike artificial neural network, support vector machine provides a unique solution to multiple regression systems to minimize errors while generating extensive computational data to improve model accuracy and predictability. On the other hand, adaptive neuro-fuzzy inference system is a hybrid algorithm that integrates the adaptability and computational power of artificial intelligence with fuzzy logic's learning ability to manage uncertainty or perturbations in process dynamics (Zaghloul and Achari 2022).

More interestingly, the complexity of data generated from biological processes can be interpreted using the multidimensional non-linearity of adaptive neuro-fuzzy inference system, where the number of fuzzy rules increased exponentially in both functions and number of input parameters. However, the high sensitivity of biomass combined with an array of parameters and frequent changes in influent characteristics can affect the stability of the operation of aerobic granular sludge reactors.

Furthermore, the combination of adaptive neuro-fuzzy inference system and support vector regression to form a two-stage prediction process as separate algorithms can be trained for individual output parameters to provide greater flexibility in tuning the discrepancies in model prediction to minimize errors. Combining machine learning-based models such as feed-forward neural network, support vector machine, and adaptive neuro-fuzzy inference system for benchmarking pharmaceutical wastewater treatment systems has yielded efficient performance.

The additional combination of feed-forward neural network increases the effectiveness of identifying complex problems or patterns using multilayer non-linearity of machine learning tools to examine the input and output parameters (Jana et al. 2022). Moreover, the three layers of feed-forward neural network are trained with the Levenberg–Marquardt algorithm (Jana et al. 2022). The first layer consists of a flattened input vector containing various input parameters.

When combined with the auto-regressive characteristics of predictive modelling, lagged data are integrated into the input vector (Negi et al. 2023). The second layer consists of hidden neurons with nonlinear activation functions (Jariwala et al. 2023). The third layer represents the output vector, which compares the predicted values with input parameters to produce targeted responses (Nourani et al. 2023). Overall, we observed that artificial neural network is prone to computational overload when handling massive datasets, which can be a challenge for wastewater treatment industries to adopt due to limited computational processing power and is primarily hardware dependent.

In addition, Fig. 3 shows the standard feed-forward neural network architecture, which consists of three layers of a computational network. Furthermore, Table 1 shows 1 shows the equations used by researchers who applied various models including the process parameters and criticize each model based on their advantages and disadvantages. Table 2 lists recent work in which error functions were used to calculate and validate the performance of models describing

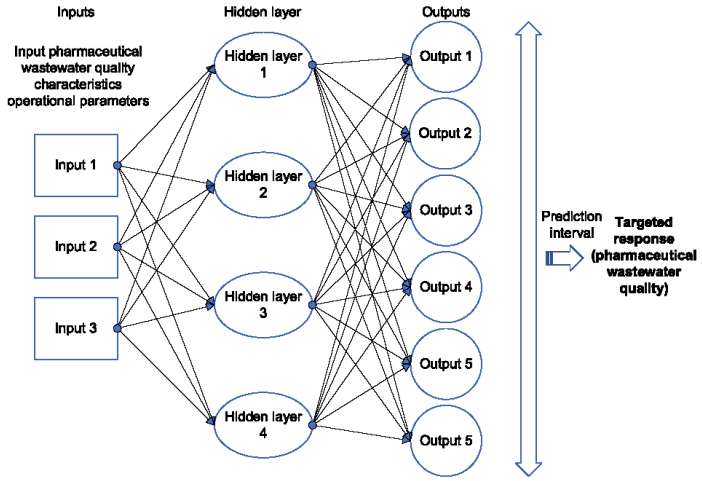


Fig. 3 Feed-forward neural network architecture involves a number of artificial neural network connections in which the flow of information is between nodes or its layers. The flow is usually in one direction or forward from the input nodes, passing through the hidden nodes to output nodes without any loops or cycles. Feed-forward neural network is trained using Marquardt's backpropagation method

Table 1 Standard functions of models applicable to pharmaceutical wastewater treatment systems

Transfer/activation function	Governing equations	Output Range	Process parameters	Advantages/disadvantages	References
Linear function	$f(x_i) = x_i$	–	pH 7.2–8.5; 48–145 mg/L (Total suspended solids); 480–1,000 mg/L (biochemical oxygen demand); 2,000–3,500 mg/L (chemical oxygen demand); 80–164 mg/L (total nitrogen); 74–116 mg/L (ammonium nitrogen); 18–47 mg/L (total phosphorus); 76–138 Nephelometric turbidity unit	Advantages: easy to implement, interpret and efficient to train Disadvantages: Prone to noise and overfitting; high sensitivity to outliers	Jiao et al. (2020), Singh et al. (2023), Wei et al. (2012)
Tan-Sigmoid function	$f(x_i) = \frac{1}{1+e^{x_i}}$ $f(x_i) = \frac{2}{(1+e^{-2x_i})-1}$	0 to 1 (but not equal to 0 or 1)	0.2–0.3 (biochemical oxygen demand/chemical oxygen demand) (biodegradability index)	Advantages: Predict the probability of outputs well Disadvantages: Vanishing gradient problem; when the gradient approaches 0, the network ceases to learn due to the gradient descent problem	Malik et al. (2019), Mohammadi et al. (2020), Singh et al. (2023)
Hyperbolic tan-sigmoid function	$f(x_i) = \tanh(x_i) = 2\sigma(2x_i) - 1$ $f(x_i) = \frac{e^{2x_i}-1}{e^{2x_i}+1}$	Any value between -1 and 1	0.2–0.4 (biochemical oxygen demand/chemical oxygen demand) (biodegradability)	Advantages: The derivative is steeper, generating more values and a wider range for faster learning and grading Disadvantages: Gradient problems towards the ends of the function	Khalaf et al. (2019), Singh et al. (2023), Wei et al. (2012)
Gaussian function	$f(x_i) = e^{x_i^2}$	Similar output for positive and negative input values	pH 3.7–6.8; 280–1113 mg/L (total suspended solids); 1770–4009 mg/L (total dissolved solids); 2135–4934 (total solids); 995–1097 mg/L (biochemical oxygen demand); 2268–3185 mg/L (chemical oxygen demand); 205–261 mg/L (chloride); 0.5–2.9 mg/L (oil and grease)	Advantages: Reliable estimation of uncertainty; usability, and flexibility Disadvantages: Overfitting; scale poorly with increasing number of measurements	Lokhande et al. (2011), Silversides et al. (2016), Singh et al. (2023)

Table 1 (continued)

Transfer/activation function	Governing equations	Output Range	Process parameters	Advantages/disadvantages	References
Ramp function	Piecewise function: $F(x) = \{x, x \geq 0, x < 0\}$	–	pH 6.2–7.0; 690–930 mg/L (total suspended solids); 600–1300 mg/L (total dissolved solids); 1300–1800 mg/L (biochemical oxygen demand); 2,500–3,200 mg/L (chemical oxygen demand); 90–180 mg/L (alkalinity); 2.2–3.0 Nephelometric turbidity unit; 95–125 mg/L (phenol)	Advantages: Better digital resolution with linearly proportional frequency. The accuracy of the digital system depends on the linearity and stability of the ramp function Disadvantages: Accuracy of the digital system often drifts and offset. Heavier penalty between two support hyperplanes	Hu et al. (2023), Saleem (2007), Zhang and Wang (2022)
	Heaviside step function: $F'(x) = H(x) for x \neq 0$ Convolution of the Heaviside step function: $F(x) = H(x) \times H(x)$				

The complex water quality characteristics require mathematical functions to design models for real-time dynamic simulations and process control systems to accurately predict and optimize a range of process parameters

pharmaceutical wastewater treatment systems. These error functions measure the deviation in a digital communication system that uses statistical computations.

Blockchain technology

The convergence of blockchain technologies and artificial intelligence in the internet of things network revolutionized intelligent network design to create sustainable processes (Mao et al. 2023). This means smart grids that use digital technologies can be connected to the network to detect and respond to local change to improve the industry's energy usage in electricity grids (Chen et al. 2021). When the electricity supply networks are equipped with internet protocol addresses, intelligent meters and energy sensors will relay the data to utility providers with information about energy usage, offering greater control over their energy consumption (Chen et al. 2021).

The emergence of blockchain technologies offers one of the most feasible solutions for decentralizing autonomous energy management in distributed energy systems using a simplified model inversion process of blockchain SM2 encryption by sending verification data of nodes with high energy distribution to improve the computational ability of the distributed energy systems (Wang et al. 2023b).

Conventional decentralized management modes have several drawbacks with respect to the high cost of communication from central controller to individual equipment, leading to single-point failures (Wang et al. 2023b). However, with the advent of new digital technology, the distributed information of blockchain provides a new vitality to the energy management of distributed energy systems. Distributed energy systems improve the permeability and utilization efficiency of renewable energy technologies, leading to high energy efficiency of pharmaceutical wastewater treatment systems.

Overall, we observed that blockchain technology has several limitations due to the scalability of software and hardware infrastructures, data security vulnerabilities, integration complexity, and high energy consumption. Innovative solutions should focus on improving energy efficiency and interoperability with existing systems.

Artificial intelligence-integrated blockchain distributed ledger technology has the potential to become one of the most critical research and development areas in the domain of renewable energy technologies and power automation (Gawusu et al. 2022). Artificial intelligence-integrated blockchain distributed ledger technology can address smart grid-based control management systems, decentralized energy management systems, power distribution, and related mechanical automation to pharmaceutical wastewater treatment plants (Junaidi et al. 2023; Khan et al. 2023).

Table 2 Error functions are used to validate artificial intelligence and machine learning models applicable to pharmaceutical wastewater treatment

Function number	Error indices	Governing equations	Proposed predictive models	Target value for best fitness	References
1	Mean square error	$\text{Mean square error} = \frac{\sum \left(Y_i - \hat{Y}_i\right)^2}{n}$	Adaptive neuro-fuzzy inference system, Artificial neural network, Response surface methodology	When the model is free from prediction error, Mean square error is approximately 0	Aghdam et al. (2023), Bhagat et al. (2020)
2	Root mean square error	$\text{Root mean square error} = \sqrt{\frac{\sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2}{n}}$	Artificial neural network, Response surface methodology, Adaptive neuro-fuzzy inference system	When the model has no prediction error, and root mean squared error reaches 0, it indicates the best fit	Shirkoohi et al. (2022)
3	Coefficient of determination R^2	$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{n} \left(Y_i - \overline{Y}_i\right)^2}$	Artificial neural network, Adaptive neuro-fuzzy inference system, Fuzzy logic model, multiple linear regression, full factorial design method, Response surface methodology	The closer the value to 1, the better the prediction model	Bhagat et al. (2020), Dadebo et al. (2023), Shirkoohi et al. (2022)
4	Mean absolute percentage error	$\text{Mean absolute percentage error} = \frac{100}{n} \sum_{i=1}^{n} \left\lvert \frac{Y_i - \hat{Y}_i}{Y_i} \right\rvert$	Adaptive neuro-fuzzy inference system	The lower the percentage of error, the better the fitness	Dadebo et al. (2023), Shirkoohi et al. (2022)
5	Residual sum of squares	$\text{Residual sum of squares} = \sum_{i=1}^{n} \left(Y_i - \overline{Y}_i\right)$	Genetic algorithm	The lower the residual sum of squares, the better the accuracy of forecast. When the value reaches 0, it indicates the best fit	Bahramian et al. (2023)
6	Normalized mean square error	$\text{Normalized mean square error} = \frac{\sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{n} Y^2}$	Multi-objective method of meta-heuristic shark smell optimization algorithm, Multilayer perceptron-particle swarm optimization, Multi-layer artificial neural network	The lower the normalized mean square error, the better the fitness	Bahramian et al. (2023)
7	Nash–Sutcliffe error	$\text{Nash} - \text{Sutcliffe error} = 1 - \frac{\sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2}{\sum_{i=1}^{n} \left(Y_i - \overline{Y}_i\right)^2}$	Auto-regressive integrated moving average plus Adaptive neuro-fuzzy inference system	The lower the Nash–Sutcliffe error, the better the accuracy of forecast. When the value reaches 0, it indicates the best fit	Bahramian et al. (2023)

Most critically, the purpose of combined technologies is to optimize power flow and process conditions to minimize energy consumption, perturbations, noise disturbances and prevent unstable working conditions (Zhu et al. 2023); promote stability, production efficiency; and reduce pollution of an industrial process using a local outlier factor-based abnormality detection logic to measure prediction statistical error (Feng et al. 2022).

The purpose of involving network reconfiguration of distributed systems is to facilitate the real-time operation of process dynamic conditions (Mishra et al. 2023), integration of cyber-physical security into software and hardware infrastructures to protect privacy and prevent external network infiltration and improve auto-generation of process control systems (Li et al. 2023a; Liu et al. 2022). Artificial intelligence-based data analysis and evolutionary learning mechanisms can diagnose water quality, facilitating autonomous decision-making and process optimization with a strong potential to establish predictive model analysis and universal process control (Li et al. 2021).

More interestingly, dynamic monitoring and controlling of smart grid technology can optimize the renewable energy used to power automation in pharmaceutical wastewater treatment plants, enabling machine learning developments and customizing executions of operational parameters to produce desired responses by screening and adjusting process parameters (Johnson et al. 2022). Integrating artificial intelligence with a power distribution network can create a real-time generation of process conditions, logistical distribution of pharmaceutical wastes by transportation, and monitoring electric power supply to facilitate wastewater treatment processes.

In remote regions, the field programmable gate array-based embedded internet of things system is one of the most preferred systems beneficial for optimizing wastewater treatment plants and leveraging logistics flow to improve the sludge management process in the future (Ding et al. 2021; Henriques et al. 2020). In addition, artificial intelligence technology in electrical automation provides fault diagnosis and troubleshooting of the process conditions, electrical control system, and electrical equipment and facilitates daily operation (Yang 2020).

In electrical process diagnosis, expert system, artificial neural network, and adaptive neuro-fuzzy inference system are three common methods of fault diagnosis, producing accurate detection results (Yang 2020). However, the main disadvantages of using artificial neural network optimization are a greater computational burden, susceptibility to overfitting, and empirical nature of the model development with minimalistic approach (Świetlicka and Kolanowski 2023). However, the existing ledger management of power distribution systems for wastewater treatment plants utilizes smart grid technology to deliver cloud scalability, optimize management, and minimize redundancy.

Integrating artificial intelligence and machine learning technologies improves data distribution efficiency and transmission across different networks, operational management, and privacy security (Khan et al. 2023). However, the most significant challenges of integrating the internet of things and blockchain technology into artificial intelligence and machine learning are related to financial, technical, environmental, organizational, and legal issues. These identified challenges are cyber-security, privacy, smart contract, trusted oracles, scalability, interoperability, lack of standardized structure, regulatory constraints, governance, fog computing, and so on (Tanha et al. 2022).

The convergence of blockchain technology, internet of things, artificial intelligence, and machine learning algorithms into cyber-security systems synergistically enhances trust, transparency, privacy, and cyber-security of overall operational systems in pharmaceutical wastewater treatment systems by providing a shared and decentralized distributed ledger (Xia et al. 2022). A blockchain technology, generally known as a distributed ledger, can store all information or data related to industry assets like a register (Thakur 2022). These data are primarily related to money and identities.

Integrating artificial intelligence and machine learning algorithms with the internet of things automates process dynamics within pharmaceutical wastewater treatment systems and related industrial networks, improving user-friendliness of business processes, which are essential for wastewater and water treatment industries (Sandner et al. 2020).

By the integration of artificial intelligence and machine learning into cyber-physical systems or its related information security infrastructure, the overall systems enhance pattern recognition, online transaction networks, supply chain management, troubleshoot information security vulnerabilities, and optimize outcomes of the wastewater treatment processes (Clark and Burstall 2018; Fernández-Caramés and Fraga-Lamas 2022; Sandner et al. 2020). In addition, Fig. 4 represents an artificial intelligence-enabled smart grid distribution network that integrates renewable energy technology, such as solar power to achieve energy efficiency and sustainability.

Overall, we observed that integrating artificial intelligence technologies in automation may help improve data analytics. Still, implementation costs are expensive and require specialized knowledge, system interoperability, and complex computational resources.

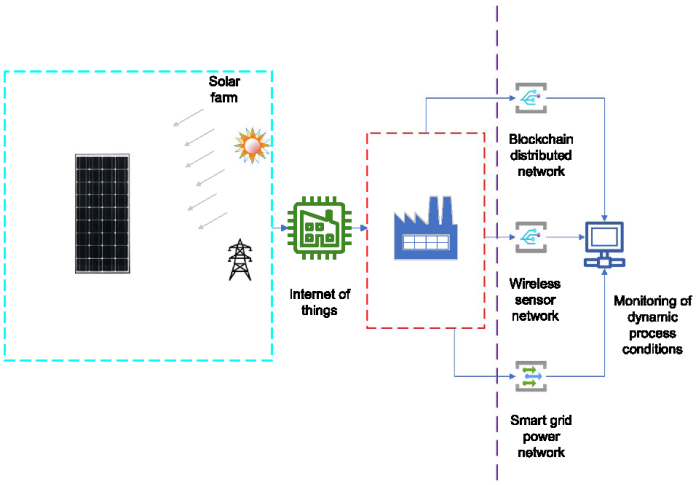


Fig. 4 Artificial intelligence-enabled smart grid distribution network for integrating renewable energy technology in pharmaceutical wastewater treatment plants to achieve process sustainability and decentralize energy management systems. This smart grid integrates energy distribution and digital communication technology to exchange two-way flow of electricity and energy usage data, offering personalized information related to the optimization of distributed energy systems and power outages or process equipment failures that impact the overall reliability of pharmaceutical wastewater treatment systems. Various electronic devices such as data concentrators, gateways, feeder meters, and aggregation meters can process data from parts of the smart grids such as consumption points, secondary substations, and so on. It helps to streamline the energy forecast across the grids to connect renewable energy technologies to large-scale pharmaceutical wastewater treatment plants

Big data

Water quality in pharmaceutical wastewater treatment plants can be optimized before discharging the effluents into the environment. The involvement of simulation models for examination of wastewater quality can be performed using databases, harmonic function, phenomenological methods, and benchmark simulation models, traditionally used to predict the behaviour and fate of wastewater constituents (Ly et al. 2022). Comprehensive knowledge and sophisticated control systems are required to facilitate model calibration and validation, making the control process a major disadvantage.

Machine learning can predict various water-related variables and wastewater constituents, unlike traditional approaches. It does not require expert knowledge to operate. It can handle and analyse large datasets and requires less processing power. In addition, complex, nonlinear variables of wastewater quality parameters can be modelled using computed autoregressive integrated moving average to forecast the levels of nitrogen, biochemical oxygen demand, chemical oxygen demand, phosphorus, ammonia, total suspended solids with relatively high accuracy ranging between 71 and 97% for the training data and low prediction errors less than 9% for the testing data (Ly et al. 2022).

Other machine learning algorithms such as random forest, support vector machine, long short-term memory, gradient tree boosting, adaptive neuro-fuzzy inference system, and so on all forming parts of deep learning architectures can be used to forecast and perform extensive data analysis of wastewater quality and its constituents. Artificial neural network and genetic algorithms can model pharmaceutical wastewater treatment systems for advanced oxidation processes to predict the operational parameters involving three-step processes such as acidification, adsorption, and photocatalysis to solve wastewater composition (Yang et al. 2021).

Data mining techniques such as artificial neural network and M5 tree model can be used to analyse a range of datasets due to its reliability, robustness, and high generalization ability to achieve a coefficient of determination greater than 0.90 for forecasting biochemical oxygen demand, chemical oxygen demand, and total suspended solids (Asami et al. 2021). However, using photocatalytic approaches has numerous limitations, such as lengthy procedures and impractically large amounts of wastewater treatment catalysts with limited resource recovery process.

The integration of sonolysis with photocatalysis could benefit the environmental remediation, maximizing the catalyst surface area and rapidly improving the production of free radicals to degrade toxic organic pollutants in

pharmaceutical wastewater (Theerthagiri et al. 2021). On the other hand, the electrocatalytic reduction of nitrogenous compounds, such as nitrate waste into ammonia, facilitates rapid removal of toxic nitrate contaminants and forming an alternative production of ammonia with secondary benefit compared to conventional Haber–Bosch process (Theerthagiri et al. 2022b).

For the advancement of photo- and electrocatalytic technologies, a fabricated electrochemical sensor based on novel zinc sulphate/gold/multi-walled carbon nanotube nanocomposites can be integrated into the process control system using big data mining technique in pharmaceutical wastewater treatment systems to improve the sensitivity of detection on toxic organic nitrogenous pollutant, which is part of human metabolites produced from the breakdown of pharmaceutical ingredients and strengthening process analytics of wastewater quality (Naik et al. 2021).

The future design and fabrication of innovative pulsed laser-assisted technologies can improve structural optimization of electrochemical sensors with electrocatalytic performance in various renewable energy and environmental remediation processes (Theerthagiri et al. 2022a). In addition, pulsed laser irradiation technologies can dechlorinate persistent organic pollutants containing chlorine-based compounds, which are by-products widely generated in industrial production (Yu et al. 2021).

Overall, we observed that data analytics processes can revolutionize wastewater treatment technologies, but operation and maintenance costs are high. Compliance concerns are also associated with reporting errors in the systems, stability of operational systems, data security vulnerabilities, data acquisition, and interoperability of existing systems.

Cyber-physical systems

The increasing interconnections and interdependencies between cyber-security systems, physical assets, humans, and environment resulted in rapid evolution of pharmaceutical wastewater treatment systems (Mohebbi et al. 2020). Technological innovation and advancement in pharmaceutical technology, environmental sustainability, economic and regulatory factors all influence wastewater treatment systems (Cui 2021).

In addition, cyber-physical framework provides an integrated approach to facilitate efficient management of technologies, improving precision in detecting wastewater constituents and optimizing output variables. Adaptable digital solutions can help various stakeholders understand the effect of pharmaceutical wastewater quality on public health and improve water governance by promoting social awareness and collaboration between wastewater treatment industries and citizens (Alexandra et al. 2023; Radini et al. 2021).

On the other hand, maintenance of cyber-physical systems in modern pharmaceutical wastewater treatment plants requires improving the cyber-resilience of information security infrastructure to complement a traditional physical resilience assessment (Colabianchi et al. 2021; Patriarca et al. 2022). To address the level of resilience, stochastic cyber-resilience metrics must be proposed and computed to assess the impact of cyber-attacks on information technology infrastructure to uncover the vulnerability of the industrial control systems and its distributed networks (Avraam et al. 2023; Chaves et al. 2017; Li et al. 2023c; Yang et al. 2022).

More critically, challenges arise from ageing information technology infrastructure, environmental impact, and sustainability of pharmaceutical wastewater treatment systems require improvement in data management, analytics and cyber-security systems, requiring knowledge and skills of experts to satisfy regulatory compliance and governmental requirements as well as supporting the decision-making process of various stakeholders (Bhandari et al. 2023).

Understanding the evolutionary process and its influences on pharmaceutical wastewater infrastructure and characteristics affects the quality of water and sanitation services, which drive socio-economic changes in industrial wastewater treatment systems. New strategies must be developed to solve health-related problems arising from pharmaceutical water pollution (Foglia et al. 2021).

Physical assets in wastewater treatment industries that involve various water infrastructures, such as hydraulic pumps, network analysis of processes, optimization of water distribution, and output variables, can influence the progression of wastewater infrastructure development.

Various stakeholders must be involved in the planning and decision-making process when configuring the cyber-physical systems of industrial wastewater treatment processes, with responsibilities assigned to federal and local governments to manage water resources and deliver sanitary drinking water and clean wastewater for public and agricultural uses (Hasan et al. 2023).

The rapid transformation of digital technology used in cyber-physical systems improves the techno-economy of wastewater treatment services, enabling the decision-making process using internet of things and assisting industry professionals to achieve a new paradigm of water resources management (Song et al. 2023).

Figure 5 outlines the evaluation criteria for appraising the artificial intelligence and machine learning-based optimization technologies recommended for pharmaceutical wastewater treatment systems. It shows a specific framework related to evolutionary artificial intelligence technologies that can be implemented into pharmaceutical wastewater treatment systems to satisfy industry standards.

In addition, Fig. 6 shows a structured analysis of various artificial intelligence and machine learning approaches and

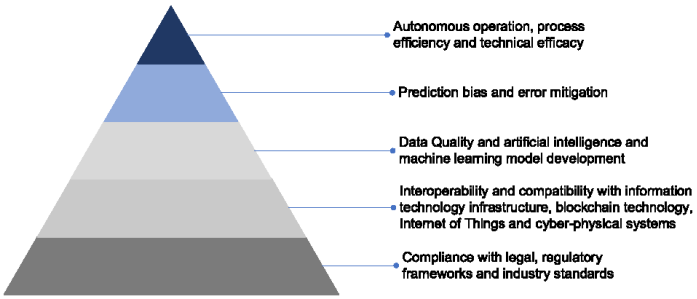


Fig. 5 Evaluation criteria for critical appraisal of artificial intelligence and machine learning technologies recommended for use in wastewater treatment systems. These evaluation criteria form the rubric for artificial intelligence tool evaluation to provide a framework for assessing the artificial intelligence tools based on a set of criteria, including interoperability, functionality, compatibility, and so on. Critical process involves the rigorous assessment of data quality and model performance, including predictive accuracy and process control capabilities. The top hierarchy represents the most critical component of artificial intelligence tool: autonomous operation, process efficiency, and technical efficiency. Last but not least, the bottommost layer in the pyramid is also important for any artificial intelligence integration into the wastewater treatment industries, but the regulatory frameworks and industry standards may differ worldwide

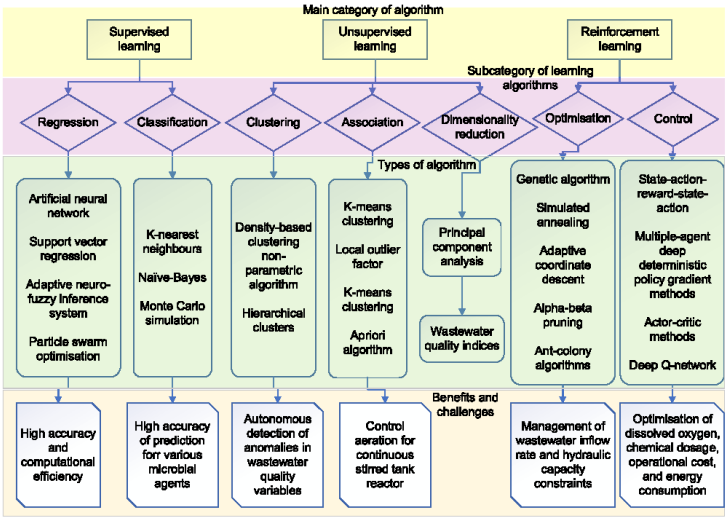


Fig. 6 A structured analysis of various artificial intelligence and machine learning approaches and their suitability for specific challenges within pharmaceutical wastewater treatment systems to facilitate autonomous process control systems and global optimization of wastewater quality characteristics and other operational conditions. The top pedigree represents the main category of algorithm in which supervised learning involves a formula generation based on input and output values. It uses labelled training datasets, whereas unsupervised learning does not. Reinforcement learning trains software to make decisions and generate the most optimal solutions. Under all subcategory of learning algorithms, clustering is the most common one. Clustering is used to detect anomalies and outliers in the dataset. Classification algorithms determine the category of an entity, object, or event in a given dataset

their suitability for addressing specific challenges encountered in pharmaceutical wastewater treatment systems. Additionally, Table 3 critically evaluates the evolutionary characteristics of the pharmaceutical wastewater treatment systems from multidimensional perspectives. Furthermore, Table 4 critically evaluates the findings of different artificial

Table 3 Evolutionary characteristics of pharmaceutical wastewater treatment systems and multidimensional changes involving the amalgamation of technologies such as artificial intelligence, blockchain, internet of things, and cyber-physical systems

Evolutionary characteristics of pharmaceutical wastewater treatment systems	Description	Benefits or vulnerabilities	Multidimensional improvement	References
Structured information technology security infrastructure	Wastewater treatment industry will slowly implement robust cyber-physical systems using advanced hardware/software computing systems to replace the ageing infrastructure to improve the resilience of information security systems and computing integrity	Regulate influent and effluent qualities by automated process control systems. Potential consequences involve disruptions by ransomware attack and restoration of systems will take several days, exposing risks to public health, causing economic losses and environmental damage	Systematic governance and a major focus on cyber-security integrity of the information technology infrastructure, industrial control systems and its distributed networks. Improvement in energy efficiency of wastewater treatment process. Enable blockchain cyber-physical systems. Fault/vulnerability detection, troubleshooting, diagnosis, and prognosis test using artificial intelligence techniques	Miller et al. (2021), Patriarca et al. (2022), Raval et al. (2024)
Effect of myriad parameters on dynamic process conditions	Internal and external factors, including governmental, regulatory, socio-economic, water security, renewable energy, climate change, technological innovation, advancement and so on, affect system integrity and infrastructure	A synergistic combination of artificial intelligence and machine learning algorithms propels sustainable solutions. Automating and optimizing real-time monitoring and simulations are highly complex and require high computational power. Prediction of system behaviour can facilitate resource allocation and accurate decision-making process. Nutrient limitations, contamination, temperature and pH fluctuations, system stability, scaling up and integration into existing infrastructure and its impact on long-term performance could be significant issues	Monitoring of dynamics changes, implementation of protocols and new digital transformation to effectuate such changes. Employment of various artificial intelligence optimization techniques to manage such changes	Nti et al. (2023), Sahu et al. (2023)

Table 3 (continued)

Evolutionary characteristics of pharmaceutical wastewater treatment systems	Description	Benefits or vulnerabilities	Multidimensional improvement	References
Enhancing interdependencies and interconnectedness	Improving human–software interface systems through interactions with cyber-physical systems. Establish governmental regulatory systems, develop environmental technology and promote deep learning using artificial intelligence technologies to deliver wastewater sanitation services to local communities	Provide comprehensive analysis of digital forensics incident response as integral part of the security of industrial internet of things. Security challenges involve secure internet of things offloading, access control, data availability and heterogeneity. Possible risks include phishing, jamming, intrusion and malware, causing leak of private data, affecting authentication, device integrity and industrial control systems. Industrial control systems include distributed control systems, programmable logic controllers, supervisory control and data acquisition	Systematic governance requires a digital transformation in corporations and organizations to improve information technology infrastructure, integrate real-time monitoring wastewater quality and advanced data analytics to minimize waste, and greenhouse gas emissions, achieve cost reduction, maximize energy efficiency and conservation, facilitate value engineering, water reusability and water resources management. Maintain normality and stability in the information systems. Collaboration between industry and research to deploy data-driven models in the wastewater sector	Bahramian et al. (2023), Binnar et al. (2024), Zhao et al. (2020)

intelligence and machine learning optimization techniques, supporting the empirical evidence with an interpretation to address the research deficiencies and outcomes.

Overall, we observed that data security vulnerabilities are significant issues due to difficulty authenticating the information data in automated systems. Coordinated cyber-attacks on critical infrastructures and industrial control systems can affect community service availability. More efficient and robust solutions are required to form a new ecosystem that involves cyber-physical systems combined with the internet of things to operate a massive and complex wastewater treatment system.

Conclusion

Sustainability of pharmaceutical wastewater treatment systems is increasingly critical in the modern world. Integrating artificial intelligence and machine learning-based models can potentially revolutionize the wastewater sectors, including public health and environment. In particular, managing wastewater quality and optimizing process parameters using artificial intelligence technologies help achieve the best removal rate of pharmaceutical pollutants to minimize the likelihood of pathogen transmission and spread of viral vectors and antimicrobial resistance genes in complex pharmaceutical wastewater environments.

Effective monitoring process dynamic conditions demand advanced process control systems to manage water resources. The application of blockchain-related technologies towards sustainable wastewater and energy management should be extended to both metropolitan and rural areas, but further technological investigation, cost, and carbon footprint assessment should be conducted to evaluate the techno-economic and financial viability of such technologies.

The technological capabilities of internet of things and cutting-edge cyber-physical systems in the digital economy to integrate decision-making processes should be incorporated into wastewater treatment industries to promote intelligent waste transportation systems, minimize carbon footprint, and remove barriers to resource recovery and energy management processes. Several points of summary for future directions are outlined as follows:

(1) Artificial intelligence and machine learning approaches are applied to develop predictive models for monitoring pharmaceutical wastewater quality and its constituents in complex wastewater matrices.
(2) Minimization of operational cost and improvement in energy efficiency of pharmaceutical wastewater treatment systems require the integration of artificial intelligence technologies.

Table 4 Systematic analysis of different artificial intelligence and/or machine learning optimization techniques, critically appraising the deficiencies and efficiencies of various artificial intelligence technologies and its impact on the optimization efficiencies

Types of artificial intelligence techniques	Synthesis, data collection or treatment method	Types of pollutants	Operating conditions	Deficiency addressed	How is the deficiency addressed	Optimum results or optimization efficiency	References
Particle swarm optimization	Green synthesis optimization of copper-gallic acid metal–organic framework	Basic red 9	Molar ratio equals to 1.8 Temperature equals to 108 °C Reaction time equals to 1.5 h	Low porosity; pore dysfunctionality; uneven pore topologies; low adsorptive capacity, and so on	Generalize constraint space to represent material characteristics; improve prediction of experimental data using metaheuristic algorithm to facilitate global optimization	Yield percentage equals to 60.6% Crystallinity percentage equals to 59.2% Dye removal efficiency equals to 96.2%	Azhar et al. (2023)
Backpropagation artificial neural network; genetic algorithm	Iron sulphate heptahydrate synthesized from amorphous FeS particle	Copper; arsenic	C_0 equals to 0.1, 0.3, 0.5, 0.7, 1 Cu/As concentration ratio Adsorbent dosage equals to 1.2 M Fe (II) solution	Traditional batch experimentation tests one variable at a time; too time-consuming and expensive	Backpropagation-artificial neural network and genetic algorithm optimization improves precision of prediction; suitable for parallel processing high global optimization efficiency	R^2 equals to 0.99 Maximum separation factor equals to 1,400	Zhang et al. (2024)
Artificial neural network-particle swarm; artificial neural network-genetic algorithm	Reduced graphene oxide/iron/cobalt nanohybrids prepared from co-precipitation method	Methylene blue	C_0 equals to 100–1,000 mg/L Methylene Blue Adsorbent dosage equals to 30 mg reduced graphene oxide/iron/cobalt nanohybrids Solution volume equals to 50 mL Contact time equals to 12 min pH_i equals to 5.0 Temperature equals to 35 °C	Expensive; low operational efficiency; low decontamination efficiency	Artificial intelligence-genetic algorithm and artificial neural network-particle swarm optimization improved decontamination efficiency of pollutants; generated models to determine optimum conditions	Experimental removal efficiency equals to 90.6% Predicted decontamination efficiency equals to 89.3% (artificial neural network-genetic algorithm) Predicted decontamination efficiency equals to 93.5% (artificial neural network-particle swarm)	Qi et al. (2020)

Table 4 (continued)

Types of artificial intelligence techniques	Synthesis, data collection or treatment method	Types of pollutants	Operating conditions	Deficiency addressed	How is the deficiency addressed	Optimum results or optimization efficiency	References
Automated machine learning; backpropagation artificial neural network	Operating data collected from wastewater treatment plants: chemical oxygen demand, NH_4^+-N, total nitrogen, total phosphorus, suspended solids, sludge yield, electricity carbon emission, chemical carbo emission	Total nitrogen; total phosphorus	Influent volume equals to 5,302–35,355 (m^3/d) Influent chemical oxygen demand equals to 90–537 mg chemical oxygen demand/L) Influent NH_4^+-N equals to 8.8–75.8 mg nitrogen/L Influent total nitrogen equals to 10.7–77.8 mg nitrogen/L Influent total phosphorus equals to 0.6–8.6 mg phosphorus/L Influent suspended solids equal to 35–236 mg/L	Total nitrogen and total phosphorus were the two main indices exceeded water standards up to 1610 national and provincial river sites	Total nitrogen and total phosphorus removal efficiencies can be optimized in current wastewater treatment plants operation using data obtained from online equipment. Prediction of pollutant concentrations in wastewater treatment plants and analysis of effluent quality can be carried out using automated machine learning and backpropagation artificial neural network	Mean square error equals to 0.42–3.82 (automated machine learning) R^2 equals to 0.57–0.62 (automated machine learning) Mean square error equals to 0.0012–6.91 (Backpropagation artificial neural network) R^2 equals to 0.43–0.89 (Backpropagation artificial neural network)	Luo et al. (2023)
Response surface methodology-artificial neural network	Forward osmosis-membrane distillation process	Lead (Pb)	C_0 equals to 60 mg/L lead Feed velocity equals to 11.57 cm/s Draw velocity equals to 7.7 cm/s	Low permeate flux; high energy consumption, high reverse solute flux; internal and external polarization of the membrane	Artificial intelligence optimization developed by machine learning algorithms improves real-time feedback and simulation; cost-effective methods of handling complex mathematical models and enhanced forward osmosis-low pressure ultrafiltration system in terms of water flux	R^2 equals to 0.99 Root mean square error equals to 0.010 Lead removal efficiency equals to 87.1%	Boubakri et al. (2023)

Table 4 (continued)

Types of artificial intelligence techniques	Synthesis, data collection or treatment method	Types of pollutants	Operating conditions	Deficiency addressed	How is the deficiency addressed	Optimum results or optimization efficiency	References
Random vector functional link networks incorporated with manta ray foraging optimizer	Activated sludge treatment process. Domestic sewage with a design capacity of 330,000 m^3/d with a dedicated area of about 325,000 m^2	Total suspended solids, volatile suspended solids, biochemical oxygen demand	Influent biochemical oxygen demand equals to 130 plus or minus 30 mg/L Influent total suspended solids equal to 130 ± 30 mg/L Effluent biochemical oxygen demand equals to 12.26 ± 6 mg/L Effluent total suspended solids equal to 16.57 ± 8 mg/L	Conventional artificial neural network problems are overfitting	Maximized prediction accuracy of biochemical oxygen demand and total suspended solids using Manta ray foraging with best random vector functional link parameters as an optimization to enhance model performance	Removal efficiency equals to 90.8% of biochemical oxygen demand Removal efficiency equals to 87.3% Total suspended solids R^2 equals to 0.84 (Train); 0.84 for Test (biochemical oxygen demand) Root mean square error equals to 2.75 (Train); 6.21 for Test (biochemical oxygen demand) R^2 equals to 0.73 (Train); 0.72 for Test (Total suspended solids) Root mean square error equals to 3.5 (Train); 10.1 for test	Elmaadawy et al. (2021)

(3) Predictive control of various contaminants, including viral vectors, antimicrobial resistance genes, severe acute respiratory syndrome coronavirus 2, water quality parameters, chemical oxygen demand, biochemical oxygen demand, phosphorus and other organics or nutrient removal, is crucial for future research.
(4) Sanitation and disinfection services are critical for pharmaceutical wastewater treatment systems, and emerging artificial intelligence technologies should be used to optimize renewable energy, process control systems, and wastewater treatment processes.
(5) Comprehensive models involving socio-economic, governmental, environmental, techno-economic, technological innovation, and so on require thorough investigation when designing pharmaceutical wastewater treatment systems.
(6) Advancements in cyber-physical systems can increase sensitivity for fault detection, troubleshoot technical issues, improve diagnosis and prognosis of vulnerability in information technology infrastructure, and help maintain distributed networks of pharmaceutical wastewater treatment systems.
(7) Different approaches should be implemented to identify and analyse vulnerabilities and risks in information technology infrastructure, but identification of complex dynamic behaviours, uncertainties, or perturbations in complex process control systems and data management processes in pharmaceutical wastewater treatment systems requires artificial intelligence technologies. Standardization of frameworks and assessment metrics will assist in computing efficiency, improving the reliability and performance of pharmaceutical wastewater treatment technologies.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10311-024-01748-w.

Authors' contributions Voravich Ganthavee involved in conceptualization, visualization, validation, investigation, formal analysis, data curation, writing—original draft. Antoine P. Trzcinski took part in supervision, review and editing, project administration, resources.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Data Availability Additional data will be provided upon request.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Aghdam E, Mohandes SR, Manu P, Cheung C, Yunusa-Kaltungo A, Zayed T (2023) Predicting quality parameters of wastewater treatment plants using artificial intelligence techniques. J Clean Prod 405:137019. https://doi.org/10.1016/j.jclepro.2023.137019

Akhil D, Lakshmi D, Senthil Kumar P, Vo D-VN, Kartik A (2021) Occurrence and removal of antibiotics from industrial wastewater. Environ Chem Lett 19:1477–1507. https://doi.org/10.1007/s10311-020-01152-0

Alexandra C, Daniell KA, Guillaume J, Saraswat C, Feldman HR (2023) Cyber-physical systems in water management and governance. Curr Opin Environ Sustain 62:101290. https://doi.org/10.1016/j.cosust.2023.101290

Aparna KG, Swarnalatha R (2023) Dynamic optimization of a wastewater treatment process for sustainable operation using multi-objective genetic algorithm and non-dominated sorting cuckoo search algorithm. J Water Process Eng 53:103775. https://doi.org/10.1016/j.jwpe.2023.103775

Asami H, Golabi M, Albaji M (2021) Simulation of the biochemical and chemical oxygen demand and total suspended solids in wastewater treatment plants: data-mining approach. J Clean Prod 296:126533. https://doi.org/10.1016/j.jclepro.2021.126533

Avraam C, Ceferino L, Dvorkin Y (2023) Operational and economy-wide impacts of compound cyber-attacks and extreme weather events on electric power networks. Appl Energy 349:121577. https://doi.org/10.1016/j.apenergy.2023.121577

Azhar B, Avian C, Tiwikrama AH (2023) Green synthesis optimization with artificial intelligence studies of copper–gallic acid metal–organic framework and its application in dye removal from wastewater. J Mol Liq 389:122844. https://doi.org/10.1016/j.molliq.2023.122844

Azuma T, Hayashi T (2021) On-site chlorination responsible for effective disinfection of wastewater from hospital. Sci Total Environ 776:145951. https://doi.org/10.1016/j.scitotenv.2021.145951

Bahramian M, Dereli RK, Zhao WQ, Giberti M, Casey E (2023) Data to intelligence: the role of data-driven models in wastewater treatment. Expert Syst Appl. https://doi.org/10.1016/j.eswa.2022.119453

Bhagat SK, Tung TM, Yaseen ZM (2020) Development of artificial intelligence for modeling wastewater heavy metal removal: State of the art, application assessment and possible future research. J Clean Prod 250:119473. https://doi.org/10.1016/j.jclepro.2019.119473

Bhandari P, Creighton D, Gong J, Boyle C, Law KMY (2023) Evolution of cyber-physical-human water systems: Challenges and gaps. Technol Forecast Soc Change 191:122540. https://doi.org/10.1016/j.techfore.2023.122540

Bhattacharya S, Abhishek K, Samiksha S, Sharma P (2023) Occurrence and transport of SARS-CoV-2 in wastewater streams and its detection and remediation by chemical-biological methods. J

Hazard Mater Adv 9:100221. https://doi.org/10.1016/j.hazadv.2022.100221
Binnar P, Bhirud S, Kazi F (2024) Security analysis of cyber physical system using digital forensic incident response. Cyber Secur Appl 2:100034. https://doi.org/10.1016/j.csa.2023.100034
Boubakri A, Elgharbi S, Dhaouadi I, Mansour D, Al-Tahar Bouguecha S (2023) Optimization and prediction of lead removal from aqueous solution using FO–MD hybrid process: Statistical and artificial intelligence analysis. J Environ Manag 337:117731. https://doi.org/10.1016/j.jenvman.2023.117731
Chaves A, Rice M, Dunlap S, Pecarina J (2017) Improving the cyber resilience of industrial control systems. Int J Crit Infrastruct Prot 17:30–48. https://doi.org/10.1016/j.ijcip.2017.03.005
Chawishborwornworng C, Luanwuthi S, Umpuch C, Puchongkawarin C (2023) Bootstrap approach for quantifying the uncertainty in modeling of the water quality index using principal component analysis and artificial intelligence. J Saudi Soc Agric Sci. https://doi.org/10.1016/j.jssas.2023.08.004
Chen C, Hu Y, Karuppiah M, Kumar PM (2021) Artificial intelligence on economic evaluation of energy efficiency and renewable energy technologies. Sustain Energy Technol Assess 47:101358. https://doi.org/10.1016/j.seta.2021.101358
Clark B, Burstall R (2018) Blockchain, IP and the pharma industry—how distributed ledger technologies can help secure the pharma supply chain. J Intellect Prop Law Pract 13:531–533. https://doi.org/10.1093/jiplp/jpy069
Colabianchi S, Costantino F, Di Gravio G, Nonino F, Patriarca R (2021) Discussing resilience in the context of cyber physical systems. Comput Ind Eng 160:107534. https://doi.org/10.1016/j.cie.2021.107534
Corominas L, Garrido-Baserba M, Villez K, Olsson G, Cortés U, Poch M (2018) Transforming data into knowledge for improved wastewater treatment operation: a critical review of techniques. Environ Model Softw 106:89–103. https://doi.org/10.1016/j.envsoft.2017.11.023
Cui X (2021) Cyber-Physical System (CPS) architecture for real-time water sustainability management in manufacturing industry. Procedia CIRP 99:543–548. https://doi.org/10.1016/j.procir.2021.03.074
Dadebo D, Obura D, Etyang N, Kimera D (2023) Economic and social perspectives of implementing artificial intelligence in drinking water treatment systems for predicting coagulant dosage: a transition toward sustainability. Groundw Sustain Dev 23:100987. https://doi.org/10.1016/j.gsd.2023.100987
Deepnarain N, Nasr M, Kumari S, Stenstrom TA, Reddy P, Pillay K, Bux F (2020) Artificial intelligence and multivariate statistics for comprehensive assessment of filamentous bacteria in wastewater treatment plants experiencing sludge bulking. Environ Technol Innov. https://doi.org/10.1016/j.eti.2020.100853
Ding X, Shi P, Li X (2021) Regional smart logistics economic development based on artificial intelligence and embedded system. Microprocess Microsyst 81:103725. https://doi.org/10.1016/j.micpro.2020.103725
Elmaadawy K, Elaziz MA, Elsheikh AH, Moawad A, Liu B, Lu S (2021) Utilization of random vector functional link integrated with manta ray foraging optimization for effluent prediction of wastewater treatment plant. J Environ Manag 298:113520. https://doi.org/10.1016/j.jenvman.2021.113520
Fan X, Li Y (2023) Energy management of renewable based power grids using artificial intelligence: digital twin of renewables. Sol Energy 262:111867. https://doi.org/10.1016/j.solener.2023.111867
Feng Z, Li Y, Xiao B, Sun B, Yang C (2022) Process monitoring of abnormal working conditions in the zinc roasting process with an ALD-based LOF-PCA method. Process Saf Environ Prot 161:640–650. https://doi.org/10.1016/j.psep.2022.03.064
Fernández-Caramés TM, Fraga-Lamas P (2022) Advances in the Convergence of Blockchain and Artificial Intelligence, 1st edn. IntechOpen, London
Foglia A, Bruni C, Cipolletta G, Eusebi AL, Frison N, Katsou E, Akyol Ç, Fatone F (2021) Assessing socio-economic value of innovative materials recovery solutions validated in existing wastewater treatment plants. J Clean Prod 322:129048. https://doi.org/10.1016/j.jclepro.2021.129048
Ganiyu SO, van Hullebusch ED, Cretin M, Esposito G, Oturan MA (2015) Coupling of membrane filtration and advanced oxidation processes for removal of pharmaceutical residues: a critical review. Sep Purif Technol 156:891–914. https://doi.org/10.1016/j.seppur.2015.09.059
Gao F, Wen H, Feng S, Li M, Zhu L, Zhang Y, Xi Y, Xiang X (2023) The elevated toxicity of the biodegradation product (guanylurea) from metformin and the antagonistic pattern recognition of combined toxicity: insight from the pharmaceutical risk assessment and the simulated wastewater treatment. Sci Total Environ 892:164747. https://doi.org/10.1016/j.scitotenv.2023.164747
Gawusu S, Zhang X, Ahmed A, Jamatutu SA, Miensah ED, Amadu AA, Osei FAJ (2022) Renewable energy sources from the perspective of blockchain integration: From theory to application. Sustain Energy Technol Assess 52:102108. https://doi.org/10.1016/j.seta.2022.102108
Ghosh S, Harsha NVMS, Singh SP, Shriwastav A (2023) Simultaneous removal of ciprofloxacin and disinfection from wastewater by combined photocatalytic reactor (PCR) and membrane bioreactor (MBR) system. J Environ Chem Eng 11:110855. https://doi.org/10.1016/j.jece.2023.110855
Guo X, Yu H, Yan Z, Gao H, Zhang Y (2018) Tracking variations of fluorescent dissolved organic matter during wastewater treatment by accumulative fluorescence emission spectroscopy combined with principal component, second derivative and canonical correlation analyses. Chemosphere 194:463–470. https://doi.org/10.1016/j.chemosphere.2017.12.023
Hai T, Alsharif S, Dhahad HA, Attia E-A, Shamseldin MA, Najat Ahmed A (2022) The evolutionary artificial intelligence-based algorithm to find the minimum GHG emission via the integrated energy system using the MSW as fuel in a waste heat recovery plant. Sustain Energy Technol Assess 53:102531. https://doi.org/10.1016/j.seta.2022.102531
Hasan MK, Habib AKMA, Shukur Z, Ibrahim F, Islam S, Razzaque MA (2023) Review on cyber-physical and cyber-security system in smart grid: Standards, protocols, constraints, and recommendations. J Netw Comput Appl 209:103540. https://doi.org/10.1016/j.jnca.2022.103540
Henriques AA, Fontes M, Camanho A, Silva JG, Amorim P (2020) Leveraging logistics flows to improve the sludge management process of wastewater treatment plants. J Clean Prod 276:122720. https://doi.org/10.1016/j.jclepro.2020.122720
Hu J, Zhang L, Tang J, Liu Z (2023) A novel transformer ordinal regression network with label diversity for wind power ramp events forecasting. Energy 280:128075. https://doi.org/10.1016/j.energy.2023.128075
Ibrahim KSMH, Huang YF, Ahmed AN, Koo CH, El-Shafie A (2022) A review of the hybrid artificial intelligence and optimization modelling of hydrological streamflow forecasting. Alex Eng J 61:279–303. https://doi.org/10.1016/j.aej.2021.04.100
Jana DK, Bhunia P, Das Adhikary S, Bej B (2022) Optimization of effluents using artificial neural network and support vector regression in detergent industrial wastewater treatment. Clean Chem Eng 3:100039. https://doi.org/10.1016/j.clce.2022.100039

Jariwala N, Putta CL, Gatade K, Umarji M, Ruhina Rahman SN, Pawde DM, Sree A, Kamble AS, Goswami A, Chakraborty P, Shunmugaperumal T (2023) Intriguing of pharmaceutical product development processes with the help of artificial intelligence and deep/machine learning or artificial neural network. J Drug Deliv Sci Technol 87:104751. https://doi.org/10.1016/j.jddst.2023.104751

Jiao S, Gao Y, Feng J, Lei T, Yuan X (2020) Does deep learning always outperform simple linear regression in optical imaging? Opt Express 28:3717–3731. https://doi.org/10.1364/OE.382319

Johnson PC, Laurell C, Ots M, Sandström C (2022) Digital innovation and the effects of artificial intelligence on firms' research and development—automation or augmentation, exploration or exploitation? Technol Forecast Soc Change 179:121636. https://doi.org/10.1016/j.techfore.2022.121636

Junaidi N, Abdullah MP, Alharbi B, Shaaban M (2023) Blockchain-based management of demand response in electric energy grids: a systematic review. Energy Rep 9:5075–5100. https://doi.org/10.1016/j.egyr.2023.04.020

Khalaf W, Zaghar D, Hashim N (2019) Enhancement of curve-fitting image compression using hyperbolic function. Symmetry (basel) 11:291. https://doi.org/10.3390/sym11020291

Khan AA, Ali Laghari A, Rashid M, Li H, Rehman Javed A, Reddy Gadekallu T (2023) Artificial intelligence and blockchain technology for secure smart grid and power distribution automation: a state-of-the-art review. Sustain Energy Technol Assess 57:103282. https://doi.org/10.1016/j.seta.2023.103282

Khawaga RI, Abdel Jabbar N, Al-Asheh S, Abouleish M (2019) Model identification and control of chlorine residual for disinfection of wastewater. J Water Process Eng 32:100936. https://doi.org/10.1016/j.jwpe.2019.100936

Kooijman G, de Kreuk MK, Houtman C, van Lier JB (2020) Perspectives of coagulation/flocculation for the removal of pharmaceuticals from domestic wastewater: a critical view at experimental procedures. J Water Process Eng 34:101161. https://doi.org/10.1016/j.jwpe.2020.101161

Li L, Rong S, Wang R, Yu S (2021) Recent advances in artificial intelligence and machine learning for nonlinear relationship analysis and process control in drinking water treatment: a review. Chem Eng J 405:126673. https://doi.org/10.1016/j.cej.2020.126673

Li G, Ren L, Fu Y, Yang Z, Adetola V, Wen J, Zhu Q, Wu T, Candan KS, O'Neill Z (2023a) A critical review of cyber-physical security for building automation systems. Ann Rev Control 55:237–254. https://doi.org/10.1016/j.arcontrol.2023.02.004

Li S, Zhu G, Li X, Wan P, Yuan F, Xu S, Hursthouse AS (2023b) Ecosystem-inspired model and artificial intelligence predicts pollutant consumption capacity by coagulation in drinking water treatment. Environ Chem Lett 21:2499–2508. https://doi.org/10.1007/s10311-023-01602-5

Li Z-L, Li P, Xia J, Yuan Z-P (2023c) Cyber-physical-social system scheduling for multi-energy microgrids with distribution network coordination. Int J Electr Power Energy Syst 149:109054. https://doi.org/10.1016/j.ijepes.2023.109054

Liu Q, Liu M, Zhou H, Yan F, Ma Y, Shen W (2022) Intelligent manufacturing system with human-cyber-physical fusion and collaboration for process fine control. J Manuf Syst 64:149–169. https://doi.org/10.1016/j.jmsy.2022.06.004

Liu Y, Ramin P, Flores-Alsina X, Gernaey KV (2023) Transforming data into actionable knowledge for fault detection, diagnosis and prognosis in urban wastewater systems with AI techniques: a mini-review. Process Saf Environ Prot 172:501–512. https://doi.org/10.1016/j.psep.2023.02.043

Lokhande R, Singare P, Pimple DS (2011) Toxicity study of heavy metals pollutants in waste water effluent samples collected from Taloja Industrial Estate of Mumbai. India 1:13–19

Luo J, Luo Y, Cheng X, Liu X, Wang F, Fang F, Cao J, Liu W, Xu R (2023) Prediction of biological nutrients removal in full-scale wastewater treatment plants using H_2O automated machine learning and back propagation artificial neural network model: optimization and comparison. Bioresour Technol 390:129842. https://doi.org/10.1016/j.biortech.2023.129842

Ly QV, Truong VH, Ji B, Nguyen XC, Cho KH, Ngo HH, Zhang Z (2022) Exploring potential machine learning application based on big data for prediction of wastewater quality from different full-scale wastewater treatment plants. Sci Total Environ 832:154930. https://doi.org/10.1016/j.scitotenv.2022.154930

Malik SN, Khan SM, Ghosh PC, Vaidya AN, Kanade G, Mudliar SN (2019) Treatment of pharmaceutical industrial wastewater by nano-catalyzed ozonation in a semi-batch reactor for improved biodegradability. Sci Total Environ 678:114–122. https://doi.org/10.1016/j.scitotenv.2019.04.097

Mao Q, Ma X, Sun Y (2023) Study of impacts of blockchain technology on renewable energy resource findings. Renew Energy 211:802–808. https://doi.org/10.1016/j.renene.2023.05.038

Matheri AN, Mohamed B, Ntuli F, Nabadda E, Ngila JC (2022) Sustainable circularity and intelligent data-driven operations and control of the wastewater treatment plant. Phys Chem Earth Parts a/b/c 126:103152. https://doi.org/10.1016/j.pce.2022.103152

Mazhar S, Ditta A, Bulgariu L, Ahmad I, Ahmed M, Nadiri AA (2019) Sequential treatment of paper and pulp industrial wastewater: prediction of water quality parameters by Mamdani Fuzzy Logic model and phytotoxicity assessment. Chemosphere 227:256–268. https://doi.org/10.1016/j.chemosphere.2019.04.022

Miller T, Staves A, Maesschalck S, Sturdee M, Green B (2021) Looking back to look forward: lessons learnt from cyber-attacks on industrial control systems. Int J Crit Infrastruct Prot 35:100464. https://doi.org/10.1016/j.ijcip.2021.100464

Mishra A, Tripathy M, Ray P (2023) A survey on different techniques for distribution network reconfiguration. J Eng Res. https://doi.org/10.1016/j.jer.2023.09.001

Mo L, Lou S, Wang Y, Liu Z, Ren P (2024) Studying the evolutions, differences, and water security impacts of water demands under shared socioeconomic pathways: a SEMs-bootstrap-ANN approach applied to Sichuan Province. J Environ Manag 349:119455. https://doi.org/10.1016/j.jenvman.2023.119455

Mohammadi F, Bina B, Karimi H, Rahimi S, Yavari Z (2020) Modeling and sensitivity analysis of the alkylphenols removal via moving bed biofilm reactor using artificial neural networks: comparison of levenberg marquardt and particle swarm optimization training algorithms. Biochem Eng J 161:107685. https://doi.org/10.1016/j.bej.2020.107685

Mohebbi S, Zhang Q, Christian Wells E, Zhao T, Nguyen H, Li M, Abdel-Mottaleb N, Uddin S, Lu Q, Wakhungu MJ, Wu Z, Zhang Y, Tuladhar A, Ou X (2020) Cyber-physical-social interdependencies and organizational resilience: a review of water, transportation, and cyber infrastructure systems and processes. Sustain Cities Soc 62:102327. https://doi.org/10.1016/j.scs.2020.102327

Muniappan A, Tirth V, Almujibah H, Alshahri AH, Koppula N (2023) Deep convolutional neural network with sine cosine algorithm based wastewater treatment systems. Environ Res 219:114910. https://doi.org/10.1016/j.envres.2022.114910

Naik SS, Lee SJ, Theerthagiri J, Yu Y, Choi MY (2021) Rapid and highly selective electrochemical sensor based on ZnS/Au-decorated f-multi-walled carbon nanotube nanocomposites produced via pulsed laser technique for detection of toxic nitro compounds. J Hazard Mater 418:126269. https://doi.org/10.1016/j.jhazmat.2021.126269

Nam S-N, Yea Y, Park S, Park C, Heo J, Jang M, Park CM, Yoon Y (2023) Modeling sulfamethoxazole removal by pump-less in-series forward osmosis–ultrafiltration hybrid processes using artificial neural network, adaptive neuro-fuzzy inference system,

and support vector machine. Chem Eng J. https://doi.org/10.1016/j.cej.2023.145821
Negi BB, Aliveli M, Behera SK, Das R, Sinharoy A, Rene ER, Pakshirajan K (2023) Predictive modelling and optimization of an airlift bioreactor for selenite removal from wastewater using artificial neural networks and particle swarm optimization. Environ Res 219:115073. https://doi.org/10.1016/j.envres.2022.115073
Niu C, Li B, Wang Z (2023) Using artificial intelligence-based algorithms to identify critical fouling factors and predict fouling behavior in anaerobic membrane bioreactors. J Membr Sci. https://doi.org/10.1016/j.memsci.2023.122076
Nourani V, Zonouz RS, Dini M (2023) Estimation of prediction intervals for uncertainty assessment of artificial neural network based wastewater treatment plant effluent modeling. J Water Process Eng 55:104145. https://doi.org/10.1016/j.jwpe.2023.104145
Nti EK, Cobbina SJ, Attafuah EE, Senanu LD, Amenyeku G, Gyan MA, Forson D, Safo A-R (2023) Water pollution control and revitalization using advanced technologies: Uncovering artificial intelligence options towards environmental health protection, sustainability and water security. Heliyon 9:e18170. https://doi.org/10.1016/j.heliyon.2023.e18170
Oruganti RK, Biji AP, Lanuyanger T, Show PL, Sriariyanun M, Upadhyayula VKK, Gadhamshetty V, Bhattacharyya D (2023) Artificial intelligence and machine learning tools for high-performance microalgal wastewater treatment and algal biorefinery: a critical review. Sci Total Environ 876:162797. https://doi.org/10.1016/j.scitotenv.2023.162797
Osman AI, El-Monaem EMA, Elgarahy AM, Aniagor CO, Hosny M, Farghali M, Rashad E, Ejimofor MI, López-Maldonado EA, Ihara I, Yap P-S, Rooney DW, Eltaweil AS (2023) Methods to prepare biosorbents and magnetic sorbents for water treatment: a review. Environ Chem Lett 21:2337–2398. https://doi.org/10.1007/s10311-023-01603-4
Osman AI, Nasr M, Farghali M, Bakr SS, Eltaweil AS, Rashwan AK, Abd El-Monaem EM (2024) Machine learning for membrane design in energy production, gas separation, and water treatment: a review. Environ Chem Lett 22:505–560. https://doi.org/10.1007/s10311-023-01695-y
Özdoğan-Sarıkoç G, Sarıkoç M, Celik M, Dadaser-Celik F (2023) Reservoir volume forecasting using artificial intelligence-based models: artificial neural networks, support vector regression, and long short-term memory. J Hydrol 616:128766. https://doi.org/10.1016/j.jhydrol.2022.128766
Park J, Lee WH, Kim KT, Park CY, Lee S, Heo T-Y (2022) Interpretation of ensemble learning to predict water quality using explainable artificial intelligence. Sci Total Environ 832:155070. https://doi.org/10.1016/j.scitotenv.2022.155070
Patriarca R, Simone F, Di Gravio G (2022) Modelling cyber resilience in a water treatment and distribution system. Reliab Eng Syst Saf 226:108653. https://doi.org/10.1016/j.ress.2022.108653
Priya AK, Gnanasekaran L, Rajendran S, Qin J, Vasseghian Y (2022) Occurrences and removal of pharmaceutical and personal care products from aquatic systems using advanced treatment—a review. Environ Res 204:112298. https://doi.org/10.1016/j.envres.2021.112298
Qi J, Hou Y, Hu J, Ruan W, Xiang Y, Wei X (2020) Decontamination of methylene Blue from simulated wastewater by the mesoporous rGO/Fe/Co nanohybrids: artificial intelligence modeling and optimization. Mater Today Commun 24:100709. https://doi.org/10.1016/j.mtcomm.2019.100709
Radini S, Marinelli E, Akyol Ç, Eusebi AL, Vasilaki V, Mancini A, Frontoni E, Bischetti GB, Gandolfi C, Katsou E, Fatone F (2021) Urban water-energy-food-climate nexus in integrated wastewater and reuse systems: cyber-physical framework and innovations. Appl Energy 298:117268. https://doi.org/10.1016/j.apenergy.2021.117268
Raval KJ, Jadav NK, Rathod T, Tanwar S, Vimal V, Yamsani N (2024) A survey on safeguarding critical infrastructures: attacks, AI security, and future directions. Int J Crit Infrastruct Prot 44:100647. https://doi.org/10.1016/j.ijcip.2023.100647
Robbins CA, Du X, Bradley TH, Quinn JC, Bandhauer TM, Conrad SA, Carlson KH, Tong T (2022) Beyond treatment technology: Understanding motivations and barriers for wastewater treatment and reuse in unconventional energy production. Resour Conserv Recycl 177:106011. https://doi.org/10.1016/j.resconrec.2021.106011
Rolbiecki D, Paukszto Ł, Krawczyk K, Korzeniewska E, Sawicki J, Harnisz M (2023) Chlorine disinfection modifies the microbiome, resistome and mobilome of hospital wastewater—a nanopore long-read metagenomic approach. J Hazard Mater 459:132298. https://doi.org/10.1016/j.jhazmat.2023.132298
Safeer S, Pandey RP, Rehman B, Safdar T, Ahmad I, Hasan SW, Ullah A (2022) A review of artificial intelligence in water purification and wastewater treatment: recent advancements. J Water Process Eng 49:102974. https://doi.org/10.1016/j.jwpe.2022.102974
Sahu S, Kaur A, Singh G, Kumar Arya S (2023) Harnessing the potential of microalgae-bacteria interaction for eco-friendly wastewater treatment: a review on new strategies involving machine learning and artificial intelligence. J Environ Manag 346:119004. https://doi.org/10.1016/j.jenvman.2023.119004
Saleem M (2007) Pharmaceutical wastewater treatment: a physico-chemical study. J Res (science) 18:125–134
Sandner P, Gross J, Richter R (2020) Convergence of Blockchain, IoT, and AI. Front Blockchain. https://doi.org/10.3389/fbloc.2020.522600
Santín I, Barbu M, Pedret C, Vilanova R (2018) Fuzzy logic for plant-wide control of biological wastewater treatment process including greenhouse gas emissions. ISA Trans 77:146–166. https://doi.org/10.1016/j.isatra.2018.04.006
Serrano-Luján L, Toledo C, Colmenar JM, Abad J, Urbina A (2022) Accurate thermal prediction model for building-integrated photovoltaics systems using guided artificial intelligence algorithms. Appl Energy 315:119015. https://doi.org/10.1016/j.apenergy.2022.119015
Shah KM, Billinge IH, Chen X, Fan H, Huang Y, Winton RK, Yip NY (2022) Drivers, challenges, and emerging technologies for desalination of high-salinity brines: a critical review. Desalination 538:115827. https://doi.org/10.1016/j.desal.2022.115827
Shirkoohi MG, Tyagi RD, Vanrolleghem PA, Drogui P (2022) A comparison of artificial intelligence models for predicting phosphate removal efficiency from wastewater using the electrocoagulation process. Digit Chem Eng 4:100043. https://doi.org/10.1016/j.dche.2022.100043
Silversides KL, Melkumyan A, Wyman D (2016) Fusing Gaussian processes and dynamic time warping for improved natural gamma signal classification. Math Geosci 48:187–210. https://doi.org/10.1007/s11004-015-9601-2
Singh NK, Yadav M, Singh V, Padhiyar H, Kumar V, Bhatia SK, Show P-L (2023) Artificial intelligence and machine learning-based monitoring and design of biological wastewater treatment systems. Bioresour Technol 369:128486. https://doi.org/10.1016/j.biortech.2022.128486
Song Z, Mishra AR, Saeidi SP (2023) Technological capabilities in the era of the digital economy for integration into cyber-physical systems and the IoT using decision-making approach. J Innov Knowl 8:100356. https://doi.org/10.1016/j.jik.2023.100356
Swanckaert B, Geltmeyer J, Rabaey K, De Buysser K, Bonin L, De Clerck K (2022) A review on ion-exchange nanofiber membranes: properties, structure and application in electrochemical (waste)water treatment. Sep Purif Technol 287:120529. https://doi.org/10.1016/j.seppur.2022.120529

Świetlicka A, Kolanowski K (2023) Homogeneous ensemble model built from artificial neural networks for fault detection in navigation systems. J Comput Appl Math 432:115279. https://doi.org/10.1016/j.cam.2023.115279

Tabatabai-Yazdi F-S, Ebrahimian Pirbazari A, Esmaeili Khalil Saraei F, Gilani N (2021) Construction of graphene based photocatalysts for photocatalytic degradation of organic pollutant and modeling using artificial intelligence techniques. Phys B Condens Matter 608:412869. https://doi.org/10.1016/j.physb.2021.412869

Tang S, Cao Y (2023) A phenomenological neural network powered by the National Wastewater Surveillance System for estimation of silent COVID-19 infections. Sci Total Environ 902:166024. https://doi.org/10.1016/j.scitotenv.2023.166024

Tanha FE, Hasani A, Hakak S, Gadekallu TR (2022) Blockchain-based cyber physical systems: comprehensive model for challenge assessment. Comput Electr Eng 103:108347. https://doi.org/10.1016/j.compeleceng.2022.108347

Tariq R, Cetina-Quiñones AJ, Cardoso-Fernández V, Daniela-Abigail H-L, Soberanis MAE, Bassam A, De Lille MV (2021) Artificial intelligence assisted technoeconomic optimization scenarios of hybrid energy systems for water management of an isolated community. Sustain Energy Technol Assess 48:101561. https://doi.org/10.1016/j.seta.2021.101561

Thakur A (2022) A comprehensive study of the trends and analysis of distributed ledger technology and blockchain technology in the healthcare industry. Front Blockchain. https://doi.org/10.3389/fbloc.2022.844834

Theerthagiri J, Lee SJ, Karuppasamy K, Arulmani S, Veeralakshmi S, Ashokkumar M, Choi MY (2021) Application of advanced materials in sonophotocatalytic processes for the remediation of environmental pollutants. J Hazard Mater 412:125245. https://doi.org/10.1016/j.jhazmat.2021.125245

Theerthagiri J, Karuppasamy K, Lee SJ, Shwetharani R, Kim H-S, Pasha SKK, Ashokkumar M, Choi MY (2022a) Fundamentals and comprehensive insights on pulsed laser synthesis of advanced materials for diverse photo- and electrocatalytic applications. Light Sci Appl 11:250–250. https://doi.org/10.1038/s41377-022-00904-7

Theerthagiri J, Park J, Das HT, Rahamathulla N, Cardoso ESF, Murthy AP, Maia G, Vo DVN, Choi MY (2022b) Electrocatalytic conversion of nitrate waste into ammonia: a review. Environ Chem Lett 20:2929–2949. https://doi.org/10.1007/s10311-022-01469-y

Tikhamarine Y, Souag-Gamane D, Najah Ahmed A, Kisi O, El-Shafie A (2020) Improving artificial intelligence models accuracy for monthly streamflow forecasting using grey Wolf optimization (GWO) algorithm. J Hydrol 582:124435. https://doi.org/10.1016/j.jhydrol.2019.124435

Verma R, Suthar S (2018) Performance assessment of horizontal and vertical surface flow constructed wetland system in wastewater treatment using multivariate principal component analysis. Ecol Eng 116:121–126. https://doi.org/10.1016/j.ecoleng.2018.02.022

Victor J, Dourado A, Angelov P (2005) On-line construction and rule base simplification by replacement in fuzzy systems applied to a wastewater treatment planT. IFAC Proc 38:61–66. https://doi.org/10.3182/20050703-6-CZ-1902.02181

Wang B, Li Z, Dai Z, Lawrence N, Yan X (2019a) A probabilistic principal component analysis-based approach in process monitoring and fault diagnosis with application in wastewater treatment plant. Appl Soft Comput 82:105527. https://doi.org/10.1016/j.asoc.2019.105527

Wang P, Yao J, Wang G, Hao F, Shrestha S, Xue B, Xie G, Peng Y (2019b) Exploring the application of artificial intelligence technology for identification of water pollution characteristics and tracing the source of water quality pollutants. Sci Total Environ 693:133440. https://doi.org/10.1016/j.scitotenv.2019.07.246

Wang D, Li X, Hu L, Qiao J (2023a) Data-driven tracking control design with reinforcement learning involving a wastewater treatment application. Eng Appl Artif Intell 123:106242. https://doi.org/10.1016/j.engappai.2023.106242

Wang L, Jiang S, Shi Y, Du X, Xiao Y, Ma Y, Yi X, Zhang Y, Li M (2023b) Blockchain-based dynamic energy management mode for distributed energy system with high penetration of renewable energy. Int J Electr Power Energy Syst 148:108933. https://doi.org/10.1016/j.ijepes.2022.108933

Wei X, Li B, Zhao S, Wang L, Zhang H, Li C, Wang S (2012) Mixed pharmaceutical wastewater treatment by integrated membrane-aerated biofilm reactor (MABR) system—a pilot-scale study. Bioresour Technol 122:189–195. https://doi.org/10.1016/j.biortech.2012.06.041

Werkneh AA, Islam MA (2023) Post-treatment disinfection technologies for sustainable removal of antibiotic residues and antimicrobial resistance bacteria from hospital wastewater. Heliyon 9:e15360. https://doi.org/10.1016/j.heliyon.2023.e15360

Wu X, Wang Y, Wang C, Wang W, Dong F (2021) Moving average convergence and divergence indexes based online intelligent expert diagnosis system for anaerobic wastewater treatment process. Bioresour Technol 324:124662. https://doi.org/10.1016/j.biortech.2020.124662

Xia W, Chen X, Song C (2022) A framework of blockchain technology in intelligent water management. Front Environ Sci. https://doi.org/10.3389/fenvs.2022.909606

Yang LB (2020) Application of artificial intelligence in electrical automation control. Proc Comput Sci 166:292–295. https://doi.org/10.1016/j.procs.2020.02.097

Yang Q, Xu R, Wu P, He J, Liu C, Jiang W (2021) Three-step treatment of real complex, variable high-COD rolling wastewater by rational adjustment of acidification, adsorption, and photocatalysis using big data analysis. Sep Purif Technol 270:118865. https://doi.org/10.1016/j.seppur.2021.118865

Yang X, Wang Y, Zhang Y, Yao W, Wen J (2022) Impact analysis of cyber system in microgrids: perspective from economy and reliability. Int J Electr Power Energy Syst 135:107422. https://doi.org/10.1016/j.ijepes.2021.107422

Yao F, Qin Z, Wang X, Chen M, Noor A, Sharma S, Singh J, Kozak D, Hunjet A (2023) The evolution of renewable energy environments utilizing artificial intelligence to enhance energy efficiency and finance. Heliyon 9:e16160. https://doi.org/10.1016/j.heliyon.2023.e16160

Ye Z, Yang J, Zhong N, Tu X, Jia J, Wang J (2020) Tackling environmental challenges in pollution controls using artificial intelligence: a review. Sci Total Environ 699:134279. https://doi.org/10.1016/j.scitotenv.2019.134279

Yu Y, Min A, Jung HJ, Theerthagiri J, Lee SJ, Kwon KY, Choi MY (2021) Method development and mechanistic study on direct pulsed laser irradiation process for highly effective dechlorination of persistent organic pollutants. Environ Pollut 291:118158–118158. https://doi.org/10.1016/j.envpol.2021.118158

Yuan S, Ajam H, Sinnah ZAB, Altalbawy FMA, Abdul Ameer SA, Husain A, Al Mashhadani ZI, Alkhayyat A, Alsalamy A, Zubaid RA, Cao Y (2023) The roles of artificial intelligence techniques for increasing the prediction performance of important parameters and their optimization in membrane processes: a systematic review. Ecotoxicol Environ Saf 260:115066. https://doi.org/10.1016/j.ecoenv.2023.115066

Zaghloul MS, Achari G (2022) Application of machine learning techniques to model a full-scale wastewater treatment plant with biological nutrient removal. J Environ Chem Eng. https://doi.org/10.1016/j.jece.2022.107430

Zahmatkesh S, Amesho KTT, Sillanpaa M, Wang C (2022) Integration of renewable energy in wastewater treatment during COVID-19

pandemic: challenges, opportunities, and progressive research trends. Clean Chem Eng 3:100036. https://doi.org/10.1016/j.clce.2022.100036

Zhan J, Li Z, Yu G, Pan X, Wang J, Zhu W, Han X, Wang Y (2019) Enhanced treatment of pharmaceutical wastewater by combining three-dimensional electrochemical process with ozonation to in situ regenerate granular activated carbon particle electrodes. Sep Purif Technol 208:12–18. https://doi.org/10.1016/j.seppur.2018.06.030

Zhang C, Wang Z (2022) Linex-RSVM: Ramp Linex Support Vector Machine. Proc Comput Sci 199:524–531. https://doi.org/10.1016/j.procs.2022.01.064

Zhang Z, Zhou Y, Han L, Guo X, Wu Z, Fang J, Hou B, Cai Y, Jiang J, Yang Z (2022) Impacts of COVID-19 pandemic on the aquatic environment associated with disinfection byproducts and pharmaceuticals. Sci Total Environ 811:151409. https://doi.org/10.1016/j.scitotenv.2021.151409

Zhang S, Jin Y, Chen W, Wang J, Wang Y, Ren H (2023a) Artificial intelligence in wastewater treatment: a data-driven analysis of status and trends. Chemosphere 336:139163. https://doi.org/10.1016/j.chemosphere.2023.139163

Zhang S, Omar AH, Hashim AS, Alam T, Khalifa HAE-W, Elkotb MA (2023b) Enhancing waste management and prediction of water quality in the sustainable urban environment using optimized algorithm of least square support vector machine and deep learning techniques. Urban Clim 49:101487. https://doi.org/10.1016/j.uclim.2023.101487

Zhang X, Lu C, Tian J, Zeng L, Wang Y, Sun W, Han H, Kang J (2024) Artificial intelligence optimization and controllable slow-release iron sulfide realizes efficient separation of copper and arsenic in strongly acidic wastewater. J Environ Sci 139:293–307. https://doi.org/10.1016/j.jes.2023.05.038

Zhao L, Dai T, Qiao Z, Sun P, Hao J, Yang Y (2020) Application of artificial intelligence to wastewater treatment: a bibliometric analysis and systematic review of technology, economy, management, and wastewater reuse. Process Saf Environ Prot 133:169–182. https://doi.org/10.1016/j.psep.2019.11.014

Zhu M, Ji Y, Zhu X, Ren K (2023) Energy consumption mode identification and monitoring method of process industry system under unstable working conditions. Adv Eng Inf 55:101893. https://doi.org/10.1016/j.aei.2023.101893

3.2 Links and implications

The convergence between blockchain technology, Internet of Things, AI, and ML in cyber-physical systems synergistically enhances trust, transparency, privacy, and cyber-security in overall operational systems in pharmaceutical wastewater treatment systems. This benefit can be extended to other types of wastewater treatment systems, such as textile dye wastewater, biological treatment or physicochemical treatment systems. For renewable energy-driven processes, the AI-enabled smart grid distribution network improves the adaptability of solar energy into existing wastewater treatment systems to achieve energy efficiency and sustainability. AI-integrated systems can also help identify and analyse security vulnerabilities or risks in IT infrastructure, monitor complex dynamic behaviours, uncertainties, or system perturbations, and manage data in the process control systems of wastewater treatment plants.

Furthermore, distributed energy systems improve the utilisation efficiency of renewable energy technologies with the support of blockchain technology. However, the main disadvantage of hybrid AI/ML ensembles generated models involves complicated design constraints and uncertainties in predicted data arising from data clustering, making it challenging to determine exact data patterns. The scalability of blockchain technology can pose a significant issue due to the incompatibility of software and hardware systems, data security vulnerabilities and interoperability of existing IT systems. System-wide compromise arising from incompatible AI software with existing operational systems can lead to a cascade of design errors, malfunctions and vulnerability to cyber-attacks on critical IT infrastructure. Nonetheless, the assistance of AI-powered technologies can help improve the reliability, computing efficiency, and overall performance of wastewater treatment systems through the standardisation of frameworks and compliance with assessment metrics.

4.1 Introduction

Among all aquatic pollutants, xenobiotic dyes and pigments are the most ubiquitous organic contaminants. These organic contaminants are highly resistant to environmental degradation due to the chemical stability of dye compounds. Xenobiotic dyes are anthropogenic contaminants often discharged into water from industrial effluents, especially from textile manufacturing, food processing, dye and paint industries. Xenobiotic dyes are carcinogenic and ecologically toxic if it finds its way into food chain. Similar to pharmaceutically active compounds, the breakdown byproducts of xenobiotic dyes can be even more toxic due to incomplete oxidation or mineralisation of intermediate transformation byproducts from its parental compounds.

Furthermore, there are several ways to treat xenobiotic dyes in contaminated water, such as adsorption, membrane filtration, chemical coagulation-flocculation, photocatalytic degradation, Fenton process, biological treatment, electrochemical treatment and many more. However, there are some drawbacks associated with different treatment methods. Membrane filtration method often leads to membrane fouling due to the buildup of contaminants on membrane surfaces. Coagulation-flocculation techniques can require chemical additives and contribute to secondary pollution if improperly managed. Biological treatment can generate excessive sludge production, resulting in high transportation and sludge management costs. Among these treatment methods, the simplest and most practical method is adsorption. It is highly effective at removing pollutants non-selectively with minimal sludge generation and does not require any pre-treatment process. There have been several attempts to regenerate the adsorptive capacity of carbon-based adsorbents using electrochemical techniques. However, activated carbon adsorbents have high porosity and low electrical conductivity, limiting its ability to regenerate its adsorptive capacity electrochemically. Thermal regeneration of activated carbon adsorbent was investigated, but the disposal of exhausted activated carbon adsorbents into landfill can pose a significant environmental issue. Incineration of activated carbon also contributes to secondary pollution. Therefore, this technical experiment used graphite intercalation compound (GIC) as a suitable alternative to activated carbon adsorbents due to its ability to electrochemically regenerate. GIC has high electrical conductivity compared to conventional activated carbon and little to no porosity. It can be electrochemically regenerated in situ by an electrochemical oxidation process. When used in a sequential batch electrochemical reactor, GIC can be electrochemically regenerated. On the other hand, organic pollutants can be further degraded with the additional support of

anodic oxidation technology in the reactor. The combined hybrid treatment method of GIC adsorbents and direct electrochemical oxidation of the reactor gives rise to a three-dimensional electrochemical reactor.

Furthermore, a three-dimensional electrochemical reactor contains GIC particle electrodes and anodic oxidation technology to synergistically enhance the electrochemical degradation of xenobiotic dye contaminants more effectively than a conventional, single wastewater treatment method. A three-dimensional electrochemical oxidation can effectively generate hydroxyl radicals and active chlorinated species to mineralise the xenobiotic dyes into inert carbon dioxide and water provided that the complete oxidation process can be achieved. There are a range of operational parameters to consider, such as adjusting current density, electrolysis time, adsorbent dosage and initial dye concentration to improve pollutant removal efficiency.

This research article comprehensively studies adsorption kinetics, isotherms and electrochemical oxidation mechanisms of a three-dimensional electrochemical oxidation technology. A range of error functions, linear and non-linear regression analyses are used to determine the accuracy and precision of process variables. Response surface methodology (Kanneganti et al.) optimisation technique was applied to GIC adsorbents to optimise the adsorption process. The salting and thermal effects of simulated, alkaline dye-contaminated wastewater on selectivity reversal of GIC adsorbents were critically evaluated using RSM optimisation technique.

International Journal of Environmental Science and Technology
https://doi.org/10.1007/s13762-024-05696-4

ORIGINAL PAPER

Removal of reactive black 5 in water using adsorption and electrochemical oxidation technology: kinetics, isotherms and mechanisms

V. Ganthavee[1] · A. P. Trzcinski[1]

Received: 15 July 2023 / Revised: 12 April 2024 / Accepted: 25 April 2024

Abstract
In this work, a novel graphite intercalation compound (GIC) particle electrode was used to investigate the adsorption of Reactive Black 5 (RB5) and the electrochemical regeneration in a three-dimensional (3D) electrochemical reactor to recover its adsorptive capacity. Various adsorption kinetics and isotherm models were used to characterise the adsorption behaviour of GIC. Several adsorption kinetics were modelled using linearised and non-linearised rate laws to evaluate the viability of the sorption process. Studies on the selective removal of RB5 dyes from binary mixture in solution were evaluated. RSM optimisation studies were integrated with ANOVA analysis to provide insight into the significance of selectivity reversal from the salting effect of textile dye solution on GIC adsorbent. A unique range of adsorption kinetics and isotherms were used to evaluate the adsorption process. Non-linear models best simulated the kinetic data in the order: Elovich > Bangham > Pseudo-second-order > Pseudo-first-order. The Redlich–Peterson isotherm was calculated to have a dye loading capacity of 0.7316 mg/g by non-linear regression analysis. An error function analysis with ERRSQ/SSE of 0.1390 confirmed the accuracy of dye loading capacity predicted by Redlich–Peterson isotherm using non-linear regression analysis. The results showed that Redlich–Peterson and SIPS isotherm models yielded better fitness to experimental data than the Langmuir type. The best dye removal efficiency achieved was ~ 93% using a current density of 45.14 mA/cm^2, whereas the highest TOC removal efficiency achieved was 67%.

Keywords Dye · Adsorption · Electrochemical oxidation · Graphite intercalation compound · Kinetics · Isotherms

Introduction

Organic contaminants such as dyes and pigments are recalcitrant pollutants highly resistant to environmental biodegradation. These pollutants are often discharged into water from industrial effluents such as dye and paint manufacturing, food processing, textile, etc. Most of these dyes are synthetic, which means they can be highly toxic and carcinogenic if it finds its way into the food chain. Therefore, there is an urgent need to remove the dyes from liquid waste until they fall below a specific concentration acceptable to environmental regulatory authorities. There are several ways to remove dyes from water, such as adsorption (El-Kammah et al. 2022; Noorimotlagh et al. 2019), membrane filtration (Ma et al. 2022; Mansor et al. 2020), chemical coagulation and flocculation (Lau et al. 2014; Szygula et al. 2009), photocatalytic degradation (Fernandes et al. 2019; Tekin 2014), electrocoagulation/electroflotation (Balla et al. 2010), Fenton process (Suhan et al. 2021), electrochemical oxidation (Kumar and Gupta 2022; Song et al. 2010) etc. Adsorption is an attractive and effective approach to removing contaminants from water, particularly when the adsorbent is cheap, has minimal sludge generation and does not require an additional pre-treatment process (Sultana et al. 2022). The adsorption method is simple and cost-effective, and high dye removal efficiency can be achieved (Chaiwichian and Lunphut 2021). Several attempts have been made to electrochemically regenerate activated carbon adsorbents (Narbaitz and McEwen 2012; Xing et al. 2023). However, low electrical conductivity of liquid media limited the energy

Editorial responsibility: Samareh Mirkia.

✉ V. Ganthavee
Voravich.Ganthavee@usq.edu.au

[1] School of Agriculture and Environmental Science, University of Southern Queensland, 487-535 West St, Darling Heights, Toowoomba, QLD 4350, Australia

Published online: 21 May 2024

required to regenerate activated carbon adsorbents (Karimi-Jashni and Narbaitz 2005). Activated carbon adsorbents have a high degree of porosity, which affects its ability to regenerate successfully. Offsite thermal regeneration is required for activated carbon. However, exhausted activated carbon is often disposed into a landfill or incinerated, which may contribute to secondary pollution. Hence, graphite intercalation compound (GIC) can be used as an alternative to removing dissolved organic contaminants from water. GIC can be electrochemically regenerated in situ by an electrochemical oxidation process. It is electrically conductive and requires a short adsorption time to reach equilibrium (Mohammed et al. 2012). However, the specific surface area of GIC is significantly smaller than activated carbon.

Previous studies have shown that the combined adsorption and electrochemical regeneration of GIC can maximise dye removal efficiency (Hussain et al. 2015b; Mohammed et al. 2012). Optimising operating conditions allows the formation of chlorinated breakdown products to be minimised during electrochemical regeneration to prevent side reactions (Hussain et al. 2015a). However, there is a lack of studies focusing on the adsorption models used to model GIC using statistical analysis and comparison of linear and non-linear regression models to describe the adsorption mechanisms of GIC. In this paper, we report the kinetics and isotherm studies used to examine the adsorption behaviour of GIC. In addition, the effect of current density in the subsequent electrochemical oxidation process was also investigated to optimise the dye and TOC removal efficiencies.

Materials and methods

Adsorbate

Reactive Black 5 (RB5) dye, with an empirical formula of $C_{26}H_{21}N_5Na_4O_{19}S_6$ was purchased from Sigma-Aldrich with the product number 306452. The chemical structure of the Reactive Black 5 dye is shown below. A stock solution of 100 mg/L was prepared from the dissolution of RB5 in water.

Adsorbent preparation

The adsorbent used in this study is an expandable graphite intercalation compound (GIC) purchased from Sigma–Aldrich (P/N: 808121). At least 75% of the flakes were larger than 300 microns (See Supplementary Material S1). GIC has no porous structure and a relatively low electroactive surface area of approximately 1 m^2/g (Hussain et al. 2016). It is highly conductive, with 0.8 S/cm (Asghar et al. 2014).

Determination of pH point of zero charge (pH_{pzc})

The point of zero charge (pzc) of the GIC adsorbent was determined using a pH meter (Eutech instruments, PC 2700). A mixture of 50 mg GIC in 40 mL of distilled water containing 0.1 M HCl and 0.1 M NaOH to adjust the pH levels ranging between 3 and 12. The pH was measured after 24 h of equilibration at an ambient temperature of 22.5 °C. The pH_{pzc} is defined as the pH at which the surface charge of the GIC adsorbent is neutral.

Adsorption studies

The adsorption kinetic studies were performed to determine the adsorption system's reaction order and identify the rate-limiting process. The adsorption studies were performed for Reactive Black 5 dye at various initial concentrations using a GIC adsorbent dosage of 23 g/L, which was within the ratios of adsorbent concentration estimated across different works of literature such as Hussain et al. (2015b), Liu et al. (2016) and Purkait et al. (2007). The mechanical agitation speed was adjusted to 300–400 rpm at 22 °C. The adsorption time was between 0 and 120 min until the adsorption equilibrium for various initial dye concentrations was reached. In order to investigate the effect of salt on the selectivity of GIC, adsorption experiments were carried out using sodium chloride (NaCl, ≥ 99%, Sigma-Aldrich, Australia) at concentration ranging from 1 to 10 g/L. The adsorbent loading, q_t (mg/g), is to be determined from the initial and final concentrations at specific time intervals as given in Eq. (1):

$$q_t = \frac{\left(C_i - C_f\right)}{m} V \tag{1}$$

where C_i and C_f are the initial and final concentrations (mg/L) of Reactive Black 5 solution, V is the volume (L) of a solution, and m is the mass (g) of the adsorbent used. Microsoft Excel Solver Add-In (2022) was used to perform graphical plots, curve fitting and mathematical models for adsorption kinetics and isotherms by comparing the experimental and theoretical data. The chi-square test is a critically important error function in determining the best fit of isotherm model applicable to the tested adsorption system. It estimates the difference in squares between the theoretical data based on the predicted model and the experimental data collected and then divides each difference by the corresponding experimental value.

The influence of the variables was examined at three different levels, and the values of the variables at each level are shown in Table 1. The total number of experimental runs was designed using the CCD method by Eq. (2):

Table 1 Range and codification of the independent variable (x_i) were used in the experimental design

Independent variables	Symbol	−α	−1	0	+1	+α
Temperature, (°C)	X_1	30	36.0809	45	53.9191	60
pH	X_2	3	4.6216	7	9.3784	11
Salt concentration (g/L)	X_3	1	2.8243	5.5	8.1757	10

$$n = 2^i + 2i + n_0 \quad (2)$$

where n is the total number of experimental runs, i is the number of independent variables, and n_0 is the number of centre points. A total of 20 experiments were designed. A quadratic polynomial equation can be used to mathematically express the relationship between the independent variables and the responses, as shown in Eq. (3):

$$Y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i=1}^{k=1} \sum_{j=i+1}^{k} \beta_{ij} x_i x_j + \gamma \quad (3)$$

where Y is the predicted response, β_0 is the constant term, β_0, β_{ij} and β_{ii} are the linear, second-order and interaction coefficients, respectively, x_i and x_j are the independent variables, and $x_i x_j$ represents the first-order interaction between x_i and x_j. k is the number of independent variables, and γ is a random error.

Adsorption and electrochemical regeneration

Electrochemical regeneration of GIC adsorbent was carried out in a batch electrochemical cell, as shown in Fig. 1.

Direct anodic oxidation of dye pollutants occurs via graphite anode, whereas indirect oxidation occurs via electrogenerated highly reactive oxidising species such as hydroxyl radicals and active chlorine species. The adsorption-electrochemical treatment process consists of three essential steps: initial adsorption-regeneration and re-adsorption via electrochemical treatment (Hussain et al. 2015b). In this experiment, 2,000 mL of the 100 mg/L of RB5 stock solution was subjected to electrochemical treatment. The dye solution was poured into a sequential batch reactor, and adsorption occurred when the air compressor was switched on. The air was sparged into the bottom of the chamber to facilitate mixing between GIC and dye solution. The active area of each electrode used in the reactor was 70 cm^2. The interelectrode distance between the anode and cathode was 6.3 cm, and about 200 g of GIC was added to fill the regeneration zone (12 cm deep and 5 cm thick) before pouring the dye solution. The volume of the reactor chamber was about 6–7 L. A 30 V DC power supply unit was connected to the graphite anode and 316-grade stainless steel cathode to facilitate an electrochemical redox reaction. The cathodic compartment was filled with 0.3% (w/v) acidified brine solution pH of 1–2 separated from the dye solution by a Daramic membrane (Daramic, USA). The current supply ranged from 0.21 to 3.16 A, corresponding to a current density of 3 to 45.14 mA/cm^2.

Analytical methods

Surface characterisation of GIC can be performed using IRAAffinity-1S, GladiaATR 10 Shimadzu Fourier Transform

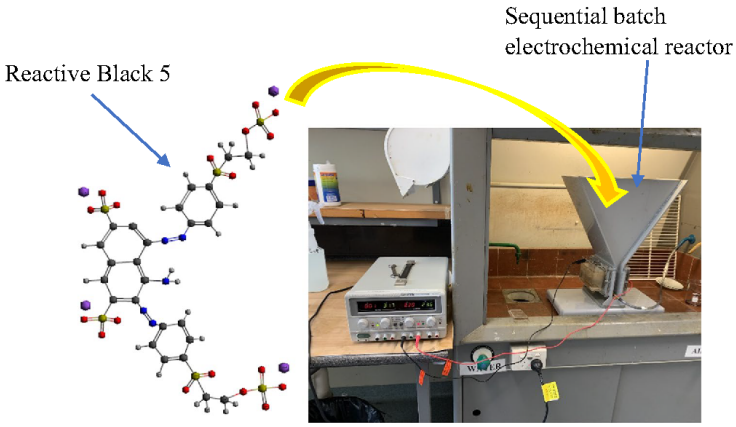


Fig. 1 RB5 textile dye wastewater in the sequential batch electrochemical reactor

Infrared Spectroscopy (FTIR) by examining its surface functional group or surface chemistry. Scanning Electron Microscopy (SEM) is used to investigate the surface morphology or textural properties of the GIC. A detailed analysis of GIC characteristics using FTIR and SEM is shown in Supplementary Material.

UV/Visible spectrophotometer (DR6000, Hach) determined the RB5 dye concentrations in solution at different time intervals. The maximum absorption occurred at a wavelength $\lambda = 596$ nm as specified by previous authors Feng et al. (2022), Saroyan et al. (2019) and Droguett et al. (2020). TOC-V CSH Shimadzu TOC analyser is used to determine the degree of mineralisation of the Reactive Black 5 dye before and after the electrochemical treatment. A Hach DR6000 UV/Visible Spectrophotometer was used to determine the RB5 concentrations in the samples. The Coefficient of Variation (de Fouchécour et al. 2022) for UV-absorbance analysis of RB5 is approximately 0.31%, whereas for TOC analysis is approximately 1.69%.

Results and discussion

Adsorption studies

The adsorption kinetics pertain to the uptake rate of RB5 adsorbates by GIC. Figure 2a shows the dye removal efficiency for various initial dye concentrations ranging from 33.81 to 136.67 mg/L. The dye removal efficiencies for different initial dye concentrations fluctuated throughout the adsorption period. This adsorption phenomenon indicated that the dye molecules adsorbed onto the GIC adsorbent and desorbed from the adsorbent surface back and forth due to weak intermolecular interaction. Dye sorption usually occurs when the dye molecules are diffused from the bulk liquid onto the adsorbent surface through the solid–liquid interface between the adsorbent surface and the bulk solution. The dye molecules interact with the active sites on the surface of the adsorbent through intermolecular interactions, which are held by either van der Waals' forces, electrostatic interactions, hydrophobic interactions, π–π electron donor–acceptor interactions, hydrogen bonding etc. The binding process is reversible if the solute-substrate interaction is held by weak intermolecular forces such as van der Waals' forces. The binding process is irreversible if there is a strong electrostatic attraction between the solute and substrate. In addition, Fig. 2a shows the dye removal efficiencies for various initial dye concentrations. It was found that the higher the initial dye concentration, the lower the overall dye removal efficiency due to the limited availability of active sites and rapid uptake of adsorbates onto active sites.

Figure 2b indicates that adsorbent loading increased within the first 0–8 min for initial RB5 dye concentrations ranging from 33.81 to 136.67 mg/L, eventually reaching an adsorption equilibrium. However, continuous adsorption–desorption occurred throughout the experimental conditions even after 100 min of adsorption time. The rise and fall of adsorbent loadings strongly indicated that the dye molecules continuously adsorbed and desorbed from the surface of the GIC adsorbent due to weak intermolecular interactions. This indicated the reversibility of the solute-substrate binding process.

On the other hand, pseudo-second-order kinetic models showed an excellent curve fitting between experimental and theoretical data (Fig. 3a), indicating that it is suitable to describe the adsorption–desorption kinetics in experimental studies. Non-linearised pseudo-first-order kinetic models yield better fitness than the linearised pseudo-second-order kinetic model (Table 1). The desorption process increased dye concentration in the bulk liquid, creating a greater concentration gradient and a greater diffusion flux of dye adsorbates re-adsorbing onto the GIC adsorbent surface. Although the adsorption kinetics of RB5 exhibited a slow rate-limiting step, the high concentration gradient improved the surface diffusivity or diffusion flux of the dye adsorbates onto the GIC adsorbent surface, reaching a rapid equilibrium point within 3 min of adsorption time.

The Bangham model is derived from the generalisation of the Weber and Morris model (Largitte and Pasquier 2016; Sumanjit et al. 2016). The kinetic data obtained from this model can be used to determine the slow limiting step in the adsorption system, as shown in Table 1. The equation indicates that q denotes the dye loading per mass of adsorbent (mg/g), θ and k represent kinetic constants. This kinetic model can be used to verify whether the rate-limiting step influences the pore or surface diffusion.

Furthermore, a novel three-stage kinetic model was used to describe the adsorption of RB5 onto GIC adsorbent. The model is based on the concept of mass conservation for RB5 combined with three distinct stages of adsorption onto GIC (Choi et al. 2007): 1) the first portion of the plot represents an instantaneous stage or external surface adsorption; 2) the second portion of the gradual stage in the plot represents the rate-limiting intraparticle diffusion; and 3) the third portion for a constant stage represents the aqueous phase which no longer interacts with GIC. The analytical three-stage kinetic model can be validated with the kinetic data from the batch experimental study involving RB5 adsorption onto the GIC.

The significance of developing a three-stage kinetic model for each initial dye concentration through modelling studies stems from strong evidence that different stages of adsorption exist in the kinetic models. The first stage involves a sharper portion of the adsorption curvature, representing the rapid decrease in the aqueous phase concentration. Dye

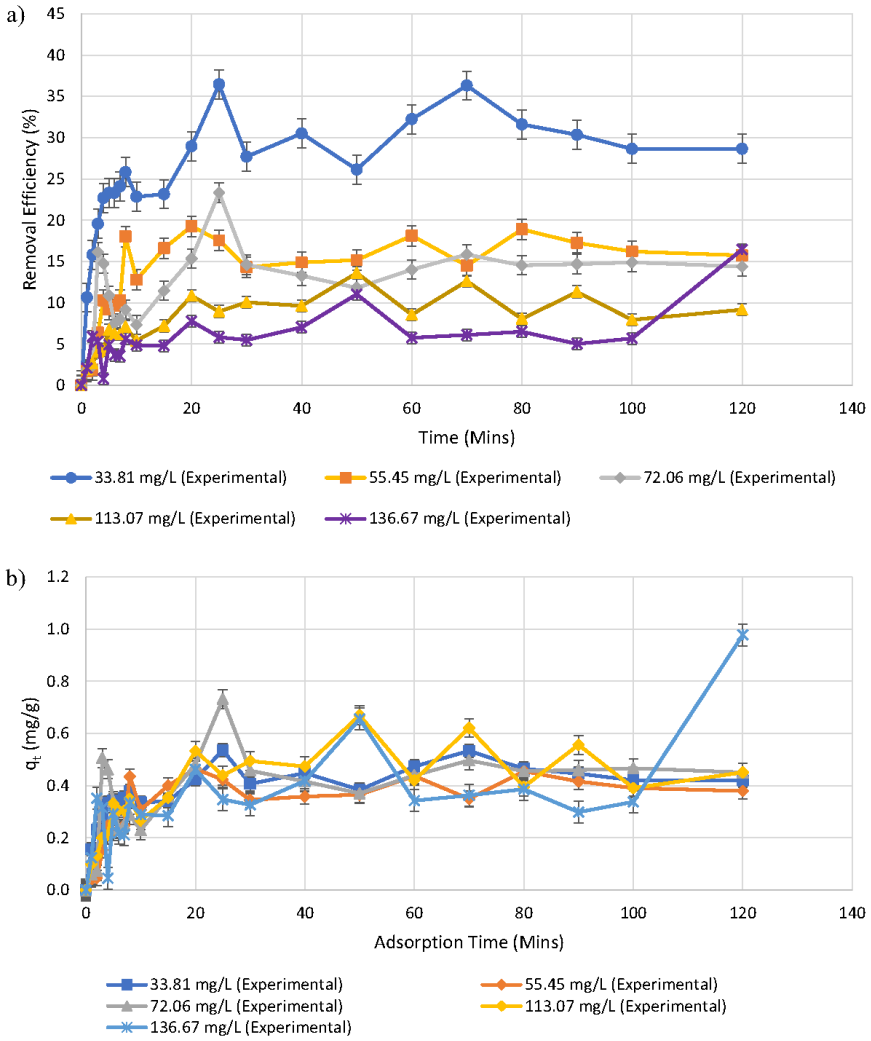


Fig. 2 **a** The changes in dye removal efficiencies over a period of time for various initial dye concentrations; **b** the adsorbent loading of Reactive Black 5 dye onto GIC over adsorption time from 0 to 120 min

molecules diffuse into the bulk solution, moving across the solid–liquid interface and binding to the active sites on the external surface of the adsorbent by instantaneous adsorption. The second gradual decline portion represents the slow rate-limiting adsorption influenced by intraparticle or surface diffusion. In contrast, the third constant line portion represents a minimal change in aqueous phase concentration or indicates equilibrium has been reached. Compared to standard kinetic expressions, pseudo-second-order equation, and Elovich kinetic equation, these three-stage kinetic models are purely empirical. They do not discriminate between the different adsorption stages, such as instantaneous versus rate-limiting adsorption (Choi et al. 2007).

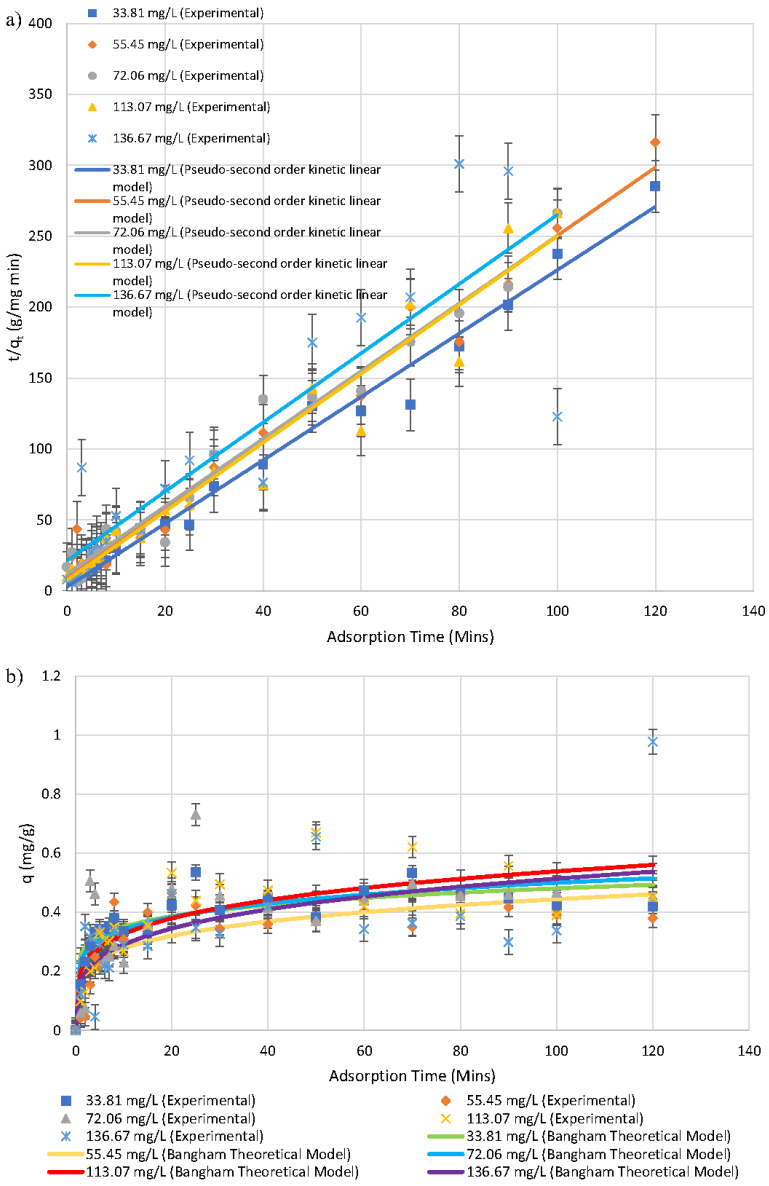


Fig. 3 **a** The linearised form of the pseudo-second-order kinetic model for Reactive Black 5 dye adsorption onto GIC adsorbent. The orange and blue lines represent linear regression analysis; **b** Bangham's kinetic model by non-linear regression method represents the slow rate-limiting step of surface diffusion and intraparticle diffusion of RB5 dye onto GIC adsorbent; **c** The pseudo-first-order kinetic model by non-linear regression analysis for Reactive Black 5 dye adsorption onto the GIC adsorbent; **d** The pseudo-second-order kinetic model by non-linear regression analysis for Reactive Black 5 dye adsorption onto the GIC adsorbent; **e** The Elovich kinetic model for Reactive Black 5 dye adsorption onto GIC adsorbent; **f** The three-stage kinetic model for RB5 dye adsorption onto GIC adsorbent

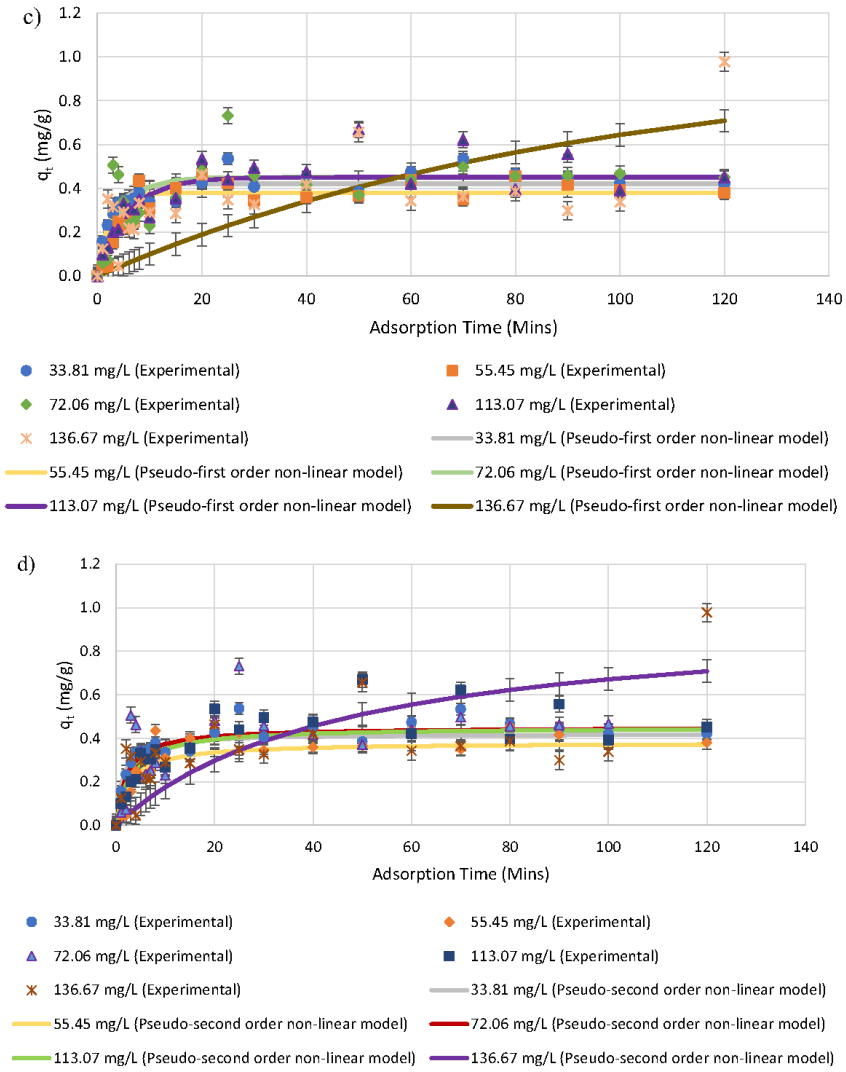


Fig. 3 (continued)

In a kinetic batch system, the adsorbate in the aqueous phase binds to the active sites on the adsorbent surface in a finite system, resulting in the decline of the aqueous phase concentration. In the first two stages of adsorption, the instantaneous adsorption in the first stage involves instantaneous adsorption followed by slow rate-limiting adsorption in the remaining stage as it moves towards the equilibrium point. The rate-limiting adsorption, which involves the entire mass transfers of adsorbate in the aqueous phase, moves towards the solid phase of the adsorbent through surface diffusion until it fills the active sites of the adsorbent with adsorbate, thereby reaching adsorption equilibrium. The second stage of the adsorption process primarily accounts for time-dependent adsorption, followed by the third stage,

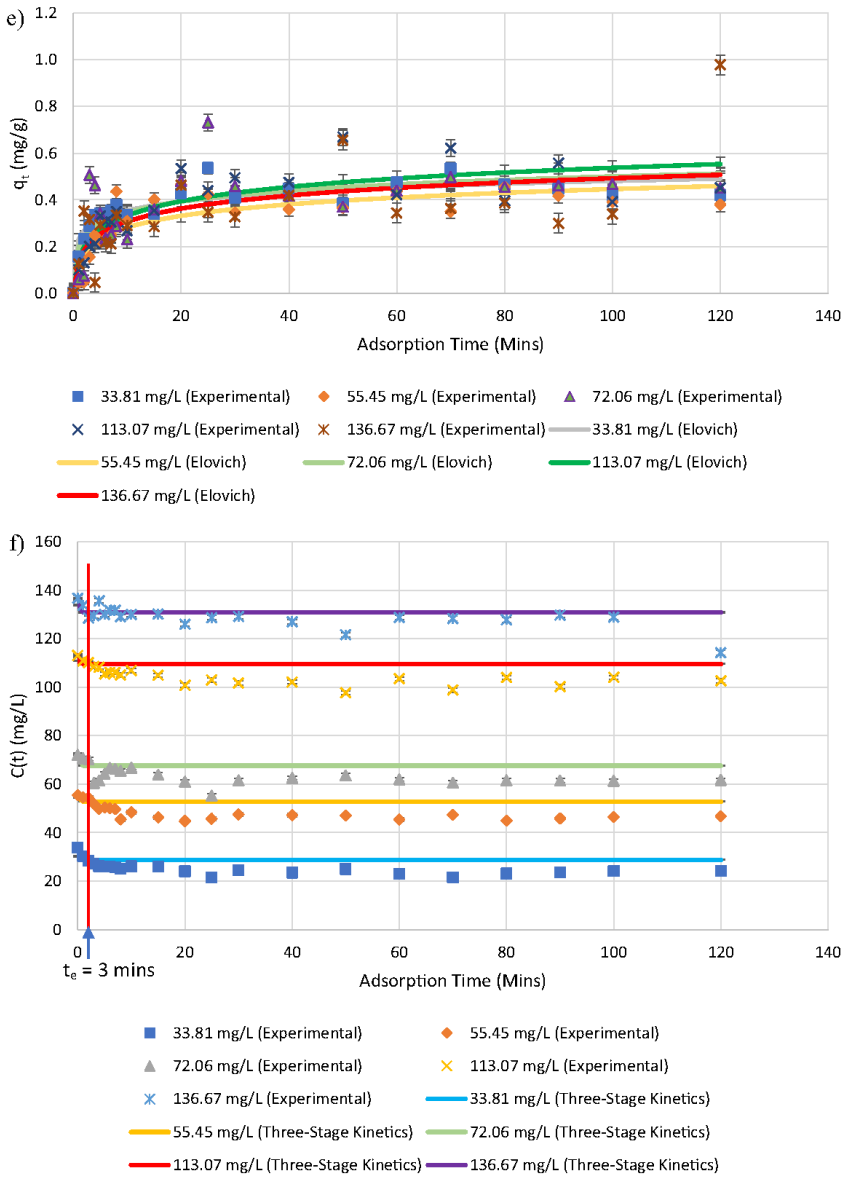


Fig. 3 (continued)

where the aqueous phase concentration remains constant throughout batch adsorption. Furthermore, it can be deduced that the existence of instantaneous adsorption (ξ_1) implies that some portion of the Reactive Black 5 adsorbate was adsorbed onto the external site of the GIC adsorbent within a transient period of time. The higher the initial adsorbate

concentration, especially when the initial dye concentration increased from 33.81 to 136.67 mg/L, resulted in a lower ξ_1 value for both two and three-stage kinetic models. The lower ξ_1 values compared to ξ_2 values, as the initial adsorbate concentration increased, indicated higher intraparticle diffusion. The continuous decrease of the aqueous phase of adsorbate concentration within 2–3 min indicated that the surface diffusion was still ongoing due to the availability of surface area or active sites on the particle size of GIC adsorbent. Due to the continuous decrease in the aqueous phase of adsorbent concentration, the final stage of kinetics had not reached the equilibrium point. Nonetheless, all three stages of kinetic models fitted well with the experimental data, with slight differences between the α and ξ_1 values for initial adsorbate concentrations ranging between 33.81 and 72.06 mg/L of RB5 dye. Moreover, the higher the initial adsorbate concentration, the more pronounced the three stages, with an especially more distinctive appearance of the second stage due to higher $q_2(\infty)$ value. This result demonstrated that different adsorption stages largely depend on the initial adsorbate concentration and the adsorbent's physicochemical properties or surface characteristics. The instantaneous adsorption (ξ_1) of GIC adsorbent decreased from 0.1064 to 0.0209 as the initial adsorbate concentration increased from 33.81 to 136.67 mg/L. As the initial adsorbate concentration increased, the fast decline in instantaneous adsorption became more pronounced. A high concentration gradient between the bulk liquid solution and the external site of the adsorbent would result in greater surface diffusivity. As the initial adsorbate concentration increased, high competition between adsorbate molecules reduced the number of active sites available for adsorption, resulting in a decrease in β values and, consequently, a reduction in adsorption rate followed by reduced α values. As the initial adsorbate concentration increased, the instantaneous adsorption portion of the parameter, ξ_1 value decreased while the two-stage parameter, ξ_2 value increased, leading to a longer equilibrium time. Conversely, the higher ξ_1 value, the lower the second-stage adsorption, which may result in a relatively long reaction time until the active sites are filled. However, integrating the third stage model improved the system's flexibility with greater β yields, resulting in a relatively short duration of the second stage. On the other hand, the three-stage kinetic model improved the flexibility of second-stage adsorption by increasing the β yields, thereby compensating the instantaneous portion of ξ_1 and shortening the equilibrium time for the complete reaction. However, a higher β value, especially at initial adsorbate concentration ranging between 33.81 and 136.67 mg/L, indicated limited aqueous mass transfer due to relatively high aqueous phase concentration. For a system with high initial adsorbate concentration and small particle sizes, the three-stage model is more suitable for determining the correct α value. The instantaneous portion, $\xi_{1,}$ increased significantly for both high and low initial adsorbate concentrations, possibly due to particle attrition. The increase in the instantaneous portion, $\xi_{1,}$ resulted in a smaller ξ_2 value, which indicated that a higher instantaneous portion inadvertently reduced the second-stage reaction duration. Moreover, the increase in the instantaneous portion, $\xi_{1,}$ strongly indicated RB5 adsorbed onto GIC. On the other hand, using a high mechanical stirring rate, significantly greater than 700 rpm, can result in particle attrition, leading to small particle sizes. Small particles have a high surface area, resulting in high instantaneous adsorption. The continuous adsorption–desorption cycles can affect the β value. In contrast, changes in γ values indicated prolonging or shortening reaction times, especially for the third stage. The balance between α and β values can be discerned through the presumption that the system may experience mass losses from the solution due to the mineralisation process.

Furthermore, the Elovich equation was initially used to examine the adsorption processes and is suitable for systems involving heterogeneous adsorbing surfaces (Wu et al. 2009). The characteristic curve of the Elovich equation resembles those of Lagergren's first-order equation and intraparticle diffusion model (Wu et al. 2009). The data obtained from Table 2 shows that Elovich kinetic model is the best fit for the experimental data. Moreover, Elovich kinetic model is the most appropriate model to describe the adsorption kinetics of RB5 with R^2 values greater than 0.9999. In addition, Fig. 3e shows the rapid rising to instant approaching equilibrium of the low-lying characteristic curve, which indicates rapid adsorption kinetics. The increasing surface diffusion flux driven by a large concentration gradient between solid–liquid interphase facilitated the rapid diffusion of RB5 adsorbate onto the surface of GIC, resulting in fast-approaching equilibrium. In addition, Fig. 4 shows the point of zero charge, pH_{pzc} relative to solution pH on the ionisation of GIC adsorbent.

Adsorption isotherms

Various isotherm models concerning Fig. 5 are stipulated in this subsection. The Langmuir adsorption isotherm assumes that a homogenous monolayer exists on the surface of the adsorbent, with no interaction between the adsorbed molecules and its neighbouring adsorption sites (Elemile et al. 2022). Moreover, Langmuir adsorption isotherm was initially developed to describe gas–solid interphase adsorption on both carbon and graphite-based adsorbent. This isotherm model is usually based on two common assumptions: 1) the forces of interaction between adsorbed molecules are negligible, and once an adsorbate occupies an active site on the surface of the adsorbent, no further sorption takes place. In addition, Langmuir isotherm refers to homogenous

Table 2 Summary of linear and non-linear kinetic parameters for adsorption of RB5 dye onto GIC adsorbent

Adsorption Kinetics parameters for Reactive Black 5 dye onto GIC adsorbent by using the pseudo-first-order non-linear equations				
C_0 (mg/L)	$q_{e,exp}$ (mg/g)	k_1 (min^{-1})	$q_{e,cal}$ (mg/g)	R^2
33.81	0.4210	0.8014	0.4210	0.9984
55.45	0.3796	1.1009	0.3796	0.9899
72.06	0.4509	0.5417	0.4509	1.0000
113.07	0.4509	0.4012	0.4509	0.9968
136.67	0.9777	0.0248	0.7089	0.9925
Pseudo-first-order equation	Non-Linear Form		Plot	
	$A = e^{\left(\frac{k_1 t}{2.303}\right)} q_e = \frac{-q_t A}{(1-A)}$		q_t versus t	
Adsorption kinetics parameters for Reactive Black 5 dye onto GIC adsorbent using the pseudo-second-order linear equations				
C_0 (mg/L)	$q_{e,exp}$ (mg/g)	k_2 (g/mg min)	$q_{e,cal}$ (mg/g)	R^2
33.81	0.4210	2.0166	0.2712	0.9869
55.45	0.3796	0.6807	0.2521	0.9781
72.06	0.4509	0.4158	0.2532	0.9754
113.07	0.4509	0.6631	0.2488	0.9500
136.67	0.9777	0.0493	0.2476	0.7603
Pseudo-second-order equations	Linear Form		Plot	
	$\frac{t}{q_t} = \frac{1}{k_2 q_e^2} + \frac{1}{q_e} t$		$\frac{t}{q_t}$ versus t	
Bangham kinetic parameters by non-linear regression method for Reactive Black 5 dye adsorption onto GIC adsorbent				
C_0 (mg/L)	$q_{e,exp}$ (mg/g)	k (min^{-1})	θ	$q_{e,cal}$ (mg/g)	R^2
33.81	0.4210	0.2574	0.1358	0.4932	0.9999
55.45	0.3796	0.1744	0.2032	0.4615	0.9710
72.06	0.4509	0.2352	0.1638	0.5154	0.9999
113.07	0.4509	0.1987	0.2167	0.5607	0.9999
136.67	0.9777	0.1651	0.2467	0.5379	0.9999
Bangham Equation	Non-Linear Form				Plot
	$q = kt^{\theta}$				$\log_e(Q)$ versus $\log_e(t)$
Adsorption kinetics parameters for Reactive Black 5 dye onto GIC adsorbent by using the pseudo-second-order non-linear equations				
C_0 (mg/L)	$q_{e,exp}$ (mg/g)	k_2 (min^{-1})	$q_{e,cal}$ (mg/g)	R2
33.81	0.4210	1.8148	0.4164	0.9945
55.45	0.3796	0.9969	0.3714	0.9968
72.06	0.4509	1.0761	0.4433	0.9996
113.07	0.4509	0.7819	0.4405	0.9873
136.67	0.9777	0.0224	0.7079	0.9957
Pseudo-second-order equation	Non-Linear		Plot	
	$q_t = \frac{q_e^2 k_2 t}{1 + q_e k_2 t}$		q_t versus t	
Kinetics parameters for adsorption of Reactive Black 5 dye on GIC by using the Three-stage kinetic model								
C_0 (mg/L)	q_1 (t) (mg/g)	q_2 (t) (mg/g)	α	β	γ	ξ_1	ξ_2	R^2
33.81	0.1564	0.4210	0.1711	0.1461	3.4830	0.1064	0.2864	0.9550
55.45	0.0414	0.3796	0.2254	0.1839	7.4240	0.0172	0.1574	0.9547
72.06	0.0598	0.4509	0.4132	0.2924	9.2199	0.0191	0.1439	0.9547
113.07	0.0989	0.4509	0.1266	0.1124	11.9101	0.0201	0.0917	0.9546
136.67	0.1242	0.9777	0.1457	0.1272	6.6720	0.0209	0.1645	0.9546

Table 2 (continued)

Kinetics parameters for adsorption of Reactive Black 5 dye on GIC by using the Three-stage kinetic model

C_0 (mg/L)	q_1 (t) (mg/g)	q_2 (t) (mg/g)	α	β	γ	ξ_1	ξ_2	R^2
Three-Stage Kinetic Equation		$q_1(t) \equiv q_1(0) \equiv q_1(\infty)\frac{\partial C}{\partial t} = -\alpha C\left[1 - \frac{q_2(t)}{q_2(\infty)}\right]$ $q(t) = q_1(t) + q_2(t)$ $\frac{\partial C}{\partial t} = -\alpha C\left[1 - \frac{q_2(t)}{\beta q_2(\infty)}\right]VC(t) + M\left[q_1 + q_2\right] = VC_0$ $\frac{C(t)}{C_0} = \frac{(1-\xi_1)(1-\beta\xi_2)}{(1-\xi_1-\beta\xi_2 exp[-\gamma t])}$ $\xi_1 = \frac{Mq_1(\infty)}{VC_0}$ $\xi_2 = \frac{Mq_2(\infty)}{VC_0}$ $\gamma = \frac{(1-\xi_1-\beta\xi_2)\alpha}{\beta\xi_2}$				Plot C(t)versust	

Elovich kinetic parameters for adsorption of Reactive Black 5 dye onto GIC adsorbent

C_0 (mg/L)	$q_{e,exp}$ (mg/g)	α	β	$q_{e,cal}$ (mg/g)	R^2
33.81	0.4210	3.6758	18.3320	0.4908	0.9999
55.45	0.3796	0.3652	13.9796	0.4592	0.9999
72.06	0.4509	1.2161	15.0065	0.5126	0.9999
113.07	0.4509	0.3677	11.2163	0.5533	0.9999
136.67	0.9777	0.35422	12.3529	0.5072	0.9999
Elovich Kinetic Equation		$q_t = \frac{1}{\beta}ln(\alpha\beta t + 1)$			Plot $q_t$versust

adsorption with a second assumption that there is no transmigration of the adsorbate molecules in the plane perpendicular to the surface of the adsorbent (Shahbeig et al. 2013). This isotherm model has the following hypotheses (Shahbeig et al. 2013):

(1) The monolayer adsorption is approximately one molecule in thickness.
(2) Adsorption usually takes at specific homogenous active sites within the surface of the adsorbent.
(3) Once an adsorbate occupies an active site, no further sorption takes place at that particular site.
(4) The free energy at the adsorption site is relatively constant and independent of the degree of adsorbate occupation on the active site of the adsorbent.

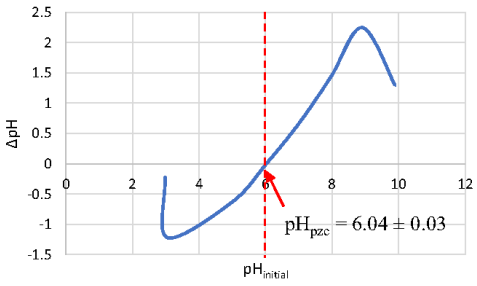


Fig. 4 Point of zero charge of GIC particle electrode

(5) The strength of attractive intermolecular forces depends on the distance between the adsorbate molecule and the active site of the adsorbent, indicating that the further away the adsorbate is from the adsorbent surface, the lower the attraction.
(6) The physicochemical structure of the adsorbent is considered to be homogeneous.

However, the Langmuir isotherm model's main limitation is its validity in low-pressure constraints. Another drawback of Langmuir isotherm is that it assumes uniform monolayer adsorption of adsorbate or solute at a specific homogenous active site. In reality, this rarely occurs in the presence of high adsorbate concentration, which results in rapid saturation of active sites caused by competition adsorption phenomena or interaction between adsorbates on different active sites. The ongoing adsorption and desorption processes affect the accuracy of the assumption for a mechanistic model.

Two significant assumptions are related to the derivation of Temkin isotherm (Chu 2021): (1) There is a uniform distribution of heterogeneous binding sites on the solid surface; (2) The binding energy varies linearly over different binding sites. Temkin isotherm is often used to characterise the environmental adsorption of contaminants, but it suffers from dimensional inconsistency (Chu 2021). The dimensionally inconsistent formulation, as shown in Table 3, may affect the accuracy of the representation of fitted theoretical data against the experimental data. Two undesirable features relate to the fitted curve: (1) any initial dye concentration beyond 100 mg/L does not accurately predict the saturation

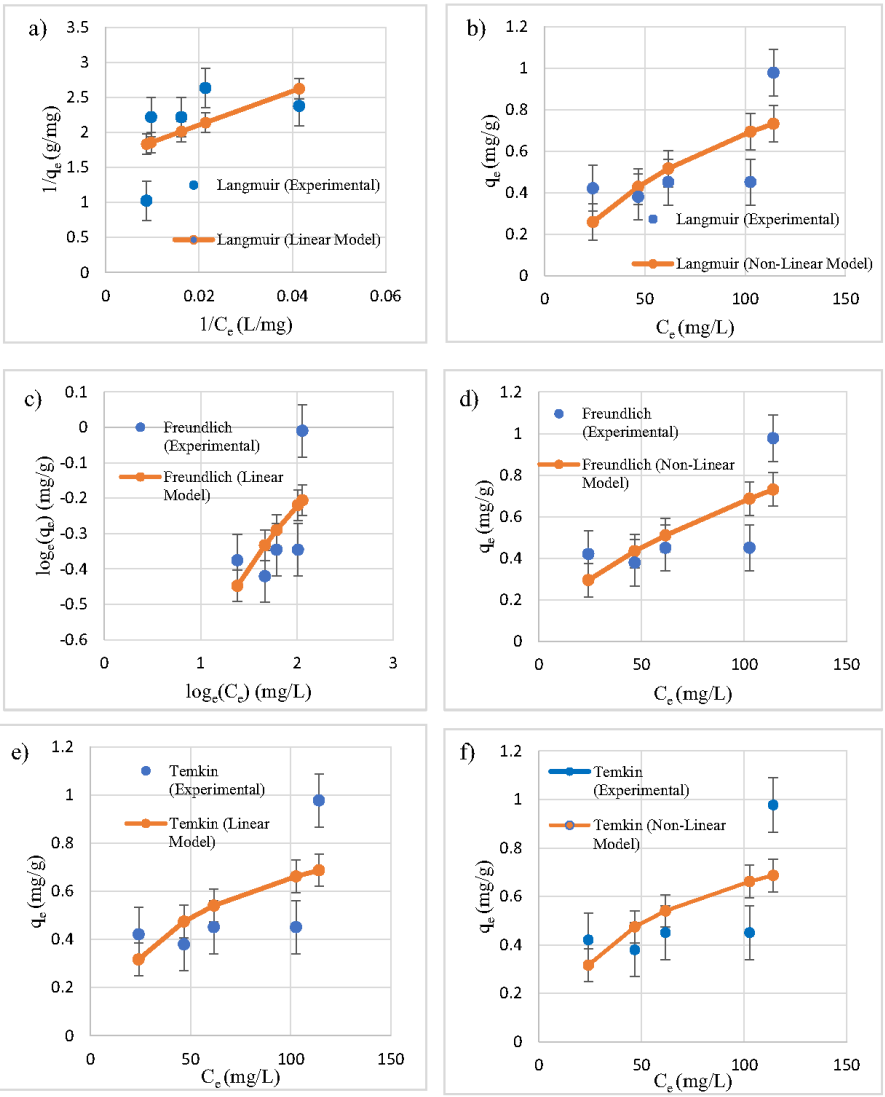


Fig. 5 **a** Langmuir adsorption isotherm for RB5 onto GIC adsorbent by linear regression method; **b** Langmuir adsorption isotherm for RB5 onto GIC adsorbent by non-linear regression method; **c** Freundlich adsorption isotherm for RB5 onto GIC adsorbent by linear regression method; **d** Freundlich adsorption isotherm for RB5 onto GIC adsorbent by non-linear regression method; **e** Temkin adsorption isotherm for RB5 onto GIC adsorbent by linear regression method; **f** Temkin adsorption isotherm for RB5 onto GIC adsorbent by non-linear regression method; **g** SIPS adsorption isotherm for RB5 onto GIC adsorbent by linear regression method; **h** SIPS adsorption isotherm for RB5 onto GIC adsorbent by non-linear regression method; **i** Redlich–Peterson isotherm modelling by linear regression method; **j** Redlich–Peterson adsorption isotherm for RB5 onto GIC adsorbent by non-linear regression method

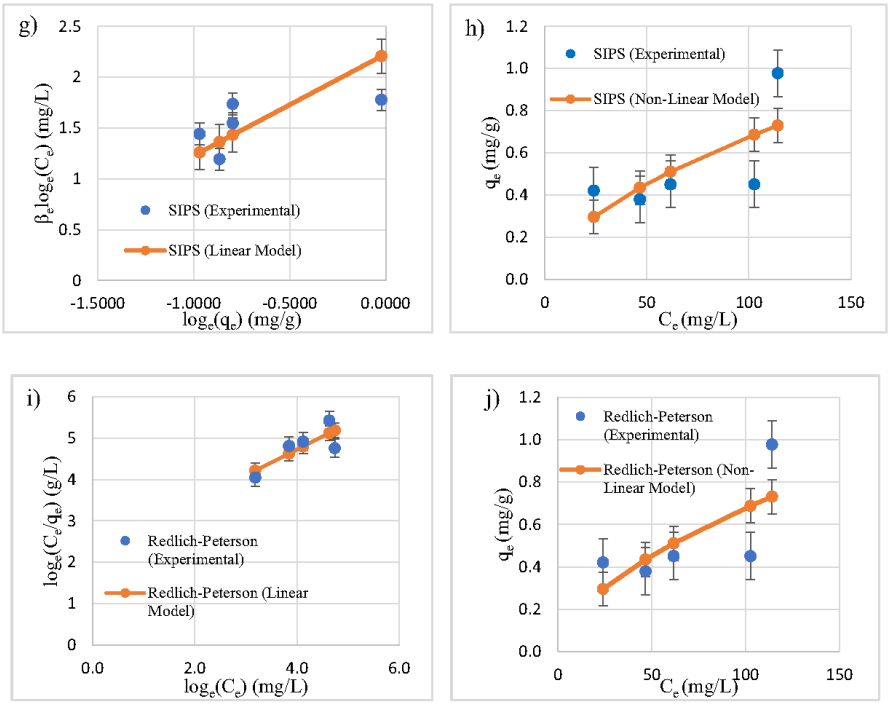


Fig. 5 (continued)

limit, especially at high concentrations, resulting in more significant disparities between theoretical and experimental data; (2) The second undesirable feature is present in both low and high regions. This indicates that the experimental values are less than the accepted values, giving the average relative errors negative values. Nonetheless, the inherent deficiencies in Temkin formulation do not entirely restrict its ability to correlate the equilibrium data to a significant degree of precision. In other words, the fitting of the Temkin equation to an equilibrium data and experimental profile is still within a validity range despite some minor deviations at both high and low concentrations. Some minor deviations could be attributed to the ongoing desorption process within the adsorption system.

Furthermore, SIPS isotherm combines Langmuir and Freundlich models and is appropriate for characterising heterogeneous adsorption systems at various pressures and temperatures (Tzabar and ter Brake 2016). The SIPS model has been found to fit liquid adsorption data remarkably well, especially for liquid–solid adsorption systems. Although the results showed that the adsorption systems are suitable for RB5 contaminant, the apparent SIPS formulation may conceal problematic application issues. The application issues are expounded as follows (de Vargas Brião et al. 2023): (1) the dubious practice of using linear versions of the SIPS equation to fit the experimental profile; (2) the inherent mismatches between the SIPS and Langmuir–Freundlich equations; and (3) the trivial practice of correlating the SIPS equation to experimental data. To overcome the inherent deficiencies, SIPS combined the characteristics of the Freundlich model to assume that the amount of adsorbate increases indefinitely with pressure, resulting in a distribution function correlated with an increase in the infinite number of active sites available for adsorption.

On the other hand, SIPS derived another distribution function to predict a finite number of active sites available on the adsorbent surface. To make this derivation method workable, SIPS incorporated exponent n, ranging between 0

Table 3 Summary of linear and non-linear parameters for isotherm models for the adsorption of RB5 dye onto GIC adsorbent

Linear and non-linear parameters of isothermal models for the sorption of Reactive Black 5 (RB5) dye onto GIC adsorbent				
Two-parameter isotherms				
Isotherm models	Linear method		Non-linear method	
Langmuir	q_m (mg/g)	0.6166	q_m (mg/g)	1.4369
	B	0.0671	b	0.0091
	R^2	1.0000	R^2	0.9992
	$\frac{1}{q_e} = \left(\frac{1}{bq_m}\right)\frac{1}{C_e} + \frac{1}{q_m}$		$q_e = \frac{bK_F C_e}{1+bC_e}$	
Freundlich	K_F (L/mg)	0.2263	K_F (L/mg)	0.0464
	n	0.7133	n	1.7179
	R^2	1.0000	R^2	0.9999
	$\log_e(q_e) = \log_e(K_F) + \frac{1}{n}\log_e(C_e)$		$q_e = K_F C_e^{\frac{1}{n}}$	
Temkin	A	0.1568	A	0.1568
	B	0.2381	B	0.2381
	R^2	1.0000	R^2	1.0000
	$q_e = \frac{RT}{b}\log_e(A) + \frac{RT}{b}\log_e(C_e)$		$q_e = B\log_e(AC_e)$	
	$\beta = \frac{RT}{b}$		$\beta = \frac{RT}{B}$	
Three-parameter isotherms				
Isotherm models	Linear method		Non-linear method	
SIPS	q_m	0.0176	q_m	84.9456
	K_s (L/g)	0.0019	K_s (L/g)	2.5727×10^{-6}
	n	0.3751	n	0.5840
	R^2	1.0000	R^2	0.9999
	$\beta_s\log_e(C_e) = -\log_e\left(\frac{K_s}{q_e}\right) + \log_e(\alpha_s)$		$q_e = q_m\left((K_sC_e)^n/(1+K_sC_e)^n\right)$	
Redlich–Peterson	A	0.1076	A	5.2680
	β (L/mg)	0.6249	B (L/mg)	113.5293
	R^2	0.9999	β	0.4176
	$\log_e\left(\frac{C_e}{q_e}\right) = \beta\log_e(C_e) - \log_e(A)$		R^2	0.9999
			$q_e = \frac{AC_e}{1+BC_e^g}$	

and 1, to make the adsorption systems more manageable at an extensive range of pressures. Since Freundlich isotherm only approximates the adsorption behaviour of the system, the value of 1/n can only range between 0 and 1; therefore, the equation only accounts for a limited range of pressure. Given the status of adsorption systems, the Langmuir–Freundlich equation is interconvertible with the SIPS formulation (de Vargas Brião et al. 2023). No meaningful insights can be gained by fitting the Langmuir–Freundlich and SIPS equations separately to similar experimental data sets and comparing their degrees of precision. Hence, the SIPS equation may lack mechanistic relevance.

More interestingly, a three-parameter Redlich–Peterson isotherm equation is introduced to amend the inaccuracies and inherent deficiencies of two-parameter Langmuir and Freundlich isotherm equations. The Redlich–Peterson isotherm equation is more accurate than Langmuir and Freundlich equations in describing the adsorption systems using GIC adsorbent, which is consistent with Wu et al. (2010). Unlike Langmuir and Freundlich isotherms, the Redlich–Peterson isotherm equation incorporates additional parameters such as A, B and β to increase the accuracy and precision of curve fitting between analytical and experimental data. Moreover, the Redlich–Peterson isotherm balances the Langmuir and Freundlich systems and incorporates the benefits of both models rather than conflicts between the two systems. In addition, the degree of curve fitting for linearity and non-linearity of isotherm equations depends on the types of experimental adsorption systems. Occasionally, a linearised form of an isotherm

equation may yield a better fit for experimental data than a non-linearised form of equation, depending on the type of adsorption system.

According to Table 3, the sum of the squares error (ERRSQ/SSE) values for Redlich–Peterson and SIPS isotherms were the lowest compared to other values of isotherm models, resulting in a better fit between theoretical and experimental data. Although ERRSQ/SSE is the most widely reported error function in isotherm modelling, the major disadvantage is its poor error prediction at high-pressure conditions, which may reflect the true nature of adsorption complexity (Serafin and Dziejarski 2023). The hybrid fractional error function (HYBRID) was initially developed to improve the curve fitting of the sum of squares error (ERRSQ/SSE) to account for low relative pressure condition by dividing it by the measured experimental values at equilibrium condition (Serafin and Dziejarski 2023). The hybrid fractional error function (HYBRID) was selected as the optimal error function to ascertain and analyse the isotherm models to characterise the finest fit between theoretical and experimental data. In addition, the sum of absolute errors (SAE) is a similar error function to ERRSQ/SSE, confirming that non-linear, three-parameter Redlich–Peterson and SIPS isotherm models provided the finest fit between theoretical and experimental data. In Table 3, it was observed that the hybrid fractional error function for Redlich–Peterson and SIPS isotherms, in conjunction with the chi-square test ($\chi 2$), yielded the overall best-fitting performance compared to other isotherm models. It can be deduced that HYBRID and $\chi 2$ error function analyses are significant tools in evaluating the isotherm modelling because they are statistically robust and well-established measurement tools for accurate curve-fitting models. The benefits of having both HYBRID and $\chi 2$ error function analyses are due to balances between the influence of both large and small error values in HYBRID measurement, whereas comparative evaluation of model deviation between predicted and experimental values by taking into account the uncertainty in experimental modelling in $\chi 2$ error function analysis. Based on this reasoning, HYBRID and $\chi 2$ error functions provided the best principal method for determining and analysing the accuracy and precision of isotherm models to yield the finest curve fitting between theoretical and experimental data. Among other error functions, Marquardt's percent standard deviation (MPSD) measures the geometric mean of the error distribution modified in accordance with the degrees of freedom in the isotherm models (Serafin and Dziejarski 2023). On the other hand, Marquardt created the average relative error (ARE) to minimise the fractional error in statistical distribution over a range of relative pressure conditions (Serafin and Dziejarski 2023). The chi-square values of Redlich–Peterson and SIPS were approximately 0.9925 compared to chi-square values of 0.2769 and 0.1472 for Langmuir and Freundlich isotherm models, respectively. Hence, the larger the chi-square values, the greater the probability that the isotherm models were statistically significant, indicating that Redlich–Peterson and SIPS isotherm models were a better fit for the experimental data compared to Langmuir and Freundlich isotherm models. The coefficient of determination was calculated using Excel Solver Add-In (2022). In terms of mechanistic phenomena, the predicted amount of adsorbed adsorbate reached an equilibrium state closely resembling the observed amount of adsorbed adsorbate at the equilibrium state, thereby validating the predictability of isotherm models to describe the experimental data. On the other hand, the G-statistic values for Redlich–Peterson and SIPS isotherm models were 0.4682 and 0.4683, respectively. Hence, it indicated that the deviance for the isotherm models from the experimental data was statistically significant compared to Langmuir and Freundlich isotherm models with G-statistic values of −6.5832 and −7.5650, respectively. Hence, this indicated that the Redlich–Peterson and SIPS isotherm models had better goodness-of-fit.

In contrast, Langmuir and Freundlich's deviance for the predicted models and experimental data was significantly smaller than expected, indicating that the comparative models were statistically insignificant. Unlike the actual goodness-of-fit test, the G-statistic test does not calculate the probability of obtaining the experimental results from something more extreme. Instead, it utilizes the experimental data to calculate a test statistic to determine how far the experimental data deviates from the theoretical results based on null expectation. To support this mathematical relationship, the chi-square calculation is then used to estimate the probability of obtaining that value of the G-test statistics. This indicated that the G-test statistic is more efficient than the chi-square test in measuring the goodness-of-fit, provided that the values between the experimental and theoretical data must be statistically significant or contingent upon the overall significance of the models (Osorio et al. 2024). In Table 4, MPSD errors for Redlich–Peterson and SIPS isotherms were the lowest compared to other isotherm models, verifying the statistical significance of HYBRID and $\chi 2$ error function tools, indicating the best curve fitting characteristics of the isotherm models. Hence, the following is the categorisation of error functions based on the findings obtained from the RB5 adsorption system, starting from the left with the error function that provides the best curve fitting between the experimental and theoretical isotherm models:

$$\chi 2 > \text{HYBRID} > \text{ERRSQ/SSQ} > \text{ARE} > \text{EABS} > \text{MPSD} > \sigma > R^2$$

In contrast to two-parameter isotherm models, the three-parameter isotherm models were characterised by relatively low HYBRID and ERRSQ/SSE values. In accordance with HYBRID, $\chi 2$ and ERRSQ/SSE, the degrees of curve fitting of all isotherm models are ranked in the following order:

Table 4 Error function analysis for six isotherm equations for comparison between experimental and theoretical data

Error functions	R^2	Coefficient of non-determination	R_{adj}^2	σ Standard Error	ERRSQ/SSE	ARE	HYBRID	EABS	MPSD	Chi-square (χ^2)	G-test
Linear approach											
Two-parameter isotherms											
Langmuir	1.0000	0.0000	1.0000	0.2620	0.2620	0.2060	0.2036	7.8964	0.2512	0.2434	0.9839
Freundlich	1.0000	0.0000	1.0000	0.2342	0.2342	0.1645	−2.2266	−5.9376	0.1225	0.3353	0.9913
Temkin	1.0000	0.0000	1.0000	0.2286	0.1568	−4.6303	−12.3474	0.0000	0.3957	0.9924	0.2514
Linear approach											
Three-parameter isotherms											
Redlich–Peterson	0.9999	0.0000	0.9999	0.2275	0.1553	−3.3954	−5.6590	0.1169	0.3303	0.9925	0.4683
SIPS	0.9999	0.0000	0.9999	0.2275	0.1553	−3.3938	−5.6563	0.1170	0.3303	0.9925	0.4683
Non-linear approach											
Two-parameter isotherms											
Langmuir	0.9992	0.0008	0.9970	0.2255	0.1530	−3.4031	−5.6718	0.0535	0.4239	0.9910	0.3792
Freundlich	0.9999	0.0001	0.9995	0.2153	0.1390	−5.0728	−8.4546	0.0199	0.3941	0.9940	0.2691
Temkin	1.0000	0.0000	1.0000	0.2286	0.1568	−7.4084	−12.3474	0.0000	0.3957	0.9924	0.2514
Non-linear approach											
Three-parameter isotherms											
Redlich–Peterson	0.9999	0.0001	0.9991	0.2637	0.1390	−5.0752	−12.6879	0.0196	0.4831	0.9938	0.2686
SIPS	0.9999	0.0001	0.9989	0.2638	0.1391	−5.0317	−12.5793	0.0215	0.4818	0.9938	0.2722

Redlich – Peterson > SIPS > Temkin > Freundlich > Langmuir

By interpreting the above sequential order, it is evident that Redlich-Peterson and SIPS isotherms are the most appropriate models for verifying research associated with RB5 adsorption systems involving GIC.

Electrochemical oxidation of reactive black 5

The experimental studies demonstrated that 12 mA/cm^2 was the minimum current density required to reduce RB5 concentration in water significantly. Both dye and TOC removal percentages increased continuously beyond 12 mA/cm^2. This indicated that more intermediate oxidation products were subsequently mineralised into CO_2 and H_2O. A higher current density of 45 mA/cm^2 could eventually eliminate residual RB5 in water, resulting in a 93% mineralisation efficiency and 67% TOC removal rate (Fig. 6). In addition, electrolysis can effectively remove RB5 from water by destroying its molecular structure through the cleavage of azo bonds. However, some RB5 molecules may not be wholly mineralised into inert CO_2 and H_2O. In addition, Fig. 6 represents the effect of current density on the mineralisation and RB5 removal efficiencies. The greater the current density, the greater the removal efficiencies of RB5 and TOC. On the other hand, Fig. 6 shows the effect of current density on the annual electricity cost for TOC, which varied due to the buildup of side reactions at a current density greater than 20 mA/cm^2. The mineralisation of RB5 molecules into CO_2 at higher current density also contributed to the loss of intermediate transformation species in water, impacting the ionic conductivity of aqueous media. In terms of technoeconomic analysis, the electricity cost per hour of electrolysis using 34 mA/cm^2 of current density was amounted to 49.59 AUD/kg of RB5 and 334.52 AUD/kg of TOC. Furthermore, Fig. 7 shows the effect of current density on the electrochemical regeneration efficiency of GIC particle electrodes in the presence of RB5-polluted wastewater. High current density led to surface roughening of the GIC, causing changes in the surface physicochemical properties of GIC. Fresh GICs initially have acidic quinone and carboxyl groups on its surface. Upon electrochemical regeneration, these functional groups increase in quantity. A sustained regeneration leads to the formation of basic lactones, which offsets the relative amounts of acidic functional groups. The increase in adsorption can be attributed to strong surface acidic functionalities. Subsequent adsorption and regeneration cycles lead to a greater formation of basic functional groups, which outnumbers the amount of surface acidic functional groups, reducing adsorption (Nkrumah-Amoako et al. 2014). The presence of anions in the solution may increase the electrostatic repulsive forces between GIC and RB5 molecules at the solid–liquid interface, reducing the ability of GIC to adsorb RB5 effectively.

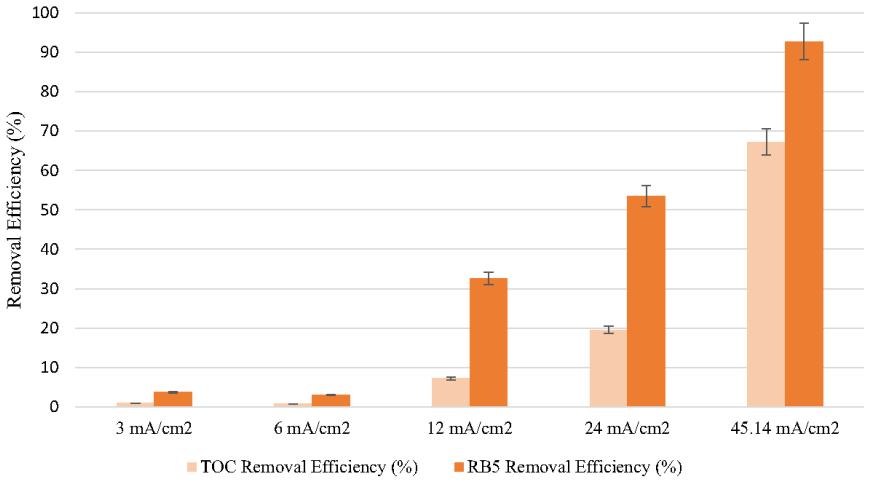


Fig. 6 Effect of current density on dye and TOC removal efficiencies. Each experiment comprises 20 min of adsorption and 10 min of regeneration at a range of current densities from 3 to 45.14 mA/cm^2 and 30 V. The error bars represent the coefficient of variation of 2.32%

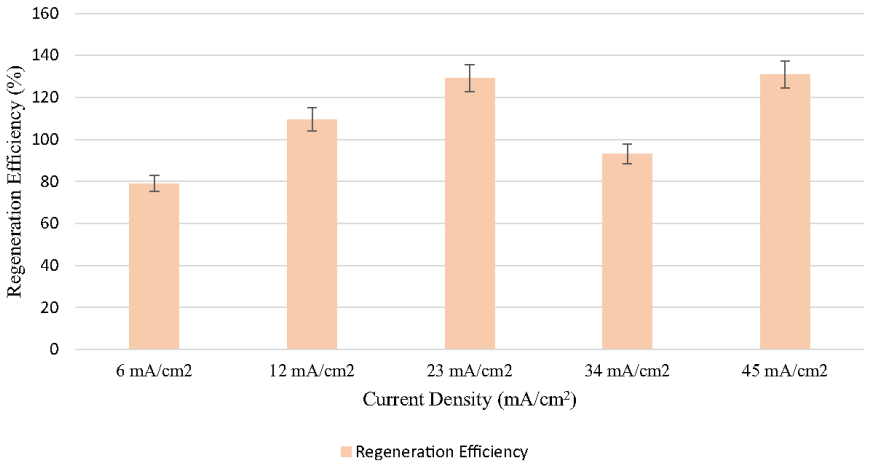


Fig. 7 Electrochemical regeneration efficiency of GIC particle electrodes in the presence of RB5-polluted wastewater

Optimisation study using three-dimensional response surface plots

This experimental study investigated the effect of operating parameters such as temperature, pH and salt concentration on the RB5 and TOC removal efficiencies using surface response methodology (Kanneganti et al. 2022) integrated with central composite design (Medeiros et al. 2022). Other constant variables included an initial RB5 concentration of 50 mg/L, an adsorption time of 10 min and an adsorbent dosage of 20 g/L. The batch experimental runs were conducted in accordance with the CCD Design of Experiment to visualize the effects of interactive variables on targeted responses using RSM optimization techniques. A general result demonstrated that when the temperature and pH or salt concentration increased, the RB5 and TOC removal efficiencies also increased as shown in RSM and contour plots of Figs. 8a–j.

The experimental results were evaluated and approximated using the mathematical expressions or functions of various targeted responses, such as dye and TOC removal efficiencies presented in regression equations in coded units:

$$Y_1 = 32.7 - 0.46x_1 - 2.56x_2 - 1.35x_3 - 0.0015x_1x_1 + 0.116x_2x_2 - 0.004x_3x_3 + 0.0493x_1x_2 + 0.0549x_1x_3 - 0.077x_2x_3 \quad (4)$$

$$Y_2 = 12.6 - 0.177x_1 - 1.06x_2 - 0.71x_3 - 0.00064x_1x_1 + 0.0484x_2x_2 - 0.0006x_3x_3 + 0.0181x_1x_2 + 0.0233x_1x_3 - 0.0190x_2x_3 \quad (5)$$

Furthermore, Table 5 shows the ANOVA analysis of RB5 removal efficiency. The F-value of pH variable was 5.60, which was significantly greater than the P-value of 0.039. On the other hand, Table 6 shows the ANOVA analysis for TOC removal efficiency with the F-value of pH variable at 5.90, which was significantly greater than the P-value of 0.036, whereas the F-value of temperature was 0.31 and the P-value was 0.589, indicating that the temperature variable had no significant interactive effect on the overall GIC adsorption efficiency. The following represents the Pareto chart that visualises the interactive effects of various operating variables on the RB5 and TOC removal efficiencies. The Pareto chart in Fig. 9a shows the bars representing one of the independent variables, pH, crosses the reference line with an absolute value of 2.228. This indicates that only pH value produced a statistically significant effect on the dye removal efficiency with less than the α-value of 0.05. Similarly, the synergistic effect of combined temperature and pH, temperature and salt concentration and pH and salt concentration (AB, AC, BC) etc., did not produce a statistically significant effect on the dye removal efficiency. A similar trend was observed for Fig. 9b, indicating that pH significantly impacted the TOC removal efficiency. Overall results showed that salt concentration had no significant effect on the selectivity reversal of GIC adsorption of RB5 compared to the pH

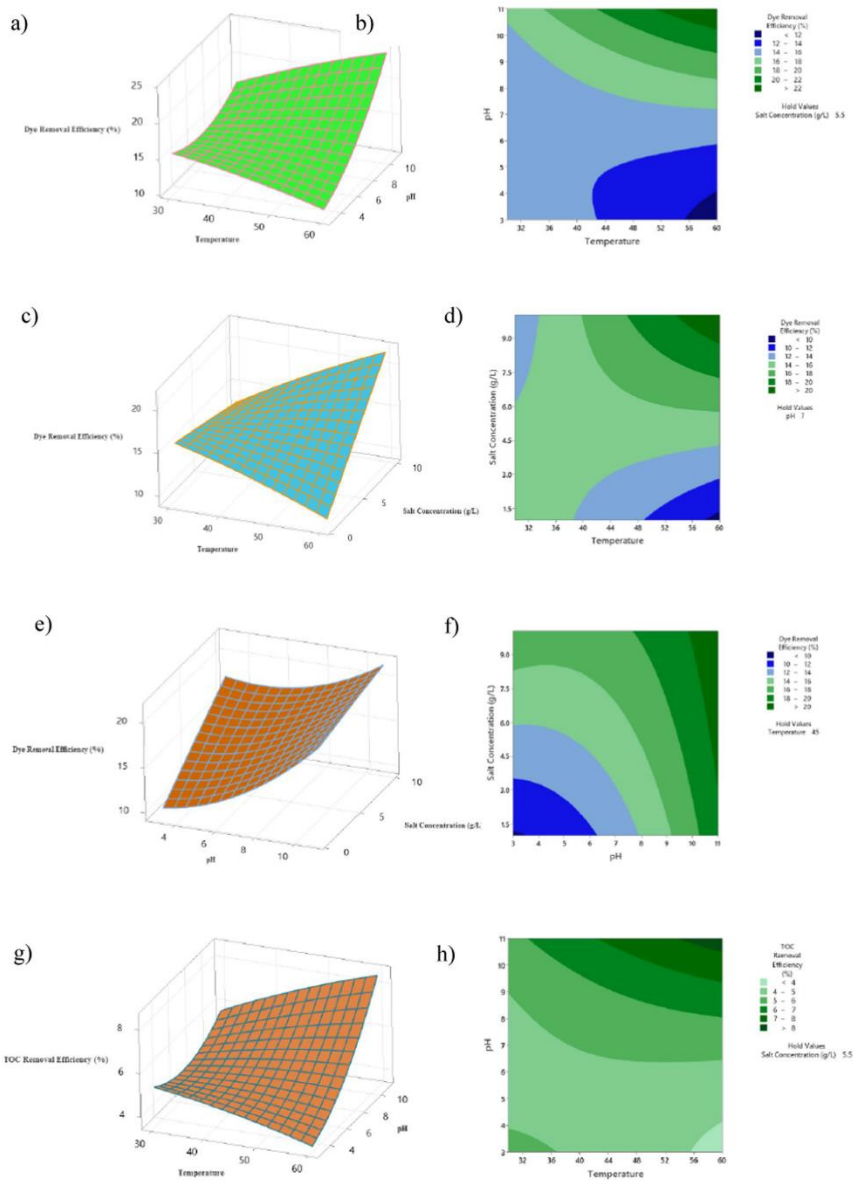


Fig. 8 3D RSM optimisation plots of the interaction effects of temperature, pH and salt concentration in a binary mixture with an initial dye concentration, adsorption time and adsorbent dosage of C_0=50 mg/L, t=10 min and adsorbent dosage of 20 g/L, respectively. The actual values of the operating parameters are mean values (coded values 0)

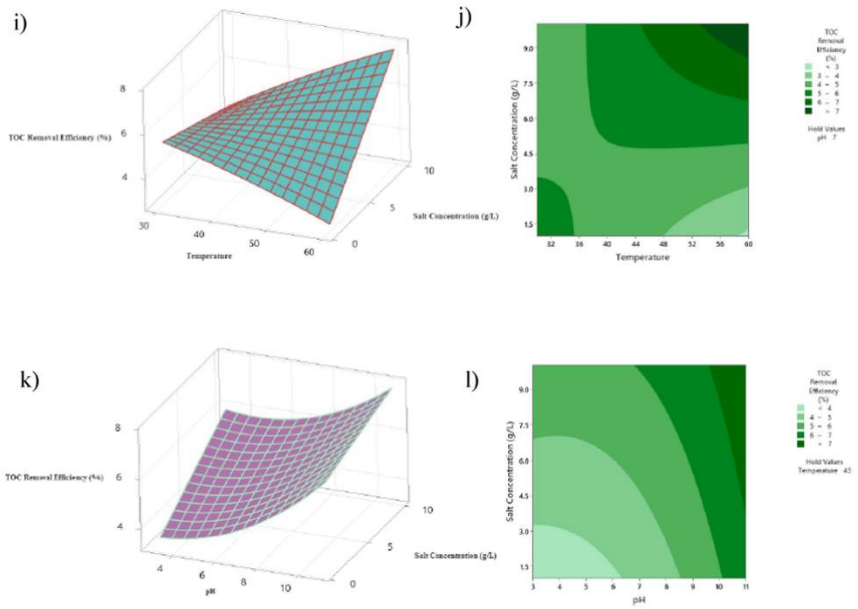


Fig. 8 (continued)

Table 5 ANOVA analysis of interactive variables on the RB5 removal efficiency

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	9	119.095	13.2328	1.30	0.344
Linear	3	87.817	29.2724	2.87	0.090
Temperature	1	2.577	2.5769	0.25	0.626
pH	1	57.194	57.1940	5.60	0.039
Salt Concentration (g/L)	1	28.046	28.0463	2.75	0.128
Square	3	6.818	2.2725	0.22	0.879
Temperature*Temperature	1	0.213	0.2125	0.02	0.888
pH*pH	1	6.227	6.2272	0.61	0.453
Salt Concentration (g/L)*Salt Concentration (g/L)	1	0.010	0.0103	0.00	0.975
2-Way Interaction	3	24.460	8.1534	0.80	0.522
Temperature*pH	1	8.761	8.7614	0.86	0.376
Temperature*Salt Concentration (g/L)	1	13.755	13.7548	1.35	0.273
pH*Salt Concentration (g/L)	1	1.944	1.9440	0.19	0.672
Error	10	102.105	10.2105		
Lack-of-Fit	5	74.423	14.8847	2.69	0.151
Pure Error	5	27.682	5.5364		
Total	19	221.200			

Table 6 ANOVA analysis for determination of the significance of interactive variables on the TOC removal efficiency

Source	DF	Adj SS	Adj MS	F-value	*P*-value
Model	9	17.6369	1.95966	1.39	0.308
Linear	3	12.6960	4.23201	3.00	0.082
Temperature	1	0.4407	0.44071	0.31	0.589
pH	1	8.3332	8.33320	5.90	0.036
Salt Concentration (g/L)	1	3.9221	3.92210	2.78	0.127
Square	3	1.1761	0.39205	0.28	0.840
Temperature*Temperature	1	0.0376	0.03764	0.03	0.874
pH*pH	1	1.0783	1.07831	0.76	0.403
Salt Concentration (g/L)*Salt Concentration (g/L)	1	0.0003	0.00026	0.00	0.989
2-Way Interaction	3	3.7648	1.25493	0.89	0.480
Temperature*pH	1	1.1836	1.18359	0.84	0.382
Temperature*Salt Concentration (g/L)	1	2.4637	2.46371	1.74	0.216
pH*Salt Concentration (g/L)	1	0.1175	0.11749	0.08	0.779
Error	10	14.1232	1.41232		
Lack-of-Fit	5	10.1346	2.02693	2.54	0.165
Pure Error	5	3.9886	0.79772		
Total	19	31.7602			

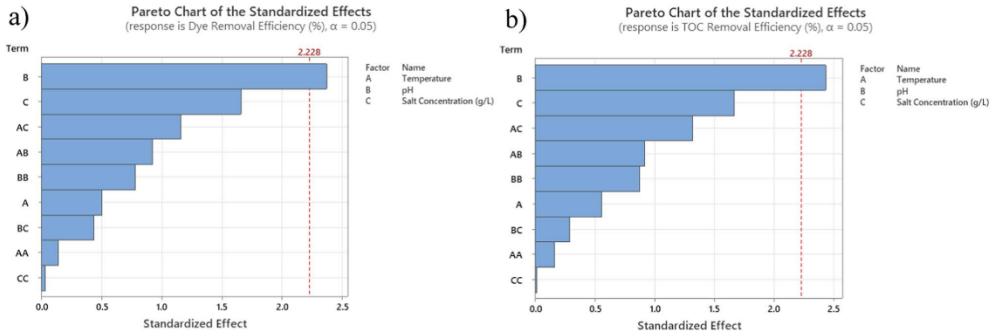


Fig. 9 **a** Pareto chart represents the significant effect of three different factors, including the interactive effects on the dye removal efficiency, and **b** Pareto chart represents the significant effect of three different factors, including the interactive effects on the TOC removal efficiency

effect. The combined effect of temperature and salt concentration selectively induced RB5 adsorption by GIC in the presence of pH changes relative to pH_{pzc}. Initially, RB5 dye solution had a pH of 5. When the solution pH was greater than pH_{pzc}, GIC adsorbent became negatively charged, increasing its ability to adsorb more RB5 dye molecules due to GIC surface ionisation. On the GIC surfaces, π–π electron donor–acceptor interactions played a significant role in the selectivity of the adsorption process. The adsorption of RB5 anionic dye on GIC surface, where an anion-π repulsive interaction was strong due to solution pH greater than pH_{pzc}. However, the presence of Na^+ ions from salt solution negated the GIC surface charges, increasing its attraction towards anionic dye via attractive cation-π and multivalent interactions. In addition, the adsorption process on GIC adsorbent was mainly dominated by π–π and electrostatic interactions. The following represents the ANOVA analysis for the significance of interactive variables on the RB5 removal efficiency:

In addition, Table 5 represents the difference between the CCD-RSM optimised and validated results. At a typical temperature of dye wastewater at 30 °C, pH 7, and NaCl concentration of 5.50 g/L, the RB5 and TOC removal efficiencies were 14.19% and 4.71%, respectively, using CCD-RSM optimisation. These targeted responses were validated with the experimental results, which showed that

Table 7 Comparison between optimised and non-optimised experimental results

CCD-RSM optimisation result				
Temperature (°C)	pH	Salt concentration (g/L)	TOC removal efficiency (%)	RB5 removal efficiency (%)
30	7	5.5	4.71	14.19
Experimental result for validation				
Temperature (°C)	pH	Salt concentration (g/L)	TOC removal efficiency (%)	RB5 removal efficiency (%)
30	7	5.5	4.64	13.99

RB5 and TOC removal efficiencies closely resembled the optimised values, verifying the validity of RSM optimisation (Table 7).

Conclusion

RB5 can be removed successfully using combined adsorption and electrochemical oxidation processes. However, the optimum removal rates needed to be evaluated using CCD-RSM optimization. The results showed that pH had the most significant effect on the GIC adsorption of RB5 compared to temperature and salt concentration. Although salt concentration had a limited impact on the selectivity of GIC towards RB5, it prevented the selectivity reversal by regulating or reinforcing the balance of surface ionization of GIC, making it more conducive to RB5 adsorption. The changes in surface physicochemical properties of GIC after several cycles of electrochemical regeneration can be compensated by regulating the surface ionization of GIC, albeit additional batches of adsorption must be performed separately. In addition, the salting effect functions as a regulatory mechanism to balance or induce the surface ionisation potential of GIC adsorbent instead of directly influencing the adsorption process, whereas pH relative to the point of zero charge had a more stabilising effect on the adsorption process. On the other hand, the electrostatic interaction from the physical process helped to strengthen the adsorption process. Although GIC adsorbents were non-porous, indicating that the adsorptive capacity can be exhausted rapidly, high electrically conductive GIC adsorbents can be electrochemically regenerated to recover its active sites. The adsorption kinetics were found to follow the Elovich kinetic model, which yielded the best fitness. On the other hand, the adsorption isotherms were found to follow both SIPS and Redlich-Peterson isotherm models with significantly fewer errors than other models in accordance with error function analyses. Almost complete removal of RB5 can be achieved within 30 min at a current density of approximately 45 mA/cm^2. Although the adsorptive capacity of GIC can be regenerated, further work is needed to evaluate the intermediate breakdown products formed during the oxidation of RB5. In addition, more characterisation tests are required to examine the changes in the surface chemistry of GIC after the prolonged period of electrochemical regeneration. More methods should be explored to minimise the likelihood of particle attrition or corrosion during electrochemical treatment.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13762-024-05696-4.

Author contributions VG: Conceptualization, visualization, validation, investigation, formal analysis, data curation, writing—original draft. APT: Supervision, review and editing, project administration, resources.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. Non-applicable.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Asghar HMA, Hussain SN, Sattar H, Brown NW, Roberts EPL (2014) Electrochemically synthesized GIC-based adsorbents for water treatment through adsorption and electrochemical regeneration. J Ind Eng Chem 20:2200–2207. https://doi.org/10.1016/j.jiec.2013.09.051

Balla W, Essadki AH, Gourich B, Dassaa A, Chenik H, Azzi M (2010) Electrocoagulation/electroflotation of reactive, disperse

and mixture dyes in an external-loop airlift reactor. J Hazard Mater 184:710–716. https://doi.org/10.1016/j.jhazmat.2010.08.097

Chaiwichian S, Lunphut S (2021) Development of activated carbon from parawood using as adsorption sheets of organic dye in the wastewater. Mater Today Proc 47:3449–3453. https://doi.org/10.1016/j.matpr.2021.03.383

Choi JW, Choi NC, Lee SJ, Kim DJ (2007) Novel three-stage kinetic model for aqueous benzene adsorption on activated carbon. J Colloid Interface Sci 314:367–372. https://doi.org/10.1016/j.jcis.2007.05.070

Chu KH (2021) Revisiting the temkin isotherm: dimensional inconsistency and approximate forms. Ind Eng Chem Res 60:13140–13147. https://doi.org/10.1021/acs.iecr.1c01788

de Fouchécour F, Larzillière V, Bouchez T, Moscoviz R (2022) Systematic and quantitative analysis of two decades of anodic wastewater treatment in bioelectrochemical reactors. Water Res 214:118142. https://doi.org/10.1016/j.watres.2022.118142

de Vargas BG, Hashim MA, Chu KH (2023) The Sips isotherm equation: often used and sometimes misused. Sep Sci Technol 58:884–892. https://doi.org/10.1080/01496395.2023.2167662

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

Elemile OO, Akpor BO, Ibitogbe EM, Afolabi YT, Ajani DO (2022) Adsorption isotherm and kinetics for the removal of nitrate from wastewater using chicken feather fiber. Cogent Eng 9:2043227. https://doi.org/10.1080/23311916.2022.2043227

El-Kammah M, Elkhatib E, Gouveia S, Cameselle C, Aboukila E (2022) Enhanced removal of Indigo Carmine dye from textile effluent using green cost-efficient nanomaterial: adsorption, kinetics, thermodynamics and mechanisms. Sustain Chem Pharm 29:100753. https://doi.org/10.1016/j.scp.2022.100753

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. Separ Purif Technol 285:120314. https://doi.org/10.1016/j.seppur.2021.120314

Fernandes A, Gągol M, Makoś P, Khan JA, Boczkaj G (2019) Integrated photocatalytic advanced oxidation system (TiO_2/UV/O_3/H_2O_2) for degradation of volatile organic compounds. Separ Purif Technol 224:1–14. https://doi.org/10.1016/j.seppur.2019.05.012

Hussain SN, Asghar HMA, Sattar H, Brown NW, Roberts EPL (2015a) Free chlorine formation during electrochemical regeneration of a graphite intercalation compound adsorbent used for wastewater treatment. J Appl Electrochem 45:611–621. https://doi.org/10.1007/s10800-015-0814-3

Hussain SN, Asghar HMA, Sattar H, Brown NW, Roberts EPL (2015b) Removal of tartrazine from water by adsorption with electrochemical regeneration. Chem Eng Commun 202:1280–1288. https://doi.org/10.1080/00986445.2014.921620

Hussain SN, Trzcinski AP, Asghar HMA, Sattar H, Brown NW, Roberts EPL (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

Kanneganti D, Reinersman LE, Holm RH, Smith T (2022) Estimating sewage flow rate in Jefferson County, Kentucky using machine learning for wastewater-based epidemiology applications. Water Supply 22:8434–8439. https://doi.org/10.2166/ws.2022.395

Karimi-Jashni A, Narbaitz RM (2005) Electrochemical reactivation of granular activated carbon: effect of electrolyte mixing. J Environ Eng 131:443–449. https://doi.org/10.1061/(ASCE)0733-9372(2005)131:3(443)

Kumar D, Gupta SK (2022) Electrochemical oxidation of direct blue 86 dye using MMO coated Ti anode: modelling, kinetics and degradation pathway. Chem Eng Process Process Intensif 181:109127. https://doi.org/10.1016/j.cep.2022.109127

Largitte L, Pasquier R (2016) A review of the kinetics adsorption models and their application to the adsorption of lead by an activated carbon. Chem Eng Res Des 109:495–504. https://doi.org/10.1016/j.cherd.2016.02.006

Lau Y-Y, Wong Y-S, Teng T-T, Morad N, Rafatullah M, Ong S-A (2014) Coagulation-flocculation of azo dye Acid Orange 7 with green refined laterite soil. Chem Eng J 246:383–390. https://doi.org/10.1016/j.cej.2014.02.100

Liu D, Roberts EPL, Martin AD, Holmes SM, Brown NW, Campen AK, de las Heras N (2016) Electrochemical regeneration of a graphite adsorbent loaded with Acid Violet 17 in a spouted bed reactor. Chem Eng J 304:1–9. https://doi.org/10.1016/j.cej.2016.06.070

Ma X-Y, Fan T-T, Wang G, Li Z-H, Lin J-H, Long Y-Z (2022) High performance GO/MXene/PPS composite filtration membrane for dye wastewater treatment under harsh environmental conditions. Compos Commun 29:101017. https://doi.org/10.1016/j.coco.2021.101017

Mansor ES, Ali H, Abdel-Karim A (2020) Efficient and reusable polyethylene oxide/polyaniline composite membrane for dye adsorption and filtration. Colloid Interface Sci Commun 39:100314. https://doi.org/10.1016/j.colcom.2020.100314

Medeiros D, Nzediegwu C, Benally C, Messele SA, Kwak JH, Naeth MA, Ok YS, Chang SX, El-Din MG (2022) Pristine and engineered biochar for the removal of contaminants co-existing in several types of industrial wastewaters: a critical review. Sci Total Environ 809:151120. https://doi.org/10.1016/j.scitotenv.2021.151120

Mohammed F, Roberts E, Campen A, Brown N (2012) Wastewater treatment by multi-stage batch adsorption and electrochemical regeneration. J Electrochem Sci Eng 2:223–236. https://doi.org/10.5599/jese.2012.0019

Narbaitz R, McEwen J (2012) Electrochemical regeneration of field spent GAC from two water treatment plants. Water Res 46:4852–4860. https://doi.org/10.1016/j.watres.2012.05.046

Nkrumah-Amoako K, Roberts EPL, Brown NW, Holmes SM (2014) The effects of anodic treatment on the surface chemistry of a graphite intercalation compound. Electrochim Acta 135:568–577. https://doi.org/10.1016/j.electacta.2014.05.063

Noorimotlagh Z, Mirzaee SA, Martinez SS, Alavi S, Ahmadi M, Jaafarzadeh N (2019) Adsorption of textile dye in activated carbons prepared from DVD and CD wastes modified with multi-wall carbon nanotubes: equilibrium isotherms, kinetics and thermodynamic study. Chem Eng Res Design 141:290–301. https://doi.org/10.1016/j.cherd.2018.11.007

Osorio F, Gárate Á, Russo CM (2024) The gradient test statistic for outlier detection in generalized estimating equations. Stat Probab Lett 209:110087. https://doi.org/10.1016/j.spl.2024.110087

Purkait MK, Maiti A, DasGupta S, De S (2007) Removal of congo red using activated carbon and its regeneration. J Hazard Mater 145:287–295. https://doi.org/10.1016/j.jhazmat.2006.11.021

Rani S, Mahajan RK (2016) Equilibrium, kinetics and thermodynamic parameters for adsorptive removal of dye Basic Blue 9 by ground nut shells and Eichhornia. Arab J Chem 9:S1464–S1477. https://doi.org/10.1016/j.arabjc.2012.03.013

Saroyan H, Ntagiou D, Rekos K, Deliyanni E (2019) Reactive black 5 degradation on manganese oxides supported on sodium hydroxide modified graphene oxide. Appl Sci 9:2167. https://doi.org/10.3390/app9102167

Serafin J, Dziejarski B (2023) Application of isotherms models and error functions in activated carbon CO_2 sorption processes.

Microporous Mesoporous Mater 354:112513. https://doi.org/10.1016/j.micromeso.2023.112513

Shahbeig H, Bagheri N, Ghorbanian S, Hallajisani A, Pourkarimi S (2013) A new adsorption isotherm model of aqueous solutions on granular activated carbon. World J Model Simul 9:243–254

Song S, Fan J, He Z, Zhan L, Liu Z, Chen J, Xu X (2010) Electrochemical degradation of azo dye C.I. Reactive red 195 by anodic oxidation on Ti/SnO_2–Sb/PbO_2 electrodes. Electrochim Acta 55:3606–3613. https://doi.org/10.1016/j.electacta.2010.01.101

Suhan MBK, Mahtab SMT, Aziz W, Akter S, Islam MS (2021) Sudan black B dye degradation in aqueous solution by Fenton oxidation process: Kinetics and cost analysis. Case Stud Chem Environ Eng 4:100126. https://doi.org/10.1016/j.cscee.2021.100126

Sultana M, Rownok MH, Sabrin M, Rahaman MH, Alam SMN (2022) A review on experimental chemically modified activated carbon to enhance dye and heavy metals adsorption. Clean Eng Technol 6:100382. https://doi.org/10.1016/j.clet.2021.100382

Szyguła A, Guibal E, Palacín MA, Ruiz M, Sastre AM (2009) Removal of an anionic dye (Acid Blue 92) by coagulation–flocculation using chitosan. J Environ Manage 90:2979–2986. https://doi.org/10.1016/j.jenvman.2009.04.002

Tekin D (2014) Photocatalytic degradation of textile dyestuffs using TiO_2 nanotubes prepared by sonoelectrochemical method. Appl Surface Sci 318:132–136. https://doi.org/10.1016/j.apsusc.2014.02.018

Tzabar N, ter Brake HJM (2016) Adsorption isotherms and Sips models of nitrogen, methane, ethane, and propane on commercial activated carbons and polyvinylidene chloride. Adsorption 22:901–914. https://doi.org/10.1007/s10450-016-9794-9

Wu F-C, Tseng R-L, Juang R-S (2009) Characteristics of Elovich equation used for the analysis of adsorption kinetics in dye-chitosan systems. Chem Eng J 150:366–373. https://doi.org/10.1016/j.cej.2009.01.014

Wu F-C, Liu B-L, Wu K-T, Tseng R-L (2010) A new linear form analysis of Redlich–Peterson isotherm equation for the adsorptions of dyes. Chem Eng J 162:21–27. https://doi.org/10.1016/j.cej.2010.03.006

Xing X, Tang J, Yao S, Chen H, Zheng T, Wu J (2023) Electrochemical regeneration of granular activated carbon saturated by p-nitrophenol in BDD anode system. Process Saf Environ Protect 170:207–214. https://doi.org/10.1016/j.psep.2022.11.082

4.2 Links and implications

Reactive Black 5 (RB5) xenobiotic dye contaminants were successfully mineralised using a three-dimensional electrochemical reactor. The advanced oxidation processes effectively degraded the xenobiotic dye pollutant to inert carbon dioxide and water. The GIC adsorption process was optimised using a central composite design-response surface methodology (CCD-RSM), which was shown to improve the accuracy and precision of estimated targeted variables to model optimum removal rates. Although salt concentration in simulated, alkaline dye-contaminated wastewater had demonstrated to have a limited impact on the selectivity reversal of GIC adsorption process, pH had a more stabilising effect on the adsorption process. The strength of adsorption was primarily governed by electrostatic interaction between GIC adsorbents and dye adsorbates. The regeneration efficiency of GIC adsorbents was improved using optimal current density, and the mineralisation efficiency of pollutants was enhanced when a three-dimensional electrochemical oxidation process was used.

The best adsorption kinetic model to characterise the adsorption phenomena was the Elovich kinetic model, which yielded the best fitness. A near complete oxidation or mineralisation efficiency was achieved within 30 mins of electrolysis time and 45 mA/cm^2 of current density. Although the adsorptive capacity of GIC adsorbents can be improved further, the adsorption or pollutant removal efficiency was suppressed by the presence of intermediate transformation oxidation byproducts when high current density was applied. The changes in surface physicochemical properties of GIC adsorbents caused by extremely high current density resulted in reduced regeneration efficiency and increased likelihood of particle attrition or corrosion. Future research should focus on maximising the regeneration efficiency and improving the physicochemical stability of adsorbents.

5.1 Introduction

This research article focuses explicitly on using a three-dimensional electrochemical reactor to treat dye-polluted wastewater, emphasising artificial intelligence and machine learning optimisations. Graphite intercalation compound was used as a particle electrode to adsorb xenobiotic dye contaminants from simulated wastewater. A more advanced, novel progressive central composite design-response surface methodology (CCD-NPRSM), hybrid artificial neural network-extreme gradient boosting (ANN-XGBoost) ensemble, and classification and regression trees (CART) were used to optimise the operational parameters of the three-dimensional electrochemical treatment of reactive black 5 (RB5) polluted wastewater.

The chemical stability of RB5 is a desirable dyeing property for textile manufacturers, making it suitable for staining cotton and other cellulose fibres. The black intensity of colouration is ideal for dyeing materials manufactured by the textile, printing, and leather industries. However, RB5 is highly recalcitrant to environmental degradation. It contributes to significant ecological toxicity when industrial effluents are discharged into marine environment without proper control measures.

Three-dimensional electrochemical oxidation technology offers an attractive method for wastewater treatment, especially when a hybrid treatment method is used to maximise pollutant removal and mineralisation efficiencies. In the electrochemical oxidation mechanism, this research article shows that there are two mechanisms to characterise the pollutant removal process: 1) direct oxidation involves the electron transfer from the organic pollutants to the electrode surface; 2) indirect oxidation involves electrogenerated oxidizing species to degrade and oxidise organic pollutants. In addition, electrically regenerative GIC particle electrodes were used to recover the adsorptive capacity when the particles were placed in the regeneration zone of the three-dimensional electrochemical reactor. Organic pollutants were adsorbed and oxidised on the surface of particle electrodes simultaneously. Highly oxidising species with strong oxygen evolution reaction potential mediated the combined oxidation process from anodic reaction and particle surface. In addition, GIC particle electrodes exhibited superior electrocatalytic potential and regeneration efficiency, restoring full adsorptive capacity after several adsorption and regeneration cycles, resulting in sustained catalytic oxidation performance.

Complex process variables warrant using AI and ML optimisation techniques to manage the combined effects of electrochemical systems. The efficacy of three-dimensional

electrochemical oxidation technology relies on the accuracy and precision of AI and ML predictive models to determine the best, optimal conditions for the electrochemical decomposition of RB5-polluted wastewater. Several novelties were associated with using different CCD-NPRSM, AI and ML ensembles to improve the prediction efficiency of targeted variables. A conventional RSM was originally used to predict targeted responses using an empirical second-order polynomial equation. More interestingly, this RSM technique was adapted to a uniquely designed transfer function to establish multilevel nested models. The optimisation procedure involved measuring maximum dye and TOC removal efficiencies, optimal current density, and electrical energy consumption for RB5 and total organic carbon (TOC) removal rates. TOC strongly indicates mineralisation efficiency associated with a particular electrical energy consumption.

On the other hand, hybrid ANN-XGBoost ensemble was an algorithm used to optimise the three-dimensional electrochemical treatment of RB5-polluted wastewater. Although XGBoost generated a relatively weaker model than ANN, combining these two hybrid ensembles significantly enhanced the prediction efficiency of targeted responses. A second-order Taylor expansion of the loss function was used to integrate a regular term to generate an optimal solution to balance the decline in the loss function. This resulted in better management of model complexity and effectively mitigating overfitting issues. In addition, the predictive analytics by CART machine learning optimisation was another significant approach used to optimise the three-dimensional electrochemical reactor. It possessed the ability to manage large datasets, messy or missing data, or any extreme outliers and other nonlinear relationships. CART-generated models effectively visualise predicted values and the interactive effects of variables to achieve optimal prediction accuracy. The predictive accuracy and precision of three different CCD-NPRSM were critically evaluated using ANOVA analysis to justify the validity of the models.

+ MODEL
ARTICLE IN PRESS
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).

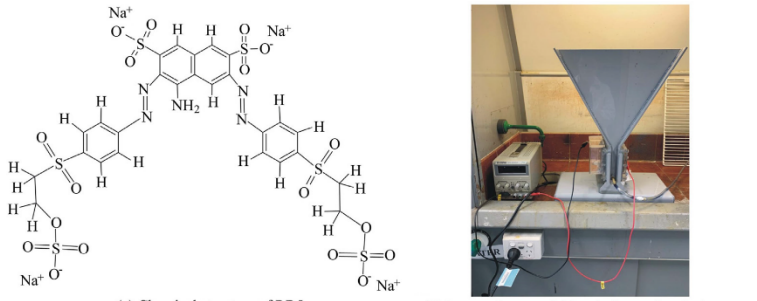
This study investigated the efficacy of 3D electrochemical treatment for removing RB5 xenobiotic dye from water, using an electrically conductive GIC. An RSM with a face-centred central composite design (CCD) was developed to construct a mathematical model for predicting dye and total organic carbon (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):

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

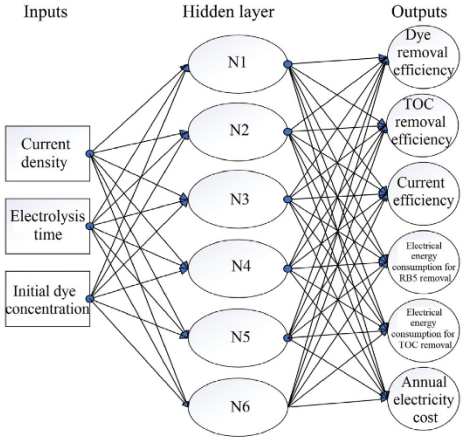


Fig. 2. ANN network with topology.

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:

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

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

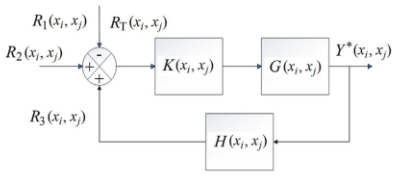


Fig. 3. Derivation of transfer function from block diagram (with R_T denoting sum of input functions and Y^* representing output function).

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

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.

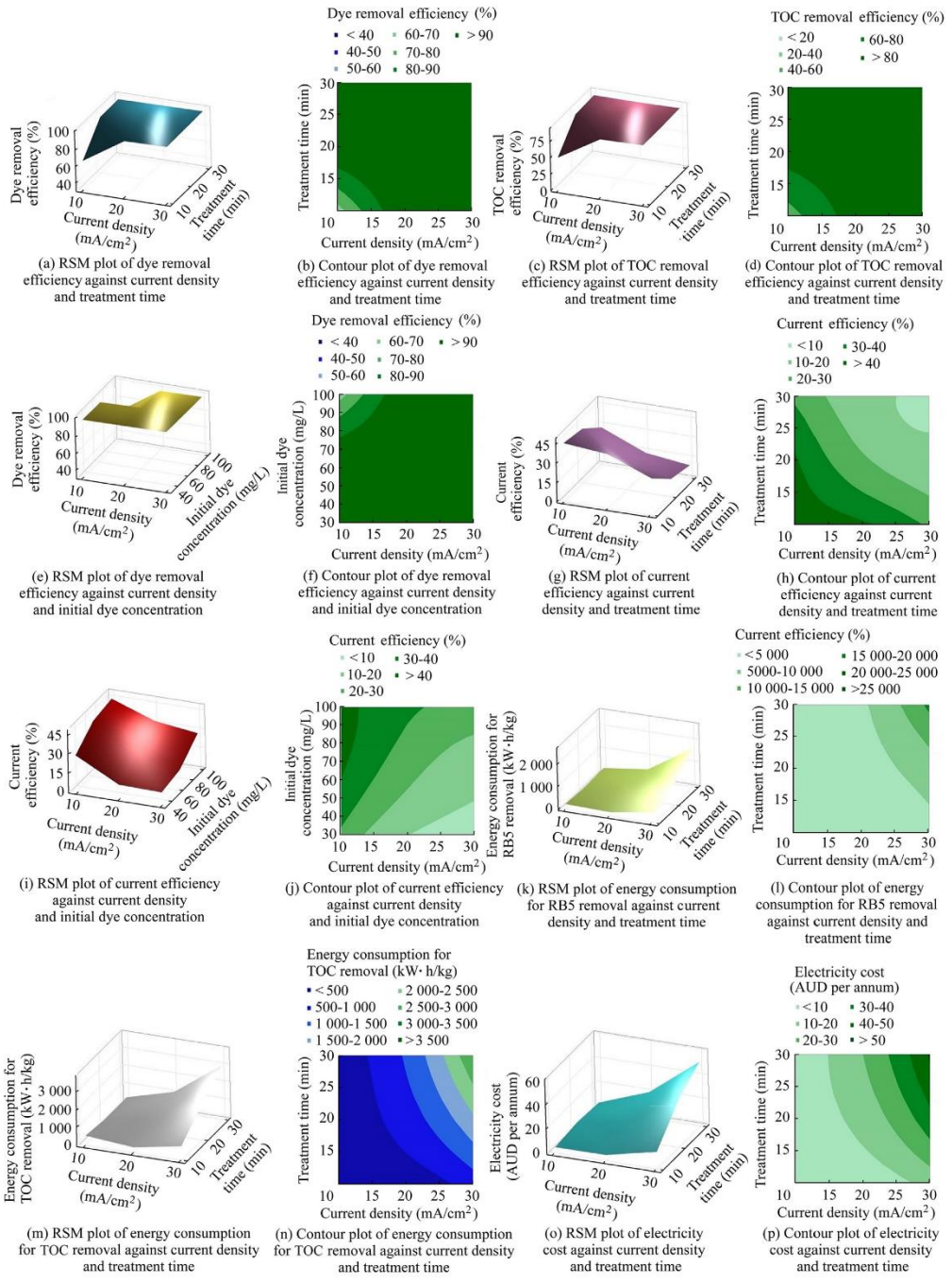
2.2.3.4. Predictive analytics by CART machine learning optimisation. CART machine learning optimisation stands out as one of the best-in-class approaches, not only fitting more accurate models when combined with experimental data but also for handling larger datasets with more variables, messy or missing 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.

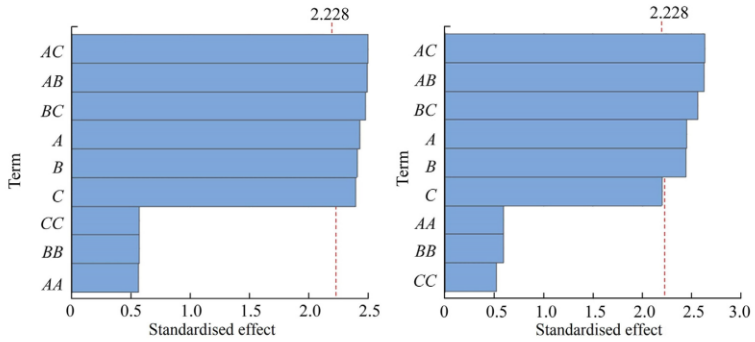
Fig. 4(h) shows that when current density exceeded 25 mA/cm^2 with an electrolysis time greater than 25 min, the current efficiency of the electrochemical reactor decreased below 10%. This result indicated that a higher current density led to side reactions due to the accumulation of intermediate transformation 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

5.2 Links and implications

The prediction efficacy of targeted responses was critically evaluated using uniquely designed CCD-NPRSM, AI and ML ensemble algorithms to optimise the three-dimensional electrochemical treatment of RB5-contaminated wastewater. The effects of operational parameters on targeted responses, such as dye and TOC removal efficiencies, current efficiency, electrical energy consumption of RB5 and TOC removals, and electricity cost were effectively optimised to achieve optimal conditions. The optimisation results showed that CCD-NPRSM, hybrid ANN-XGBoost ensemble and CART algorithms, along with its approximating functions, successfully led to satisfactory convergence of solutions, resulting in the best fitness of modelling. The key optimisation results showed that hybrid ANN-XGBoost ensemble generated the best optimal solution with TOC mineralisation efficiency of 90.47% compared to 89.76% and 89.68% by CCD-NPRSM and CART optimisations. Although CART optimisation was effective, the prediction errors of response variables were significantly higher than other optimisation techniques. Overall, the predictive efficiency of hybrid ANN-XGBoost ensemble exceeded expectations compared to other optimisation methods.

6.1 Introduction

This article is an extension of Paper 4 research, specifically emphasising the advanced combination of AI and ML ensembles to optimise the process conditions of three-dimensional electrochemical treatment of xenobiotic dye wastewater. The synergistic performance of three-dimensional electrochemical process of sequential batch reactor allowed a rapid degradation of methyl orange (MO) dye pollutant in xenobiotic dye wastewater. In the absence of AI and ML-optimised models, the targeted variables, such as MO and TOC removal efficiencies, electrical energy consumption and current efficiency could not be accurately predicted, optimised and estimated. One of the most remarkable findings from this research was that the mineralisation efficiency of 50 mg/L MO was rapidly degraded using a current density of 15 mA/cm^2. The electrochemical degradation mechanisms were mediated by direct and indirect oxidation processes involving highly oxidizing species, such as hydroxyl radicals and active chlorine species. Novel electro-regenerative and electro-degradative GIC particle electrodes were incorporated into the three-dimensional electrochemical reactor to facilitate the electrochemical degradation and adsorption of MO dye pollutant.

Furthermore, unique combination of AI and ML ensembles was formulated involving artificial neural networks, support vector machine and random forest and finally combining it with Monte Carlo simulations, to conduct sensitivity analysis and manage the system perturbations, process variabilities and estimation uncertainties in the process conditions of three-dimensional electrochemical reactor. In addition, the prediction efficiency of each AI/ML technique was compared with multiple regression analyses and ranked in terms of superiority of the predictive models. More importantly, AI and ML ensembles were used to balance the current efficiency and electrical energy consumptions of three-dimensional electrochemical reactor, while maintaining its mineralisation and energy efficiencies.

Environmental Processes (2024) 11:41
https://doi.org/10.1007/s40710-024-00719-1

RESEARCH

Monte Carlo Simulation, Artificial Intelligence and Machine Learning-based Modelling and Optimization of Three-dimensional Electrochemical Treatment of Xenobiotic Dye Wastewater

Voravich Ganthavee[1] · Merenghege M. R. Fernando[1] · Antoine P. Trzcinski[1]

Received: 16 January 2024 / Accepted: 10 July 2024

Abstract

The present study investigates the synergistic performance of the three-dimensional electrochemical process to decolourise methyl orange (MO) dye pollutant from xenobiotic textile wastewater. The textile dye was treated using electrochemical technique with strong oxidizing potential, and additional adsorption technology was employed to effectively remove dye pollutants from wastewater. Approximately 98% of MO removal efficiency was achieved using 15 mA/cm^2 of current density, 3.62 kWh/kg of energy consumption and 79.53% of current efficiency. The 50 mg/L MO pollutant was rapidly mineralized with a half-life of 4.66 min at a current density of 15 mA/cm^2. Additionally, graphite intercalation compound (GIC) was electrically polarized in the three-dimensional electrochemical reactor to enhance the direct electrooxidation and·OH generation, thereby improving synergistic treatment efficiency. Decolourisation of MO-polluted wastewater was optimized by artificial intelligence (AI) and machine learning (ML) techniques such as Artificial Neural Networks (ANN), Support Vector Machine (SVM), and random forest (RF) algorithms. Statistical metrics indicated the superiority of the model followed this order: ANN > RF > SVM > Multiple regression. The optimization results of the process parameters by artificial neural network (ANN) and random forest (RF) approaches showed that a current density of 15 mA/cm^2, electrolysis time of 30 min and initial MO concentration of 50 mg/L were the best operating parameters to maintain current and energy efficiencies of the electrochemical reactor. Finally, Monte Carlo simulations and sensitivity analysis showed that ANN yielded the best prediction efficiency with the lowest uncertainty and variability level, whereas the predictive outcome of random forest was slightly better.

Highlights

- In-depth analysis of various artificial intelligence optimization techniques.
- Prediction efficiency of artificial intelligence and machine learning algorithms.
- 98% dye removal and 100% regeneration of graphite intercalation compound.
- Advanced statistical analysis of targeted responses and data fitting techniques.
- Analysis of uncertainties and variability using Monte Carlo simulation.

Keywords Dye removal · Adsorption and electrochemical treatment · Artificial neural network · Support vector machine · Random forest · Monte Carlo simulation

Extended author information available on the last page of the article

Published online: 01 August 2024

1 Introduction

The textile, printing and dyeing industries are some of the largest producers of dye wastewater, contributing up to about 0.7 million metric tons of chemical dyes produced annually, accounting for 17 to 20% of water pollution worldwide (Pavlović et al. 2014). In Bangladesh, the textile sector currently exports nearly 28 billion USD annually, up to 82% of the country's total export earnings (Hossain et al. 2018). In 2021, the textile industries in Bangladesh produced approximately 2.91 million metric tons of fabrics and around 349 million metric tons of wastewater generated from conventional dyeing practices (Hossain et al. 2018). Figure 1 represents the water and chemical consumption of the textile processing industry in Bangladesh.

1.1 Types of Textile Wastewater Treatment

Globally, about 60% of the annual output of synthetic dyes consists of azo compounds (Liu et al. 2022). These azo dyes possess stable azo function groups (N=N) and aromatic rings, which make it very chemically stable and highly resistant to environmental biodegradation and UV photolysis (Cui et al. 2021). These azo dyes have strong chromaticity, ecotoxicity and carcinogenicity, which pose a significant health risk and environmental hazard (Kumar and Gupta 2022). Residual dyestuffs are characterized by intense colour, high organic content, and highly stable chemical structure, with strong potential to cause serious environmental pollution (El-Kammah et al. 2022).

Various treatment methods are commonly employed to remove textile dyes from wastewater using biological degradation (Singh et al. 2022), coagulation-flocculation (Lau et al.

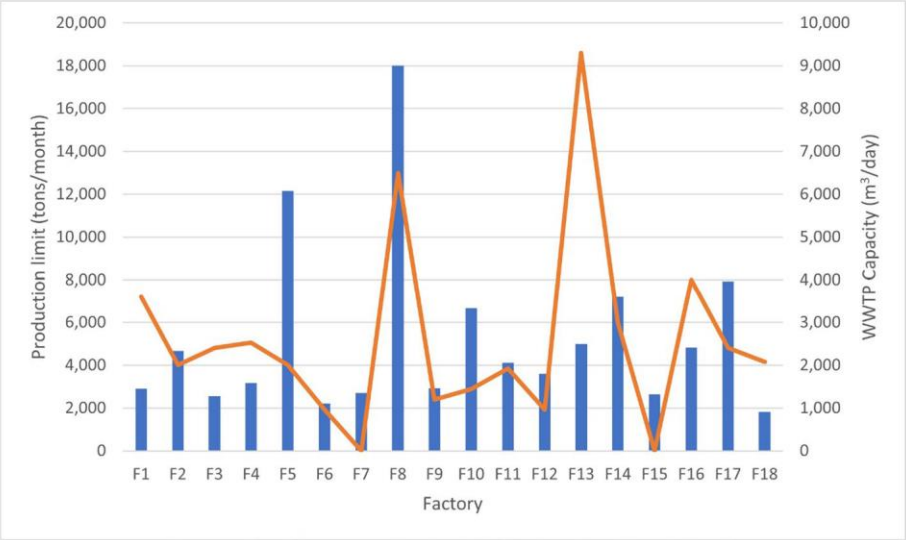


Fig. 1 Water and chemical consumption of textile processing industry in Bangladesh with a production capacity of 1,812 tons to 18,000 tons annually, amounting to 24-h shifts and 25 working days (Uddin et al. 2023)

2014), membrane filtration (Wu et al. 2022a), Fenton reagent (Badmus et al. 2020) and photocatalytic degradation (Chairungsri et al. 2022). Still, they are ineffective due to excessive sludge production, secondary pollution and membrane fouling (Nidheesh et al. 2018). Excessive sludge production from biological treatment process requires additional post-treatment and waste management processes, resulting in large energy consumption and financial expenditure (Shoukat et al. 2019). On the other hand, membrane fouling is a significant issue in filtration process, which hampers filtration effectiveness (Wu et al. 2022a). Coagulation-flocculation process will contribute to secondary pollution due to chemical reagents used to remove the pollutants (Januário et al. 2021; Tahraoui et al. 2023). The recovery of chemical reagents is challenging, resulting in a loss of energy and resources (Ihaddaden et al. 2022). Catalyst poisoning and electron–hole recombination are significant issues in advanced oxidation processes, leading to reduced oxidation potential (Fu et al. 2023; Kanjal et al. 2023). The catalysts used in advanced oxidation processes are costly (Saravanan et al. 2022). Fenton's reagent leads to a significant issue with iron sludge generation due to combined flocculation with the reagent and organic compounds (Mechati et al. 2023; Suhan et al. 2021). Additional pH adjustment is needed for Fenton's reagent to facilitate oxidation, increasing operational costs (Can-Güven 2021; Kebir et al. 2023). Furthermore, a fluidized three-dimensional electrochemical oxidation process was used to treat MO wastewater and achieved a removal efficiency of 99.9% in 30 min, whereas the original adsorption capacity of activated carbon was maintained at 64.5% after 8 cycles of adsorption-electrochemical regeneration (Liu et al. 2022). Therefore, the literature recommended using electrochemical process as an advanced wastewater treatment technique to remove dyes from industrial effluent.

The electrochemical treatment is commonly used to eliminate pollutants on the anodic surface, via generation of·OH oxidants and active chlorine species (Hamida et al. 2022). However, few studies have been conducted on AI and machine learning-based optimization of three-dimensional electrochemical treatment of textile wastewater involving graphite intercalation compound (GIC) particle electrodes. The current disadvantages of using anodic oxidation technology such as boron-doped diamond and mixed metal oxide electrodes are due to poisoning of electrodes and buildup of biofilm or thin-oxide layer, which can decrease its electrocatalytic efficiency and service life by 10–90% (El Aggadi et al. 2021). These electrodes could not overcome the issues associated with mass transfer resistance and short half-life of oxidizing species with approximately $10^{-6} \sim 10^{-3}$ s in wastewater media (Chen et al. 2023; Xie et al. 2022). The current disadvantages of adsorbent materials made from agricultural sources and carbon-based substituents are non-regenerative and susceptible to heat stress or other physicochemical degradation (Vinayagam et al. 2022). On the other hand, granular activated carbon (GAC) has 10–20% lower regeneration efficiency than GIC, making it unsuitable for use in electrochemical reactor (Narbaitz and McEwen 2012; Narbaitz and Karimi-Jashni 2012). Hence, we aimed to improve the electrocatalytic efficiency of three-dimensional electrochemical reactor using an electrically regenerative particle electrode to achieve high mineralization efficiency of dye pollutants in wastewater.

More critically, there is a potential gap in comparing the prediction efficiency of various artificial intelligence and machine learning-based optimization approaches specific to three-dimensional electrochemical treatment process. Intelligent control of electrochemical nitrate removal was based on artificial neural network whereas electrochemical sensors were applied to monitor and remove azo dyes and food colorant substances. None of the past research explored the artificial intelligence and machine learning-based optimization techniques on three-dimensional electrochemical treatment of xenobiotic dye removal (Meng et al. 2022; Wu et al. 2022b). On the other hand, past research focused on using electrochemical conversion of ammonia into harmless nitrogen gas by utilizing granular activated carbon

as three-dimensional particle electrode which was poorly regenerative or of low electrical conductivity compared to graphite intercalation compound (Zhang et al. 2024). Moreover, system perturbations, uncertainties and variability of operating parameters and their impact on targeted responses specific to three-dimensional electrochemical treatment process are never accounted for in the current literature. This involved exploring the uncertainties in AI optimization effect of operating parameters such as applied current density, electrolysis time and initial dye concentration, to improve the electrooxidation efficiency of the three-dimensional electrochemical reactor. Most significantly, the novelty of this research lies in finding the best artificial intelligence-based models to improve the prediction efficiency of complex phenomena by applying them to large physical, chemical and biological processes. Secondly, the research aims to develop accurate artificial intelligence-based models which can be integrated into the upscaled conventional wastewater treatment systems to enhance value engineering, water resources management, energy efficiency, real-time process dynamics, data controllability and streamlining distributed network of process control systems. Unlike other past research, this research also aims to scrutinize the prediction efficiency of different artificial intelligence and machine learning-based models by analysing the level of uncertainties or the effect of various operating parameters on system perturbations and variability to enhance the accuracy and precision of predictive model platforms.

2 Materials and Methods

2.1 Experimental Equipment and Materials

Methyl orange ($C_{14}H_{14}N_3NaO_3S$) was a chemical reagent grade obtained from Chem-Supply, Australia. Commercial GIC was purchased from Sigma-Aldrich, Australia. The particle size of GIC was greater than 300 µm (50 mesh). GIC has an electrical conductivity of approximately 0.8 S/cm. The MO solution was prepared using high purity distilled water. UV/Visible spectrophotometer (DR6000, Hach) was used to determine the MO dye concentrations in solution at different time intervals. The maximum absorption occurred at a wavelength $\lambda = 463$ nm. The coefficient of variation (COV) for the UV-absorbance analysis of MO was approximately 3.08%, whereas for the TOC analysis (TOC-V CSH, Shimadzu), it was approximately 0.55%.

The experiment was performed in a 6–7 L electrochemical reactor equipped with anode and cathode. A more detailed description can be found in Trzcinski and Harada (2023). Graphite plate anode with approximately 70 cm^2 of electroactive surface area and stainless steel 316 cathode were connected with a 60 V DC power supply unit (Model GPR-6030D, GW INSTEK, Taiwan) to form a closed-looped electrical circuit. Compressed air at 2 bar was sparged into the anodic compartment of the reactor to mix GIC and contaminated water. A solution of 0.3% (w/v) of NaCl adjusted to pH 2 using HCl was used as the supporting electrolyte.

2.2 MO Adsorption and Electrochemical Oxidation Process

In 3D electrochemical process, 1-L of MO-contaminated water was first added into the reactor, and air pressure was set at 2 bar to start the adsorption process. After the adsorption process was over at 20 min, GIC particle electrodes were allowed to settle down in

the regeneration zone between the cathode and the anode. The regeneration zone is located within the anodic compartment where the GIC particle electrodes are electrochemically regenerated when subjected to electrolysis. The current supply ranged from 1.05 to 3.16 A, corresponding to a current density of 15 to 35 mA/cm^2 applied for 10 min. Mathematical equations used for characterising three-dimensional process are outlined in subsection 2.2.1.

2.2.1 Mathematical Equations for the Electrochemical Process

In the study of 3D electrochemical process, a pseudo-first-order kinetic model (Eq. 1) was used to describe the change of concentration over time, and the pseudo-first-order reaction rate constant represents the electrooxidation kinetics of MO removal by 3D process. Alternatively, $t_{1/2}$ represents the half-life of mineralization rate for 50% of MO pollutants to degrade in an aqueous solution. The combined adsorption and electrochemical oxidation process synergistically maximise the dye and TOC removal efficiencies of MO in aqueous solutions. The pseudo-first-order kinetic rate constant representing the electrooxidation kinetics can be determined from the following equation (Liu et al. 2022):

$$-\log_e\left(\frac{C_t}{C_0}\right) = kt \tag{1}$$

where k represents the kinetic rate constant in min^{-1}; C_0 represents the initial dye concentration from 50 mg/L to 125 mg/L; C_t represents the final dye concentration changes according to time after a period of adsorption and electrochemical oxidation; and t is the time in min.

The regeneration efficiency, RE, can be calculated from the following equation:

$$RE = \frac{q_r}{q_i} \times 100\% \tag{2}$$

where q_i represents the initial loading of MO (mg/g) onto fresh GIC adsorbent; and q_r represents the final loading (mg/g) on the regenerated GIC adsorbent under identical adsorption conditions:

$$q_i = \frac{(C_0 - C_i)V}{m} \tag{3}$$

$$q_r = \frac{(C_0 - C_t)V}{m} \tag{4}$$

where C_0 denotes the initial dye concentration (mg/L), C_i represents the dye concentration (mg/L) after adsorption but before electrochemical regeneration, C_t represents the dye concentration (mg/L) after electrochemical regeneration, and t is the regeneration time.

The charge passed per gram of GIC adsorbent is given by the following relationship:

$$Q_t = \frac{It}{m} \tag{5}$$

where I is the applied current (A); t is the electrolysis time (min); and m is the mass of GIC adsorbent (g).

To calculate the applied current density, J_{EO}, the following equation is used:

$$J_{EO} = \frac{I}{SA} \tag{6}$$

where I denotes the current applied (A); and SA denotes the surface area of the anode, which was 70 cm^2.

To calculate the electrical energy consumption per kg of adsorbed MO, the equation is shown below:

$$EC\left(\frac{kWh}{kg}\right) = \frac{I \times V \times t}{(C_0 - C_t)V} \tag{7}$$

where I is the applied current (A); U_t is the cell potential at time t (V); and V is the MO solution volume, which was 1.0 L.

To calculate the electrical energy consumption per kg TOC of adsorbed MO, the following equation is used:

$$EC\left(\frac{kWh}{kg\ TOC}\right) = \frac{I \times V \times t}{(TOC_0 - TOC_t)V} \tag{8}$$

where I is the applied current (A); U_t is cell potential at time t (V); V is MO solution volume (1.0 L); t is the time (min); and TOC_0, TOC_t are TOC concentrations initial and final total organic carbon concentrations in mg/L, respectively at time t (min).

Based on the actual charge passed per gram for the equation above, the theoretical equation is as follows:

$$Q_{th} = \frac{n(C_0 - C_f)VF}{M_w} \tag{9}$$

where C_i and C_f are the initial and final MO concentrations in solution taken before and after the adsorption-electrochemical regeneration for 5 cycles; V is the solution volume (L); F is Faraday's constant (96,487 C mol^{-1}); M_w is the molecular weight of MO is 327.33 g mol^{-1}; and n represents the number of electrons, which is 90 for complete oxidation and 36 for incomplete oxidation (see Sect. 3.1).

The current efficiency equation is as follows:

$$Current efficiency(\%) = \left(\frac{Theoretical\ Charge}{Actual\ Charge\ Passed}\right) \times 100\% \tag{10}$$

After the graphs were generated using the experimental data, various AI and machine learning optimisation techniques were applied to compare any deviation between the experimental and optimised values. Subsection 2.3. briefly summarises the data analysis methods used for AI and machine learning optimisation techniques.

2.3 Data Analysis Methods

Before performing the optimization of experimental data using Artificial Neural Networks (ANN), Support Vector Machine (SVM) and random forest (RF), an approximate model must be developed in preparation for training and testing. To configure the dataset,

designing the network architecture of any AI model is critical to incorporating activation functions, transfer functions, nodes or layers, etc. Once the model architecture is created, model training and testing procedures must be performed to train and test the network architecture to evaluate the model performance. The training and testing procedures are critical to improve the generalization of predictive performance. In these procedures, the input operating parameters from the experimental data were transferred into the activation functions of the modelled network architecture to generate output response variables. The tested experimental results were compared with the AI or machine learning optimised results to derive any error deviation using statistical analysis for data fitting purposes involving the use of either ANN, SVM classifier model or random forest decision trees to enhance the predictive outcomes.

2.4 Electrochemical Regeneration of GIC

This electrochemical regeneration experiment was subdivided into three phases:

1) *Initial adsorption*: Air was sparged into the reactor containing 200 g of GIC particle electrodes for 20 min. The air pressure was 2 bar to facilitate the mass transfer of dye molecules onto the particle electrodes. After 20 min, the air supply was turned off to allow the GIC particle electrodes to settle onto the bottom of the anodic compartment.
2) *Adsorption-electrochemical regeneration phase*: A DC power source supplied a fixed current through the cell during electrochemical regeneration. The electric field was turned on for 10 min to facilitate the electrochemical regeneration of GIC.
3) *Next cycles of adsorption-electrochemical regeneration*: The air was turned off, allowing the particle electrodes to settle onto the regeneration zone of the electrochemical cell. The remaining electrochemically treated solution was drained off. A fresh dye solution was added for the next round of adsorption-electrochemical regeneration.

2.5 Analytical Methods

In the following experiment, 1,000 mL of the 50–250 mg/L of MO stock solution was subjected to electrochemical treatment. Experiments were carried out at a temperature of 22 ^{0}C, and the dye solutions were filtered using a 5 µm filter funnel. A 5,000 µL "Eppendorf" syringe was used to take the liquid samples from the dye solutions at intervals ranging from 0 to 30 min. These liquid samples were analysed using a UV/Visible Spectrophotometer (λ_m = 463 nm, Hach DR6000) and a TOC analyser (Shimadzu TOC-V CSH) to determine the dye and TOC concentrations throughout the electrochemical treatment. The coefficient of variation (COV) for the UV-absorbance analysis of MO is approximately 3.08%, whereas for the TOC analysis, it is approximately 0.55%.

2.6 AI Modelling and Optimization

2.6.1 ANN Procedure

ANN is widely used to solve complex, multivariate and non-linear problems via classification and regression modelling (Khan et al. 2022). ANN optimization method was applied to model and predict responses influenced by operational variables. ANN is a subset of

machine learning algorithms (Oruganti et al. 2023). It mimics the behaviour of human brain and nervous system with outstanding learning ability. ANN is a black-box model that employs a gradient descent propagation technique to predict a target output variable (Picos-Benítez et al. 2020). It is structured into three layers, each node connected by inputs and outputs, as shown in Fig. 2. The ANN processes involve one or more hidden layers connected by input parameters consisting of current density, electrolysis time and initial MO concentration, and output layers consisting of MO removal efficiency, current efficiency, electrical energy consumption of MO and TOC, which is known as the multilayer perceptron (MLP) structure (Asgari et al. 2020). The number of neurons in each input and output layer can be as many as the number of input and response variables. In this study, a three-layer ANN model with a hidden layer was designed, in which the tangential sigmoid function was used at the hidden layer, whereas a linear transfer function was used at the output layer. The Levenberg–Marquardt backpropagation algorithm with 1000 epochs was employed for training the network. The number of neurons located in the hidden layer was a range of 1–20 to give the best optimum values based on minimum mean squared error (MSE). The ANN analysis was performed using MATLAB R2023a. The performance of ANN modelling can be statistically evaluated using the MSE and the correlation coefficient, in accordance with the following Eqs. (11) and (12), respectively (Khan et al. 2022; Özdoğan-Sarıkoç et al. 2023):

$$\mathrm{MSE} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}}\left(\left|\mathrm{y}_{\mathrm{pred,i}} - \mathrm{y}_{\mathrm{exp,i}}\right|\right)^2 \tag{11}$$

$$\mathrm{R}^2 = 1 - \sum_{\mathrm{i}=1}^{\mathrm{N}}\frac{\left(\mathrm{y}_{\mathrm{i,cal}} - \mathrm{y}_{\mathrm{i,exp}}\right)^2}{\left(\overline{\mathrm{y}_{\mathrm{exp}}} - \mathrm{y}_{\mathrm{i,exp}}\right)^2} \tag{12}$$

$\mathrm{y}_{\mathrm{pred,i}}$ and $\mathrm{y}_{\mathrm{exp,i}}$ denote predicted and experimental i^{th} values in scalar unit such as dye or TOC removal efficiency, respectively. $\overline{\mathrm{y}_{\mathrm{exp}}}$ represents an average experimental value of either dye or TOC removal efficiency. MSE and R^2 are mean square error and coefficient of determination, respectively.

2.6.2 SVM Procedure

The SVM method is built upon the fundamental concept that involves applying either a linear or non-linear mapping function to map the experimental or actual data into a higher dimensional feature space and search for an optimum hyperplane in the new space to achieve classification of samples (Ding et al. 2023). The SVM algorithm and its regression models have faster training time and are more advantageous than the ANN models in finding the universal optimal solutions for a given experimental dataset (Özdoğan-Sarıkoç et al. 2023). The support vector regression (SVR) algorithm can be extracted from the SVM algorithm to predict response variables. However, given the limited predictability of the ANN algorithm, the radial basis of ANN function was still dominant compared to the SVM algorithm (Safeer et al. 2022). Moreover, SVM helps to identify patterns and/or classify the specific dataset. It compares the differences between the predicted and experimental values, providing information on the degree of fitness. The primary goal of SVM algorithm is to identify the hyperplane in an N-dimensional space that classifies distinct datasets (Singh et al. 2023). There are a number of features that define the hyperplane. However, as the number of features increases, the complexity of model also increases, making it more challenging to comprehend. When combined with ANN model, the

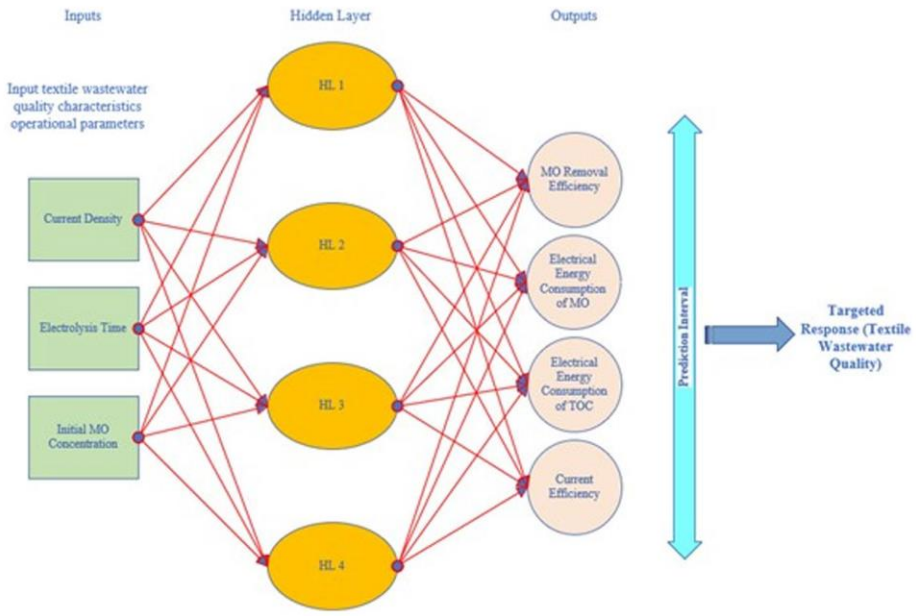


Fig. 2 ANN network with topology. ANN operates like a human brain and nervous system. It possesses one or more hidden layers, input and output layers, which are known as multilayer perceptron (MLP) structures. The neurons in the input are feedforwarded through the hidden layers to the output layers, representing the response variables. The Levenberg–Marquardt backpropagation algorithm is adopted to train the network

interpretation of complex model becomes more manageable. The predictive performance indicator of SVM model is used in AI optimization as follows (Khan et al. 2022):

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(y_{i,cal} - y_{i,exp}\right)^2 \tag{13}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(Z_i - Y_i\right)^2}{n}} \tag{14}$$

$y_{i,cal}$ and $y_{exp,i}$ denote calculated and experimental i^{th} values in scalar unit, such as dye or TOC removal efficiency, respectively. Z_i and Y_i denote predicted and experimental i^{th} values, such as dye or TOC removal efficiency.

SVM is a regression model that requires a decision boundary involving a maximum-margin hyperplane to solve a learning sample (Wang et al. 2022). To perform curve fitting, the conceptual relationship of SVM and Lagrange multiplier method involves regression analysis of the data. This relationship can be described using a functional equation of the regression as follows (Wang et al. 2022):

$$f(x) = \omega \cdot \phi(x) + b \tag{15}$$

where x represents the input vector; ω, b: the parameter vector; ϕ(x): the characteristic function. In addition, $\phi : X \rightarrow \phi(X) \in R^H$ is any non-linear function that maps the input experimental data into a high-dimensional feature space (Rodriguez-Galiano et al. 2015).

The model optimisation was subjected to the soft-margin constraint involving hyperplane, distinguishing the training data with the maximum margin. The optimization problem can be solved using the Lagrange multipliers method, which is the Kernel function defined as the inner product of the transformed input feature vectors (Rodriguez-Galiano et al. 2015):

$$K(x_i, x_j) = \langle \phi(x_i) | \phi(x_j) \rangle \tag{16}$$

2.6.3 Random Forest

Random forest is essentially a Classification and Regression Trees (CART) algorithm, which is part of a machine learning-based approach with the potential to capture complex non-linear relationships between selected models (Wang et al. 2022). Random forest utilizes multiple trees with nodes to train and predict samples, with representation by decision trees. The chosen training data are randomly returned, and newly learned data is continuously constructed, resulting in newly established decision trees to increase the overall effect of accuracy and stability of predictions. For solving regression problems, the random forest generates a final prediction result for each decision tree based on the mean of the predicted data.

2.7 Statistical Analysis and Data Fitting Using AI Models

2.7.1 Development of ANN Architecture

All operational parameters used in ANN approach were adopted from the experimental data. In addition, the desired output responses were MO removal efficiency, electrical energy consumption of MO and TOC, and current efficiency of the electrochemical reactor. Firstly, it was assumed that artificial neurons are arranged in sequential layers. Secondly, the neurons within the same layers do not interact with one another. Thirdly, all input operating parameters entering the network architecture must pass from the input layer through the hidden layer to the output layer. All hidden layers must have a similar activation or transfer function. Once the output variables are generated, they are compared with the input variables using statistical analyses involving MSE, RMSE, R^2, etc.

The proposed mathematical equation representing the ANN model can be written as follows (Asgari et al. 2020):

$$Y_n = f_0\left\{ a_0 + \sum_{k=1}^{h} \left[w_k \times f_h\left(a_{hk} + \sum_{i=1}^{m} j_{ik} X_{ni} \right) \right] \right\} \tag{17}$$

where Y_n represents the normalized response variable, f_0 denotes the transfer function in the output layer, b_0 is the bias value in the output layer, w_k is the weights between the output and hidden layers, f_h is the transfer function representing the tan-sigmoid function in a specific study in the hidden layer, a_{hk} is bias value in the hidden layer, j_{ik} represents the weights involved between the hidden and input layers, and X_{ni} denotes the normalized input variables ranging between 0.1 to 0.9 for a specific study.

2.7.2 Multiple Regression Analysis

Multiple regression analysis is one of the statistical techniques used to analyse the relationship between a single dependent variable and a range of independent variables. The primary purpose of using multiple regression analysis is to use independent variables to predict the value of a single dependent variable (Wagner et al. 2006). Each predictor is weighed, with total weights contributing to the overall prediction. The following represents the equation for describing the overall prediction (Wagner et al. 2006):

$$Y = a + b_1X_1 + b_2X_2 + \cdots + b_nX_n \tag{18}$$

where Y denotes the dependent variable; X_1 and X_n represent the number of independent variables; b_1 and b_n represent the weights to ensure maximum prediction of dependent variable from the set of independent variables.

3 Results and Discussion

3.1 Effect of the Operational Parameters on the Electrochemical Process

Current density was one of the most influential parameters affecting the overall electrochemical treatment efficiency. The experiment studied the effect of 15 mA/cm^2 of current density on the degradation efficiency of MO by 3D electrochemical process. In addition, Fig. 3 shows that when the current density of 15 mA/cm^2 was applied for at least 30 min of electrolysis time to treat a range of initial MO concentrations ranging from 50 to 125 mg/L, the MO removal rate constants changed from 0.149 to 0.036 min^{-1} while MO removal efficiency decreased from 98.8% to 66.0%. Approximately 70% (0.046 min^{-1}) and 90% (0.241 min^{-1}) of removal efficiencies and removal rate constants were achieved in 2D and 3D electrochemical treatment by Liu et al. (2022). The results indicate that the higher the initial MO concentration, the lower the MO removal efficiency and removal rate constants due to competitive reaction between·OH and dye pollutants. There were two types of oxidation reaction: 1) direct anodic oxidation of MO pollutants via anode; 2) indirect oxidation of MO via powerful oxidants such as hydroxyl radical and active chlorine species electrogenerated in bulk solution, anode and particle electrode surfaces. In addition, Fig. 3 shows that the applied current density of 15 mA/cm^2 increased the regeneration efficiency of GIC particles beyond 100% after a few adsorption-regeneration cycles. This means that after 5 cycles, all the adsorbed MO was degraded, leaving the GIC with fully recovered active sites. In addition, the propagation of error was calculated to determine the effects of function by variable uncertainty to provide a more accurate measurement of uncertainty. In this case, the uncertainty propagation for regeneration efficiencies was approximately 15.3%. This value indicates that the effect of electrochemical regeneration on active site recovery on GIC particle electrodes was not significantly different in each cycle of adsorption-regeneration, and almost equal proportion or approximately 100% of active sites can be recovered after electrochemical treatment. Secondly, the result also indicated that the effect of electrochemical regeneration on the surface roughening of GIC particle electrodes was minimal as prolonged regeneration can affect the physicochemical properties of particle electrodes, offsetting the recovery of active sites for better regeneration and adsorption efficiencies.

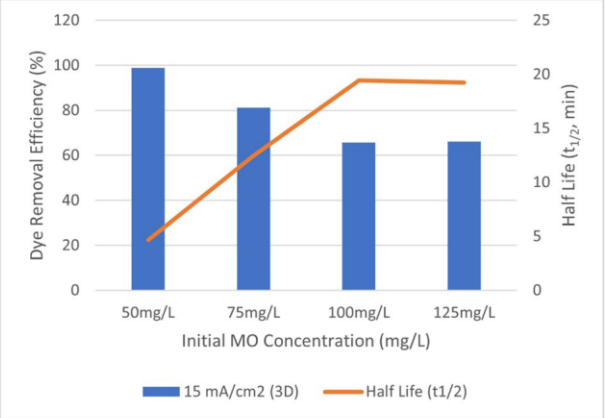


Fig. 3 Effect of 15 mA/cm^2 of current density on the mineralisation rate and MO removal efficiency. The significance of this result is that higher current density is required to completely mineralise large amount of MO pollutants in higher concentrations. The higher half-life of MO pollutants indicates that not all dyes are completely mineralised, leaving them in the aqueous solution

This phenomenon was attributed to surface roughening, which led to changes in surface chemistry or physicochemical properties of GIC (Nkrumah-Amoako et al. 2014). Past researchers showed that the surface area of GIC was expanded during the electrochemical regeneration process (Nkrumah-Amoako et al. 2014). During the electrochemical oxidation process, MO pollutants were adsorbed onto the GIC particle electrodes and oxidized on its electroactive surface into intermediate byproducts. The electrogenerated hydroxyl radicals from the water-splitting process and active chlorine species formed during the electrolytic process helped degrade the MO pollutants through indirect oxidation. In contrast, the direct oxidation of MO pollutants occurred on the surface of anodic material (Martínez-Huitle and Ferro 2006). Hydroxyl radicals formed on the surface of the anodic material by physisorption were released into the bulk liquid media to degrade the MO pollutants.

Notwithstanding the effect of physisorption, Fig. 4 shows that the regeneration efficiency of GIC adsorbent also played a significant role in recovering the surface-active sites. The regeneration efficiency was influenced by surface roughening of the GIC particle electrodes, resulting in changes in physicochemical properties. Figure 5 shows that MO and TOC concentrations decreased significantly when the current density increased from 15 to 35 mA/cm2. However, when 15 to 35 mA/cm^2 of current densities were applied to treat the initial MO concentration of 50 mg/L, this significantly decreased MO and TOC concentrations. The competitive reaction of·OH oxidants and active chlorine species with MO pollutants affected the amount of highly reactive oxidizing species available to mineralise the organic pollutants completely.

The electrochemical oxidation mechanism of organic pollutants via highly reactive hydroxyl radicals using anode is the following:

$$M(s) + H_2O(l) \rightarrow M(\cdot OH) + H^+(aq) + e^- \quad (19)$$

$$M(\cdot OH) \rightarrow MO(s) + H^+(aq) + e^- \quad (20)$$

$$MO(s) + P \rightarrow M(s) + PO \quad (21)$$

where P denotes pollutants; M denotes metal oxide electrodes.

The hydroxyl radical is one of the most potent oxidants in an aqueous solution with $E^0 = 2.73$ V/SHE, which can be electrogenerated on the surface of the electrode (Serrano

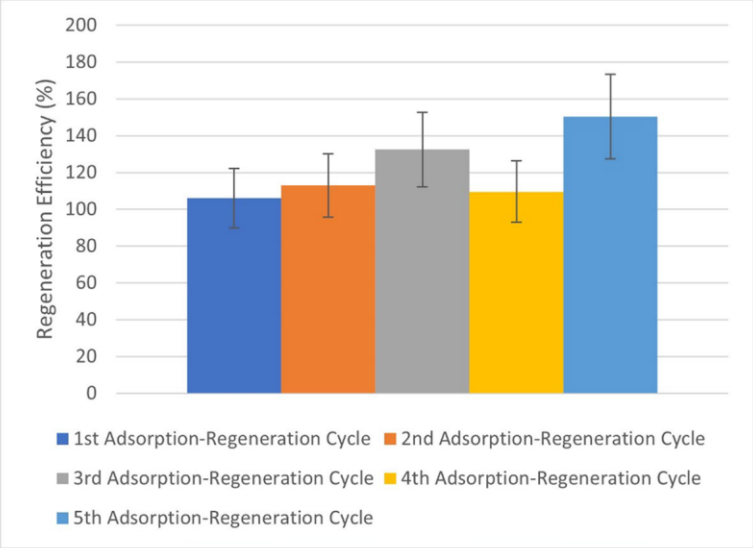


Fig. 4 Effect of 15 mA/cm^2 of current density on the regeneration efficiency of GIC particle electrodes. The significance of this result is that electrolysis leads to changes in GIC physicochemical properties, causing surface roughening and surface area recovery or availability for more adsorption due to high regeneration efficiency, thereby improving the uptake of MO pollutants

2021). It is desirable to have a weak interaction between the radical and electrode surface to make reactivity with the nearby pollutant species possible. The physisorption process depends on the strength of the interaction of hydroxyl radicals with the electrode surface. Attractive electrostatic forces mainly involve van der Waals' forces, which are more vulnerable than a covalent bond. Although the radical species is highly reactive, it has a half-life of approximately 10 ns (Serrano 2021). Hydroxyl radicals can be either physisorbed or chemisorbed onto the electrode. If the chemisorption is predominantly strong, it will hinder the mass transfer of hydroxyl radicals into the bulk solution, reducing the oxidation potential of the electrochemical system.

Active chlorine species are often present with hydroxyl radicals, especially in an electrochemical system that uses NaCl as an electrolyte species. H^+ ions lead to increased acidity of treated wastewater, but it positively affects sustaining hydroxyl radicals and active chlorine species. On the other hand, high current density exacerbates the side reactions, resulting in reduced treatment efficiency.

In addition, the cathodic half-reaction for active chlorine species and water electrolysis for an electrochemical reaction is as follows:

$$2H_2O\,(l) + 4e^- \rightarrow 2OH^-(l) + H_2(g) \tag{22}$$

$$^-OCl(l) + H_2O(l) + 2e^- \rightarrow Cl^-(l) + 2OH^-(l) \tag{23}$$

The presence of chloride and hydroxide species increases alkalinity in the catholyte solution.

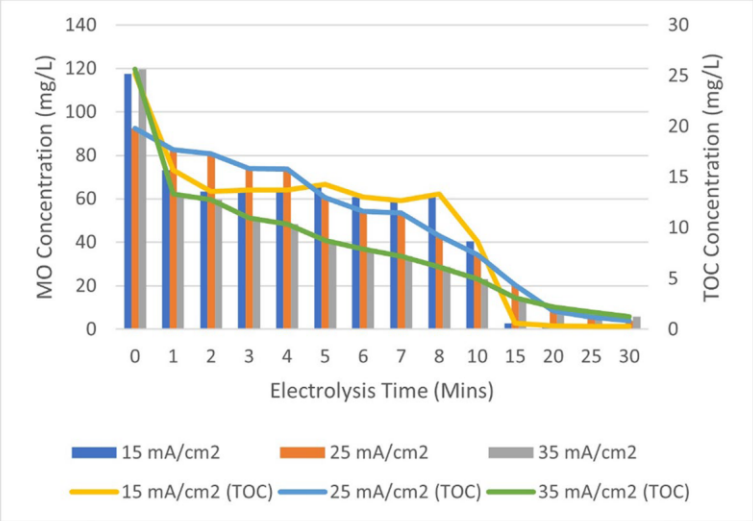


Fig. 5 Effect of changing current densities and electrolysis time on different MO and TOC concentrations. The significance of this result is that the greater the current density, the greater the mineralisation efficiency, resulting in a significant decrease in MO and TOC concentrations over electrolysis time

During electrochemical process, assuming that the nitrogen and sulfur atoms in MO are converted into nitrate and sulfate, the complete electrochemical oxidation reaction of MO is given by the equation as follows:

$$C_{14}H_{14}N_3NaO_3S(s) + 38H_2O(1) \rightarrow 14CO_2(g) + 90H^+(aq) + Na^+(aq) + SO_4^{2-}(aq) + 3NO_3^-(aq) + 90e^- \tag{24}$$

For incomplete electrochemical oxidation reaction of MO due to the influence of side reactions, the equation is as follows:

$$C_{14}H_{14}N_3NaO_3S(s) + 11H_2O(1) \rightarrow 14CO(g) + 36H^+(aq) + Na^+(aq) + S^{2-}(aq) + \frac{3}{2}N_2(aq) + 36e^- \tag{25}$$

Judging from Eqs. (24) and (25), both complete and incomplete oxidation reactions influence the MO and TOC removal efficiencies. Figure 6 shows the differences between the effects of complete and incomplete oxidation reactions on current efficiency of 3D electrochemical process. Complete oxidation reaction of MO byOH oxidants resulted in higher current efficiency than incomplete oxidation reaction. This phenomenon was caused by greater utilization efficiency of current to generate powerful oxidants such as hydroxyl radicals and active chlorine species to degrade MO pollutants in aqueous solutions. However, when the current density was increased from 15 to 35 mA/cm^2, the current efficiency decreased significantly for all initial dye concentrations. The result indicated that the formation of side reactions produced a significant amount of intermediate transformation byproducts, which offset the current efficiency.

In electrochemical process, electrical energy consumption is a critically important parameter. Electrical conductivity of the MO solution and GIC particle electrode directly influenced the energy consumption of the 3D electrochemical reactor. Therefore, enhancing electrical conductivity by integrating the electrochemical reactor with electrically conductive GIC particle

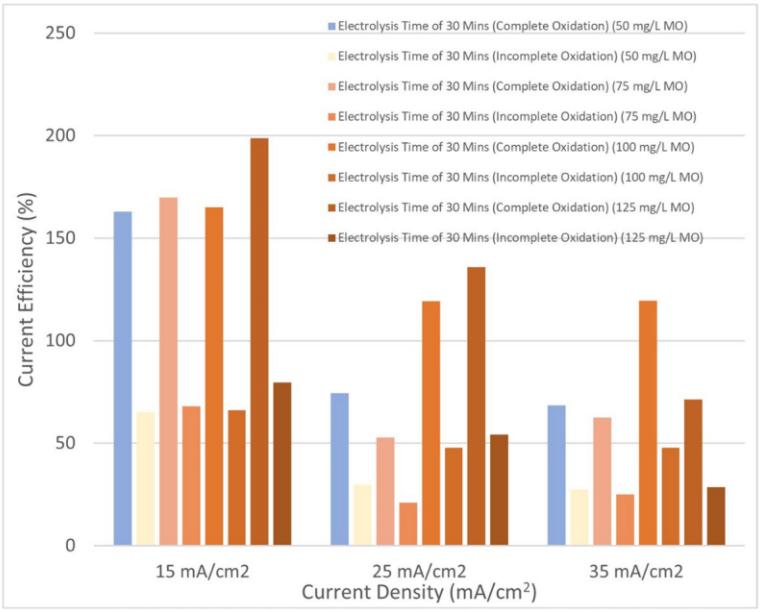


Fig. 6 Effect of different current densities on current efficiency of the 3D electrochemical reactor. The significance of this result is that not all current potentials are utilised efficiently to mineralise the MO pollutants. Some currents were lost through the buildup of side reactions or quenching effect of surrounding media and interference from intermediate transformation byproducts

electrodes can decrease the solution's electrical and mass transfer resistances, leading to better voltage utilization when fixing an electric current. The existence of ions such as nitrate, ammonium and sulphate ions provided electrical conductivity in the solution. Mineralisation of MO pollutants was accompanied by the evolution of NH_4^+, $NO_3^-$and SO_4^{2-}. In addition, the electrogenerated oxidant species may lead to corrosion of the electrodes, inadvertently increasing the electrical resistance. The increase in ohmic resistance of the electrode due to corrosion may result in additional maintenance and repair costs after prolonged electrochemical treatment. Moreover, the results from Fig. 6 showed that the current efficiency significantly impacted the utilisation efficiency of current, directly influencing the amount of energy channelled into the degradation of dye contaminants. The differences between the complete and incomplete oxidation reactions were due to differences in the number of coulombic electrons produced. Complete oxidation reactions were considered ideal reactions with more electrons yielded as presented by Eq. (24). On the other hand, incomplete oxidation reactions involved some loss of electrons due to inefficient reactions and desirable uptake of electrons due to the quenching effect of surrounding media. In addition, Fig. 7a shows that when the current density increased from 15 to 35 mA/cm^2 for all initial MO concentrations, the electrical energy consumption increased from 5 kWh/kg MO to greater than 30 kWh/kg MO. On the other hand, Fig. 7b shows that the electrical energy consumption for TOC removal increased more significantly than the electrooxidation of MO pollutants due to greater electrical energy required to achieve complete mineralisation efficiency. In addition, the values of electrical energy consumption of TOC were more critical and reflective of the actual breakdown of dye contaminants into CO_2 and H_2O, representing the complete reduction of dye contaminants to prevent it from forming aromatic amines, which could be more ecotoxic than its original organic compound.

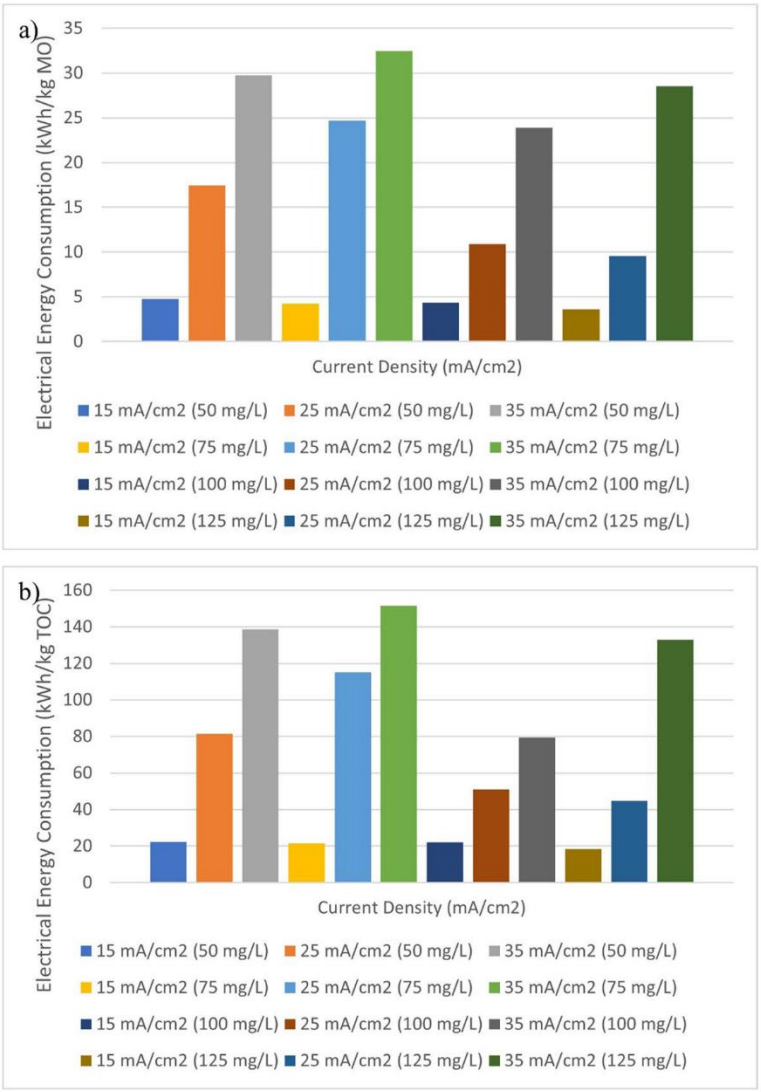


Fig. 7 a) Effect of different current densities and initial MO concentrations on electrical energy consumption (kWh/kg MO) of 3D electrochemical reactor; b) Effect of different current densities and initial MO concentrations on electrical energy consumption (kWh/kg TOC) of 3D electrochemical reactor. This result shows that higher electrical energy is required to completely mineralise the MO pollutants compared to the lower electrical energy needed to break down or convert the MO pollutants into intermediate transformation byproducts through different oxidation pathways. Side reactions may influence the amount of electrical energy consumption

3.1.1 The Prediction Efficiency of Multi-regression Analysis, ANN and SVM models

To assess the prediction efficiency of ANN model in relation to multiple regression analysis and SVM models, 14 experimental runs were conducted for each set of current density, initial MO concentration and electrolysis time. The ANN prediction results for the experimental and predicted removal efficiencies demonstrated that the models

yielded a promising result, with experimental values remarkably close to the predicted data as shown in Fig. 8a-d. Similarly, the prediction results for electrical energy consumption of MO and TOC and current efficiency showed high R^2 values between the experimental and predicted values, highlighting the robustness of ANN optimization power to provide accurate predictions. Figures 8a-d showed different training, validation, and testing proportions, and all data were randomly segregated and imported into the ANN model. The efficiency of MSE calculation depended on the number of neurons applied in the hidden layer so that the statistical metric could be evaluated. The statistical analyses were based on the parameterized hypotheses between the experimental and AI-generated data. The value of MSE trained network was 22.44, along with the correlation coefficient ($R^2 = 0.992$), as shown in Fig. 8a and e. The degree of curve fitting and its relationship between experimental and predicted responses were demonstrated by R^2. The R^2 values obtained for the training, validation and testing were 0.992, 0.965 and 0.845, respectively as shown in Fig. 8a. The R^2-value close to 1 indicates a satisfactory relationship between outputs and target values. The linear fitting model attained plotting regression outputs of ANN, which were given in Fig. 8a-d. The plot of validation outputs and targets created the model (output = 0.72*Target + 15) in Fig. 8a. Figures 8a-d show a good correlation between the experimental and theoretical results obtained using the training function. Furthermore, the ANN topology was examined by varying the number of layers and neurons at the hidden layer to yield an optimal solution. In other words, the number of hidden layers was determined by trial-and-error methodology. Statistical metrics were used as evaluation criteria to determine the best optimal result with minimal deviation between the response variables in the experimental and theoretical results. The prediction capability of ANN did not increase with the number of neurons in the hidden layer due to overfitting of data, leading to increased error deviation and variability. In addition, one of the prediction results of MO removal efficiency showed that the R^2 values for training, validation, testing and all data were 0.992, 0.965, 0.845 and 0.909 in Fig. 8a-d. These results indicated remarkable compatibility between the experimental and predicted results using the ANN model. Furthermore, Levenberg Marquardt Post-Diffusion Algorithm (LMPA) was used to train the network. The performance plot of the trained network is shown in Fig. 8e-h, which showed that the training stopped at 0.0713 at epoch 100 in Fig. 8f, which was close to the acceptable range. In general, the function estimation with network parameters less than 100, the LMP tends to show higher efficiency and speed of calculation. However, high accuracy is still significantly prominent in the majority of cases. The benefit of using this algorithm is due to minimal error. During the data training, the output predicted by the model was comparatively better than the expected value, which can be observed when the MSE values are calculated.

During the first phase, the error training decreased until the network approached a minimal error, and by supplying more data, the error increased again. The network training was halted at this stage, and weights were returned to the minimum error. In addition, Fig. 8e to h showed the statistical significance and error distribution (MSE) of MO removal efficiency, electrical energy consumption of MO and TOC, and current efficiency, predicted by ANN model. The MSE values were significantly low coupled with high R^2 values determined the goodness of measured and predicted results. Although linear relationships between the experimental and predicted results showed a good fit, it provided limited information on the model prediction efficiency due to the absence of non-linear multiple regression analysis. Furthermore, Fig. 8f and h showed that the MSE values were the lowest compared to other statistical metrics in the number

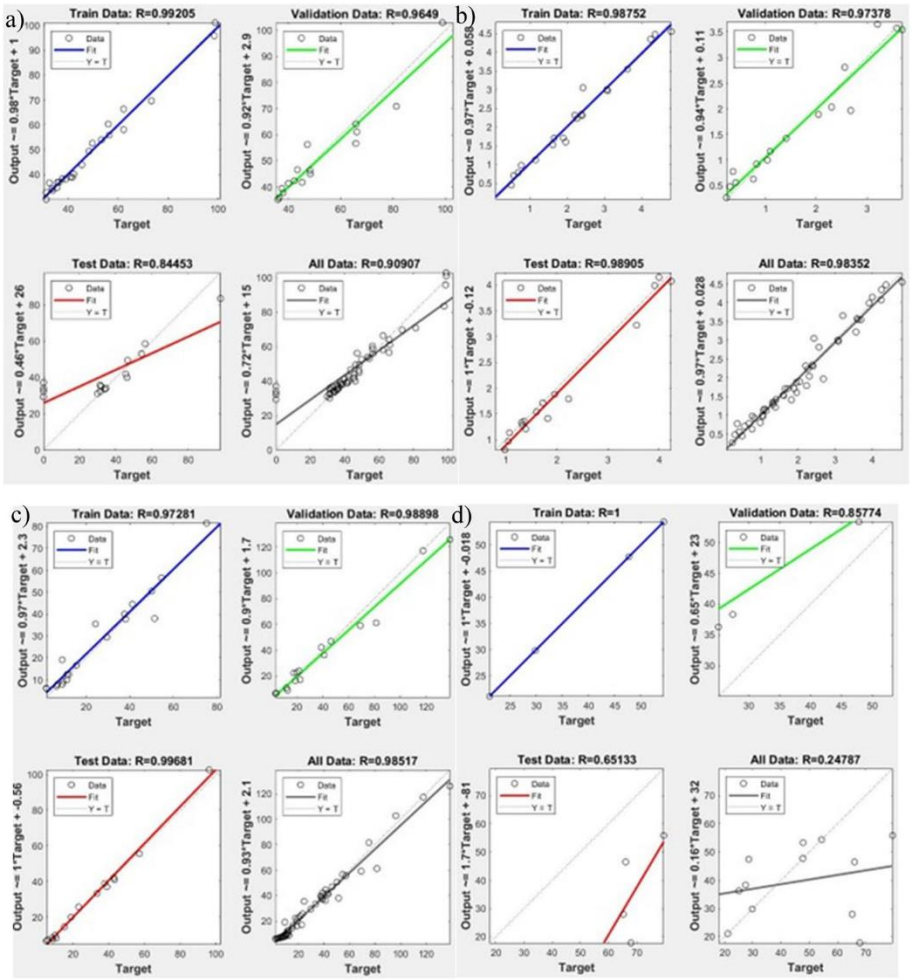


Fig. 8 The performance of ANN models with topology for training, validation, test and all data for **a**) MO removal efficiency; **b**) electrical energy consumption (kWh/kg MO); **c**) electrical energy consumption (kWh/kg TOC); **d**) current efficiency; **e**) mean square error of validation performance for MO removal efficiency; **f**) mean square error of validation performance for electrical energy consumption (kWh/kg MO); **g**) mean square error of validation performance for electrical energy consumption (kWh/kg TOC); **h**) mean square error of validation performance for current efficiency; **i**) SVM prediction efficiency between the experimental and predicted data for MO removal efficiency; **j**) SVM prediction efficiency between the experimental and predicted data for electrical energy consumption of MO; **k**) SVM prediction efficiency between the experimental and predicted data for electrical energy consumption of TOC; **l**) ANN prediction efficiency between the experimental and predicted data for MO removal efficiency; **m**) RF prediction efficiency between the experimental and predicted data for MO removal efficiency; **n**) ANOVA analysis prediction efficiency between the experimental and predicted data for MO removal efficiency. The significance of the result is that when the ANN algorithm yields robust prediction efficiency of response variables such as MO and TOC removal efficiencies that can be improved significantly, rectifying the imprecise or complicated data to derive and extract patterns by controlling the operating parameters to minimise system perturbations or errors

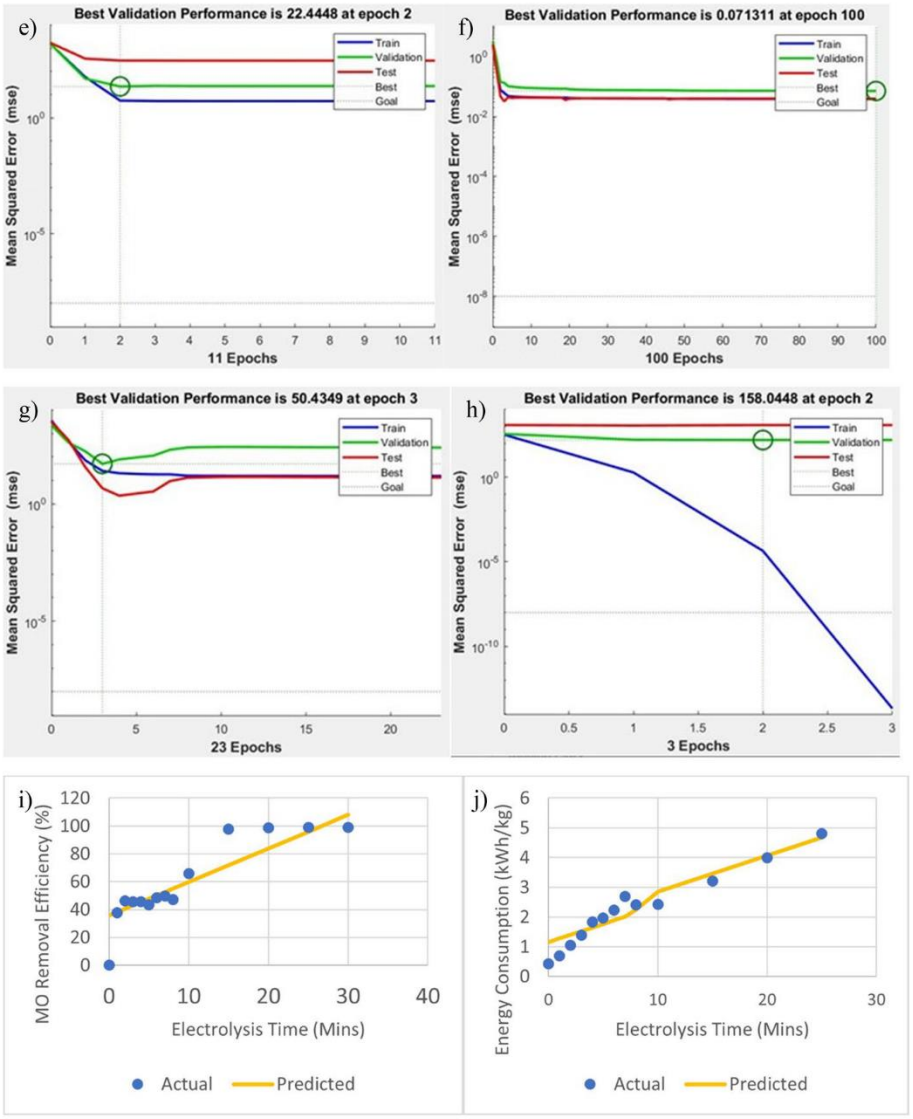


Fig. 8 (continued)

of neurons contained within the hidden layer. The relationship between the experimental and predicted data can be evaluated using the correlation coefficient. On the other hand, Fig. 8i-k showed the equivalent prediction results between the experimental and predicted values, indicating that SVM algorithm can be used to strengthen the optimization power of ANN model. In other words, the SVM model yielded one of the best fitness between the experimental and predicted values.

The nature of surrounding media and quenching effect of ions on hydroxyl radicals and active chlorine species within the bulk solution influenced the synergistic adsorption

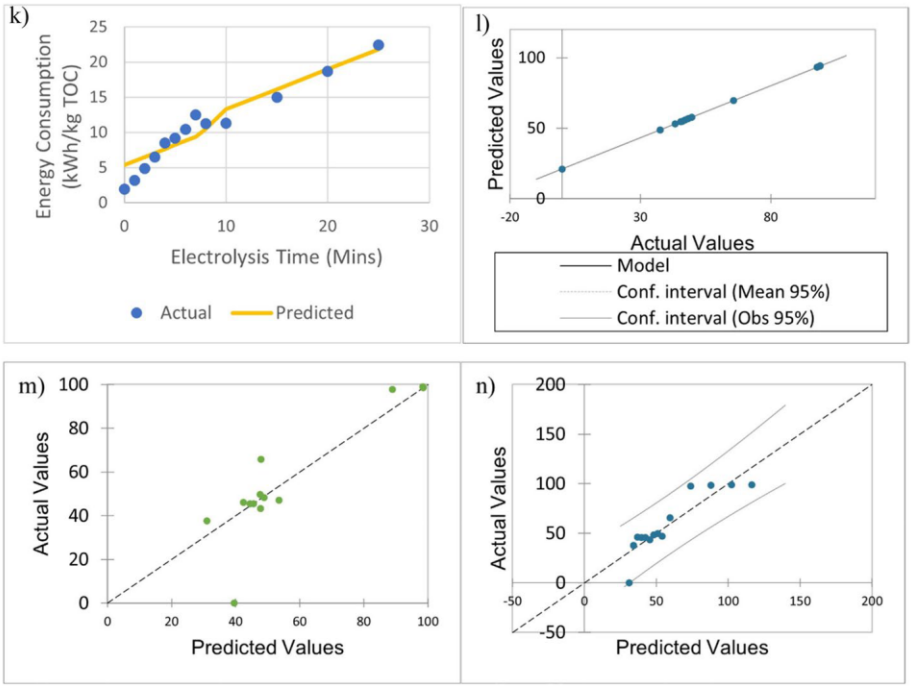


Fig. 8 (continued)

and electrochemical oxidation of MO pollutants in wastewater. Some minor offset in the response variables predicted by the ANN models and experimental data stemmed from side reactions and interference from the immediate transformation byproducts from different oxidation pathways, resulting in slightly reduced correlation coefficients. In addition, the electrolytic effect of anode and cathode on the surface physicochemical properties of the GIC adsorbent can impact the adsorption capacity, increasing the surface area availability for further uptake of MO pollutants from the bulk solution. The surface functionalisation of GIC adsorbent also played a critical role in imparting electrostatic attraction between the MO pollutants and adsorbents. Moreover, the normalisation of experimental and ANN-predicted data shown in Fig. 8a-d indicated that the trained network was applied throughout the dataset, evidencing no misleading interpretation of the results. The minor deviation between the experimental and ANN-predicted data was partly due to experimental variability. Still, the entire experimental dataset yielded a high correlation coefficient with small MSE values, indicating that the ANN optimisation technique was efficient.

Table 1 lists the MSE/RMSE and R^2 obtained from Fig. 8 and compares them with values from the literature. It can be observed that MSE/RMSE and R^2 values for the response variable MO removal efficiency were approximately similar to the best AI optimization results achieved by other researchers, albeit experimenting on different pollutants. This result shows that AI optimization techniques can be applied to the electrochemical treatment of dye wastewater, which was not previously shown. More interestingly, the ANN-optimised results were similar to the values in Table 1, especially when compared with the conventional wastewater treatment plants. Minor differences were attributed to the side reactions and buildup of intermediate transformation byproducts from different oxidation pathways,

Table 1 Comparison of ANN versus SVM models using different wastewater treatment technologies

Types of wastewater treatment system	Types of AI optimization technique	MSE/RMSE	Correlation coefficient (R^2)	Reference
Methyl orange-contaminated wastewater	ANN, SVM	~1.0 (ANN); 172.74, (MSE, SVM); 149.015 (MSE, RF)	0.909 (ANN); 0.898 (SVM); 0.907 (RF)	This study
Trihalomethanes (THMs)	ANN, SVM	0.09 (ANN); 0.70 (SVM)	0.998 (ANN); 0.998 (SVM)	(Safeer et al. 2022)
Biological wastewater treatment systems (WWTs)	ANN, SVM, MLR, fuzzy logic (FL), random forest (RF). Long short-term memory (LSTM)	1.10 (ANN, training); 0.9912 (ANN, validation)	0.9998 (ANN)	(Khan et al. 2022)
Conventional wastewater treatment plants	ANN, Bootstrap methods, principal component analysis (PCA)	0.94 (ANN, training); 3.20 (SVR, training); 0.86 (ANN, testing); 3.20 (SVR, testing)	0.993 (ANN, training); 0.948 (ANN, training); 0.993 (ANN, testing); 0.949 (SVR, testing)	(Chawishborworn-worng et al. 2023)
Detergent industrial wastewater treatment	ANN, SVM	0.238 (MSE, MLP); 0.488 (RMSE, MLP)	0.999 (MLP, training)	(Jana et al. 2022)
WWTPs (activated sludge process)	FFNN (ANN), WOA, IFFNN, GWO, IFFNN-LSSVM	0.2425 (MSE, FFNN); 0.15 (RMSE, FFNN); 0.3120 (MSE, WOA); 0.3120 (RMSE, WOA)	0.9749 (FFNN); 0.9502 (WOA); 0.9685 (IFFNN); 0.9758 (GWO); 0.9921 (IFFNN-LSSVM)	(Zhu et al. 2022)

ANN = Artificial neural network; *FFNN* = Feed-forward neural network; *GWO* = Grey wolf optimization; *IFFNN* = Enhanced feed-forward neural network; *IFFNN-LSSVM* = Upgraded feed-forward neural network with least square support vector machine; *MLP* = Feed-forward backpropagation; *MSE* = Mean square error; *RMSE* = Root mean square error; *SVM* = Support vector machine; *WOA* = Whale optimization algorithm

affecting the pollutant removal efficiency. The discrepancies in results are attributed to the type of pollutant under treatment, unit operations, and other laboratory parameters.

To compare and validate ANN, the Random Forest optimization technique using the CART algorithm was used to evaluate this work. The following Fig. 9 presents the general process of random forest by CART algorithm (Wang et al. 2022):

Figure 9 shows the random forest process and computational procedure for generating the regression trees or optimal tree diagrams. Figure SM1a shows that the optimal tree diagrams using random forest can be used to analyse the energy efficiency of three-dimensional electrochemical process by finding the optimum current density for electrolysing the MO textile wastewater. The optimal tree diagram in Figure SM1a in Supplementary File shows that when the current density dropped below 20 mA/cm^2, the predictive analytics showed terminal node 1 with percentages around 7.1% for a range of calculations for electrical energy consumption of MO. The optimization results indicated that any current density below 20 mA/cm^2 can achieve better energy efficiency than higher current density. When the current density was between 20 and 30 mA/cm^2 at terminal node 2, the prediction results indicated that the electrical energy consumption of MO was higher than terminal node 1, indicating lower energy efficiency when the current density increased beyond 20 mA/cm^2. However, when the current density increased beyond 30 mA/cm^2, the energy efficiency decreased more significantly. The patterns of electrical energy consumption of TOC for Figure SM1b in Supplementary File were similar to Figure SM1a except that the amounts of energy consumption of TOC were higher than the typical energy consumption of MO due to greater electrical energy required to mineralize the dye contaminants in aqueous solutions. The efficiency analysis tree approach can optimize or monitor the energy usage in the electrochemical process within the WWTPs to substantially benefit people and the environment, reducing operational costs and greenhouse gas emissions significantly (Maziotis and Molinos-Senante 2023; Maziotis et al. 2023).

In conjunction with ANN and SVM models, multiple regression analyses in Fig. 10a-d showed variations between the fitness of experimental and predicted values. In addition, multi-regression analysis results presented in Fig. 10b shows a small residual error between the experimental and predicted values. The result indicates a slight variation between the experimental and predicted values when determining the adequate current efficiency required to facilitate oxidation reactions. The benefits of optimisation using multi-regression analysis are due to more controllability over the process parameters while maintaining the energy efficiency of the oxidation reactions. The results from multiple regression analysis in Fig. 10d showed minimal residuals between the experimental and predicted results, indicating that the prediction efficiency was a good fit for optimising the process parameters. On the other hand, the results from Fig. 10c and d show very minimal residuals between the experimental

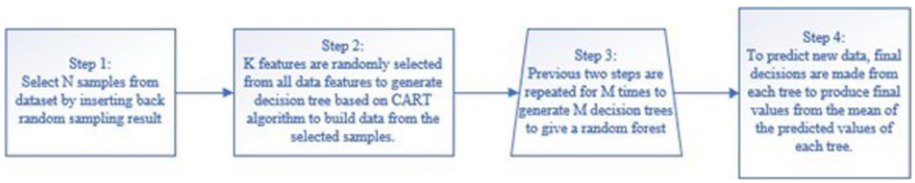


Fig. 9 Random forest process is an ensemble learning method or algorithm for classification and regression by operating a multitude of decision trees at different training times. The step-by-step procedure used to construct the decision trees is stipulated

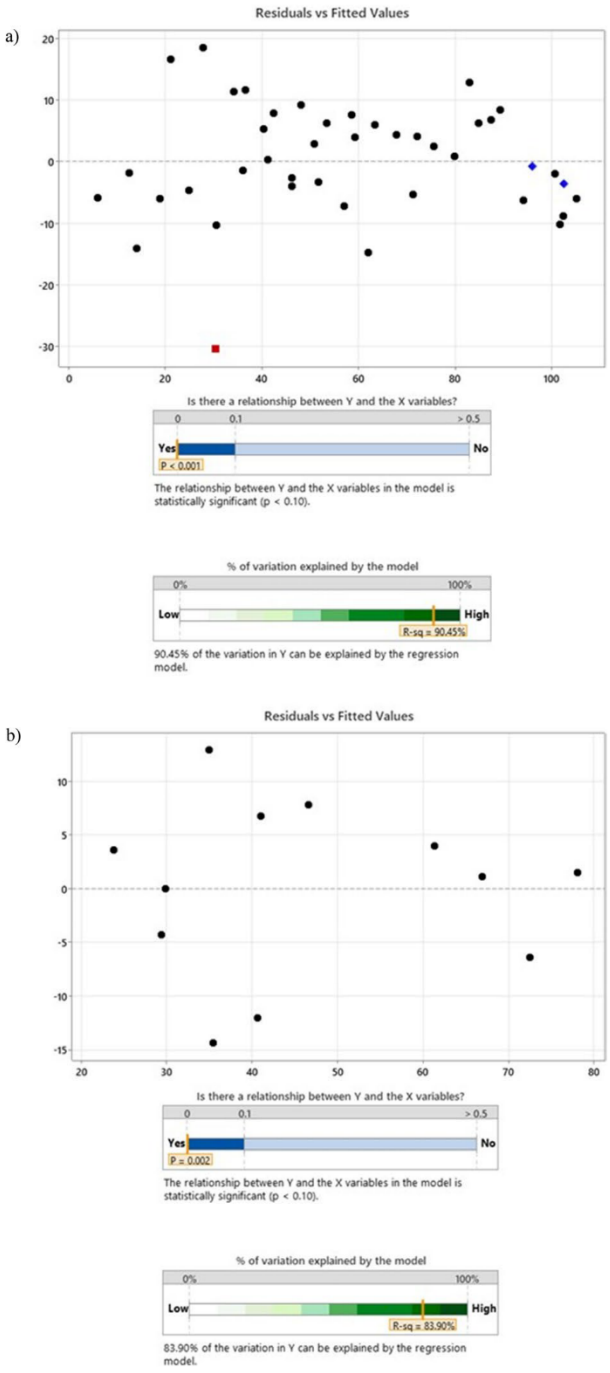
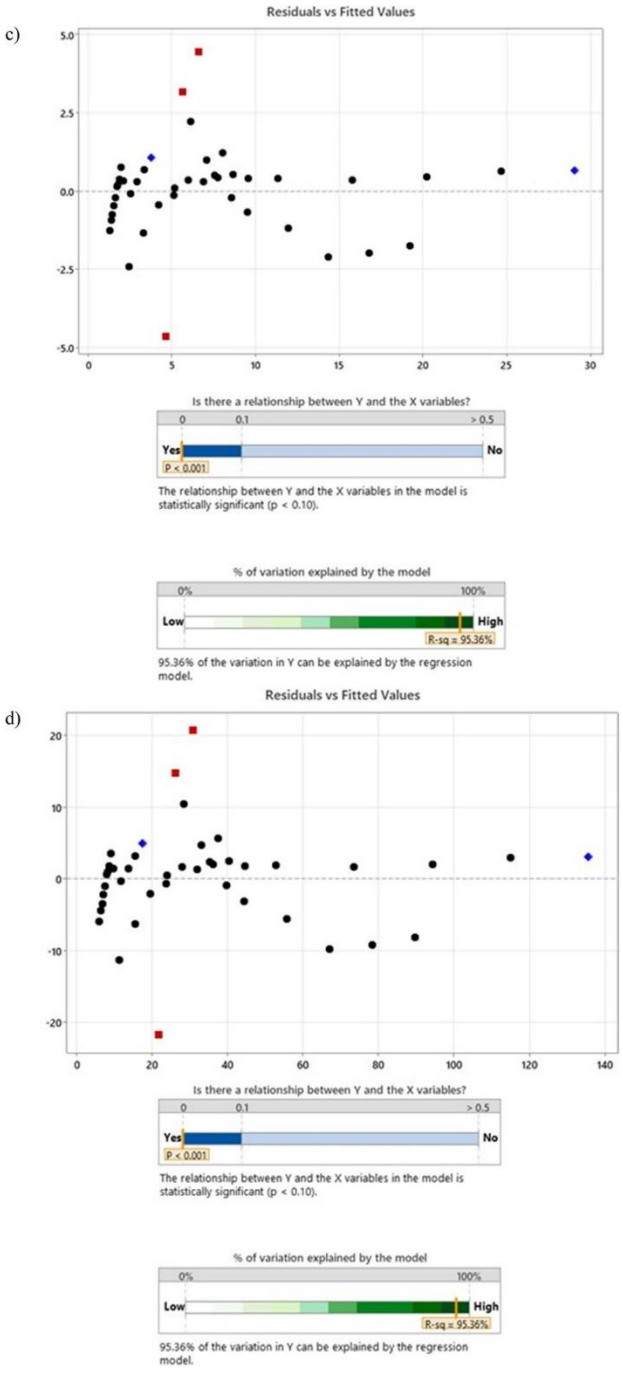


Fig. 10 **a**) Multiple regression analysis for optimisation of MO removal efficiency at 50 mg/L of initial MO concentration; **b**) Multiple regression analysis for optimisation of current efficiency using current densities ranging from 15 to 35 mA/cm^2; **c**) Multiple regression analysis for optimisation of electrical energy consumption of MO using current densities ranging from 15 to 35 mA/cm^2; **d**) Multiple regression analysis for optimisation of electrical energy consumption of TOC using current densities ranging from 15 to 35 mA/cm^2. The significance of using multiple regression analysis is to analyse the relationship between dependent variables and several independent variables to predict the outcome of the dependent values. In this case, multiple regression is used to compare the effects of initial MO concentrations and current densities on MO removal efficiency and electrical energy consumption of MO and TOC, respectively. The predictive outcomes of multiple regression analyses can be compared with the prediction efficiency of AI and machine learning techniques for validation

Fig. 10 (continued)



and predicted values, indicating that the prediction efficiency was a good fit with R^2 values greater than 0.95. However, both ANN and SVM models yielded high R^2 values, highlighting its superior optimisation power over multiple regression functions. In addition, a two-layer feed-forward network imparted with hidden sigmoid neurons and linear output neurons can solve problems for multidimensional mapping to improve curve fitting to match the data. MLP network (3:4:4) was trained with the Levenberg–Marquardt backpropagation algorithm (LMBPA). MSE varied in validation samples, and the training automatically stopped or adjusted to improve the generalisation. The network was trained for 4 replications to find the best number of neurons for the hidden layer.

Finally, the multi-regression analysis in Fig. 10c and d shows minor residual errors or variations between the predicted and experimental values, indicating that the optimisation technique can achieve more robust process conditions by adjusting parameters. However, the ANN optimisation method provided the best prediction result over control of process parameters compared to multi-regression analysis. In addition, Table 2 summarises model validation by ANOVA analysis to compare differences between AI/ML optimisation techniques and multiple regression fit to actual versus predicted values for examining the three-dimensional electrochemical treatment of 50 mg/L MO using a current density of 15 mA/cm^2.

3.1.2 Monte Carlo Simulations

The uncertainties associated with different ML predicted models were estimated using Monte Carlo simulations. The uncertainty and variability of input parameters influence the estimation of uncertainty. Compared to actual data, the predicted variables have an inherent uncertainty in estimating response variables. The Monte Carlo simulation was based on

Table 2 Model validation by comparing AI/ML optimisation and ANOVA analysis through multiple linear regression fit to actual versus predicted data

Actual values	MSE	MSE	MSE	MSE
	22.44	172.74	149.015	183.655
	R^2	R^2	R^2	R^2
	0.909	0.898	0.907	0.806
	ANN (Predicted values)	SVM (Predicted values)	Random forest (Predicted values)	ANOVA (Model validation)
0.000	21	35.743	39.571	31.170
37.626	48.84302	38.151	31.038	34.012
46.088	55.10498	40.559	42.382	36.853
45.426	54.61491	42.967	45.616	39.694
45.426	54.61491	45.375	44.503	42.535
43.341	53.07211	47.783	47.750	45.376
48.320	56.75668	50.191	48.805	48.218
49.644	57.73682	52.599	47.625	51.059
47.143	55.88545	55.007	53.584	53.900
65.759	69.66176	59.823	47.895	59.582
97.694	93.29384	71.862	88.889	73.788
98.577	93.94727	83.902	98.546	87.994
98.896	94.18322	95.942	98.395	102.200

the repeated random sampling (n = 1,000 simulated samples) of the probability distributions defined for principal response variables of certain variation and uncertainty of each input parameter. The Monte Carlo approach allows the approximate estimation of variation and uncertainty stemming from system perturbation associated with specific input parameters and incorporating them into the estimates of response variables. In addition, simulations with 1,000 iterations were used to construct the distributions to calculate the level of uncertainties in different predictive model platforms. The simulated parameters can be extended beyond the current number of operating parameters. Uncertainty analyses in wastewater treatment systems compare the reliability of results, which is subject to variability that leads to significant imprecision in the predictive model platforms. The quality of wastewater treatment standards is based on rigorous regulation of water quality criteria to monitor the risk of adverse effects on the receiving bodies. This research aims to apply Monte Carlo simulation to assess the probability of adverse effects of xenobiotic dye wastewater in meeting environmental standards for effluent discharge. The achievable limits for textile dyeing effluent standards can be evaluated based on the simulated models to adhere to water quality standards.

Appropriate selection of certain input distributions to estimate uncertainty and variability between the actual and predicted models from ML optimisation helps facilitate probabilistic analysis of optimized results. A distribution is determined based on how well it represents a certain dataset from the actual experimental results. The best representation of probabilistic distributions can be empirical or take any form of parametric distributions such as normal, logarithmic normal, uniform, triangular etc. All parameters in this study were assumed to be normal or logarithmic normal distributions. When a certain number of random variables influences the dataset, the result tends to form a normal distribution as shown in Figure SM2 in Supplementary File. A theoretical criterion for selecting a certain normal distribution is based on a central limit theorem (CLT). In addition, Figure SM2 shows the probabilistic density distributions of actual and optimized models. The uniformity of probabilistic distributions and then lack of skewness or heavy-tailedness highlight the prediction efficiency of AI/ML optimized models with limited uncertainty or variability. However, the probability distribution for different artificial intelligence and machine learning-based models showed that the higher the efficiency of targeted responses, the greater the uncertainty, which impacted the accuracy and precision of predictive models. The random forest algorithm generated greater uncertainty than other artificial neural network and support vector machine algorithms, indicating greater instability of system perturbation predicted by random forest. The simulated normal distribution seen in all artificial intelligence and machine learning-based models showed that most of the targeted responses achieved efficiencies within the 90% and 100% range. The strong correlation between the current density and targeted responses based on the sensitivity analysis indicated that current density had the most significant effect on the pollutant removal efficiency. When the predicted models were combined, especially between ANN and SVM, the level of uncertainties or system perturbation increased, leading to greater variability of combined predictive models. Similarly, if the ANN and RF models were integrated into a single predictive model platform, the level of uncertainties in prediction efficiency was less than the ANN-SVM model.

4 Conclusions

The results above confirmed that 3D electrochemical treatment integrated with graphite intercalation compound (GIC) particle electrodes and anodic oxidation technology is a very efficient technique to degrade methyl orange (MO) pollutants in simulated wastewater, achieving

greater than 98% removal rate within 30 min of electrolysis time. The GIC particle electrodes in the 3D electrochemical process act as an electrocatalytic adsorbent material to effectively improve mineralisation efficiency and generation of·OH oxidants, demonstrating the effectiveness of combined adsorption and electrochemical oxidation. However, the strength of electrolysis in this experiment was limited by the type of electrocatalytic material used, and the acidified salt concentration was also limited, resulting in slightly reduced electrical conductivity of solution, and less active chlorine species available to degrade MO pollutants. Nonetheless, the research results justify the application potential of green and efficient 3D electrochemical treatment of complex industrial wastewater. The synergistic effect of 3D electrochemical process resulted in high MO removal and current efficiencies, reducing overall electrical energy consumption. In addition, GIC particle electrodes consistently maintained high regeneration efficiency beyond 100% throughout several consecutive cycles of adsorption and regeneration, highlighting the potential for reusability of particle electrodes. The AI optimisation power of multi-regression analysis, ANN, SVM and random forest ranked in the following order: ANN > RF > SVM > multiple regression analysis. The probabilistic distributions and scatterplots from Monte Carlo simulations indicated limited uncertainty and variability between actual and optimised models, highlighting the prediction efficiency of AI/ML optimisation approaches that are potentially applicable to water resources engineering and wastewater remediation in WWTP. Most interestingly, the overall critical findings of the research showed that RF is intrinsically suited for analysing multiclass problems, while SVM is only suited for two-class problems. In this research, the predictive performance of RF versus SVM was approximately comparable due to almost equal uncertainties. In contrast, the ANN algorithm yielded significantly better prediction efficiency than the other two algorithms with fewer uncertainties. Although RF is considered robust to overfitting and excellent in handling extensive nonlinear data, SVM can effectively operate at high dimensional spaces or hyperplanes and is versatile in handling multiple data types. The predictive performance of these algorithms is primarily influenced by the sample size, the complex nature of the dataset and the type of problem being addressed. The subsequent studies should focus on evaluating other equally robust classifiers for optimising the electricity costs from industrial operation and greenhouse gas emissions of WWTP to identify the potential gap between pollutant generation and discharge sources to improve the efficacy and broaden the applicability of optimised models.

Author contributions **Voravich Ganthavee:** Conceptualization, Visualization, Methodology, Validation, Investigation, Formal Analysis, Data Curation, Writing-original draft; **Antoine P. Trzcinski:** Supervision, Review and Editing, Project Administration, Resources; **Merenghege M R Fernando:** Investigation, Data Curation, Visualization.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. Non-applicable.

Data Availability Additional data will be provided upon request.

Declarations

Ethical Approval Non-applicable.

Consent to Participate Non-applicable.

Consent to Publish Non-applicable.

Competing Interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

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

Badmus KO, Irakoze N, Adeniyi OR, Petrik L (2020) Synergistic advance Fenton oxidation and hydrodynamic cavitation treatment of persistent organic dyes in textile wastewater. J Environ Chem Eng 8:103521. https://doi.org/10.1016/j.jece.2019.103521

Can-Güven E (2021) Advanced treatment of dye manufacturing wastewater by electrocoagulation and electro-Fenton processes: effect on COD fractions, energy consumption, and sludge analysis. J Environ Manage 300:113784. https://doi.org/10.1016/j.jenvman.2021.113784

Chairungsri W, Subkomkaew A, Kijjanapanich P, Chimupala Y (2022) Direct dye wastewater photocatalysis using immobilized titanium dioxide on fixed substrate. Chemosphere 286:131762. https://doi.org/10.1016/j.chemosphere.2021.131762

Chawishborwornworng C, Luanwuthi S, Umpuch C, Puchongkawarin C (2023) Bootstrap approach for quantifying the uncertainty in modeling of the water quality index using principal component analysis and artificial intelligence. J Saudi Soc Agric Sci. https://doi.org/10.1016/j.jssas.2023.08.004

Chen D, Zhao L, Chen D, Hou P, Liu J, Wang C, Aborisade MA, Yin M, Yang Y (2023) Fabrication of a SnO_2–Sb electrode with TiO_2 nanotube array as the middle layer for efficient electrochemical oxidation of amaranth dye. Chemosphere 325:138380. https://doi.org/10.1016/j.chemosphere.2023.138380

Cui MH, Liu WZ, Tang ZE, Cui D (2021) Recent advancements in azo dye decolorization in bio-electrochemical systems (BESs): Insights into decolorization mechanism and practical application. Water Res 203:117512. https://doi.org/10.1016/j.watres.2021.117512

Ding C, Xia Y, Yuan Z, Yang H, Fu J, Chen Z (2023) Performance prediction for a fuel cell air compressor based on the combination of backpropagation neural network optimized by genetic algorithm (GA-BP) and support vector machine (SVM) algorithms. Therm Sci Eng Prog 44:102070. https://doi.org/10.1016/j.tsep.2023.102070

El Aggadi S, Kaichouh, G, El Abbassi, Z, Fekhaoui M, El Hourch A (2021) Electrode material in electrochemical decolorization of dyestuffs wastewater: a review. E3S Web of Conferences ICIES, vol 234. p. 00058. https://doi.org/10.1051/e3sconf/202123400058. Accessed 5 July 2024

El-Kammah M, Elkhatib E, Gouveia S, Cameselle C, Aboukila E (2022) Enhanced removal of Indigo Carmine dye from textile effluent using green cost-efficient nanomaterial: adsorption, kinetics, thermodynamics and mechanisms. Sustain Chem Pharm 29:100753. https://doi.org/10.1016/j.scp.2022.100753

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

Hamida M, Dehane A, Merouani S, Hamdaoui O, Ashokkumar M (2022) The role of reactive chlorine species and hydroxyl radical in the ultrafast removal of Safranin O from wastewater by CCl_4/ultrasound sono-process. Chem Eng Process 178:109014. https://doi.org/10.1016/j.cep.2022.109014

Hossain L, Sarker SK, Khan MS (2018) Evaluation of present and future wastewater impacts of textile dyeing industries in Bangladesh. Environ Dev 26:23–33. https://doi.org/10.1016/j.envdev.2018.03.005

Ihaddaden S, Aberkane D, Boukerroui A, Robert D (2022) Removal of methylene blue (basic dye) by coagulation-flocculation with biomaterials (bentonite and Opuntia ficus indica). J Water Process Eng 49:102952. https://doi.org/10.1016/j.jwpe.2022.102952

Jana DK, Bhunia P, Das Adhikary S, Bej B (2022) Optimization of effluents using artificial neural network and support Vector regression in detergent industrial wastewater treatment. Clean Chem Eng 3:100039. https://doi.org/10.1016/j.clce.2022.100039

Januário EFD, Vidovix TB, Bergamasco R, Vieira AMS (2021) Performance of a hybrid coagulation/flocculation process followed by modified microfiltration membranes for the removal of solophenyl blue dye. Chem Eng Process - Process Intensif 168:108577. https://doi.org/10.1016/j.cep.2021.108577

Kanjal MI, Muneer M, Jamal MA, Bokhari TH, Wahid A, Ullah S, Amrane A, Hadadi A, Tahraoui H, Mouni L (2023) A study of treatment of reactive red 45 dye by advanced oxidation processes and toxicity evaluation using bioassays. Sustainability 15:7256

Kebir M, Benramdhan I-K, Nasrallah N, Tahraoui H, Bait N, Benaissa H, Ameraoui R, Zhang J, Assadi AA, Mouni L, Amrane A (2023) Surface response modeling of homogeneous photo Fenton Fe(III) and Fe(II) complex for sunlight degradation and mineralization of food dye. Catal Commun 183:106780. https://doi.org/10.1016/j.catcom.2023.106780

Khan H, Wahab F, Hussain S, Khan S, Rashid M (2022) Multi-object optimization of navy-blue anodic oxidation via response surface models assisted with statistical and machine learning techniques. Chemosphere (Oxford) 291:132818–132818. https://doi.org/10.1016/j.chemosphere.2021.132818

Kumar D, Gupta SK (2022) Electrochemical oxidation of direct blue 86 dye using MMO coated Ti anode: modelling, kinetics and degradation pathway. Chem Eng Process - Process Intensif 181:109127. https://doi.org/10.1016/j.cep.2022.109127

Lau Y-Y, Wong Y-S, Teng T-T, Morad N, Rafatullah M, Ong S-A (2014) Coagulation-flocculation of azo dye Acid Orange 7 with green refined laterite soil. Chem Eng J 246:383–390. https://doi.org/10.1016/j.cej.2014.02.100

Liu X, Chen Z, Du W, Liu P, Zhang L, Shi F (2022) Treatment of wastewater containing methyl orange dye by fluidized three dimensional electrochemical oxidation process integrated with chemical oxidation and adsorption. J Environ Manage 311:114775–114775. https://doi.org/10.1016/j.jenvman.2022.114775

Martínez-Huitle CA, Ferro S (2006) Electrochemical oxidation of organic pollutants for the wastewater treatment: direct and indirect processes. Chem Soc Rev 35:1324–1340. https://doi.org/10.1039/B517632H

Maziotis A, Molinos-Senante M (2023) A comprehensive eco-efficiency analysis of wastewater treatment plants: estimation of optimal operational costs and greenhouse gas emissions. Water Res 243:120354. https://doi.org/10.1016/j.watres.2023.120354

Maziotis A, Sala-Garrido R, Mocholí-Arce M, Molinos-Senante M (2023) A comprehensive assessment of energy efficiency of wastewater treatment plants: an efficiency analysis tree approach. Sci Total Environ 885:163539. https://doi.org/10.1016/j.scitotenv.2023.163539

Mechati S, Zamouche M, Tahraoui H, Filali O, Mazouz S, Bouledjemer INE, Toumi S, Triki Z, Amrane A, Kebir M, Lefnaoui S, Zhang J (2023) Modeling and optimization of hybrid fenton and ultrasound process for crystal violet degradation using AI techniques. Water 15:4274

Meng G, Fang L, Yin Y, Zhang Z, Li T, Chen P, Liu Y, Zhang L (2022) Intelligent control of the electrochemical nitrate removal basing on artificial neural network (ANN). J Water Process Eng 49:103122. https://doi.org/10.1016/j.jwpe.2022.103122

Narbaitz RM, Karimi-Jashni A (2012) Electrochemical reactivation of granular activated carbon: Impact of reactor configuration. Chem Eng J (Lausanne, Switzerland : 1996) 197:414–423. https://doi.org/10.1016/j.cej.2012.05.049

Narbaitz R, McEwen J (2012) Electrochemical regeneration of field spent GAC from two water treatment plants. Water Res 46:4852–4860. https://doi.org/10.1016/j.watres.2012.05.046

Nidheesh PV, Zhou M, Oturan MA (2018) An overview on the removal of synthetic dyes from water by electrochemical advanced oxidation processes. Chemosphere (Oxford) 197:210–227. https://doi.org/10.1016/j.chemosphere.2017.12.195

Nkrumah-Amoako K, Roberts EPL, Brown NW, Holmes SM (2014) The effects of anodic treatment on the surface chemistry of a graphite intercalation compound. Electrochim Acta 135:568–577. https://doi.org/10.1016/j.electacta.2014.05.063

Oruganti RK, Biji AP, Lanuyanger T, Show PL, Sriariyanun M, Upadhyayula VKK, Gadhamshetty V, Bhattacharyya D (2023) Artificial intelligence and machine learning tools for high-performance microalgal wastewater treatment and algal biorefinery: A critical review. Sci Total Environ 876:162797. https://doi.org/10.1016/j.scitotenv.2023.162797

Özdoğan-Sarıkoç G, Sarıkoç M, Celik M, Dadaser-Celik F (2023) Reservoir volume forecasting using artificial intelligence-based models: artificial neural networks, support vector regression, and long short-term memory. J Hydrol 616:128766. https://doi.org/10.1016/j.jhydrol.2022.128766

Pavlović MD, Buntić AV, Mihajlovski KR, Šiler-Marinković SS, Antonović DG, Radovanović Ž, Dimitrijević-Branković SI (2014) Rapid cationic dye adsorption on polyphenol-extracted coffee grounds—a response surface methodology approach. J Taiwan Inst Chem Eng 45:1691–1699. https://doi.org/10.1016/j.jtice.2013.12.018

Picos-Benítez AR, Martínez-Vargas BL, Duron-Torres SM, Brillas E, Peralta-Hernández JM (2020) The use of artificial intelligence models in the prediction of optimum operational conditions for

the treatment of dye wastewaters with similar structural characteristics. Process Saf Environ Prot 143:36–44. https://doi.org/10.1016/j.psep.2020.06.020
Rodriguez-Galiano V, Sanchez-Castillo M, Chica-Olmo M, Chica-Rivas M (2015) Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. Ore Geol Rev 71:804–818. https://doi.org/10.1016/j.oregeorev.2015.01.001
Safeer S, Pandey RP, Rehman B, Safdar T, Ahmad I, Hasan SW, Ullah A (2022) A review of artificial intelligence in water purification and wastewater treatment: recent advancements. J Water Process Eng 49:102974. https://doi.org/10.1016/j.jwpe.2022.102974
Saravanan A, Deivayanai VC, Kumar PS, Rangasamy G, Hemavathy RV, Harshana T, Gayathri N, Alagumalai K (2022) A detailed review on advanced oxidation process in treatment of wastewater: mechanism, challenges and future outlook. Chemosphere 308:136524. https://doi.org/10.1016/j.chemosphere.2022.136524
Serrano KG (2021) A critical review on the electrochemical production and use of peroxo-compounds. Curr Opin Electrochem 27:100679. https://doi.org/10.1016/j.coelec.2020.100679
Shoukat R, Khan SJ, Jamal Y (2019) Hybrid anaerobic-aerobic biological treatment for real textile wastewater. J Water Process Eng 29:100804. https://doi.org/10.1016/j.jwpe.2019.100804
Singh A, Pal DB, Mohammad A, Alhazmi A, Haque S, Yoon T, Srivastava N, Gupta VK (2022) Biological remediation technologies for dyes and heavy metals in wastewater treatment: new insight. Biores Technol 343:126154. https://doi.org/10.1016/j.biortech.2021.126154
Singh NK, Yadav M, Singh V, Padhiyar H, Kumar V, Bhatia SK, Show P-L (2023) Artificial intelligence and machine learning-based monitoring and design of biological wastewater treatment systems. Biores Technol 369:128486. https://doi.org/10.1016/j.biortech.2022.128486
Suhan MBK, Mahtab SMT, Aziz W, Akter S, Islam MS (2021) Sudan black B dye degradation in aqueous solution by Fenton oxidation process: kinetics and cost analysis. Case Stud Chem Environ Eng 4:100126. https://doi.org/10.1016/j.cscee.2021.100126
Tahraoui H, Belhadj A-E, Triki Z, Boudellal NR, Seder S, Amrane A, Zhang J, Moula N, Tifoura A, Ferhat R, Bousselma A, Mihoubi N (2023) Mixed coagulant-flocculant optimization for pharmaceutical effluent pretreatment using response surface methodology and Gaussian process regression. Process Saf Environ Prot 169:909–927. https://doi.org/10.1016/j.psep.2022.11.045
Trzcinski AP, Harada K (2023) Adsorption of PFOS onto graphite intercalated compound and analysis of degradation by-products during electro-chemical oxidation. Chemosphere 323:138268. https://doi.org/10.1016/j.chemosphere.2023.138268
Uddin MA, Begum MS, Ashraf M, Azad AK, Adhikary AC, Hossain MS (2023) Water and chemical consumption in the textile processing industry of Bangladesh. PLOS Sustain Transform 2:e0000072. https://doi.org/10.1371/journal.pstr.0000072
Vinayagam V, Murugan S, Kumaresan R, Narayanan M, Sillanpää M, Vo DVN, Kushwaha OS, Jenis P, Potdar P, Gadiya S (2022) Sustainable adsorbents for the removal of pharmaceuticals from wastewater: A review. Chemosphere 300:134597. https://doi.org/10.1016/j.chemosphere.2022.134597
Wagner MM, Moore AW, Aryel RM (2006) Handbook of biosurveillance. Academic Press, Amsterdam
Wang R, Yu Y, Chen Y, Pan Z, Li X, Tan Z, Zhang J (2022) Model construction and application for effluent prediction in wastewater treatment plant: Data processing method optimization and process parameters integration. J Environ Manage 302:114020. https://doi.org/10.1016/j.jenvman.2021.114020
Wu L, Li Q, Ma C, Li M, Yu Y (2022) A novel conductive carbon-based forward osmosis membrane for dye wastewater treatment. Chemosphere 308:136367. https://doi.org/10.1016/j.chemosphere.2022.136367
Wu Y, Al-Huqail A, Farhan ZA, Alkhalifah T, Alturise F, Ali HE (2022) Enhanced artificial intelligence for electrochemical sensors in monitoring and removing of azo dyes and food colorant substances. Food Chem Toxicol 169:113398. https://doi.org/10.1016/j.fct.2022.113398
Xie J, Zhang C, Waite TD (2022) Hydroxyl radicals in anodic oxidation systems: generation, identification and quantification. Water Res 217:118425. https://doi.org/10.1016/j.watres.2022.118425
Zhang Z, Shen S, Xu Q, Cui L, Qiu R, Huang Z (2024) Electrochemical oxidation of ammonia in a granular activated carbon/peroxymonosulfate/chlorine three-dimensional electrode system. Sep Purif Technol 342:127038. https://doi.org/10.1016/j.seppur.2024.127038
Zhu J, Jiang Z, Feng L (2022) Improved neural network with least square support vector machine for wastewater treatment process. Chemosphere 308:136116. https://doi.org/10.1016/j.chemosphere.2022.136116

Authors and Affiliations

Voravich Ganthavee[1] · Merenghege M. R. Fernando[1] · Antoine P. Trzcinski[1]

✉ Voravich Ganthavee
Voravich.Ganthavee@unisq.edu.au

[1] School of Agriculture and Environmental Science, Darling Heights, University of Southern Queensland, Toowoomba, QLD 4350, Australia

6.2 Links and implications

The synergistic performance of three-dimensional electrochemical process enhanced the mineralisation efficiency of MO while improving the energy efficiency of the electrochemical reactor. With the support of electrocatalytic adsorbent material, such as GIC particle electrodes, the pollutant removal efficiency was maximised. The optimisation power of multiple regression analysis, AI and ML was ranked in the order: ANN > RF > SVM > Multiple Regression. On the other hand, the probabilistic distributions and scatterplots generated from Monte Carlo simulations showed limited uncertainty and variability between the actual and optimised models, justifying the remarkable prediction efficiency of AI/ML ensembles in optimising the three-dimensional electrochemical process.

Most interestingly, the critical findings demonstrated that RF is more suited at managing overfitting and better at handling extensive nonlinear data, whereas SVM can effectively operate at high dimensional spaces or hyperplanes. When the ANN and RF models were combined into a single predictive model platform, as predicted by the Monte Carlo simulations, the level of system perturbations or uncertainties in prediction efficiency was less than the ANN-SVM model. However, the predictive performance of these algorithms is primarily influenced by the complexity of datasets, nature of datasets, sample size and the type of problem being addressed. The future studies should focus on evaluating other robust classifiers, AI and ML ensembles, to determine its optimisation efficiency in enhancing the energy efficiency of the industrial operation and reducing the greenhouse gas emission of WWTP, especially when the three-dimensional electrochemical reactor is used.

COMPUTATIONAL MODELLING OF INDIGO CARMINE ADSORPTION ONTO BONE CHAR: APPLICATION OF MONTE CARLO SIMULATION, BAYESIAN NETWORKS, ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING-BASED OPTIMISATION APPROACHES

7.1 Introduction

This paper is an extension of three-dimensional electrochemical technology, emphasizing the benefits of using a green, renewable carbon-based adsorbent material, i.e., bone char, to remove xenobiotic dye pollutants from wastewater. This research specifically focused on using computational modelling techniques involving the application of Monte Carlo simulations, a range of combined AI and ML optimization ensembles to predict the targeted variables, such as dye and TOC removal efficiencies by optimizing operational parameters to achieve optimal conditions. This research investigated uniquely designed ensembles to optimise the operational conditions to improve the response variables. Sensitivity analysis was conducted using Monte Carlo simulation to determine the levels of uncertainty in the predictive models. The causal relationships between the operational variables were determined using Bayesian inference network analysis. Computational fluid dynamics modelling was conducted to design a packed bed reactor system to integrate bone char adsorbent material into the reactor to remove xenobiotic dye pollutants in wastewater. The thermal distribution effects, thermodynamic parameters and heat fluxes within the wastewater treatment system were explored.

Abstract

Various ensembles of artificial intelligence (AI)c and machine learning (ML)-based optimization methods have been evaluated to examine the prediction efficiencies of Indigo Carmine (IC) adsorption onto a bone char adsorbent to ensure adequate wastewater remediation quality. In this study, AI and ML ensembles were developed to optimize the model parameters, and Monte Carlo simulation models were used to estimate the uncertainties of different model parameters. Following the simulation-based comparisons, Bayesian predictive network analysis was performed to determine the probability distribution to reveal the causal interrelationships between variables to represent the impact of uncertainties on the simulation models. In particular, the Indigo Carmine (IC) removal efficiency and dye loading capacity were optimized using a range of AI and ML optimisation techniques such as adaptive neuro-fuzzy inference system (ANFIS), support vector machine (SVM), random forest (RF) and eXtreme Boosting Gradient (XGBoost). The prediction efficiencies of combined AI and ML ensembles were evaluated for ANN-GA-SVM, ANN-GA-RF and ANN-GA-XGBoost and ANN-GA optimisations that achieved a coefficient of determination (R^2) of 0.730, 0.445, 0.625 and 0.526, respectively, for IC removal within 20 minutes using bone char (BC). The calculations from the correlation matrix indicated that the combination of AI/ML models, such as ANN-GA-XGBoost and ANN-GA-RF, yielded the most significant impact on prediction accuracy with a positive Spearman correlation of 0.726 and 0.694, respectively, relative to ANN-GA. In contrast, ANFIS and ANN-GA-SVM models generated significantly less uncertainty in prediction efficiency with low mean square error (MSE) of 0.13 and 3.072, respectively, compared to other AI/ML ensembles.

Keywords: Artificial intelligence; machine learning; adaptive neuro-fuzzy inference system; Monte Carlo simulation

Water is an essential element for sustaining life. Increasing water scarcity and pollution are the greatest challenges the world is facing. A specific category of organic

chemicals deemed unsafe for human consumption is synthetic textile dyes used by various industries worldwide, such as cosmetics, paper and pulp, plastics, etc. Approximately 15% of synthetic dyes escape through the industrial effluent system and find its ways into the marine environment. Textile dyes are xenobiotics with strong physicochemical stability, fixation efficiency to the natural substrates and the ability to resist chemical and environmental biodegradation or UV photolysis (Kekes & Tzia, 2020). Xenobiotic dyes are common environmental contaminants in wastewater effluent systems discharged from textile industry. Wastewater-containing dyes are a significant environmental contaminant that affects human health as textile industries generate large amounts of coloured dyeing effluent into the aquatic environment (Al-Tohamy et al., 2022). Among these dyes, Indigo Carmine (Acid Blue 74) found widespread industrial applications. It is a highly toxic chemical substance with a strong dyeing colourant and has various applications in the textile, cosmetics, plastics, pharmaceutical, leather and food manufacturing industries (Adel et al., 2021; El-Kammah et al., 2022). Nowadays, Indigo carmine dye is synthetic, whereas its natural form was widely used in the 17^{th} century due to its antibacterial and insect-repellent properties. Synthetic indigo carmine is often used to dye blue jeans and other household fabrics. However, these indigo fabrics and clothes are usually disposed into landfill. Indigo carmine dye can be harmful if ingested, especially when the dye finds its way into food chain and may have the potential to cause cancer.

The annual production capacity of dyeing wastewater from textile industry is approximately 0.7 million metric tonnes (Bilal et al., 2022). Indigo Carmine (IC) dye is extensively used in textile industries due to its colour intensity, which stems from auxochromes, a chromophore functional group presents in significant amount of synthetic dyes (Shabir et al., 2022). Indigo dyes are highly hazardous and carcinogenic; they can cause skin and eye irritation upon direct contact and may lead to chronic damage to the cornea and

conjunctiva (Guezzen et al., 2023). For this reason, Indigo Carmine is considered a major contaminant that must be treated or removed prior to discharging dye-contaminated wastewater into marine environment (Fu et al., 2019).

Many treatment methods have been developed to remove synthetic dyes from wastewater. These treatment methods produced different removal rates and can be categorized into physical, chemical and biological technologies such as 99% of dyes can be removed using membrane filtration (Ma et al., 2022; Mansor et al., 2020), 92-99% dyes removed by chemical or solvent precipitation (Alikarami et al., 2022), 85-95% by adsorption (Aqdam et al., 2021; Hassani et al., 2015; Obayomi et al., 2023), > 90% of dyes removed by electrochemical degradation process (Salazar et al., 2018; Tang et al., 2020; Tang et al., 2022), 100% by advanced oxidation processes (Dadban-Shahamat et al., 2022; Nidheesh et al., 2022; Saravanan et al., 2022), 98% by Fenton process (Arslan-Alaton et al., 2009; Esteves et al., 2016; Suhan et al., 2021), greater than 50% by ion-exchange (Hassan & Carr, 2018; Lu et al., 2022), 100% by coagulation-flocculation (Januário et al., 2021; Li et al., 2017; McYotto et al., 2021), 49-76% by photocatalysis (Chairungsri et al., 2022; Rosa et al., 2015), 96% by ozonation (El Hassani et al., 2019; Hu et al., 2016) and > 70% by microbiological degradation (Singh et al., 2022; Wang et al., 2019). Although chemical and biological approaches are effective at removing toxic dyes from wastewater, they require special equipment, and the process is energetically intensive and the addition of chemical reagents often lead to large amounts of byproducts, whereas the biological process leads to excessive sludge generation, which requires additional financial expenditures for waste management (Kuo et al., 2008). On the other hand, other treatment techniques, such as ion-exchange, photocatalysis, membrane filtration, coagulation-flocculation, chemical oxidation etc., can be employed to remove dyes from aqueous solutions. Still, these techniques have significant drawbacks due to high operational costs, undesirable production of toxic

intermediate byproducts, large volumes of sludge formation, and formation of numerous heavy metals during synthesis or preparation, which may have adverse health consequences on humans and aquatic organisms (Obayomi et al., 2023). Moreover, treating reactive azo dyes is rather challenging due to their diversity of physicochemical properties and structural complexity (Aqdam et al., 2021). Hence, no standalone treatment method is considered suitable to remove dye wastewater at high efficiency. Nonetheless, adsorption is the most frequently used technique to remove dyes from wastewater due to its cost-effectiveness, environmental friendliness, chemical stability, selectivity, high decolourisation efficiency, simplicity, flexibility and ease of operation (Aqdam et al., 2021; Aysan et al., 2016; Obayomi et al., 2023). No matter how complex the process is, the large-scale adsorption process requires thorough optimisation of IC removal by varying the initial dye concentration, contact time, solution temperature and dye loading capacity to reduce the industrial wastewater treatment costs while meeting the required quality of pollutant removal efficiency. The more complex the wastewater treatment process is, the more challenging it is to control the process parameters in industrial operations, which rely on the operators. Without adequate control of the process conditions, the large-scale adsorption process may lead to suboptimal design and operation, resulting in reduced quality of wastewater treatment.

Computational modelling is a common approach to solving complex problems using experimental data and performing parameter estimation based on the predicted data to optimise the process parameters. To validate the prediction efficiency of the optimised parameters, Monte Carlo simulation can determine the level of uncertainty within the process variables, which may cause a model mismatch to improve robust design and operation of real processes and make the complex industrial operation relatively more straightforward to control. To manage the complexity of processes, ensemble forecasting is a useful modelling approach that combines data sources and models of different types supported with alternative

assumptions to discern distinct patterns within the models without being restricted by arbitrary choices and dependencies from a single statistical, artificial intelligence or machine learning approach limited to a single functional form or dataset (Wu & Levinson, 2021). Past research had shown that a combination of artificial neural network, particle swarm optimization, and Monte Carlo simulation was used to investigate the optimum condition of co-combustion of coal and peanut hull, achieving a multilayer perceptron model with a R^2 of 0.99994 (Buyukada, 2016). For example, an ensemble machine learning approach can examine process parameters and other scale-up opportunities of microbial electrochemical systems for determining hydrogen peroxide production (Chung et al., 2023). More interestingly, XGBoost ensemble machine learning approach is essential to achieve safe maintenance of drinking water supply system that requires an interpretable analysis of model prediction, such as predicting the concentration of algae (Park et al., 2022). However, complex industrial processes have numerous variables to control, and the interactions between the variables are largely unknown. To discern the interrelationships between the causal variables, Bayesian distribution network analysis can be used to estimate the levels of interactions or interrelationships between the process variables, finding the significance of impact between the input and output variables. These computational methods require adequate assumptions, such as approximating probability distribution functions to normal distributions to facilitate this process. However, such methods do not guarantee the accuracy of the estimation, but they eliminate the drawbacks of using a single approach, thereby avoiding costly computational processes.

The uncertainty quantification based on Monte Carlo simulation can be validated using Bayesian inference, which utilizes the prior knowledge of simulation (Ighnih et al., 2023). Monte Carlo simulation has been used successfully by other wastewater treatment technologies. To manage the complexity of chemical engineering process, the combination of

Monte Carlo simulation and Bayesian network analysis can minimise lengthy computational time, reducing the computational effort where alternative algorithms can support the complex computational process. Nonetheless, the combination of optimization approaches, Monte Carlo simulations and Bayesian distribution network analysis, have yet to be applied to industrial adsorption process.

In this study, process parameters are optimized using a range of AI and ML algorithms and subjected to Monte Carlo simulations to determine the level of uncertainty within the variables. By implementing parallel computing, Bayesian network analysis was applied to determine the probability distribution; Bayesian inference on parameters can then be performed quickly. The artificial data were evaluated by simulation through generation of models. The remainder of his paper evaluated mathematical models of the targeted IC removal process using various AI and ML optimization techniques. It was subjected to validation using Monte Carlo simulations and Bayesian inference systems.

2.1. Indigo carmine dye

Indigo carmine dye powder was purchased from Sigma-Aldrich (Merck KGaA), Australia with a purity of 99.9% and molecular weight of 466.36 g/mol. It has the chemical formula, $C_{16}H_8N_2Na_2O_8S_2$ with a dusky, purplish-blue appearance in powder form. It is soluble in water and alcohol and partially soluble in organic solvents.

2.2. Adsorbent material

Bovine bone char (BOV) was purchased from Charcoal House, Australia. It is a precursor for the adsorbent material sourced commercially and relatively inexpensive adsorbents that can be obtained from waste products in the food industry. The preparation techniques used by the industry involve pyrolysis, which can influence the yields and textural features. Factors such as temperature, time and flow rate of gasifying agents are carefully

adjusted to facilitate the production of bone char to desired features such as particle size, specific surface area and other physicochemical properties unique to the adsorbent. The bone char adsorbent has 841 × 250 microns, a specific surface area of 200 m^2/mg and a bulk density of 2.497 kg/m^3. The pyrolysis of animal bones under limited oxygen conditions can maximise carbon contents, enlarging surface area and increasing ash contents.

2.3. Adsorption studies

Adsorption studies were conducted by varying the process parameters, such as contact time and initial dye concentration. 1-L of the dye solution with appropriate concentration was taken into a 250 mL Erlenmeyer flask. The adsorbent dosage was carefully weighed on analytical balance, added to the flasks, and subjected to mechanical agitation. The agitation speed was approximately between 150 and 200 rpm, and samples were collected at room temperature of 22.5°C. The samples were filtered and analysed by a UV-visible spectrophotometer (HACH DR600) at 610 nm (the maximum absorption wavelength for IC) for residual dye concentration in the aqueous solutions. The IC removal efficiency in the aqueous solution by BC was computed using the following equation:

$$\text{IC removal } (\%) = \left(\frac{C_i - C_f}{C_i}\right) \times 100\% \quad (1)$$

Where C_i and C_f are the initial and final IC concentrations (mg/L), respectively (Pavlović et al., 2014).

Adsorption capacity q_t (mg/g) was computed by the following equation:

$$q_t = \frac{(C_i - C_e)V}{m} \quad (2)$$

Where m is the adsorbent dosage (g), V is the volume of solution (L). C_i is the initial dye concentration (mg/L), and C_e is the residual dye concentration (mg/L) at different time intervals (Pavlović et al., 2014).

2.4. Dubinin-Radushkevich isotherm model

Dubinin-Radushkevich isotherm model is an empirical adsorption isotherm model which is generally applied to adsorption mechanism with Gaussian energy distribution on heterogeneous surfaces (Maamoun et al., 2021). This isotherm is used to determine the thermodynamics of the adsorption process. In addition, the model is a semi-empirical equation involving an ion-exchange mechanism under the assumption that multilayer adsorption occurs by van der Waals' forces, which applies to physisorption process (Pandiarajan et al., 2018). At low pressure, this isotherm model is unable to account for Henry's physicochemical laws and describes the sorption process due to impracticality (Maamoun et al., 2021). The isotherm model also describes the adsorption of gases and vapours on microporous surface of adsorbents. The most distinctive function of this isotherm model is temperature-dependent, whereby the adsorption data can be plotted based on the differences in temperature as a function of dye concentration changes or the amount of dye adsorbed onto the adsorbent surface, resulting in changes in potential energy distribution. Dubinin-Radushkevich isotherm model is defined as (Tan et al., 2009):

$$\log_e q_e = \log_e q_m - \beta E^2 \quad (3)$$

Rearranging the Eq. (3) to give:

$$q_e = q_m \exp(-\beta \varepsilon^2) \quad (4)$$

where ε can be correlated as:

$$\varepsilon = RT\log_e\left(1 + \frac{1}{C_e}\right) \quad (5)$$

where R is the universal gas constant (8.314 J/mol.K), ε is Polanyi potential, β is Dubinin-Radushkevich constant, E is mean adsorption energy and T is the absolute temperature (K). A plot of $\log_e q_e$ versus ε^2 enables q_m to be determined from the slope. The constant β gives the average free energy E of adsorption per molecule of adsorbate when one mole of ion is transferred onto the surface of the adsorbent from infinity to the surface of the adsorbent and can be computed as (Pandiarajan et al., 2018):

$$E = \frac{1}{\sqrt{2\beta}} \quad (6)$$

For mean free energy values between 8 and 16 kJ mol^{-1}, the adsorption process is considered to be influenced by an ion-exchange mechanism and for values greater than 16 kJ mol^{-1}, the adsorption process is considered to be dictated by a particle diffusion mechanism (Pandiarajan et al., 2018).

2.5. Physicochemical characteristics of bone char

The classification of bone char involved the findings from the characterisation test results. The characterisation of the adsorbent material included determining particle size, surface functionalities and surface morphologies using Fourier Transform Infrared Spectroscopy (FTIR) and scanning electron microscopy (SEM), respectively.

The particle size of the porous bone char was determined to assess its ability to uptake the amount of dye adsorbate from the aqueous solution. The particle size has a direct impact on the maximum adsorptive capacity. In addition, mesopores (2-50 nm), macropores (>50 nm) and micropores (<2 nm) play a significant role in imparting the adsorbent materials with different adsorptive capacities. Therefore, the larger the particle size, the greater the availability of pore volume to accommodate a significant amount of dye adsorbate. Aside from that, the bulk volume of adsorbent encompasses internal and external features whereby the internal volume involves the region within the enclosed pores, residing within the vicinity of the adsorbent mass, whereas the external area involves primarily fissures and cracks that extend deep into the bulk volume of adsorbent with microchannels or networks of varying widths or sizes (Guo et al., 2022). In addition, the size of the pores on the adsorbent materials determines how much the dye adsorbate can be adsorbed on its surface or diffused into the materials. The surface charge or functionalities of the adsorbent materials determine whether there are any electrostatic repulsive or attractive forces to enhance the adsorption efficiency. The clogging of pores could be a significant issue when using adsorbents to treat heavily

polluted wastewater with high turbidity levels. However, particle attrition is another significant issue in the presence of mechanical disturbances, which cause the adsorbent particles to break up into smaller particles. When the pore size is very small, the diffusion rate is influenced by the rate-limiting step, which means the amount of dye adsorbate diffused into the adsorbent materials is limited.

2.6. AI optimization techniques

1.1.1 Adaptive neuro-fuzzy inference system (ANFIS)

ANFIS is a hybrid algorithm where nodes of a feedforward neural network handle fuzzy parameters (Zaghloul et al., 2020). It is used to model a complex system with high uncertainty. The model is characterised by a first-order Takagi-Sugeno fuzzy inference system (FIS), which transforms the input properties into membership values using input membership functions (MF) through a fuzzification process. It is also a hybrid model which comprises an artificial neural network (ANN) with fuzzy logic reasoning. Hence, it can learn from training data, and the computed solution is mapped onto a FIS, giving a rule-based system with three essential components: fuzzification, fuzzy database and defuzzification. The strength of FIS is its non-linear mapping performance between input and output variables by managing linguistic concepts (Nam et al., 2023). However, one of the limitations of ANFIS is the increase in the ANFIS input numbers, which may result in lengthening computational time and rule numbers. Under some circumstances, insufficient data used for training may lead to failure of the ANFIS model (Heydari et al., 2021). Thus, a standard ANFIS could not be employed to model an output variable. In this case, the data are strategically arranged with input parameters randomly divided into smaller groups to reduce criteria and visual representation to provide a converged solution. This requires considering constraints such as the number of input parameters, which can be more than 5 or the number of observations, which can be more than 19 (Heydari et al., 2021). To train the data, the input neuron values were normalised in the range of 0.1-0.9 by Eq. (7):

$$x_{i,norm} = 0.8 \times \left(\frac{x_i - x_{i,min}}{x_{i,max} - x_{i,min}}\right) + 0.1 \quad (7)$$

The accuracy of the ANFIS models was determined with the root mean square error (RMSE) and the determination of coefficient (R^2) in accordance with Eq. (8) and Eq. (9) (Hadi et al., 2020):

$$R^2 = \frac{\sum_{i=1}^{n}\left(q_{e,pred} - \overline{q_{e,exp}}\right)^2}{\sum_{i=1}^{n}\left(q_{e,pred} - \overline{q_{e,exp}}\right)^2 + \sum_{i=1}^{n}\left(q_{e,pred} - q_{e,exp}\right)^2} \quad (8)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_{pred,i} - y_{exp,i}\right)^2} \quad (9)$$

Where N is the number of experimental runs, $y_{pred,i}$ is the predicted value, $y_{exp,i}$ is the experimental value, and $q_{e,pred}$ and $q_{e,exp}$ are the predicted and experimental adsorption equilibrium capacity, respectively.

The pilot scale adsorption system can be upscaled by managing multiple physicochemical parameters and variables, but the process is highly sophisticated and not well-defined. Moreover, it is also challenging to use specific equations to ascertain quantities of determinants to define the upscaled adsorption system. Therefore, FIS with fuzzy logic rules is considered a feasible technological application. A set of fuzzy logic rules is developed to infer a particular model output. Afterwards, a set of output MFs is used for defuzzification of the inference output to real output values. This algorithm utilises the input membership functions to separate the input variables to minimise the search space. It uses error optimisation algorithms equivalent to the back-propagation feature of neural networks to adjust the fuzzy parameters. The general ANFIS architecture consists of five distinct layers (Nam et al., 2023): (i) fuzzy layer, (ii) product layer, (iii) normalisation (or rule) layer, (iv) defuzzification layer, (v) overall output summation layer. Each layer performs a specific mathematical process using a first-order Sugeno-type model, which consists of 4 input variables with fuzzy IF-THEN rules employed in this study. The following outlines different layer functions with specific mathematical processes:

Layer 1: Adaptive nodes that map the inputs (x or y) utilising the membership function μ into linguistic variables such as A and B, respectively. Membership functions take different forms or shapes, including trapezoidal, triangular, generalised bell-shaped curve, and Gaussian. The output of layer 1 for specific input variables is as follows (Nam et al., 2023; Zaghloul et al., 2020):

$$O_{1,i} = \mu_{Ai}(x) \tag{10}$$

Rule 1: If x is A_1 and y is B_1, then $f_1 = p_1x + q_1y + r_1$ (11)

Rule 2: If x is A_2 and y is B_2, then $f_2 = p_2x + q_2y + r_2$ (12)

Layer 2: Fixed nodes that apply different logical rules, such as AND/OR by multiplying the node input signals. The output layer is known as the firing strength, which can be mathematically expressed as follows (Nam et al., 2023; Zaghloul et al., 2020):

$$O_{2,i} = w_i = \mu_{Ai}(x) \times \mu_{Bi}(y) \tag{13}$$

μ is the membership function, A is the linguistic label related to the membership function and output, O represents the parameter to which x belongs to A.

Layer 3: Fixed nodes are implemented to normalise the firing strengths, which involve the calculation of the firing strength ratio at each node and the combination of all firing strengths. The output layer 3 is known as the normalised firing strength (Nam et al., 2023; Zaghloul et al., 2020):

$$O_{3,i} = w_i = \frac{w_i}{w_1 + w_2} \tag{14}$$

Layer 4: Adaptive nodes derived from the following output function (Nam et al., 2023; Zaghloul et al., 2020):

$$O_{4,i} = w_i f_i = w_i(p_i x + q_i y + r_i) \tag{15}$$

Where f represents the IF-THEN rule, p, q, and r denote the consequent parameters. IF-THEN rules can be expressed as: IF x is A AND y is B, it can be written as THEN f=px+qy+r.

Layer 5: The sum of all incoming signals from a single fixed node generates an overall output of the model (Nam et al., 2023; Zaghloul et al., 2020):

$$O_{5,i} = \sum_i w_i f_i \propto \frac{\sum_i w_i f_i}{\sum_i w_i} \quad (16)$$

The output of any node I in layer 1 is denoted as $O_{1,I}$, x is the input parameter.

1.1.1 Support vector machine

Support vector regression (SVR) is a machine learning algorithm that can be used for statistical analysis of regression or classification of non-linear models using the structural risk minimization (SRM) principle (Nam et al., 2023). SVR has been widely used as a machine learning algorithm for solving complex regression problems associated with wastewater treatment simulations. In particular, SVR is used to solve non-linear correlation issues. Given the training dataset, D = {(x_1, y_1), (x_2, y_2), …., (x_m, y_m)}, the original dataset used in the regression problem can be mapped by SVR into a higher dimensional space involving a kernal function to enable fitting of the original dataset with a theoretical linear regression function as expressed in the following form (Nam et al., 2023):

$$f(x) = w^T \varphi(x) + b \quad (17)$$

Where f(x) is the simulated value obtained from SVR using input variable x, w, and b denote the coefficient weight vector and bias term achieved by minimizing the upper bound of the generalization error, respectively. For dataset of $x_i \in \boldsymbol{\mathcal{R}^d}$ as the training input vector with d dimensions and $y_i \in \boldsymbol{\mathcal{R}}$ as the training target vector. i = 1: N represents a number of data pairs, and SVR training data are mapped into a higher dimensional space to generate a linear model f (w, x) to predict the target vector (Zaghloul et al., 2020). $\varphi(x)$ is the non-linear mapping function, which is also known as kernel function.

1.1.2 Random forest algorithm

Random forest (RF) is an ensemble machine learning method for classification or regression that operates on a multitude of decision trees (Wu et al., 2023). It involves a

modified version of bagging that uses an improved bootstrapping method. RF consists of a series of decision trees (DT) classifiers, which can be characterized by their high interpretability, especially in problems with strong physical knowledge, but can be increasingly unstable and relatively inaccurate when analysing irregular patterns of data and prone to overfitting the training datasets (Bellamoli et al., 2023). The DT classifiers can be expressed as h (X, θ_k), k=1, 2, 3… with X as an independent input vector and k denotes the number of decision trees, and θ_k is a random vector (Wu et al., 2023). The RF algorithm starts by selecting bootstrap datasets $\psi\ (x_1, \dots, x_n) \in D(x_1, \dots, x_N)$ with a random set of features m ∈ M for t estimators in parallel to nominate node-splitting variables (Qambar & Khalidy, 2022). For multiclass classification, the output variables can be generated from RF selected from most decision trees by averaging multiple deep decision trees and trained on different datasets to reduce variance and overfitting. Most importantly, the hyperparameters of the RF involve a number of decision trees, the maximum depth of trees and the maximum number of leaves. Only a fraction of features can be utilized for a tree, and another fraction of bootstrapped samples can be finely tuned (Bellamoli et al., 2023).

1.1.3 XGBoost algorithm

Extreme Gradient Boosting (XGBoost) is an efficient ensemble of decision tree-based algorithm using the gradient boost algorithm. The purpose of using this algorithm is to minimise the regularized equation as follows (Qambar & Khalidy, 2022):

$$L(\emptyset) = \sum_{i=1}^{N} L(y_i, y_i) + \gamma T + \frac{1}{2}\lambda\left\|\omega_j\right\|^2 \quad (18)$$

Where L (y_i, y_i) is the regression loss function and equals 0.5 $(y_i, \gamma)^2$. γ is the model complexity penalizing term, and T represents the number of terminal nodes, whereas λ is the regularization parameter and ω_j denotes the node j predicted value.

The model output parameter for a sample can be calculated by adding the sum of the leaves assigned for each regression tree as follows (Ching et al., 2022):

$$\widehat{y_i} = \sum_{i=1}^{K} f_k(x_i)$$

(19)

The regression trees can be added to the ensemble as f_t for iteration (t) to form new regression trees to minimise the learning objective. It can be optimized in an Euclidean space in a pre-defined structure rather than a singular model to five the following form (Ching et al., 2022):

$$\mathcal{L} = \sum_{i=1}^{n} l\left(y_i, \widehat{y_i^{t-1}}\right) + f_t(x_i) + \Omega(f_t)$$

(20)

For benchmarking purposes, XGBoost model was compared with SVM and RF models.

1.1.4 Hybrid artificial neural network and genetic algorithm

Single models have some limitations in the learning process. The current learning rate of a configurable hyperparameter used in the artificial neural network training has a small positive value ranging between 0.0 and 1.0. This learning rate controls how quickly the predictive model adapts to the problem. Notwithstanding its benefits, hybrid methods tend to perform better than a standalone optimization technique to enhance accuracy or minimise error. This work applies a hybrid combination of artificial neural network (ANN) and genetic algorithm (GA) to the adsorption data to model and predict IC concentration changes and removal efficiency based on the operational variables. ANN operates like a human brain and nervous system with an outstanding ability to learn, classify and optimise data (Asgari et al., 2020). An ANN consists of an input layer, one or more hidden layers and an output layer that is known as a multi-layer perceptron (MLP) structure (Özdoğan-Sarıkoç et al., 2023). Improvement in ANN modelling is based on the number of hidden layers, neurons and the types of transfer functions used (Mohammad et al., 2020; Özdoğan-Sarıkoç et al., 2023). The Levenberg-Marquardt backpropagation algorithm with 1000 epochs is considered one of the

best algorithms to train the network (Asgari et al., 2020). The number of neurons in the hidden layer is arranged within a range of 1-20 to determine the optimum number of neurons required to generate a minimum mean squared error (MSE). Optimal learning rate is adjusted between 0.001 and 0.01, starting from the low side and increasing incrementally by 0.0005 to avoid overshooting the target. A default batch size of 32 is used first, requiring fewer epochs to converge while avoiding a large batch size, typically requiring a long computation time to complete an epoch. Approximately 70% of data is allocated for training, whereas 30% of validation and testing data are used to classify the random variables.

On the other hand, GA emulates the natural evolution process, operating on three operators: selection, crossover, and mutation (Wahidna et al., 2024). GA functions well in determining optimal global solutions. The designated parameter values and functions for GA involve crossover fraction, mutation function and migration function of 0.8, Gaussian and 0.2, respectively. GA generally leads to better forecast accuracy when combining ANN with an iterative optimization algorithm (Wahidna et al., 2024). The hybrid optimization of the operational variables was further conducted using three standard AI and ML approaches, such as RF, SVM, and XGBoost, to generate a new population. All ANN-GA, ANN-GA-RF, ANN-GA-SVM and ANN-GA-XGBoost optimized data are subjected to Monte Carlo simulation to determine the levels of uncertainties in the simulated models.

3.1. Mechanism of Indigo Carmine dye adsorption on the solid-liquid interface

The surface phenomena of adsorption process are influenced by any conditions such as pressure and temperature changes. The adsorption of dye adsorbate can be affected by thermodynamic conditions of the aqueous solution. High temperature of the solution can disrupt the intermolecular forces between the dye adsorbate and adsorbent materials, loosening up the adhesion at the contact surface within the solid-liquid interphase, resulting in the desorption of dye adsorbate from the adsorbent surface. Throughout the experimental

studies, the temperature of the dye solution was varied to examine the adsorption behaviour of bone char to remove Indigo Carmine dye pollutants. Bone char has a high degree of porosity, a strong tendency to uptake significant amounts of dye pollutants from solutions, and the adsorption equilibrium time takes longer to reach. During the adsorption process, the concentration of dye adsorbate varies as the adsorbent attracts the dye pollutants using its surface functionalities. The dye pollutants can either accumulate on the adsorbent surface or diffuse into the pores of the adsorbent materials as shown in Figure 1.

The thermodynamic conditions of the aqueous solution have a significant effect on the adsorption efficiency of the adsorbent material. When the pores of the adsorbent material are solvated by the water molecules, it takes a considerable amount of energy to displace the water molecules to facilitate the adsorption of dye molecules. If the free energy of the water molecules is too high and the activation energy of the adsorption process is too large, it would take a greater amount of energy from the dye molecules to displace the water molecules to be diffused or adsorbed into the pores of the adsorbent material. The adsorption process must be thermodynamically favourable to facilitate the adsorption of dye pollutants onto the adsorbent materials. Therefore, the balance between the Gibbs free energy, activation energy, enthalpy and other thermodynamic parameters is very critical to facilitate the adsorption process. In addition, the interactive forces between the adsorbent surface and dye adsorbate must be strong to facilitate the mass transfer of solute onto a substrate. The adsorption equilibrium is reached when the adsorbent surface is saturated with dye pollutants. The isotherm models are used to characterize the adsorption phenomena using detailed analysis and estimation between the experimental and theoretical models. The nature of the adsorbate-adsorbent complex is influenced by the variation of operating conditions, which directly impacts the adsorption process. The physicochemical characteristics of the adsorbent materials are significant and play major roles in facilitating the adsorption process.

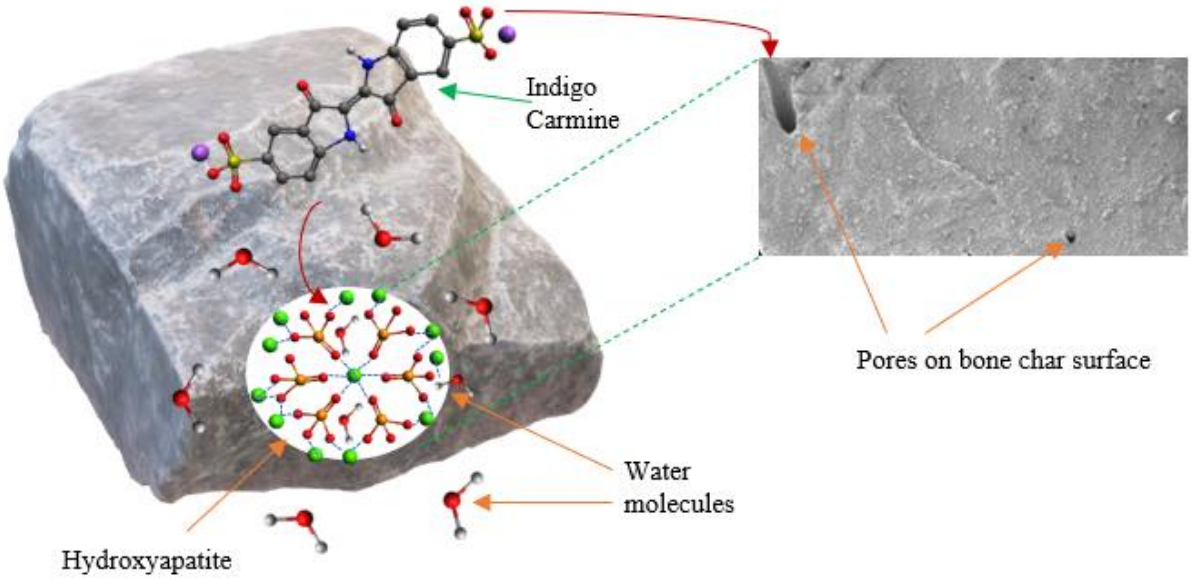


Figure 1. Inner pore region of the bone char adsorbent.

3.2. Adsorption data

The batch adsorption studies were performed to examine the uptake amount of IC dye pollutant using bone char adsorbent. Figure 2a shows the removal ratio of IC pollutants across different initial IC concentrations. The dye sorption usually occurs when the IC molecules diffuse from the bulk liquid onto the adsorbent surface through the solid-liquid interface. On the other hand, Figure 2b represents the adsorbent loading of IC adsorbates onto the bone char adsorbent surface. The IC adsorbates interacted with the active sites on the adsorbent surface through intermolecular interactions such as van der Waals' interaction, electrostatic interaction, hydrophobic interaction, π-π electron donor-acceptor interaction, and hydrogen bonding.

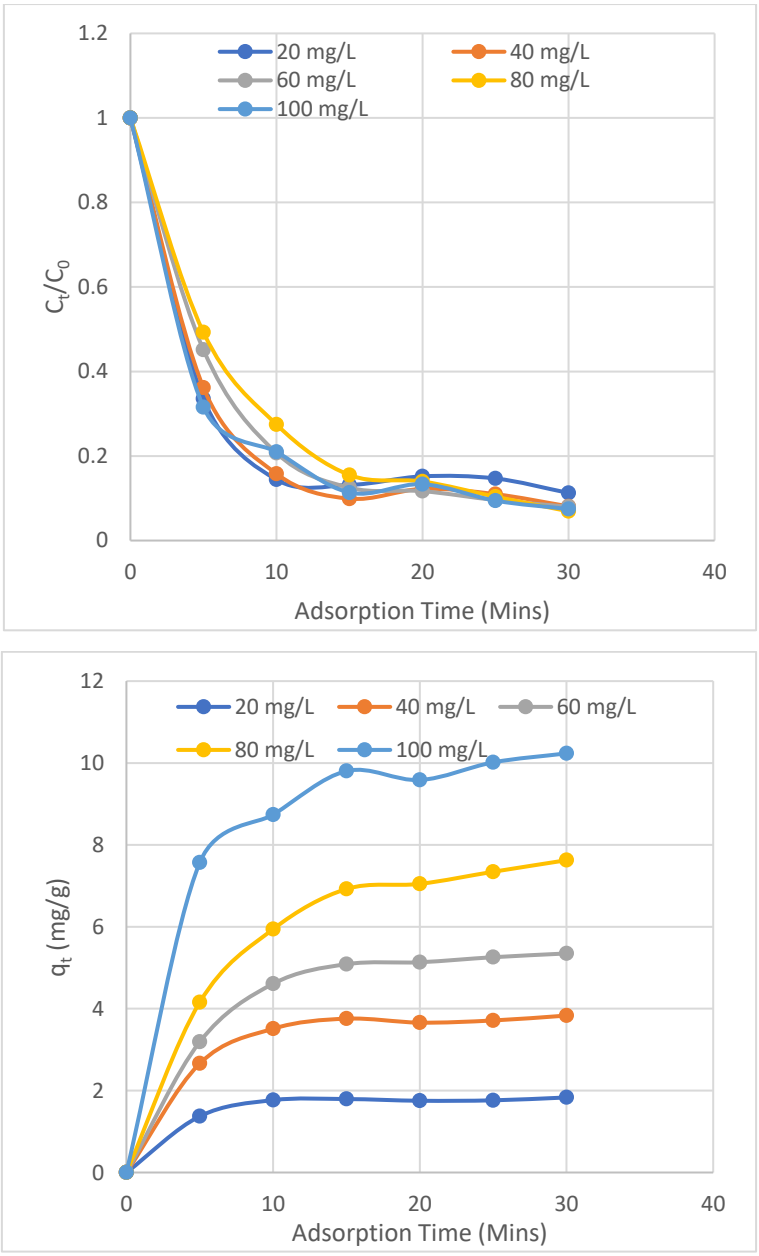


Figure 2. a) IC removal efficiencies of bone char adsorbent across different initial dye concentrations at 30°C; b) Adsorbent loading of bone char adsorbent across different initial dye concentrations at 30°C.

3.3. Thermodynamics of adsorption system

The surface phenomena of adsorption process are influenced by any conditions such as pressure and temperature changes. The adsorption of dye adsorbate can be affected by thermodynamic conditions of the aqueous solution. High temperature of the solution can disrupt the intermolecular forces between the dye adsorbate and adsorbent materials, loosening up the adhesion at the contact surface within the solid-liquid interphase, resulting in the desorption of dye adsorbate from

the adsorbent surface. Throughout the experimental studies, the temperature of the dye solution was varied to examine the adsorption behaviour of bone char to remove Indigo Carmine dye pollutants. Bone char has a high degree of porosity, a strong tendency to uptake significant amounts of dye pollutants from solutions, and the adsorption equilibrium time takes longer to reach. During the adsorption process, the concentration of dye adsorbate varies as the adsorbent attracts the dye pollutants using its surface functionalities. The dye pollutants can either accumulate on the adsorbent surface or diffuse into the pores of the adsorbent materials. The thermodynamic conditions of the aqueous solution have a significant effect on the adsorption efficiency of the adsorbent material. When the pores of the adsorbent material are solvated by the water molecules, it takes a significant amount of energy to displace the water molecules to facilitate the adsorption of dye molecules. If the free energy of the water molecules is too high and the activation energy of the adsorption process is too large, it would take a greater amount of energy from the dye molecules to displace the water molecules to be diffused or adsorbed into the pores of the adsorbent material. The adsorption process must be thermodynamically favourable to facilitate the adsorption of dye pollutants onto the adsorbent materials. Therefore, the balance between the Gibbs free energy, activation energy, enthalpy and other thermodynamic parameters is very critical to facilitate the adsorption process.

Furthermore, thermodynamic conditions had a significant effect on the adsorptive performance of bone char. A set of thermodynamic parameters evaluated in Figures 3 and 4 show that Gibbs free energy and activation energy of the IC adsorption equilibrium system were the lowest at 1.68 kJ/mol and -0.099 kJ/mol (Table 1), respectively, indicating that minimal free energy from dye molecules was required to displace the water molecules within interior adsorption sites of the adsorbent, resulting in strong adsorption taking place. Low Gibbs free energy indicated that the IC adsorption system at 30°C was thermodynamically

favourable. When coupled with low activation energy, the adsorption or reaction process was spontaneous, leading to strong adsorption efficiency. In addition, ANSYS FLUENT software was used innovatively to investigate the computational fluid dynamics and thermodynamic effects of a packed bed reactor integrated with bone char adsorbents to filter xenobiotic dye wastewater, as shown in Figure 5. The path lines represented both flow velocity vectors and temperature distribution within the packed bed reactor to indicate the process dynamics of the upscaled adsorption process technology.

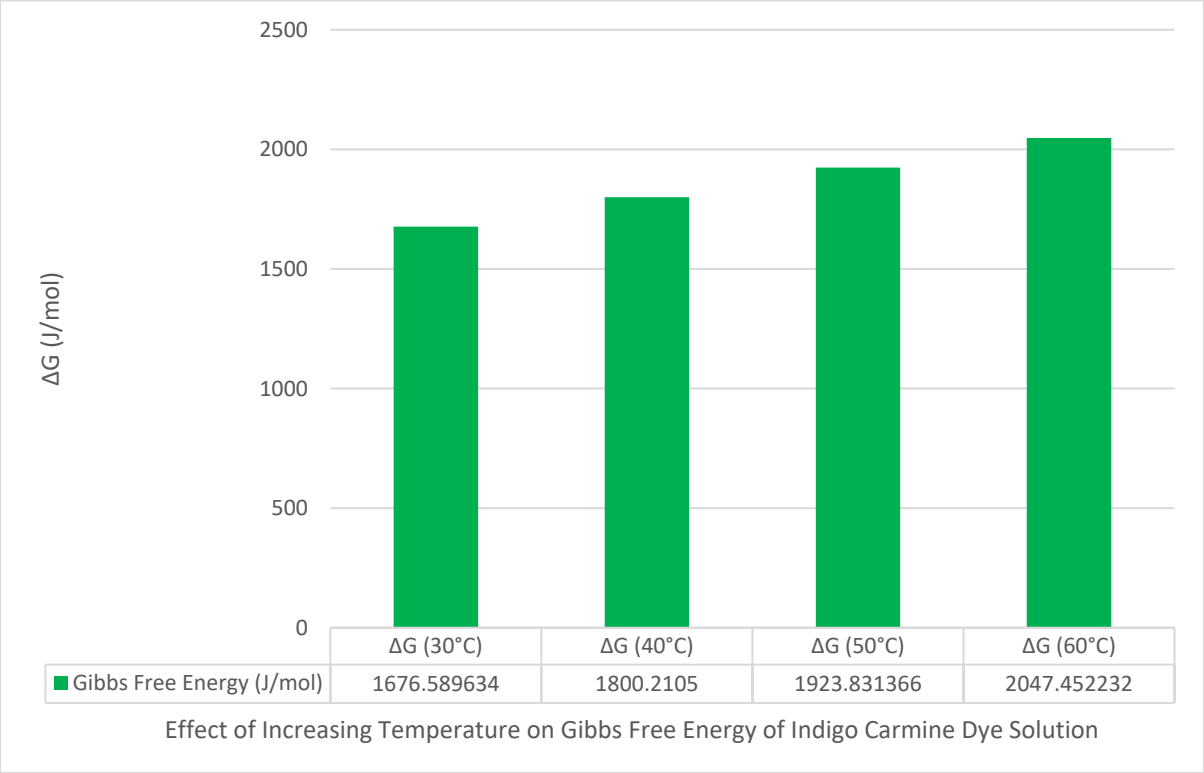


Figure 3. Determination of Gibbs free energy of the solution with temperature ranging from 30°C to 60°C.

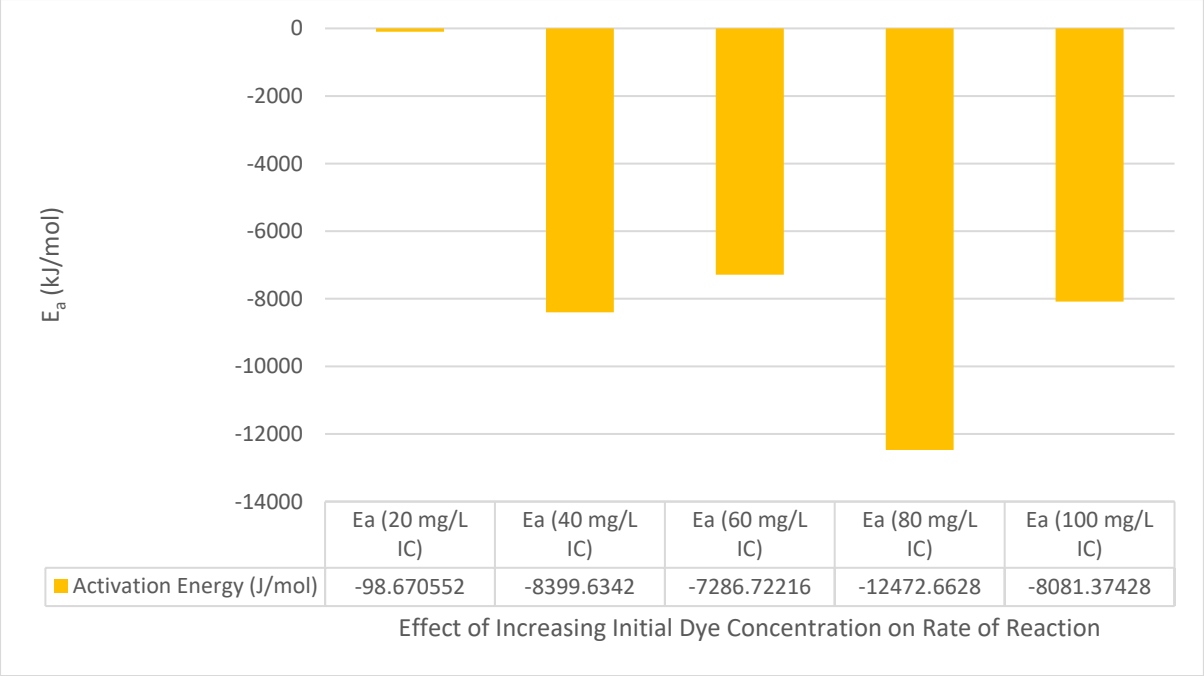


Figure 4. Determination of activation energy of the solution with increasing initial dye concentration at temperatures ranging from 30°C to 60°C.

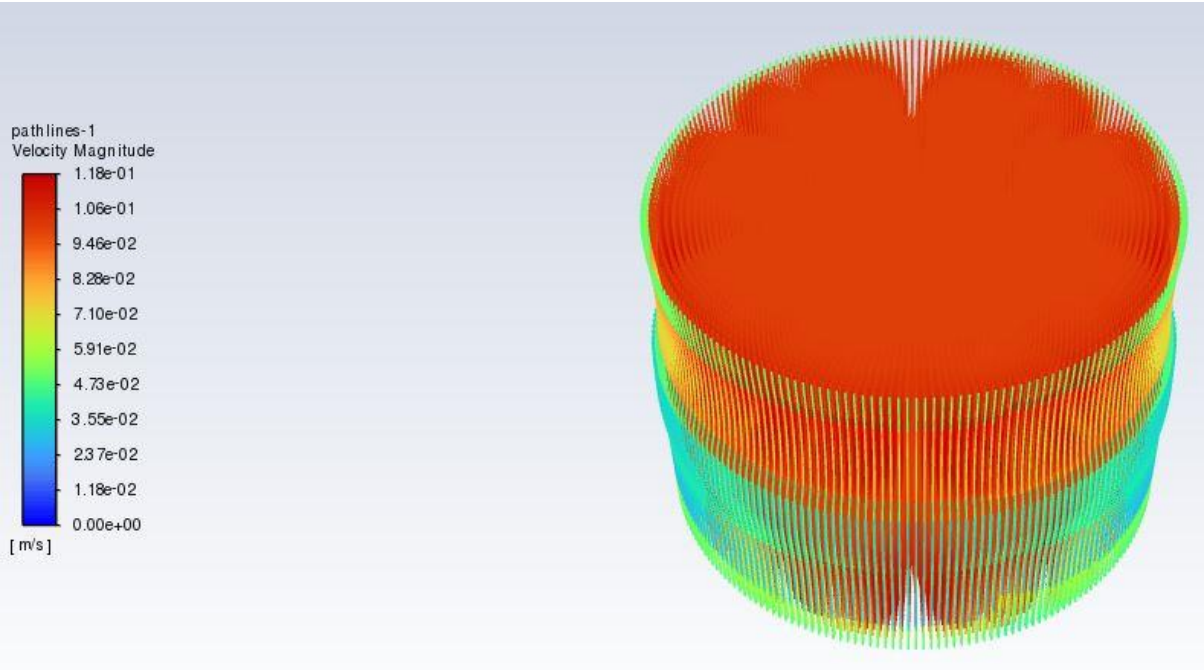


Figure 5. Computational fluid dynamics (CFD) simulation of a packed bed reactor containing bone char adsorbents for industrial filtration of xenobiotic dye wastewater.

Thermodynamic parameters of IC adsorption using bone char.

T (K)	ΔG (kJ/mol)	ΔH (kJ/mol)	ΔS (kJ/mol.K)
303.15	1676.59	-2071.10	-12.36
313.15	1800.21		
323.15	1923.83		
333.15	2047.45		

1.1.5 3.4. AI optimisation of adsorption system***Adaptive neuro-fuzzy inference system (ANFIS)***

ANFIS is a hybrid algorithm where nodes of a feedforward neural network handle fuzzy parameters (Zaghloul et al., 2020). ANFIS is specifically used to model the IC adsorption system, which is characterized by high uncertainty, perturbation, or complexity. In particular, Figure 4a shows the ratio of IC removal versus initial IC concentration in aqueous solution over time using bone char adsorbent. On the other hand, Figure 4b shows that the adsorbent loading reached an adsorption equilibrium between 10 to 15 minutes, indicating that it took considerable time to remove IC from aqueous solution using bone char adsorbent. Similarly, for modelling artificial neural network (ANN), ANFIS is programmed to train dataset. Training ANN means determining the input parameters, such as IC concentration removal, using an optimisation algorithm, as shown in Figure 6a. In this approach, the premise parameters are determined using gradient descent (GD), and consequence parameters are generated using the least square estimation (LSE) method.

The coefficient of determination (R^2) for determining the degree of curve fitting is illustrated in Figure 6b. This technique is known as learning with samples. When learning is completed, an appropriate ANN model can be generated. More critically, the test dataset is utilised to measure the success of the developed model as shown in Figure 6c. The difference between the predicted and actual values is the error. In this case, Figure 6d shows that the correlation coefficient of the test dataset is approximately 0.707. The lower the predictive error, the better the ANN model is. In particular, ANFIS has adopted some proportion of learning ability and relational structure to ANN with the decision-making mechanism of fuzzy logic combined in the system. This way, the most ideal ANFIS architecture is generated to solve the related problem. The obtained structure is subjected to the test process to evaluate the level of impact on samples. The lower the predictive error, the more suitable

the ANFIS model is for optimisation. Unlike ANFIS optimisation, the weight values generated from ANN could not be defined (Karaboga & Kaya, 2019). This disadvantage is addressed by fuzzy inference system, which is the critical component of ANFIS. This core architecture can be exploited to solve various real-world problems.

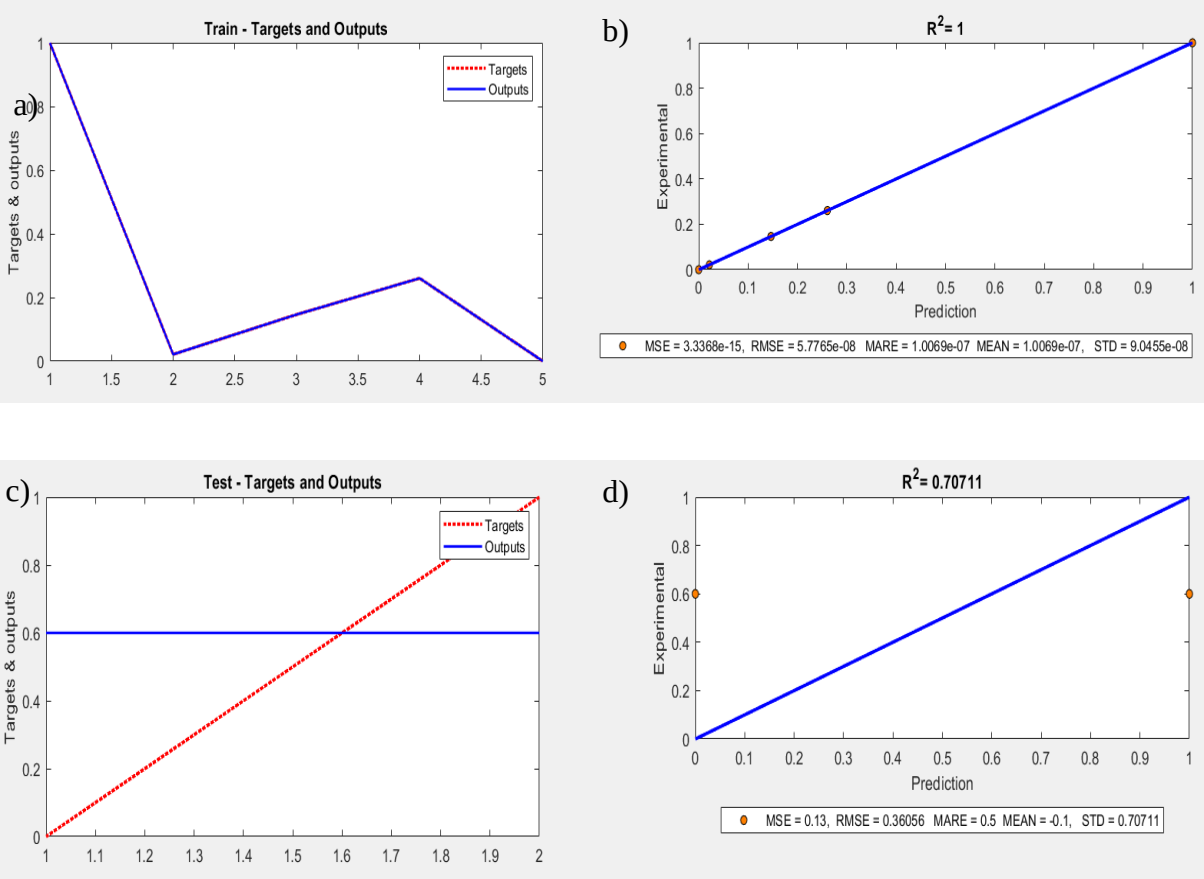


Figure 6. a) ANFIS trained input and output data; b) ANFIS generated regression coefficient of input and output data; c) ANFIS tested input and output data and d) Regression coefficient of all input and output data.

3.5. Prediction efficiency of the AI and Monte Carlo simulation models

Simulation-based approaches were formulated using Monte Carlo generated data. The Monte Carlo simulation provides a variety of possible outcomes determined from the probability of random variables. It offers a succinct visualisation through deterministic forecasts and prediction of variables based on uncertainties. By determining the true probability distribution, the performance of optimised models can be quantified by employing formal and objective statistical analysis based on criteria such as approximating the

probability density functions to normal distributions. Table 2 shows the prediction efficiencies of different optimised models based on the actual experimental values. To validate the optimised models, one of the requirements was to compare the uncertainties between the prediction efficiencies of different machine learning-based approaches to evaluate the estimated uncertainties by studying different shapes of P-P and Q-Q plots, which represent theoretical cumulative distribution versus empirical cumulative distribution curves as shown in Figures 7a-h. Figure 7a shows the P-P plot of actual experimental values, which was not significantly different from the P-P plot of optimised models as shown in Figure 7b. In contrast, the Q-Q plot in Figure 7e compares the quantiles of standardised normal data distribution versus actual IC concentration values, which assessed whether the dataset plausibly came from some theoretical distribution, such as normal data distribution. Judging from the Q-Q plot of Figure 7e, the standardised normal data distribution was not significantly different from Figure 7a, indicating that the actual data could be over or underestimated. The pattern of optimised data in Figure 7b shows the ANN-GA-RF algorithm resembled all other P-P plots, indicating no significant deviation in prediction. Compared with Figure 7f, the Q-Q plot shows a significant tendency for ANN-GA-RF-optimised models to exhibit a slight overfitting data pattern based on the standardised normal data distribution. When cross-referenced with the ANN-GA-RF-optimised values in Table 2, there are some minor fluctuations in predicted values. On the other hand, Figure 7c exhibits an oscillating pattern of optimised data, indicating that the ANN-GA-SVM algorithm tends to fit the data remarkably well, leading to significant uncertainty or inaccuracy in the predicted values. Similarly, the Q-Q plot in Figure 7g shows that the standardised normal distribution of predicted IC concentration values exhibited no overfitting data pattern under ANN-GA-SVM optimisation. More interestingly, Figure 7d P-P plot shows a more stable pattern of optimised data, more aligned with the actual experimental values. It exhibits little to no

fluctuation in predicted values, indicating that the ANN-GA-XGBoost algorithm generated a remarkable curve fitting between actual and optimised data. Moreover, the Q-Q plot in Figure 7h shows the ANN-GA-XGBoost optimised model exhibited a well-aligned normal data distribution to the actual data, indicating slightly better prediction efficiency than the ANN-GA-RF optimised model. Furthermore, Figure 7i shows the probability density function, which represents a continuous version of a histogram with densities, and it specifies how the probability density is distributed over a range of actual final IC aqueous concentrations as random variables. On the other hand, Figure 7j shows a moderate cumulative relative frequency of approximately 0.5 or 50%, which was achieved for an actual IC aqueous concentration of 85 mg/L. In contrast, at a higher initial IC concentration, especially greater than 85 mg/L, the cumulative relative frequency increased from 0.5 or 50% to 1.0 or 100%, indicating that the predictive performance of ANN-GA tends to fit better if a moderately high IC concentration parameter was considered. Figure 7i shows an increase in probability density at 30 mg/L, indicating that the ANN-GA-RF algorithm was more suited to accurately predict a low range of IC concentrations due to less noise or reduced tendency to overfit data. On the other hand, Figure 7m-n shows that the prediction efficiency of ANN-GA-SVM optimisation was better when actual IC aqueous concentration ranging between 40 and 70 mg/L, indicating that this optimisation technique was capable of encapsulating the actual range of IC concentration more accurately compared to ANN-GA-RF ensemble. More interestingly, ANN-GA-XGBoost optimisation generated a strong probability density level, giving a more accurate prediction of moderately high IC aqueous concentration ranging between 55 and 85 mg/L. In addition, Table 2 ANN-GA optimised values show that the data more closely resembled the actual values, achieving a stabilised IC solution equilibrium concentration of 4.712 mg/L after 20 mins of contact time onwards. In addition, Table 3 provides a more thorough analysis of the actual performance of simulations using descriptive

statistics to analyse the level of uncertainty in AI/ML-optimised models. Comparing the skewness versus standard error of skewness, the results showed that ANN-GA-SVM achieved a skewness of 0.018 versus 0.077 of standard error of skewness, indicating that the values were sufficiently close, achieving a significant reduction in prediction error. On the other hand, the skewness values of ANN-GA, ANN-GA-RF and ANN-GA-XGBoost optimised models were largely different from the standard errors of skewness, indicating more significant deviations compared to ANN-GA-SVM models, resulting in higher prediction errors. Similarly, the values of Kurtosis of the ANN-GA-XGBoost and ANN-GA-SVM optimised models were significantly smaller than the standard errors of Kurtosis, indicating significantly less variance or extreme deviations compared to ANN-GA and ANN-GA-RF optimised models. In contrast, the optimised Kurtosis value of ANN-GA-XGBoost was minimal. At the same time, Table 4 shows the MSE value of this ensemble was 21.407, indicating that the normal distribution was lightly tailed compared to other ensembles. Still, the prediction accuracy was limited to a moderately high IC concentration range between 55 and 85 mg/L. The result indicated that the ANN-GA-SVM algorithm achieved better prediction efficiency than ANN-GA-RF and ANN-GA-XGBoost. Regarding sensitivity analysis, Figure 8a-i shows that both ANN-GA-XGBoost optimisation has the least significant impact on the prediction efficiency due to fewer contributions or sensitivities to the changes in actual IC aqueous concentration as shown in Tornado diagrams. In contrast, the spider diagrams show that ANN-GA-RF exhibited the highest range of impact from minimal slope between the actual IC aqueous concentration values versus predicted values, indicating that ANN-GA-RF optimisation has a much greater tendency to overfit the data. A stronger sensitivity of ANN-GA-RF-predicted models than the actual model indicated greater overall uncertainty. However, ANN-GA-XGBoost-optimised models demonstrated the least sensitivity impact compared to ANN-GA-RF-predicted models. Still, the ANN-GA-XGBoost

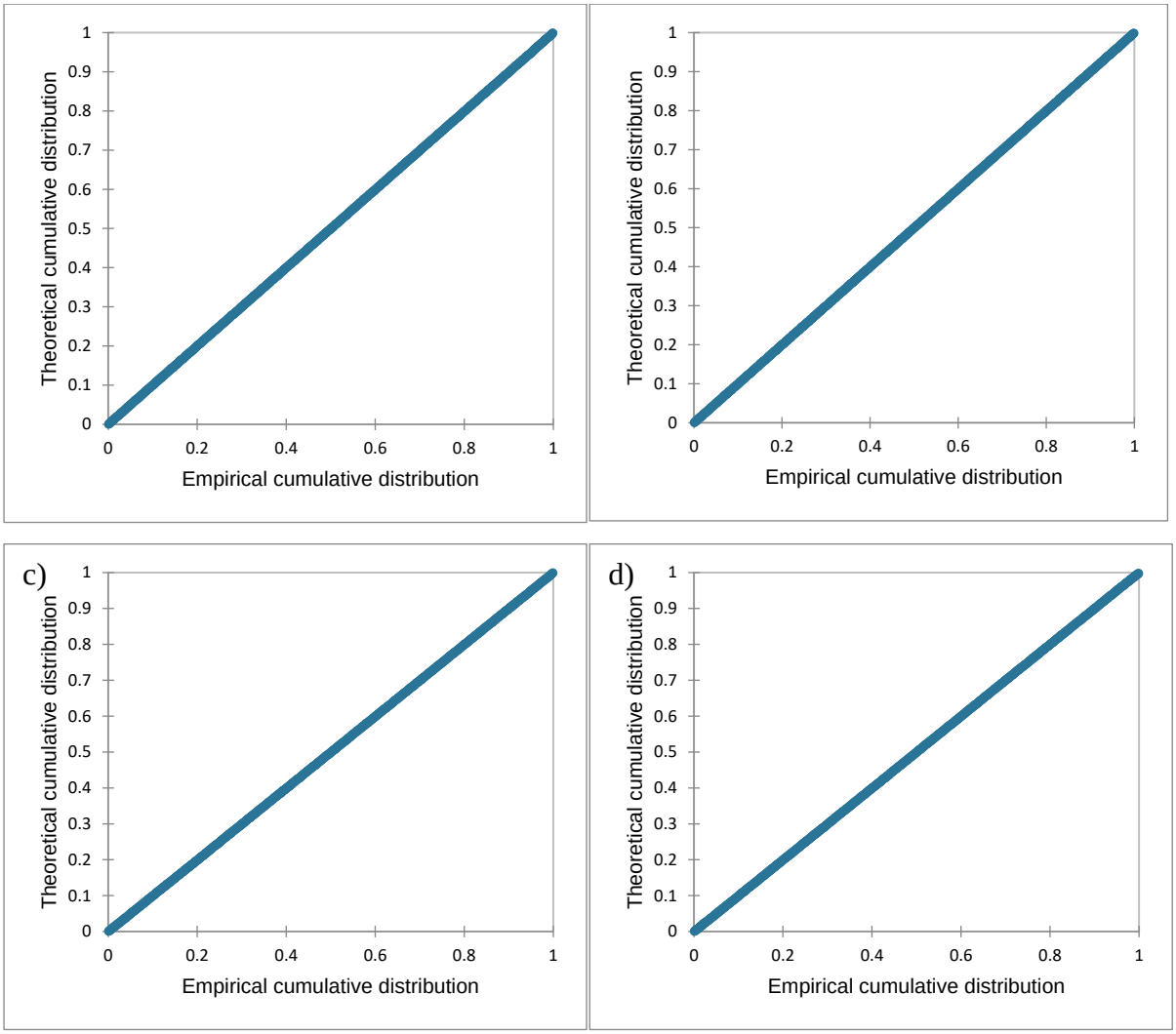
tended to underestimate the values of IC concentration removed from the aqueous solution. On the other hand, ANN-GA-SVM-optimised models demonstrated a well-balanced sensitivity impact compared to its counterparts, indicating that it tends to predict the target IC concentration more accurately.

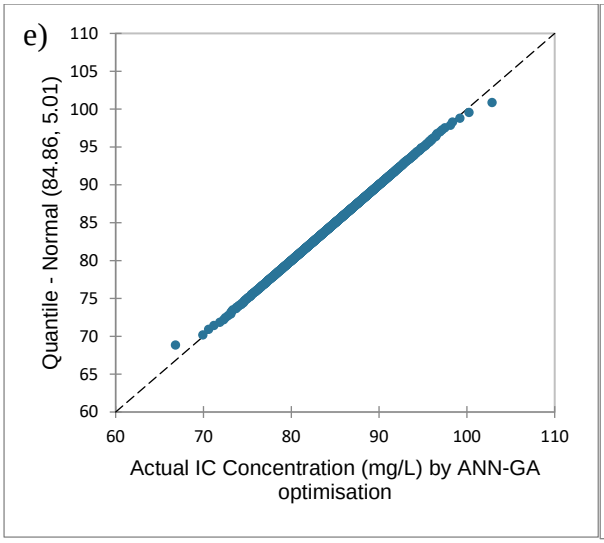
Furthermore, Table 5 represents the Spearman correlation matrix, which measures the monotonic association between the variables regarding ranks. It measures the effect that increasing one variable has on other variables. Occasionally, the relationship among the variables is non-linear or bivariate normal. If the Spearman correlation coefficient is 1, it indicates a strong correlation between the variables, whereas 0 indicates a neutral relationship, and -1 indicates no significant correlation. For example, the simulation results from three AI/ML optimised models showed that there is a slight negative coefficient of -0.063 between the predicted results between ANN-GA-XGBoost and ANN-GA-SVM optimisation techniques, which indicated that these optimisation methods yielded the best prediction result relative to ANN-GA and ANN-GA-RF ensembles. However, there was a less positive correlation coefficient of 0.694 and 0.726 between the predicted results obtained between ANN-GA versus ANN-GA-RF and ANN-GA versus ANN-GA-XGBoost, indicating that when ANN-GA-RF and ANN-GA-XGBoost optimisation methods were used in combination resulted in higher positive deviations between the actual and predicted datasets, overshooting the targeted values of IC concentration. In addition, Table S1 in the Supplementary Material shows further detail on the simulation results of distributions and result variables. After simulations, Table S1 results validated that the ANN-GA-SVM algorithm produced the most accurate prediction compared to ANN-GA, ANN-GA-XGBoost and ANN-GA-RF algorithms based on its prediction efficiency. In addition, Figure 9 represents the Monte Carlo simulation details of uncertainty estimation and prediction efficiency of IC removal efficiency in aqueous systems. Improvement in skewness or heavy-

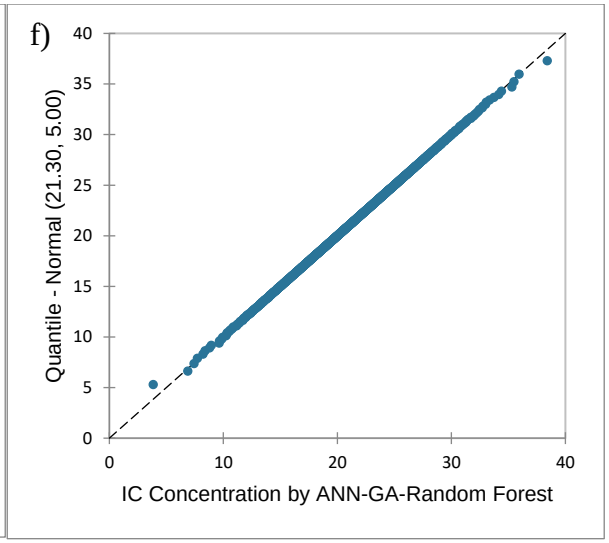
tailed distributions in Figure 9 regression models indicated strong predictive density forecasts of IC adsorption process.

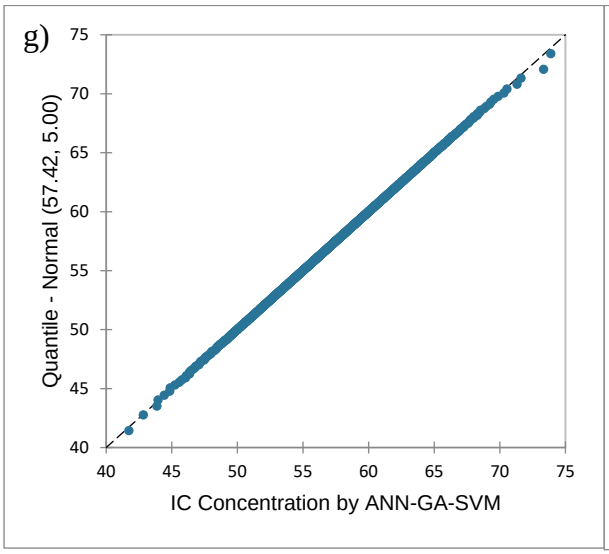
Table 2
Comparison between the prediction efficiencies for removal efficiency of xenobiotic IC dye using Bone Char at 30°C

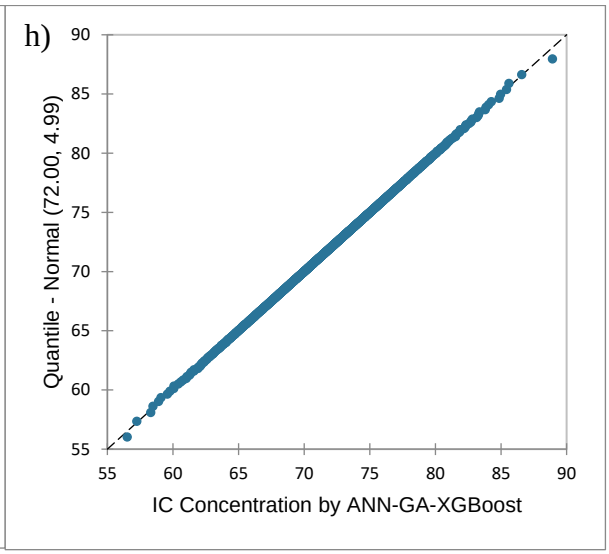
Contact Time in Mins	Experimental IC Concentration (mg/L)	ANN-GA Optimisation (mg/L)	ANN-GA-SVM optimisation (mg/L)	ANN-GA-RF optimisation (mg/L)	ANN-GA-XGBoost optimisation (mg/L)
0	30.980	29.802	12.746	7.488	19.633
5	10.392	10.861	9.943	12.436	19.633
10	4.471	5.413	6.933	10.242	5.498
15	4.078	5.052	5.124	6.105	5.498
20	4.706	5.629	4.712	5.560	5.498
25	4.549	5.485	4.997	5.368	5.498
30	3.490	4.511	5.428	5.893	5.498

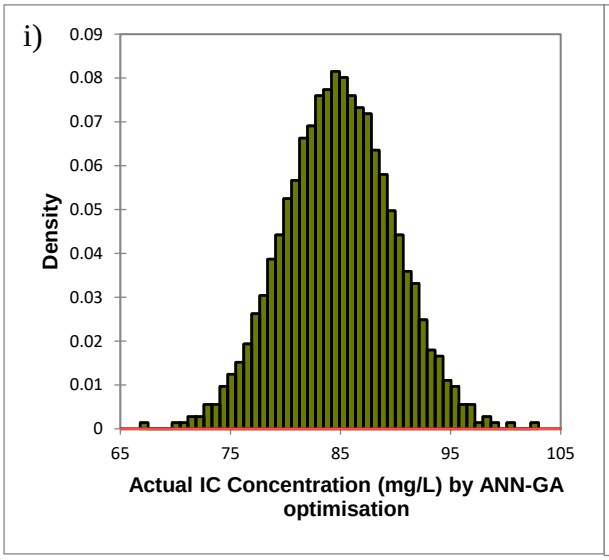


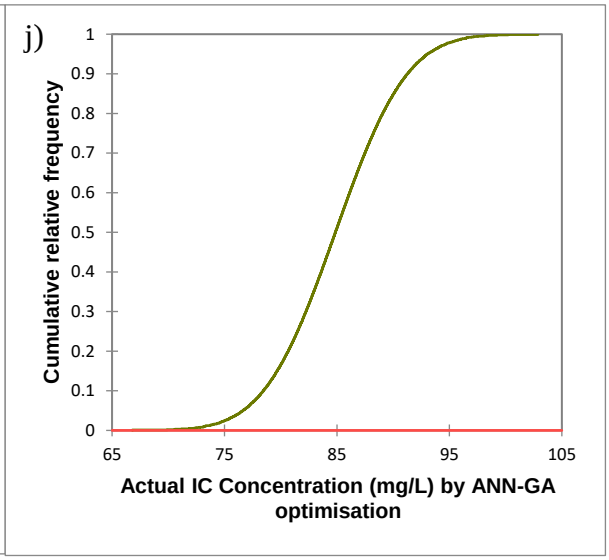

e)
Quantile - Normal (84.86, 5.01)
Actual IC Concentration (mg/L) by ANN-GA optimisation

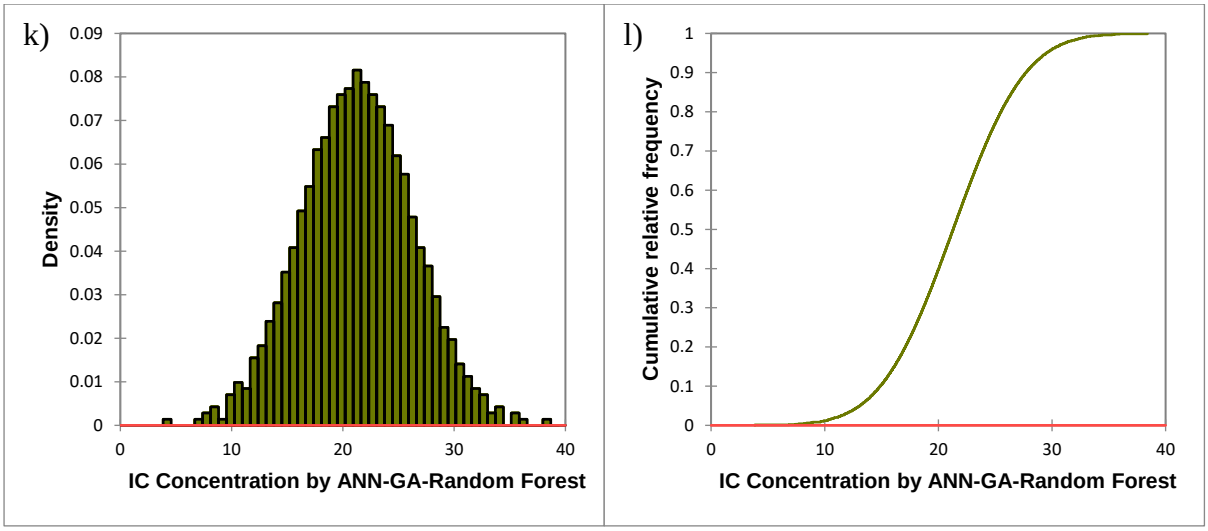

f)
Quantile - Normal (21.30, 5.00)
IC Concentration by ANN-GA-Random Forest

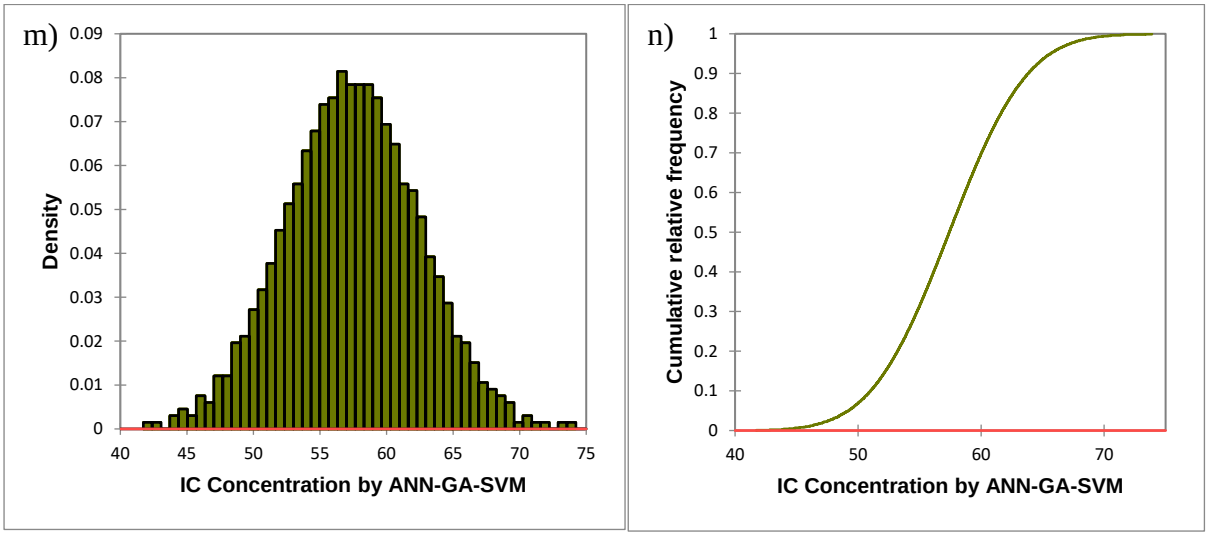

g)
Quantile - Normal (57.42, 5.00)
IC Concentration by ANN-GA-SVM

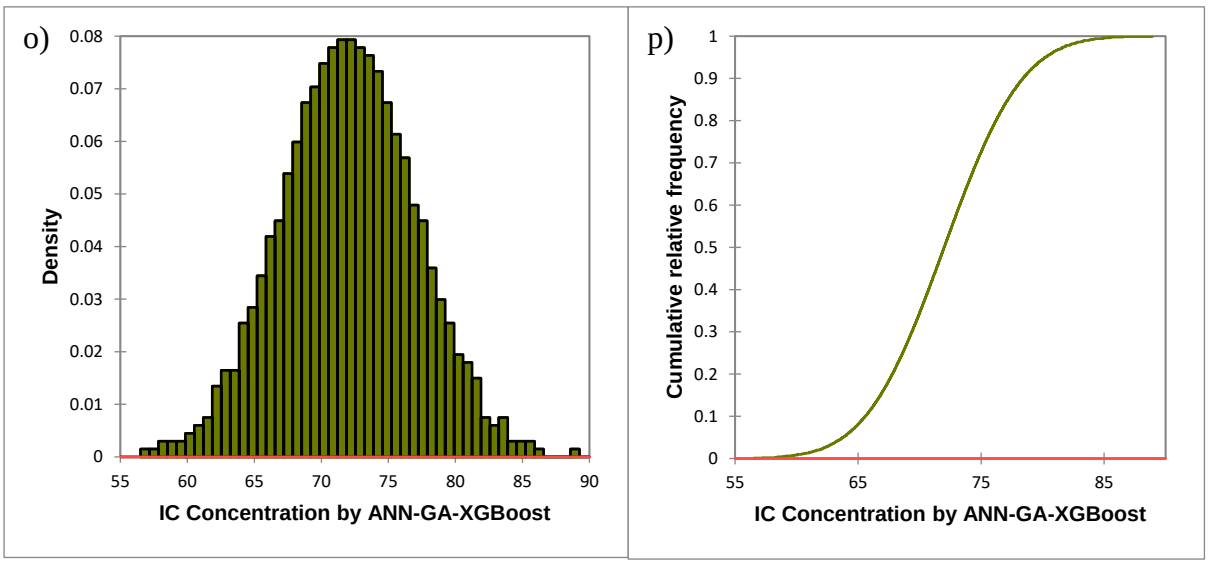

h)
Quantile - Normal (72.00, 4.99)
IC Concentration by ANN-GA-XGBoost


i)
Density
Actual IC Concentration (mg/L) by ANN-GA optimisation


j)
Cumulative relative frequency
Actual IC Concentration (mg/L) by ANN-GA optimisation


k)
Density
IC Concentration by ANN-GA-Random Forest
l)
Cumulative relative frequency
IC Concentration by ANN-GA-Random Forest


m)
Density
IC Concentration by ANN-GA-SVM
n)
Cumulative relative frequency
IC Concentration by ANN-GA-SVM


o)
Density
IC Concentration by ANN-GA-XGBoost
p)
Cumulative relative frequency
IC Concentration by ANN-GA-XGBoost

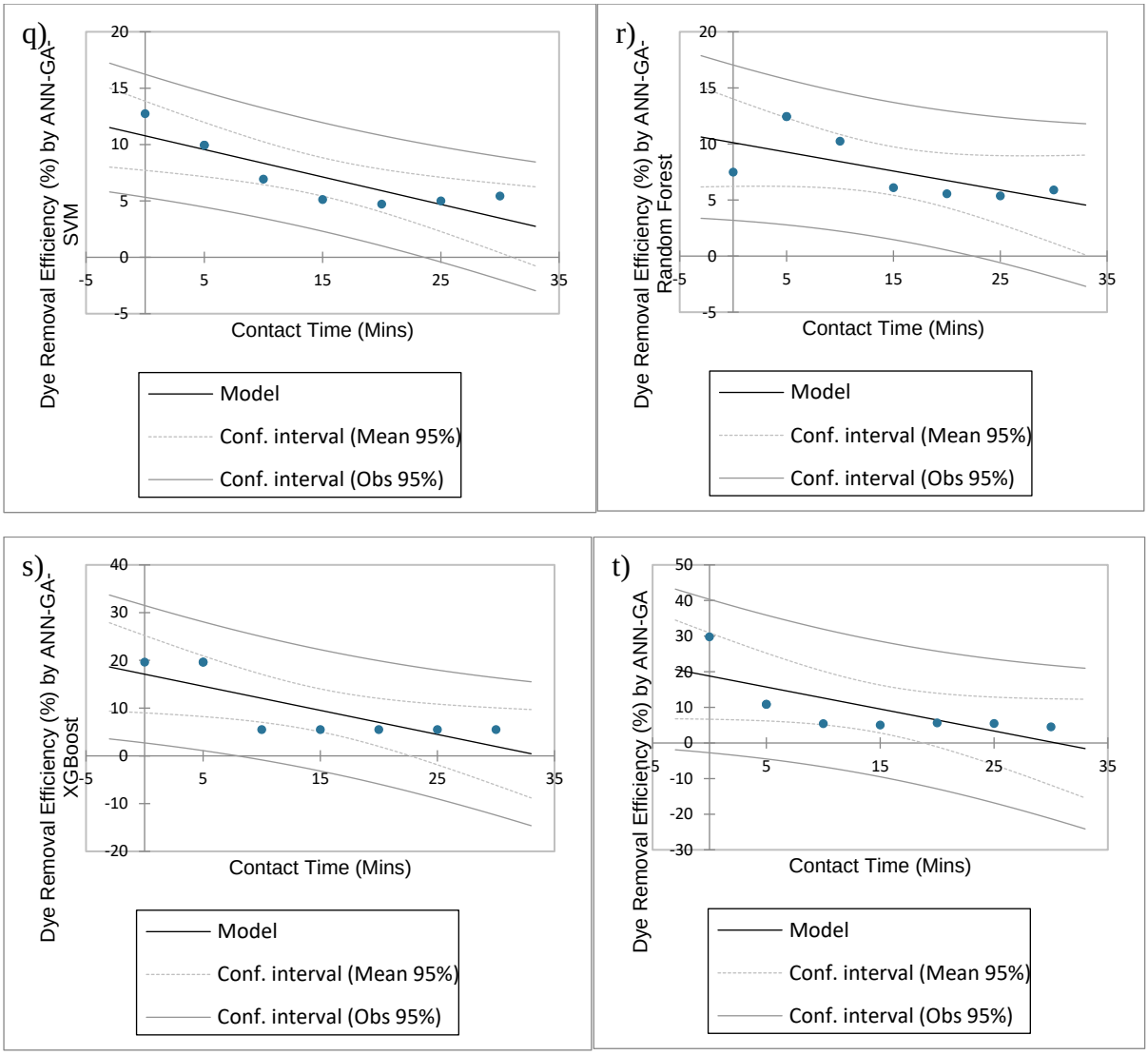


Figure 7. Monte Carlo simulations are used to predict the probability of IC removal outcomes based on the random variables to evaluate the level of uncertainties in prediction models; a) P-P plot based on actual IC aqueous concentration by ANN-GA optimisation; b) P-P plot based on predicted final IC aqueous concentration by ANN-GA-RF optimisation; c) P-P plot based on predicted final IC aqueous concentration by ANN-GA-SVM optimisation; d) P-P plot based on predicted final IC aqueous concentration by ANN-GA-XGBoost optimisation; e) Q-Q plot based on actual IC concentration by ANN-GA optimisation; f) Q-Q plot based on predicted IC aqueous concentration by ANN-GA-RF optimisation; g) Q-Q plot based on predicted IC concentration removed by ANN-GA-SVM optimisation; h) Q-Q plot based on predicted IC aqueous concentration by ANN-GA-XGBoost; i) Histogram of actual final IC aqueous concentration by ANN-GA optimisation; j) Empirical cumulative distribution of actual IC concentration by ANN-GA optimisation; k) Histogram of predicted IC aqueous concentration by ANN-GA-RF optimisation; l) Empirical cumulative distribution of predicted IC aqueous concentration by ANN-GA-RF optimisation; m) Histogram of predicted IC aqueous concentration by ANN-GA-SVM optimisation; n) Empirical cumulative distribution of predicted final IC aqueous concentration by ANN-GA-SVM optimisation; o) Histogram of predicted final IC aqueous concentration by ANN-GA-XGBoost optimisation; p) Empirical cumulative distribution of predicted final IC aqueous concentration by ANN-GA-XGBoost optimisation; q) ANOVA regression analysis of IC removal efficiency versus contact time by ANN-GA-SVM; r) ANOVA regression analysis of

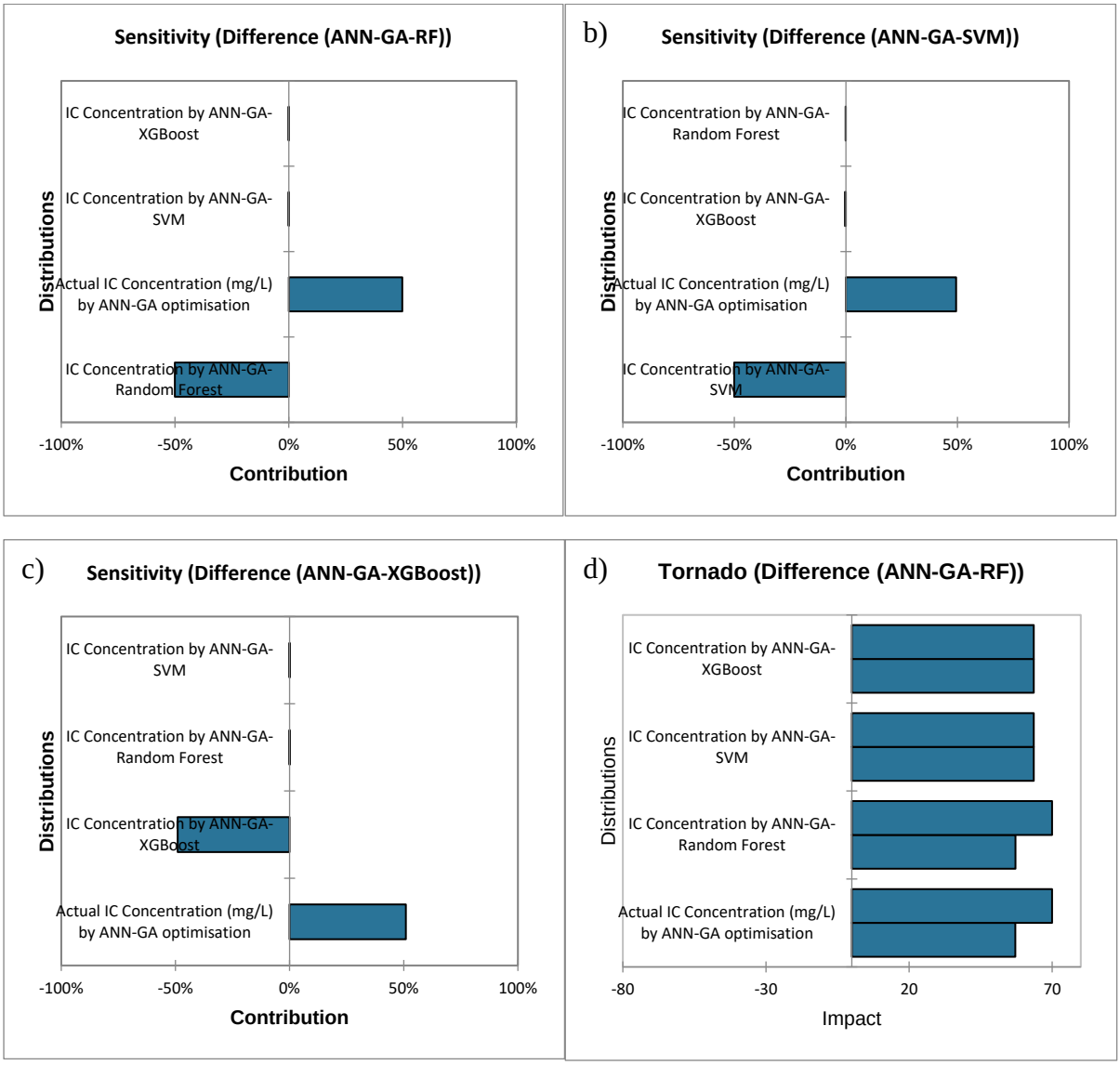
IC removal efficiency versus contact time by ANN-GA-RF; s) ANOVA regression analysis of IC removal efficiency versus contact time by ANN-GA-XGBoost; t) ANOVA regression analysis of IC removal efficiency versus contact time by ANN-GA.

Table 3. Summary of simulation results using descriptive statistics to evaluate the uncertainties in predicting IC removal efficiencies by bone char adsorbent.

Statistic	Actual IC Concentration (mg/L) by ANN-GA optimisation	IC Concentration by ANN-GA-Random Forest	IC Concentration by ANN-GA-SVM	IC Concentration by ANN-GA-XGBoost
Nbr. of observations	1000	1000	1000	1000
Nbr. of missing values	0	0	0	0
Sum of weights	1000	1000	1000	1000
Minimum	66.802	3.846	41.726	56.510
Maximum	102.872	38.401	73.882	88.923
Freq. of minimum	1	1	1	1
Freq. of maximum	1	1	1	1
Range	36.070	34.555	32.157	32.413
1st Quartile	81.489	17.928	54.041	68.628
Median	84.866	21.309	57.414	71.989
3rd Quartile	88.233	24.663	60.777	75.364
Sum	84863.967	21301.449	57421.236	71997.289
Mean	84.864	21.301	57.421	71.997
Variance (n)	25.097	25.025	24.966	24.924
Variance (n-1)	25.123	25.050	24.991	24.949
Standard deviation (n)	5.010	5.002	4.997	4.992
Standard deviation (n-1)	5.012	5.005	4.999	4.995
Variation coefficient	0.059	0.235	0.087	0.069
Skewness (Pearson)	0.002	0.001	0.018	0.004
Skewness	0.002	0.001	0.018	0.004
Skewness (Bowley)	-0.002	-0.004	-0.001	0.002
Kurtosis (Pearson)	0.065	0.009	-0.028	-0.047
Kurtosis	0.071	0.015	-0.023	-0.042
Standard error of the mean	0.159	0.158	0.158	0.158
Lower bound on mean (XXXX%)	84.553	20.991	57.111	71.687
Upper bound on mean (XXXX%)	85.175	21.612	57.731	72.307
Standard error (Skewness (Fisher))	0.077	0.077	0.077	0.077
Standard error (Kurtosis (Fisher))	0.155	0.155	0.155	0.155
Mean absolute deviation	3.992	3.990	3.989	3.987
Median absolute deviation	3.375	3.376	3.379	3.368
Geometric mean	84.715	20.662	57.202	71.823
Geometric standard deviation	1.061	1.292	1.092	1.072
Harmonic mean	84.566	19.928	56.980	71.648

Table 4. Error function analysis of prediction efficiencies by various AI/ML ensembles.

	ANN-GA-SVM	ANN-GA-RF	ANN-GA-XGBoost	ANN-GA
R^2	0.730	0.445	0.625	0.526
Adjusted R^2	0.676	0.334	0.550	0.432
MSE	3.072	4.955	21.407	47.967
RMSE	1.753	2.226	4.627	6.926
MAPE	20.884	21.117	48.800	64.370



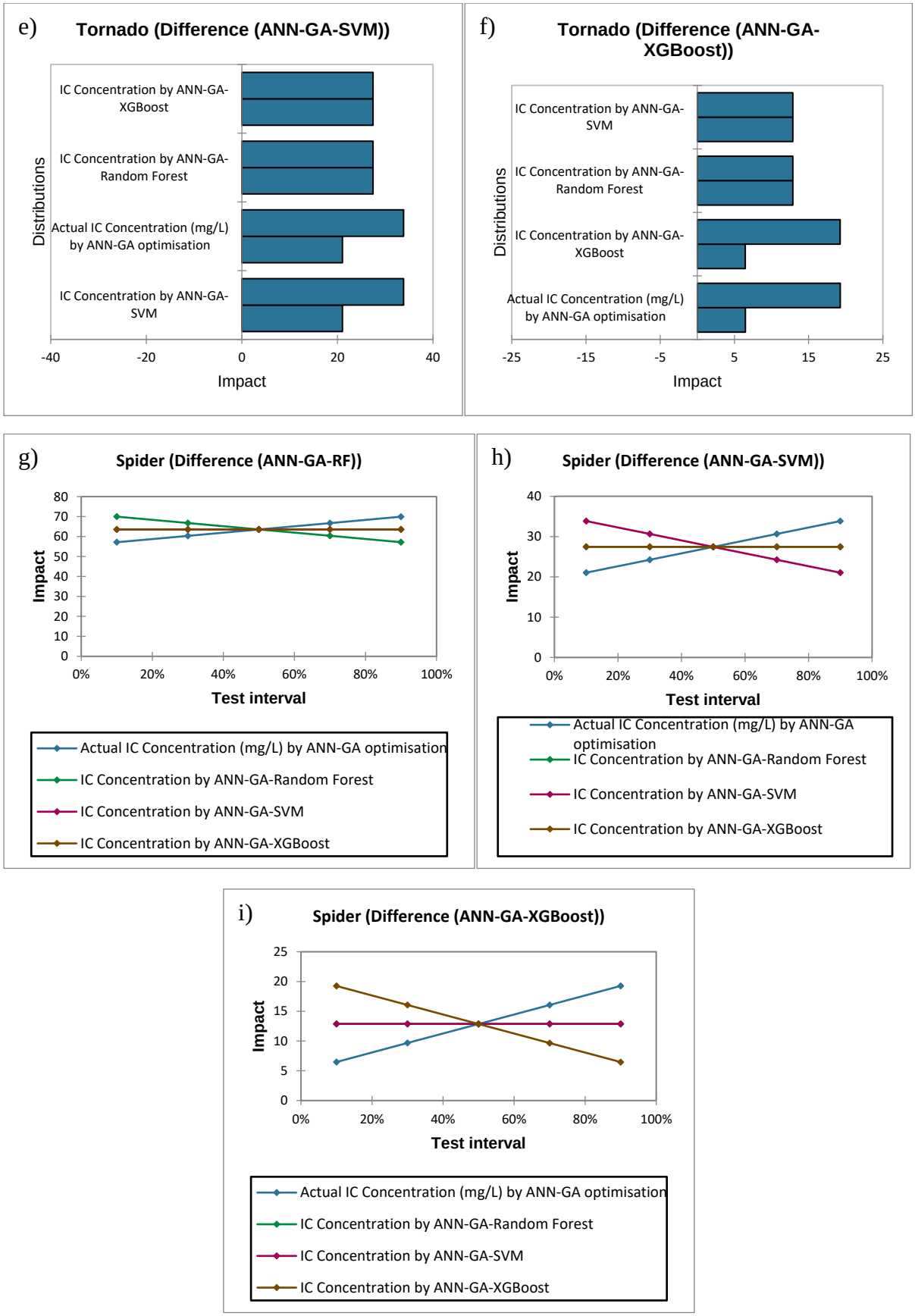


Figure 8. Monte Carlo simulation models based on the empirical distribution of input and output variables employing repeated simulation of random variables to determine the

uncertainties; a) Sensitivity analysis of the difference between the actual and predicted IC removal efficiencies by ANN-GA-RF; b) Sensitivity analysis of the difference between the actual and predicted IC removal efficiencies by ANN-GA-SVM; c) Sensitivity analysis of the difference between the actual and predicted IC removal efficiencies by ANN-GA-XGBoost; d) Tornado analysis of the difference between the actual and predicted IC removal efficiencies by ANN-GA-RF; e) Tornado analysis of the difference between the actual and predicted IC removal efficiencies by ANN-GA-SVM; f) Tornado analysis of difference between the actual and predicted IC removal efficiencies by ANN-GA-XGBoost; g) Spider analysis of difference between the actual and predicted IC removal efficiencies by ANN-GA-RF; h) Spider analysis of difference between the actual and predicted IC removal efficiencies by ANN-GA-SVM; i) Spider analysis of difference between the actual and predicted IC removal efficiencies by ANN-GA-XGBoost.

Table 5. Simulation results by Spearman correlation matrix.

Variables	Actual IC Concentration (mg/L) by ANN-GA optimisation	IC Concentration by ANN-GA-Random Forest	IC Concentration by ANN-GA-SVM	IC Concentration by ANN-GA-XGBoost	Difference by ANN-GA-RF	Difference by ANN-GA-SVM	Difference by ANN-GA-XGBoost
Actual IC Concentration (mg/L) by ANN-GA optimisation	**1**	0.013	0.013	**-0.077**	**0.694**	**0.682**	**0.726**
IC Concentration by ANN-GA-Random Forest	-0.013	**1**	0.036	-0.037	**-0.696**	-0.037	0.018
IC Concentration by ANN-GA-SVM	0.013	0.036	**1**	0.017	-0.022	**-0.687**	0.002
IC Concentration by ANN-GA-XGBoost	**-0.077**	-0.037	0.017	**1**	-0.015	**-0.063**	**-0.712**
Difference (ANN-GA-RF)	**0.694**	**-0.696**	-0.022	-0.015	**1**	**0.505**	**0.485**
Difference (ANN-GA-SVM)	**0.682**	-0.037	**-0.687**	**-0.063**	**0.505**	**1**	**0.504**
Difference (ANN-GA-XGBoost)	**0.726**	0.018	0.002	**-0.712**	**0.485**	**0.504**	**1**

Please note that hybrid AI/ML differences represent the difference between the actual and predicted IC removal efficiencies by hybrid algorithms.

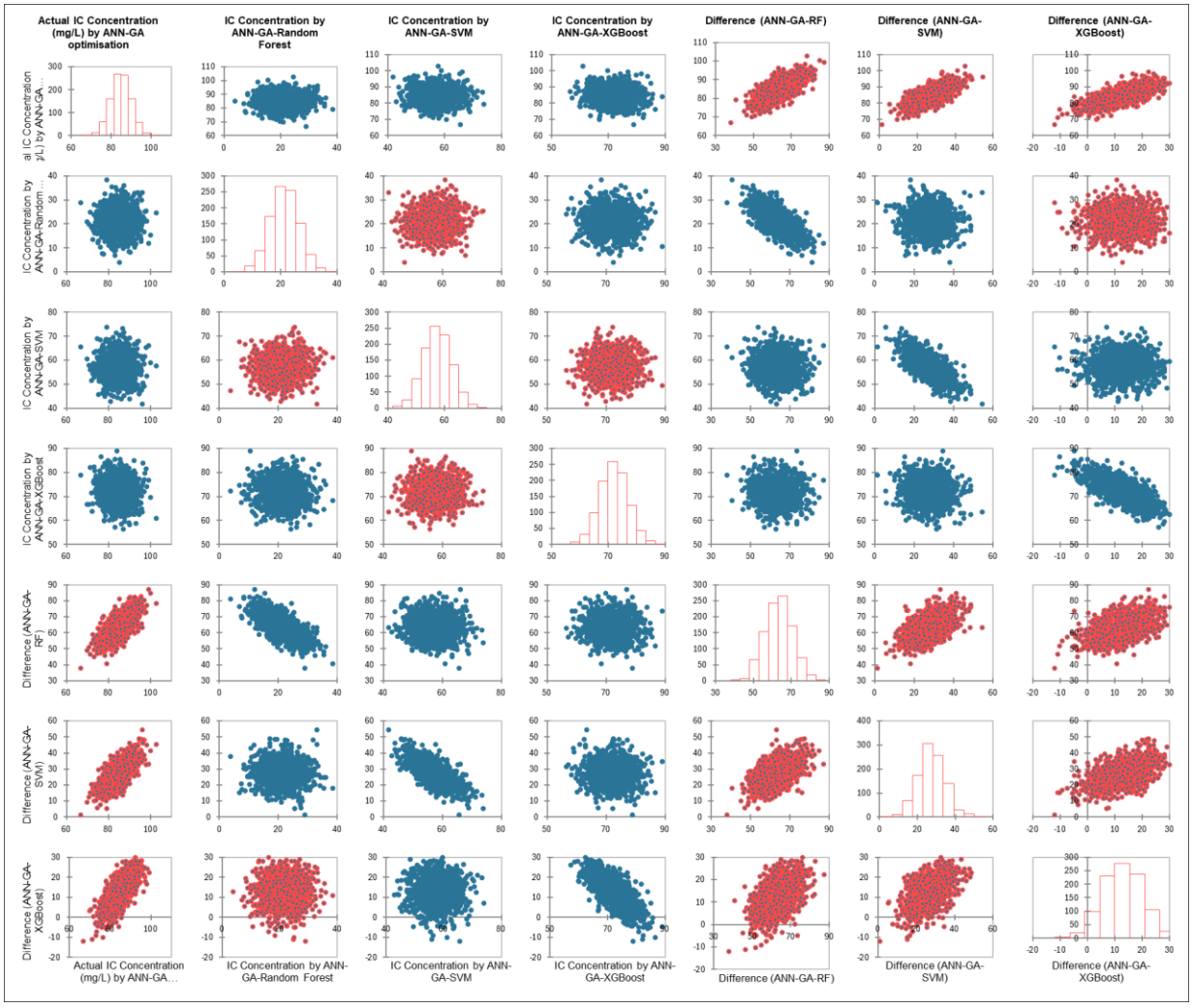


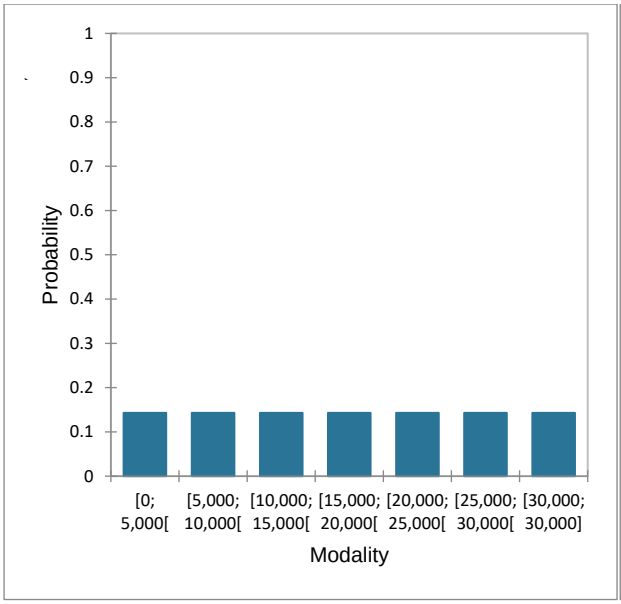
Figure 9. Monte Carlo simulation details are represented by scatter plots and correlation maps.

3.6. Bayesian distribution network analysis

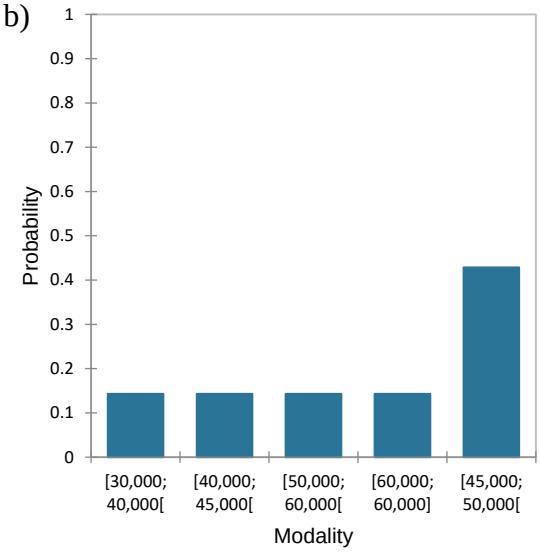
A Bayesian network analysis involves a mathematical model for representing a causal relationship between random variables by estimating the conditional probability outcome (Wang et al., 2022). Bayesian network is an appropriate tool for evaluating the uncertainty with real applications. Through representation, a Bayesian network projects a probabilistic graphical model to illustrate knowledge about an uncertain domain. Each node corresponds to a random variable, and each edge represents a conditional probability for the corresponding random variable (Cui et al., 2024). The Bayesian approach helps to identify the causal relationship with the IC adsorption process by estimating the conditional probability. Figure 10a shows that one of the random

variables, contact time, has approximately 15% conditional probability based on its impact on the IC adsorption from 0 to 30 mins.

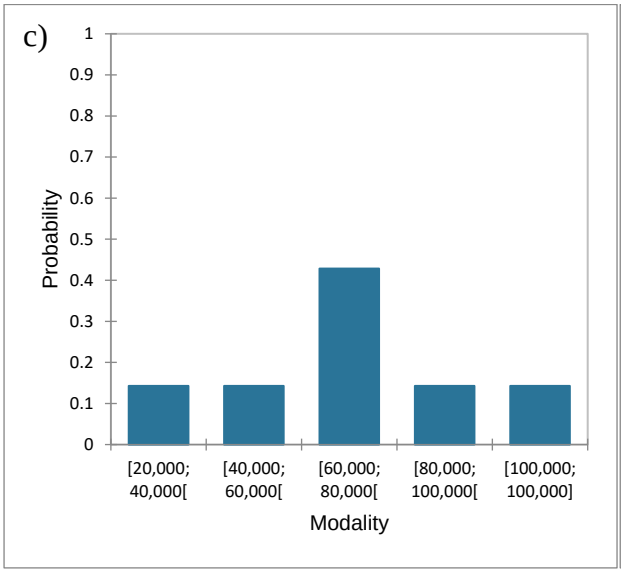
On the other hand, Figure 10b shows that a temperature range of 45 to 50ºC yielded approximately 43% conditional probability compared to all different temperature ranges, indicating that this temperature range may significantly impact IC removal efficiency. This result corresponded to Figure 3 of Gibbs free energy ranging between 1.800 and 1.924 kJ/mol, indicating an optimal condition with a high probability of an IC removal event occurring at the particular temperature range. Figure 10c shows the highest probability of an IC removal event occurring at an initial IC concentration of 60 to 80 mg/L. In contrast, other random variables, such as adsorbent loading (Figure 10d) and IC removal efficiency (Figure 10e) yielded a conditional probability of 15%, indicating no significant causal relationship between random variables with IC removal efficiency other than solution temperature and initial IC concentration at a specific range. Although the results from Bayesian network analysis showed that adsorbent loading and contact time may have a marginal probabilistic effect on the IC removal efficiency, solution temperature and initial IC concentration at a specific range have more impact on the IC removal rate which could not be estimated from other data analyses other than using Bayesian approach. In addition, Table 6 summarises Bayesian results that determine the causal relationships among random variables by using conditional probabilities to represent knowledge about the uncertainty in the variables. According to Table 6, the Bayesian network analysis shows that solution temperature between 45 and 50°C and initial IC concentration of 60 to 80 mg/L significantly impacted the IC removal rate compared to other variables. The Bayesian results also indicated that there could be some uncertainties in other AI/ML optimisation techniques, which may affect the prediction efficiency. In addition, Table 7 provides comparative analysis of AI/ML predictive efficiency between primary and secondary sources.


Probability
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
[0; 5,000[
[5,000; 10,000[
[10,000; 15,000[
[15,000; 20,000[
[20,000; 25,000[
[25,000; 30,000[
[30,000; 30,000]
Modality

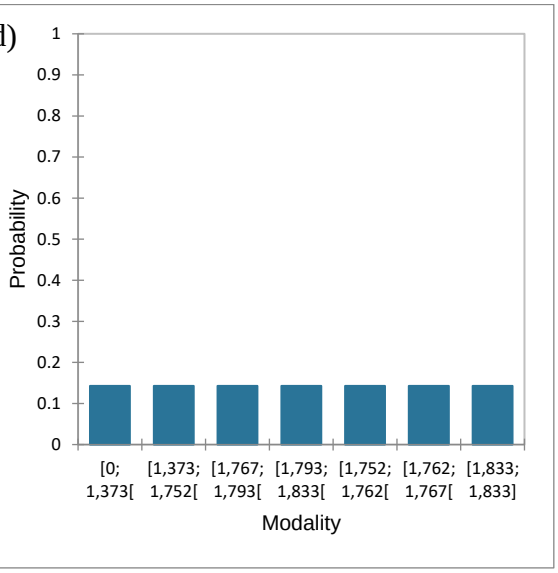
b)


Probability
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
[30,000; 40,000[
[40,000; 45,000[
[50,000; 60,000[
[60,000; 60,000]
[45,000; 50,000[
Modality

c)


Probability
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
[20,000; 40,000[
[40,000; 60,000[
[60,000; 80,000[
[80,000; 100,000[
[100,000; 100,000]
Modality

d)


Probability
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
[0; 1,373[
[1,373; 1,752[
[1,767; 1,793[
[1,793; 1,833[
[1,752; 1,762[
[1,762; 1,767[
[1,833; 1,833]
Modality

e)

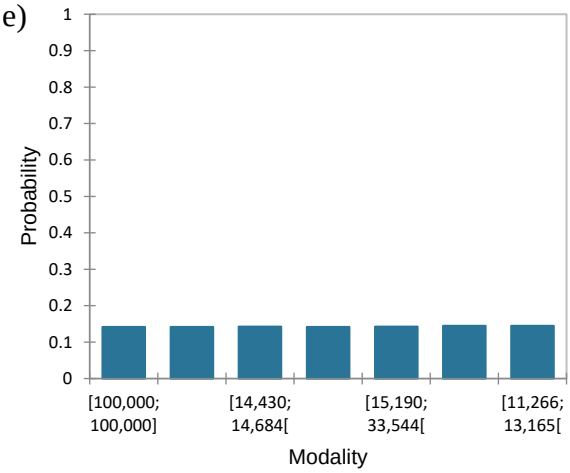

Probability
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
[100,000; 100,000]
[14,430; 14,684[
[15,190; 33,544[
[11,266; 13,165[
Modality

Figure 10. Bayesian distribution network analysis is a statistical tool based on an acyclic-oriented graph and a probability table, which is commonly used in artificial intelligence to represent data and its uncertainties: a) Marginal probability distribution of the contact time node; b) Marginal probability distribution of the solution temperature node; c) Marginal probability distribution of the initial IC concentration node; d) marginal probability distribution of the adsorbent loading node; e) Marginal probability distribution of the IC removal efficiency.

Table 6

Summary of causal relationships among random variables by using conditional probabilities to represent knowledge about the uncertainty in the variables.

Contact Time in Mins		Solution Temperature (°C)		Initial Dye Concentration (mg/L)		Adsorbent Loading (mg/g)		Dye Removal Efficiency (%)	
Modality	Probability	Modality	Probability	Modality	Probability	Modality	Probability	Modality	Probability
[0; 5,000[	0.143	[30,000; 40,000[	0.143	[20,000; 40,000[	0.143	[0; 1,373[	0.143	[100,000; 100,000]	0.142
[5,000; 10,000[	0.143	[40,000; 45,000[	0.143	[40,000; 60,000[	0.143	[1,373; 1,752[	0.143	[33,544; 100,000[	0.142
[10,000; 15,000[	0.143	[50,000; 60,000[	0.143	[60,000; 80,000[	0.429	[1,767; 1,793[	0.143	[14,430; 14,684[	0.143
[15,000; 20,000[	0.143	[60,000; 60,000]	0.143	[80,000; 100,000[	0.143	[1,793; 1,833[	0.143	[13,165; 14,430[	0.142
[20,000; 25,000[	0.143	[45,000; 50,000[	0.429	[100,000; 100,000]	0.143	[1,752; 1,762[	0.143	[15,190; 33,544[	0.143
[25,000; 30,000[	0.143	[30,000; 40,000[	0.143			[1,762; 1,767[	0.143	[14,684; 15,190[	0.145
[30,000; 30,000]	0.143	[40,000; 45,000[	0.143			[1,833; 1,833]	0.143	[11,266; 13,165[	0.145

Table 7

Comparison of AI/ML prediction efficiency between experimental results and secondary sources from literature.

Type of adsorbent	Type of pollutant	Operating Conditions	Prediction efficiency	Reference
Polymer/graphene/clay/MgFeAl-LTH nanocomposite	Methyl orange and crystal violet	C_0 = 5 to 500 mg/L C_t = 5 mins to 20 hours	R^2 (RF) = 0.92 MSE (RF)= 6636.84 RMSE (RF) = 81.47 MAE (RF) = 57.57	BinMakhashen et al. (2024)
Sugarcane-derived carbon dots and TiO2 based chitosan composite	Methyl red, Brilliant green	C_0 = 50 to 1,000 mg/L	R^2 (SVM-RBF kernal function) = 0.7419	Momina et al. (2024)

		C_t = 5-60 mins and 10-240 mins Adsorbent mass = 0.02-0.05 g pH = 2-10 Temperature = 30-60°C Stirring speed = 60, 80, 100 rpm	R^2 (SVM-polynomial) = 0.5576	
ZIF-60	Crystal violet	C_0 = 200 mg/L C_t = >24 hours Adsorbent dosage = 0.025 g/L pH = 8.0 Temperature = 40°C	R^2 (SVM) = 0.9812 RAE (%) (SVM) = 15.94 RRSE (%) (SVM) = 19.92	Ismail et al. (2024)

C_0 = Initial dye concentration; C_t = Contact time; MAE = Mean absolute error; MSE = Mean squared error; RAE = Relative absolute error; RF = Random Forest; RMSE = Root mean squared error; RRSE = Root relative squared error; SVM = Support vector machine

In this study, we developed Monte Carlo simulations, Bayesian networks and AI/ML optimisation techniques to evaluate the prediction efficiency of bone char adsorption process using IC as a model pollutant. The data analyses demonstrated that the proposed approaches could improve the prediction efficiency and estimation accuracy using computational modelling. Although the model parameters were deduced to impact the IC adsorption process directly, solution temperature and initial IC concentration may have a more significant impact on the adsorption efficiency of bone char, especially when a specific optimal condition was satisfied. The prediction efficiency of ANFIS optimisation yielded appropriate curve fitting but with minor uncertainty. On the other hand, AI/ML optimisation techniques had some

uncertainties, especially for ANN-GA-RF, ANN-GA-XGBoost and ANN-GA due to overfitting or inability to predict noisy regression problems. In contrast, the ANFIS and ANN-GA-SVM algorithms yielded minimal prediction error and converged more efficiently with high accuracy. Therefore, the prediction efficiency of AI/ML algorithms ranked in the order: ANN-GA-SVM > ANFIS > ANN-GA-XGBoost > ANN-GA-RF > ANN-GA. However, some issues need to be considered in future work. Increasing noises or uncertainties in AI/ML optimisation techniques often hamper estimation accuracy. More challenging upscaled industrial processes require sizeable computational processing power to model the adsorption efficiency of bone char in a packed bed reactor. Model validation tests are needed to verify whether the computational models are congruent with the upscaled adsorption process, especially when there are multiple inlets, outlets and recycling streams, solution thermodynamics and the overall impact of physicochemical phenomena on the industrial adsorption process.

Statements & Declarations

Funding

Not applicable

Conflict of interest/Competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethics approval

Not applicable

Consent to participate

Not applicable

Consent for publication

All authors confirm their assent to publication.

Availability of data and material

Additional data will be provided upon request.

Code availability

Not applicable

List of References

Abujazar, M. S. S., Karaağaç, S. U., Abu Amr, S. S., Alazaiza, M. Y. D., & Bashir, M. J. K. (2022). Recent advancement in the application of hybrid coagulants in coagulation-flocculation of wastewater: A review. Journal of Cleaner Production, 345, 131133. https://doi.org/https://doi.org/10.1016/j.jclepro.2022.131133

Adel, M., Ahmed, M. A., & Mohamed, A. A. (2021). Effective removal of indigo carmine dye from wastewaters by adsorption onto mesoporous magnesium ferrite nanoparticles. Environmental Nanotechnology, Monitoring & Management, 16, 100550. https://doi.org/https://doi.org/10.1016/j.enmm.2021.100550

Al-Tohamy, R., Ali, S. S., Li, F., Okasha, K. M., Mahmoud, Y. A. G., Elsamahy, T., . . . Sun, J. (2022). A critical review on the treatment of dye-containing wastewater: Ecotoxicological and health concerns of textile dyes and possible remediation approaches for environmental safety. Ecotoxicology and Environmental Safety, 231, 113160-113160. https://doi.org/10.1016/j.ecoenv.2021.113160

Alikarami, S., Soltanizade, A., & Rashchi, F. (2022). Synthesis of CdS–SnS photocatalyst by chemical co-precipitation for photocatalytic degradation of methylene blue and rhodamine B under irradiation by visible light. Journal of Physics and Chemistry of Solids, 171, 110993. https://doi.org/https://doi.org/10.1016/j.jpcs.2022.110993

Aqdam, S. R., Otzen, D. E., Mahmoodi, N. M., & Morshedi, D. (2021). Adsorption of azo dyes by a novel bio-nanocomposite based on whey protein nanofibrils and nano-clay: Equilibrium isotherm and kinetic modeling. Journal of colloid and interface science, 602, 490-503. https://doi.org/https://doi.org/10.1016/j.jcis.2021.05.174

Arslan-Alaton, I., Tureli, G., & Olmez-Hanci, T. (2009). Treatment of azo dye production wastewaters using Photo-Fenton-like advanced oxidation processes: Optimization by response surface methodology. Journal of Photochemistry and Photobiology A: Chemistry, 202(2), 142-153. https://doi.org/https://doi.org/10.1016/j.jphotochem.2008.11.019

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. Environmental Research, 184, 109367. https://doi.org/https://doi.org/10.1016/j.envres.2020.109367

Aysan, H., Edebali, S., Ozdemir, C., Celik Karakaya, M., & Karakaya, N. (2016). Use of chabazite, a naturally abundant zeolite, for the investigation of the adsorption kinetics and mechanism of methylene blue dye. Microporous and Mesoporous Materials, 235, 78-86. https://doi.org/https://doi.org/10.1016/j.micromeso.2016.08.007

Bellamoli, F., Di Iorio, M., Vian, M., & Melgani, F. (2023). Machine learning methods for anomaly classification in wastewater treatment plants. Journal of environmental management, 344, Article 118594. https://doi.org/10.1016/j.jenvman.2023.118594

Bello, M. M., Abdul Raman, A. A., & Asghar, A. (2019). A review on approaches for addressing the limitations of Fenton oxidation for recalcitrant wastewater treatment. Process safety and environmental protection, 126, 119-140. https://doi.org/10.1016/j.psep.2019.03.028

Bilal, M., Ihsanullah, I., Hassan Shah, M. U., Bhaskar Reddy, A. V., & Aminabhavi, T. M. (2022). Recent advances in the removal of dyes from wastewater using low-cost adsorbents. Journal of environmental management, 321, 115981. https://doi.org/https://doi.org/10.1016/j.jenvman.2022.115981

BinMakhashen, G. M., Bahadi, S. A., Al-Jamimi, H. A., & Onaizi, S. A. (2024). Ensemble meta machine learning for predicting the adsorption of anionic and cationic dyes from aqueous solutions using Polymer/graphene/clay/MgFeAl-LTH nanocomposite.

Chemosphere, 349, 140861. https://doi.org/https://doi.org/10.1016/j.chemosphere.2023.140861

Buyukada, M. (2016). Co-combustion of peanut hull and coal blends: Artificial neural networks modeling, particle swarm optimization and Monte Carlo simulation. Bioresource Technology, 216, 280-286. https://doi.org/https://doi.org/10.1016/j.biortech.2016.05.091

Chairungsri, W., Subkomkaew, A., Kijjanapanich, P., & Chimupala, Y. (2022). Direct dye wastewater photocatalysis using immobilized titanium dioxide on fixed substrate. Chemosphere, 286, 131762. https://doi.org/https://doi.org/10.1016/j.chemosphere.2021.131762

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. Environmental Research, 210, 112953. https://doi.org/https://doi.org/10.1016/j.envres.2022.112953

Chung, T. H., Shahidi, M., Mezbahuddin, S., & Dhar, B. R. (2023). Ensemble machine learning approach for examining critical process parameters and scale-up opportunities of microbial electrochemical systems for hydrogen peroxide production. Chemosphere (Oxford), 324, 138313-138313. https://doi.org/10.1016/j.chemosphere.2023.138313

Cui, J., Kong, Y., Liu, C., Cai, B., Khan, F., & Li, Y. (2024). Failure probability analysis of hydrogen doped pipelines based on the Bayesian network. Engineering Failure Analysis, 156, 107806. https://doi.org/https://doi.org/10.1016/j.engfailanal.2023.107806

Dadban-Shahamat, Y., Masihpour, M., Borghei, P., & Hoda-Rahmati, S. (2022). Removal of azo red-60 dye by advanced oxidation process O3/UV from textile wastewaters using Box-Behnken design. Inorganic chemistry communications, 143, 109785. https://doi.org/https://doi.org/10.1016/j.inoche.2022.109785

El-Kammah, M., Elkhatib, E., Gouveia, S., Cameselle, C., & Aboukila, E. (2022). Cost-effective ecofriendly nanoparticles for rapid and efficient indigo carmine dye removal from wastewater: Adsorption equilibrium, kinetics and mechanism. Environmental Technology & Innovation, 28, 102595. https://doi.org/https://doi.org/10.1016/j.eti.2022.102595

El Hassani, K., Kalnina, D., Turks, M., Beakou, B. H., & Anouar, A. (2019). Enhanced degradation of an azo dye by catalytic ozonation over Ni-containing layered double hydroxide nanocatalyst. Separation and purification technology, 210, 764-774. https://doi.org/10.1016/j.seppur.2018.08.074

Esteves, B. M., Rodrigues, C. S. D., Boaventura, R. A. R., Maldonado-Hódar, F. J., & Madeira, L. M. (2016). Coupling of acrylic dyeing wastewater treatment by heterogeneous Fenton oxidation in a continuous stirred tank reactor with biological degradation in a sequential batch reactor. Journal of environmental management, 166, 193-203. https://doi.org/https://doi.org/10.1016/j.jenvman.2015.10.008

Fu, J., Zhu, J., Wang, Z., Wang, Y., Wang, S., Yan, R., & Xu, Q. (2019). Highly-efficient and selective adsorption of anionic dyes onto hollow polymer microcapsules having a high surface-density of amino groups: Isotherms, kinetics, thermodynamics and mechanism. Journal of colloid and interface science, 542, 123-135. https://doi.org/https://doi.org/10.1016/j.jcis.2019.01.131

Guezzen, B., Medjahed, B., Benhelima, A., Guendouzi, A., Didi, M. A., Zidelmal, S., . . . Adjdir, M. (2023). Improved pollutant management by kinetic and Box-Behnken design analysis of HDTMA-modified bentonite's adsorption of indigo carmine dye. Journal of Industrial and Engineering Chemistry, 125, 242-258. https://doi.org/https://doi.org/10.1016/j.jiec.2023.05.034

Guo, J.-T., Wei, Y.-Q., Chen, S.-L., Sun, W., Fan, T.-T., Xu, M.-R., & Zhang, C.-C. (2022). Measurement of pore diffusion factor of porous solid materials. Petroleum Science, 19(4), 1897-1904. https://doi.org/https://doi.org/10.1016/j.petsci.2022.04.008

Hadi, S., Taheri, E., Amin, M. M., Fatehizadeh, A., & Aminabhavi, T. M. (2020). Synergistic degradation of 4-chlorophenol by persulfate and oxalic acid mixture with heterogeneous Fenton like system for wastewater treatment: Adaptive neuro-fuzzy inference systems modeling. Journal of environmental management, 268, 110678. https://doi.org/https://doi.org/10.1016/j.jenvman.2020.110678

Hassan, M. M., & Carr, C. M. (2018). A critical review on recent advancements of the removal of reactive dyes from dyehouse effluent by ion-exchange adsorbents. Chemosphere (Oxford), 209, 201-219. https://doi.org/10.1016/j.chemosphere.2018.06.043

Hassani, A., Soltani, R. D. C., Karaca, S., & Khataee, A. (2015). Preparation of montmorillonite–alginate nanobiocomposite for adsorption of a textile dye in aqueous phase: Isotherm, kinetic and experimental design approaches. Journal of Industrial and Engineering Chemistry, 21, 1197-1207. https://doi.org/https://doi.org/10.1016/j.jiec.2014.05.034

Heydari, B., Abdollahzadeh Sharghi, E., Rafiee, S., & Mohtasebi, S. S. (2021). Use of artificial neural network and adaptive neuro-fuzzy inference system for prediction of biogas production from spearmint essential oil wastewater treatment in up-flow anaerobic sludge blanket reactor. FUEL, 306, 121734. https://doi.org/https://doi.org/10.1016/j.fuel.2021.121734

Hu, E., Wu, X., Shang, S., Tao, X.-m., Jiang, S.-x., & Gan, L. (2016). Catalytic ozonation of simulated textile dyeing wastewater using mesoporous carbon aerogel supported copper oxide catalyst. Journal of Cleaner Production, 112, 4710-4718. https://doi.org/https://doi.org/10.1016/j.jclepro.2015.06.127

Ighnih, H., Haounati, R., Ouachtak, H., Regti, A., El Ibrahimi, B., Hafid, N., . . . Addi, A. A. (2023). Efficient removal of hazardous dye from aqueous solutions using magnetic kaolinite nanocomposite: Experimental and Monte Carlo simulation studies. Inorganic chemistry communications, 153, 110886. https://doi.org/https://doi.org/10.1016/j.inoche.2023.110886

Ismail, U. M., Onaizi, S. A., & Vohra, M. S. (2024). Crystal violet removal using ZIF-60: Batch adsorption studies, mechanistic & machine learning modeling. Environmental Technology & Innovation, 33, 103456. https://doi.org/https://doi.org/10.1016/j.eti.2023.103456

Januário, E. F. D., Vidovix, T. B., Bergamasco, R., & Vieira, A. M. S. (2021). Performance of a hybrid coagulation/flocculation process followed by modified microfiltration membranes for the removal of solophenyl blue dye. Chemical Engineering and Processing - Process Intensification, 168, 108577. https://doi.org/https://doi.org/10.1016/j.cep.2021.108577

Karaboga, D., & Kaya, E. (2019). Adaptive network based fuzzy inference system (ANFIS) training approaches: a comprehensive survey. ARTIFICIAL INTELLIGENCE REVIEW, 52(4), 2263-2293. https://doi.org/10.1007/s10462-017-9610-2

Kekes, T., & Tzia, C. (2020). Adsorption of indigo carmine on functional chitosan and β-cyclodextrin/chitosan beads: Equilibrium, kinetics and mechanism studies. Journal of environmental management, 262, 110372. https://doi.org/https://doi.org/10.1016/j.jenvman.2020.110372

Kuo, C.-Y., Wu, C.-H., & Wu, J.-Y. (2008). Adsorption of direct dyes from aqueous solutions by carbon nanotubes: Determination of equilibrium, kinetics and thermodynamics parameters. Journal of colloid and interface science, 327(2), 308-315. https://doi.org/https://doi.org/10.1016/j.jcis.2008.08.038

Li, R., Gao, B., Guo, K., Yue, Q., Zheng, H., & Wang, Y. (2017). Effects of papermaking sludge-based polymer on coagulation behavior in the disperse and reactive dyes wastewater treatment. Bioresource Technology, 240, 59-67. https://doi.org/https://doi.org/10.1016/j.biortech.2017.02.088

Lu, C., Yang, J., Khan, A., Yang, J., Li, Q., & Wang, G. (2022). A highly efficient technique to simultaneously remove acidic and basic dyes using magnetic ion-exchange microbeads. Journal of environmental management, 304, 114173. https://doi.org/https://doi.org/10.1016/j.jenvman.2021.114173

Ma, X.-Y., Fan, T.-T., Wang, G., Li, Z.-H., Lin, J.-H., & Long, Y.-Z. (2022). High performance GO/MXene/PPS composite filtration membrane for dye wastewater treatment under harsh environmental conditions. Composites Communications, 29, 101017. https://doi.org/https://doi.org/10.1016/j.coco.2021.101017

Maamoun, I., Eljamal, R., Falyouna, O., Bensaida, K., Sugihara, Y., & Eljamal, O. (2021). Insights into kinetics, isotherms and thermodynamics of phosphorus sorption onto nanoscale zero-valent iron. Journal of Molecular Liquids, 328, 115402. https://doi.org/10.1016/j.molliq.2021.115402

Mansor, E. S., Ali, H., & Abdel-Karim, A. (2020). Efficient and reusable polyethylene oxide/polyaniline composite membrane for dye adsorption and filtration. Colloid and Interface Science Communications, 39, 100314. https://doi.org/https://doi.org/10.1016/j.colcom.2020.100314

McYotto, F., Wei, Q., Macharia, D. K., Huang, M., Shen, C., & Chow, C. W. K. (2021). Effect of dye structure on color removal efficiency by coagulation. Chemical engineering journal (Lausanne, Switzerland : 1996), 405, 126674. https://doi.org/10.1016/j.cej.2020.126674

Merck KGaA. (1999). Chemical Oxygen Demand (COD) by Dichromate Oxidation and Photometry. In. Darmstadt.

Mohammad, A. T., Al-Obaidi, M. A., Hameed, E. M., Basheer, B. N., & Mujtaba, I. M. (2020). Modelling the chlorophenol removal from wastewater via reverse osmosis process using a multilayer artificial neural network with genetic algorithm. Journal of Water Process Engineering, 33, 100993. https://doi.org/https://doi.org/10.1016/j.jwpe.2019.100993

Momina, M., Qurtulen, Q., Salimi Shahraki, H., Ahmad, A., & Zaheer, Z. (2024). Machine learning approaches to predict adsorption performance of sugarcane derived-carbon dot −based composite in the removal of dyes. Separation and purification technology, 351, 127937. https://doi.org/https://doi.org/10.1016/j.seppur.2024.127937

Nam, S.-N., Yea, Y., Park, S., Park, C., Heo, J., Jang, M., . . . Yoon, Y. (2023). Modeling sulfamethoxazole removal by pump-less in-series forward osmosis–ultrafiltration hybrid processes using artificial neural network, adaptive neuro-fuzzy inference system, and support vector machine. Chemical Engineering Journal, 145821. https://doi.org/https://doi.org/10.1016/j.cej.2023.145821

Nidheesh, P. V., Divyapriya, G., Ezzahra Titchou, F., & Hamdani, M. (2022). Treatment of textile wastewater by sulfate radical based advanced oxidation processes. Separation and purification technology, 293, 121115. https://doi.org/https://doi.org/10.1016/j.seppur.2022.121115

Obayomi, K. S., Lau, S. Y., Ibrahim, O., Zhang, J., Meunier, L., Aniobi, M. M., . . . Rahman, M. M. (2023). Removal of Congo red dye from aqueous environment by zinc terephthalate metal organic framework decorated on silver nanoparticles-loaded biochar: Mechanistic insights of adsorption. Microporous and Mesoporous Materials, 355, 112568. https://doi.org/https://doi.org/10.1016/j.micromeso.2023.112568

Özdoğan-Sarıkoç, G., Sarıkoç, M., Celik, M., & Dadaser-Celik, F. (2023). Reservoir volume forecasting using artificial intelligence-based models: Artificial Neural Networks,

Support Vector Regression, and Long Short-Term Memory. Journal of Hydrology, 616, 128766. https://doi.org/https://doi.org/10.1016/j.jhydrol.2022.128766

Pandiarajan, A., Kamaraj, R., Vasudevan, S., & Vasudevan, S. (2018). OPAC (orange peel activated carbon) derived from waste orange peel for the adsorption of chlorophenoxyacetic acid herbicides from water: Adsorption isotherm, kinetic modelling and thermodynamic studies. Bioresource Technology, 261, 329-341. https://doi.org/10.1016/j.biortech.2018.04.005

Park, J., Lee, W. H., Kim, K. T., Park, C. Y., Lee, S., & Heo, T.-Y. (2022). Interpretation of ensemble learning to predict water quality using explainable artificial intelligence. Science of The Total Environment, 832, 155070. https://doi.org/https://doi.org/10.1016/j.scitotenv.2022.155070

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. Journal of the Taiwan Institute of Chemical Engineers, 45(4), 1691-1699. https://doi.org/10.1016/j.jtice.2013.12.018

Qambar, A. S., & Khalidy, M. M. A. (2022). Prediction of municipal wastewater biochemical oxygen demand using machine learning techniques: A sustainable approach. Process safety and environmental protection, 168, 833-845. https://doi.org/https://doi.org/10.1016/j.psep.2022.10.033

Rosa, J. M., Fileti, A. M. F., Tambourgi, E. B., & Santana, J. C. C. (2015). Dyeing of cotton with reactive dyestuffs: the continuous reuse of textile wastewater effluent treated by Ultraviolet / Hydrogen peroxide homogeneous photocatalysis. Journal of Cleaner Production, 90, 60-65. https://doi.org/https://doi.org/10.1016/j.jclepro.2014.11.043

Salazar, R., Ureta-Zañartu, M. S., González-Vargas, C., Brito, C. d. N., & Martinez-Huitle, C. A. (2018). Electrochemical degradation of industrial textile dye disperse yellow 3: Role of electrocatalytic material and experimental conditions on the catalytic production of oxidants and oxidation pathway. Chemosphere (Oxford), 198, 21-29. https://doi.org/10.1016/j.chemosphere.2017.12.092

Saravanan, A., Deivayanai, V. C., Kumar, P. S., Rangasamy, G., Hemavathy, R. V., Harshana, T., . . . Alagumalai, K. (2022). A detailed review on advanced oxidation process in treatment of wastewater: Mechanism, challenges and future outlook. Chemosphere, 308, 136524. https://doi.org/https://doi.org/10.1016/j.chemosphere.2022.136524

Shabir, M., Yasin, M., Hussain, M., Shafiq, I., Akhter, P., Nizami, A.-S., . . . Park, Y.-K. (2022). A review on recent advances in the treatment of dye-polluted wastewater. Journal of Industrial and Engineering Chemistry, 112, 1-19. https://doi.org/https://doi.org/10.1016/j.jiec.2022.05.013

Singh, A., Pal, D. B., Mohammad, A., Alhazmi, A., Haque, S., Yoon, T., . . . Gupta, V. K. (2022). Biological remediation technologies for dyes and heavy metals in wastewater treatment: New insight. Bioresource Technology, 343, 126154. https://doi.org/https://doi.org/10.1016/j.biortech.2021.126154

Suhan, M. B. K., Mahtab, S. M. T., Aziz, W., Akter, S., & Islam, M. S. (2021). Sudan black B dye degradation in aqueous solution by Fenton oxidation process: Kinetics and cost analysis. Case Studies in Chemical and Environmental Engineering, 4, 100126. https://doi.org/https://doi.org/10.1016/j.cscee.2021.100126

Tan, I. A. W., Ahmad, A. L., & Hameed, B. H. (2009). Adsorption isotherms, kinetics, thermodynamics and desorption studies of 2,4,6-trichlorophenol on oil palm empty fruit bunch-based activated carbon. Journal of hazardous materials, 164(2), 473-482. https://doi.org/10.1016/j.jhazmat.2008.08.025

Tang, Y., He, D., Guo, Y., Qu, W., Shang, J., Zhou, L., . . . Dong, W. (2020). Electrochemical oxidative degradation of X-6G dye by boron-doped diamond anodes: Effect of operating parameters. Chemosphere, 258, 127368. https://doi.org/https://doi.org/10.1016/j.chemosphere.2020.127368

Tang, Y., Liu, M., He, D., Pan, R., Dong, W., Feng, S., & Ma, L. (2022). Efficient electrochemical degradation of X-GN dye wastewater using porous boron-doped diamond electrode. Chemosphere, 307, 135912. https://doi.org/https://doi.org/10.1016/j.chemosphere.2022.135912

Wahidna, A., Sookia, N., & Ramgolam, Y. K. (2024). Performance evaluation of artificial neural network and hybrid artificial neural network based genetic algorithm models for global horizontal irradiance forecasting. Solar Energy Advances, 4, 100054. https://doi.org/https://doi.org/10.1016/j.seja.2024.100054

Wang, L., Zhou, J., Wei, J., Pang, M., & Sun, M. (2022). Learning causal Bayesian networks based on causality analysis for classification. Engineering Applications of Artificial Intelligence, 114, 105212. https://doi.org/https://doi.org/10.1016/j.engappai.2022.105212

Wang, W.-L., Cai, Y.-Z., Hu, H.-Y., Chen, J., Wang, J., Xue, G., & Wu, Q.-Y. (2019). Advanced treatment of bio-treated dyeing and finishing wastewater using ozone-biological activated carbon: A study on the synergistic effects. Chemical Engineering Journal, 359, 168-175. https://doi.org/https://doi.org/10.1016/j.cej.2018.11.059

Wu, H., & Levinson, D. (2021). The ensemble approach to forecasting: A review and synthesis. Transportation Research Part C: Emerging Technologies, 132, 103357. https://doi.org/https://doi.org/10.1016/j.trc.2021.103357

Wu, X., Zheng, Z., Wang, L., Li, X., Yang, X., & He, J. (2023). Coupling process-based modeling with machine learning for long-term simulation of wastewater treatment plant operations. Journal of environmental management, 341, 118116. https://doi.org/https://doi.org/10.1016/j.jenvman.2023.118116

Xie, M., Shon, H. K., Gray, S. R., & Elimelech, M. (2016). Membrane-based processes for wastewater nutrient recovery: Technology, challenges, and future direction. Water Res, 89, 210-221. https://doi.org/10.1016/j.watres.2015.11.045

Zaghloul, M. S., Hamza, R. A., Iorhemen, O. T., & Tay, J. H. (2020). Comparison of adaptive neuro-fuzzy inference systems (ANFIS) and support vector regression (SVR) for data-driven modelling of aerobic granular sludge reactors. Journal of environmental chemical engineering, 8(3), Article 103742. https://doi.org/10.1016/j.jece.2020.103742

7.2 Links and implications

This research specifically examined the thermodynamic effects and pollutant removal efficiencies of bone char adsorbents using uniquely developed AI and ML ensembles. Monte Carlo simulations, Bayesian inference network analysis and computational fluid dynamics were performed to determine the influential variables and its causal relationships by subjecting it optimisation to achieve optimal conditions for maximum pollutant removal efficiency. The results from AI and ML ensembles provided insights into the development of clean, renewable wastewater treatment system with minimal greenhouse gas emissions and carbon footprints. The energy efficiency of adsorption process can be improved without compromising the pollutant removal efficiency and renewability of the adsorbent material.

Most importantly, the AI and ML optimisation techniques can be applied to other priority areas of interest, such as the removal of pharmaceutical and PFAS/PFOS-contaminated wastewater to achieve optimal conditions, especially in highly complex wastewater treatment plants, where this adsorption technology can be integrated into industrial processes without incurring significant capital costs.

This chapter presents a comprehensive overview of the key findings of this PhD research project and provides a summary of recommendations for future research directions.

8.1. Discussion and conclusions

This PhD research project was initially conceived to investigate specific anthropogenic pollutants of interest, such as pharmaceutically active compounds and textile dyes, to be discussed in literature reviews. However, attention was redirected to emphasize the novelty of AI/ML technologies applied to wastewater treatment more than highlighting the types of pollutants suitable for wastewater treatment. There are some ethical dilemmas regarding whether pharmaceutical wastewater treatment should be progressed into the technical experimentation stage due to hazardous aspects of chemical properties and its ability to give rise to antimicrobial resistance genes if not properly managed to avoid accidental discharge of endocrine-disrupting compounds into natural waters.

Our central focus was to bridge related themes between Papers 1 and 2, giving rise to AI optimisation techniques applicable to three-dimensional electrochemical reactor and adsorption technology, which can be practically used to treat model dye pollutants in contaminated wastewater, as part of technical experimentation. Moreover, the presence of xenobiotic dye pollutants in our natural waters is more ubiquitous than pharmaceutical contaminants due to the unregulated nature of textile substances, thereby justifying the focus on xenobiotic dyes to be used as model pollutants and satisfying overlapping interests of researchers. Nonetheless, the literature reviews included a thorough discussion and critical analysis of critical operational parameters and its impacts on targeted responses, setting a foundation to extend beyond the scope of discussion and application, shifting focus onto xenobiotic dye wastewater treatment, the intended theme of technical experimentation.

The technical reliability and technoeconomic aspects of three-dimensional electrochemical oxidation technology were critically evaluated by accounting for its ability to remove dye and TOC from wastewater, optimising the current efficiency, electrical energy consumption of dye and TOC and annual electricity cost. On the other hand, the effectiveness of GIC particle electrodes as adsorbent materials was evaluated using a range of adsorption kinetics and isotherm models to characterise the adsorption phenomena. The regeneration efficiency of GIC particle electrodes was equally significant. GIC particle electrodes are different from a conventional carbon-based adsorbent due to its ability to regenerate in the presence of an applied electric field electrochemically. The adsorptive capacity of GIC

particle electrodes can be continuously recovered. However, prolonged electrochemical regeneration can lead to the depletion of unique surface physicochemical properties or critical chemical compositions crucial for establishing electrostatic interaction between adsorbates and adsorbent materials. On the other hand, when intensely high current density is applied, intermediate transformation oxidation byproducts from the breakdown of parental compounds can produce undesirable effects on the dye and TOC removal rates. Therefore, a range of optimisation techniques were uniquely developed to achieve the optimal conditions for enhancing the pollutant removal efficiency of three-dimensional electrochemical oxidation technology.

Management of complex operational variables requires AI and ML technologies to enhance the targeted responses. Although there were some drawbacks associated with some AI and ML optimisation techniques, more advanced techniques, such as uniquely designed hybrid ensembles and algorithms, were used to improve the prediction accuracy and precision of the models. ANOVA and error function analyses were developed to evaluate the adsorption efficiency and other targeted responses. The salting, pH and thermal effects on the selectivity reversal of GIC particle electrodes in a binary mixture were evaluated using RSM optimisation techniques. Nonlinear models were the best kinetic models in the order: Elovich > Bangham > Pseudo-second > Pseudo-first order. An error function analysis confirmed that the Redlich-Peterson isotherm model was the best nonlinear regression model due to the estimation accuracy of dye-loading capacity. The best dye removal efficiency achieved was approximately 93% using a current density of 45.14 mA/cm^2, whereas the TOC removal efficiency was 67%. These estimated variables were determined based on the RSM optimisation studies. However, the interactive effects of process variables could be underestimated based on results from more advanced AI and ML ensembles. The production of intermediate transformation oxidation byproducts may offset the accuracy and precision of estimated values due to the extremely high current density used.

Figure 3. Electrochemical degradation of Reactive Black 5 xenobiotic dye solution after 5 cycles of treatment.

Another research study was conducted with full emphasis on the use of uniquely designed CCD-NPRSM, AI, and ML ensembles to improve the accuracy and precision of estimated targeted responses. Similarly, the experiment was conducted on RB5-polluted wastewater to investigate the efficacy of AI and ML optimisation techniques to enhance the pollutant removal rates, using global optimal solutions to achieve optimal conditions for superior decomposition of RB5. The optimised decolouration efficiencies were 99.30%, 96.63% and 99.14%, for CCD-NPRSM, hybrid ANN-XGBoost ensemble and CART, respectively, using an applied current density of 20 mA/cm^2, 20 mins of electrolysis time and 65 mg/L of RB5 as initial dye concentration. The optimisation results of CCD-NPRSM, AI and ML ensembles were significantly better in terms of accuracy and precision than a single optimisation result in another study. The optimised AI and ML models were validated using analysis of variance (ANOVA), which revealed that hybrid ANN-XGBoost ensemble had the lowest mean square error (MSE) and best coefficient of determination (R^2) of 0.014 and 0.998 compared to other optimisation techniques. Overall, the final research justified that hybrid ANN-XGBoost ensemble approach is the most feasible optimisation technique for RB5 dye wastewater treatment. Additional recommendations, such as different combination of AI and ML variants or higher dimensional order of RSM using nested transfer function

should be investigated in more detail to maximise the prediction efficiency of models with potential applications in achieving optimal wastewater treatment conditions.

In comparison to more advanced AI and ML optimisation techniques, approximately 98% of MO removal efficiency was achieved using 15 mA/cm^2 of current density, 3.62 kWh/kg of electrical energy consumption and 79.53% of current efficiency. The statistical metrics showed the superiority of different AI and ML optimisation techniques were ranked in the order: ANN > RF > SVM > Multiple Regression. The sensitivity analysis from Monte Carlo simulations of unique combination of AI and ML ensembles showed that ANN-RF ensembles yielded slightly less system perturbations, prediction variability and levels of estimation uncertainty compared to ANN-SVM model.

Furthermore, the future directions are to investigate the effect of renewable energy-driven three-dimensional electrochemical processes on pollutant removal efficiency, using uniquely developed electrically conductive adsorbent materials fabricated from green, renewable agricultural sources. The physicochemical stability of adsorbent material is critical to endure sustained electrolysis and protect it against particle attrition. Advanced AI and ML optimisation approaches should be applied to develop more efficient predictive models for monitoring complex wastewater matrices and effectively managing process dynamic conditions using process control systems. The technological capabilities of cutting-edge cyber-physical systems, blockchain -related technologies and the Internet of Things in the digital economy can be integrated into the AI and ML software and hardware, improving the compatibility and facilitating decision-making processes for wastewater treatment industries. AI-powered three-dimensional electrochemical technology can help to improve energy efficiency, value engineering, minimise carbon footprint and remove barriers to resource recovery and energy management processes. Operational costs can be reduced significantly by streamlining the AI and ML advanced computing techniques to create efficient energy management processes for renewable energy-driven electrochemical technology, thereby improving the technoeconomic aspects and financial viability of such technologies.

8.2. Recommendations and future directions

This PhD research project successfully improved the electrochemical degradation of xenobiotic dye contaminants in wastewater, using uniquely designed AI and ML ensembles to optimise the three-dimensional electrochemical oxidation technology. The following research recommendations and future directions are proposed:

1) Different combination of AI and ML variants or higher dimensional order of RSM using nested transfer function should be applied in wastewater management systems
2) The effect of renewable energy-driven three-dimensional electrochemical processes on pollutant removal efficiency should be investigated and applied in current wastewater treatment plants.
3) Technoeconomic feasibility and environmental impact of renewable energy-driven three-dimensional electrochemical oxidation technology powered by AI and ML advanced computing technology should be investigated.
4) Fabrication of electrically conductive adsorbent materials derived from green, renewable agricultural sources
5) Electrode doping and surface morphology tuning using nanoengineering techniques should be applied to improve the electrocatalytic efficiency of anodic material in the three-dimensional electrochemical reactor.
6) Advanced AI and ML computing techniques should be applied to manage data generated from process control systems in wastewater treatment plants.
7) The technological capabilities of cutting-edge cyber-physical systems, blockchain - related technologies and the Internet of Things in the digital economy should be integrated into the AI and ML software and hardware, improving energy management processes and reducing operational costs.

Alrobei, H., Prashanth, M. K., Manjunatha, C. R., Kumar, C. B. P., Chitrabanu, C. P., Shivaramu, P. D., . . . Raghu, M. S. (2021). Adsorption of anionic dye on eco-friendly synthesised reduced graphene oxide anchored with lanthanum aluminate: Isotherms, kinetics and statistical error analysis. Ceramics International, 47(7), 10322-10331. https://doi.org/10.1016/j.ceramint.2020.07.251

Bilal, M., Ihsanullah, I., Hassan Shah, M. U., Bhaskar Reddy, A. V., & Aminabhavi, T. M. (2022). Recent advances in the removal of dyes from wastewater using low-cost adsorbents. Journal of environmental management, 321, 115981. https://doi.org/https://doi.org/10.1016/j.jenvman.2022.115981

Bosio, M., de Souza-Chaves, B. M., Saggioro, E. M., Bassin, J. P., Dezotti, M. W. C., Quinta-Ferreira, M. E., & Quinta-Ferreira, R. M. (2021). Electrochemical degradation of psychotropic pharmaceutical compounds from municipal wastewater and neurotoxicity evaluations. Environmental science and pollution research international, 28(19), 23958-23974. https://doi.org/10.1007/s11356-020-12133-9

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

Gadekar, M. R., & Ahammed, M. M. (2019). Modelling dye removal by adsorption onto water treatment residuals using combined response surface methodology-artificial neural network approach. Journal of environmental management, 231, 241-248. https://doi.org/https://doi.org/10.1016/j.jenvman.2018.10.017

Kanneganti, D., Reinersman, L. E., Holm, R. H., & Smith, T. (2022). Estimating sewage flow rate in Jefferson County, Kentucky using machine learning for wastewater-based epidemiology applications. WATER SUPPLY, 22(12), 8434-8439. https://doi.org/10.2166/ws.2022.395

Liu, L., Chen, Z., Zhang, J., Shan, D., Wu, Y., Bai, L., & Wang, B. (2021). Treatment of industrial dye wastewater and pharmaceutical residue wastewater by advanced oxidation processes and its combination with nanocatalysts: A review. Journal of Water Process Engineering, 42, 102122. https://doi.org/https://doi.org/10.1016/j.jwpe.2021.102122

Liu, X., Chen, Z., Du, W., Liu, P., Zhang, L., & Shi, F. (2022). Treatment of wastewater containing methyl orange dye by fluidized three dimensional electrochemical oxidation process integrated with chemical oxidation and adsorption. Journal of environmental management, 311, 114775-114775. https://doi.org/10.1016/j.jenvman.2022.114775

McYotto, F., Wei, Q., Macharia, D. K., Huang, M., Shen, C., & Chow, C. W. K. (2021). Effect of dye structure on color removal efficiency by coagulation. Chemical engineering journal (Lausanne, Switzerland : 1996), 405, 126674. https://doi.org/10.1016/j.cej.2020.126674

Nidheesh, P. V., Zhou, M., & Oturan, M. A. (2018). An overview on the removal of synthetic dyes from water by electrochemical advanced oxidation processes. Chemosphere (Oxford), 197(210-227.), 210-227. https://doi.org/10.1016/j.chemosphere.2017.12.195

Obayomi, K. S., Lau, S. Y., Ibrahim, O., Zhang, J., Meunier, L., Aniobi, M. M., . . . Rahman, M. M. (2023). Removal of Congo red dye from aqueous environment by zinc terephthalate metal organic framework decorated on silver nanoparticles-loaded biochar: Mechanistic insights of adsorption. Microporous and Mesoporous Materials, 355, 112568. https://doi.org/https://doi.org/10.1016/j.micromeso.2023.112568

Park, J., Lee, W. H., Kim, K. T., Park, C. Y., Lee, S., & Heo, T.-Y. (2022). Interpretation of ensemble learning to predict water quality using explainable artificial intelligence. Science of The Total Environment, 832, 155070. https://doi.org/https://doi.org/10.1016/j.scitotenv.2022.155070

Samal, K., Mahapatra, S., & Hibzur Ali, M. (2022). Pharmaceutical wastewater as Emerging Contaminants (EC): Treatment technologies, impact on environment and human health. Energy Nexus, 6, 100076. https://doi.org/10.1016/j.nexus.2022.100076

Saravanan, A., Kumar, P. S., Jeevanantham, S., Anubha, M., & Jayashree, S. (2022). Degradation of toxic agrochemicals and pharmaceutical pollutants: Effective and alternative approaches toward photocatalysis. Environmental Pollution, 298, 118844. https://doi.org/https://doi.org/10.1016/j.envpol.2022.118844

Sun, H., Liu, Z. G., Wang, Y., & Li, Y. S. (2013). Electrochemical in situ regeneration of granular activated carbon using a three-dimensional reactor. Journal of environmental sciences (China), 25, S77-S79. https://doi.org/10.1016/S1001-0742(14)60630-6

Tang, Y., He, D., Guo, Y., Qu, W., Shang, J., Zhou, L., . . . Dong, W. (2020). Electrochemical oxidative degradation of X-6G dye by boron-doped diamond anodes: Effect of operating parameters. Chemosphere, 258, 127368. https://doi.org/https://doi.org/10.1016/j.chemosphere.2020.127368

Uddin, M. A., Begum, M. S., Ashraf, M., Azad, A. K., Adhikary, A. C., & Hossain, M. S. (2023). Water and chemical consumption in the textile processing industry of Bangladesh. PLOS Sustainability and Transformation, 2(7), e0000072. https://doi.org/10.1371/journal.pstr.0000072

Wu, H., & Levinson, D. (2021). The ensemble approach to forecasting: A review and synthesis. Transportation Research Part C: Emerging Technologies, 132, 103357. https://doi.org/https://doi.org/10.1016/j.trc.2021.103357

Xu, T., Fu, L., Lu, H., Zhang, M., Wang, W., Hu, B., . . . Yu, G. (2023). Electrochemical oxidation degradation of Rhodamine B dye on boron-doped diamond electrode: Input mode of power attenuation. Journal of Cleaner Production, 401, 136794. https://doi.org/https://doi.org/10.1016/j.jclepro.2023.136794

Ye, X., Cai, W., Lu, D., Liu, R., Wu, Y., & Wang, Y. (2022). Electrochemical regeneration of granular activated carbon using an AQS (9,10- anthraquinone-2-sulfonic acid)/PPy modified graphite plate cathode. Chemosphere, 308, 136189. https://doi.org/https://doi.org/10.1016/j.chemosphere.2022.136189

Yuan, J., Chen, Z., Yu, Q., Zhu, W., Li, S., Han, L., . . . Chen, B. (2022). Enhanced electrochemical removal of dye wastewater by PbO_2 anodes using halloysite nanotubes with different surface charge properties. Journal of Electroanalytical Chemistry, 923, 116816. https://doi.org/https://doi.org/10.1016/j.jelechem.2022.116816

Zhang, S., Jin, Y., Chen, W., Wang, J., Wang, Y., & Ren, H. (2023). Artificial intelligence in wastewater treatment: A data-driven analysis of status and trends. Chemosphere, 336, 139163. https://doi.org/https://doi.org/10.1016/j.chemosphere.2023.139163

Supplementary materials related to published papers.

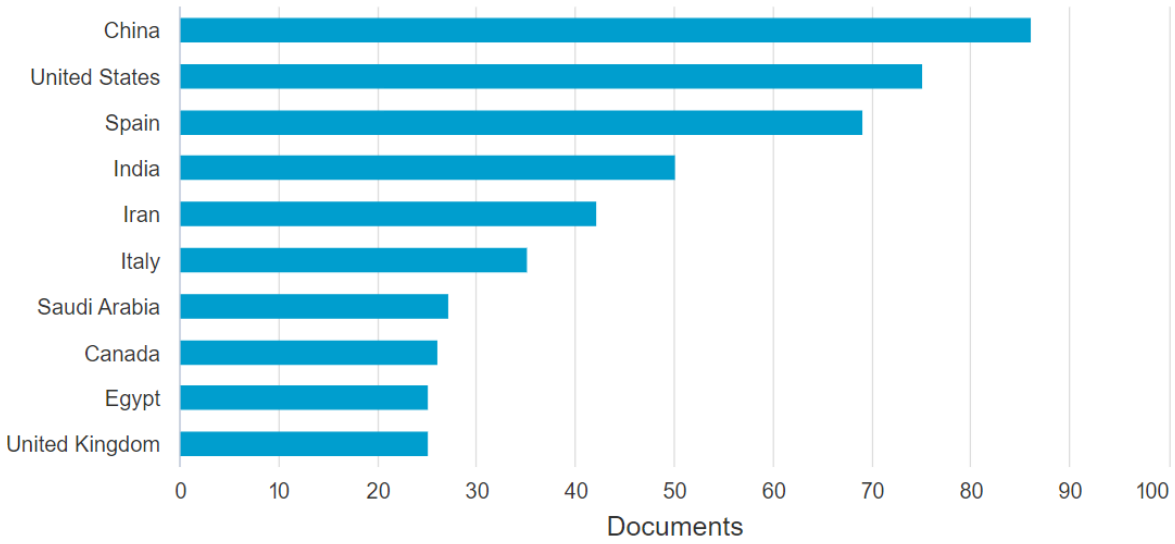


Figure A1. Number of publications by different countries in artificial intelligence technologies applied to wastewater treatment in 2023 (Scopus Citation Index).

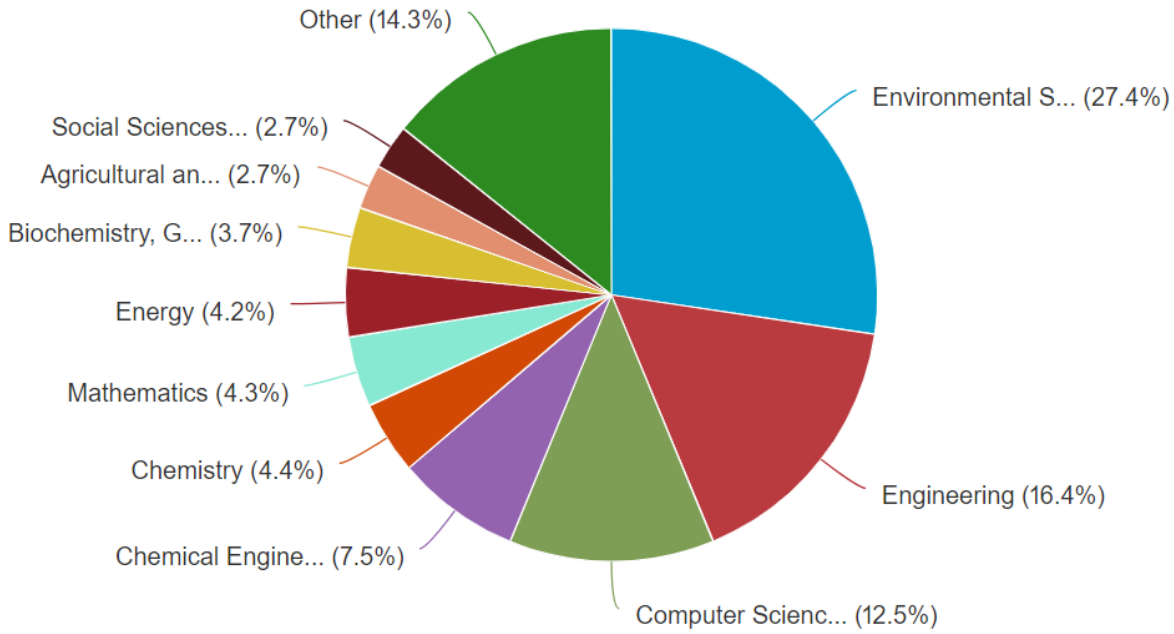


Figure A2. Subject fields of publications in artificial intelligence technologies applied to wastewater treatment in 2023 (Scopus Citation Index).

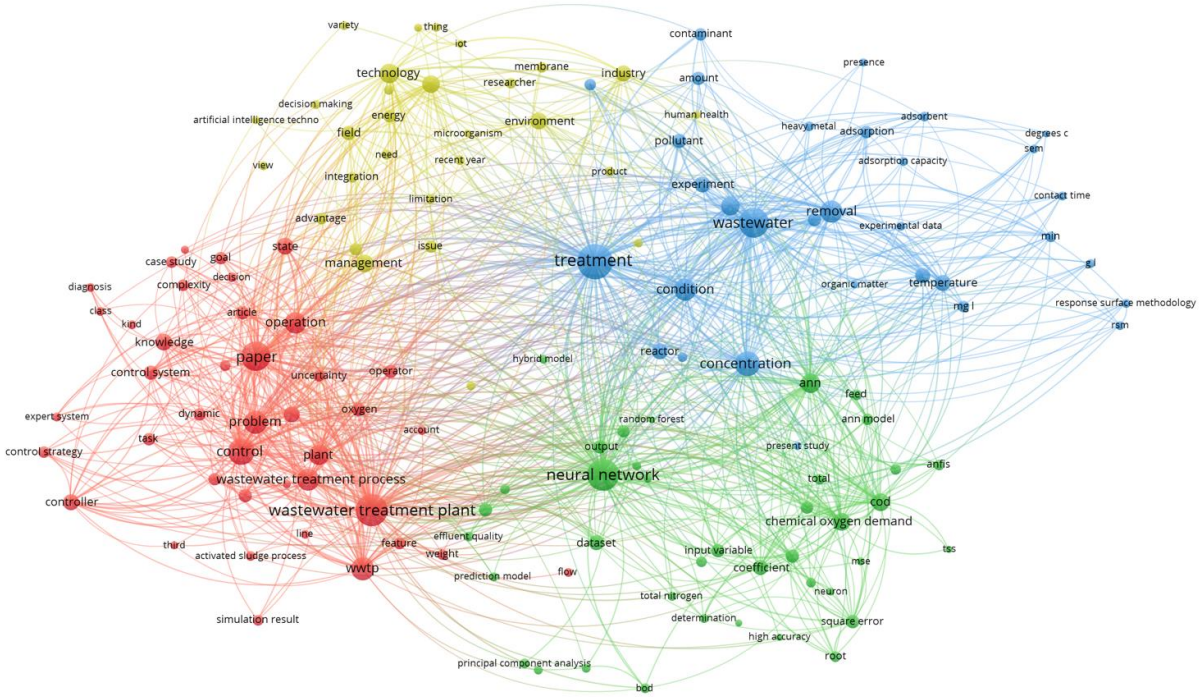


Figure A3. Network visualisation of publications on artificial intelligence applied to wastewater treatment during 1990-2023 (Web of Science, Science Citation Index).