Environmental Science
Water Research & Technology

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Check for updates

Cite this: *Environ. Sci.: Water Res. Technol.*, 2025, **11**, 2935

Optimizing sodium percarbonate oxidation for wastewater treatment with artificial intelligence

Luyu Guo,[a] Jing Zhang,[a] Yahan Cao,[a] Jiayu Zhang,[a] Zhengyang Li,[a] Xiaofei Chen,[b] Lei Ma*[a] and Xiaowei Liu *[c]

Received 24th July 2025,
Accepted 22nd September 2025

DOI: 10.1039/d5ew00689a

rsc.li/es-water

The effective removal of toxic pollutants like *m*-cresol from wastewater remains challenging despite technological advancements. This study optimized total organic carbon (TOC) removal from *m*-cresol-contaminated wastewater using sodium percarbonate (SPC) oxidation through artificial neural network (ANN) and response surface methodology (RSM) modeling. TOC was selected as the optimization target due to its comprehensive representation of organic pollution levels. Six operational parameters were evaluated: initial pH, reaction time, SPC dosage, temperature, catalyst dosage, and initial *m*-cresol concentration. The ANN model demonstrated superior performance over RSM, achieving near-perfect R^2 values with significant improvement in predictive accuracy. Under optimal ANN-derived conditions (pH 2.3, 35.7 min, 2.9 g L^{-1} SPC, 45.7 °C, 12.9 g L^{-1} catalyst, 75 mg L^{-1} *m*-cresol), maximum experimental TOC removal reached 67.8%, significantly exceeding RSM's 38.2%. These findings demonstrate ANN's superior capability to model complex, nonlinear relationships in advanced oxidation processes, providing a robust optimization framework for enhancing wastewater treatment efficiency.

Water impact

AI-optimized sodium percarbonate oxidation offers enhanced removal of toxic organic pollutants from industrial wastewater, addressing critical environmental contamination challenges. This method provides treatment facilities with significantly improved efficiency for protecting water quality, reducing chemical discharge into aquatic ecosystems, and advancing sustainable wastewater management practices.

1. Introduction

m-Cresol, a representative phenolic compound, is widely found in industrial wastewater from coal-based chemical, petrochemical, and pharmaceutical industries. Its existence poses significant challenges for wastewater treatment due to its resistance to biodegradation, corrosive nature and strong irritancy. As a result, *m*-cresol has been classified by the Environmental Protection Agency of the United States as one of eleven refractory phenolic compounds, and is listed among China's priority control pollutants for water pollution. Advanced oxidation processes (AOP) are commonly adopted to treat *m*-cresol, converting it into small aliphatic compounds or achieving complete mineralization within a relatively short timeframe. These processes lay a critical foundation for the treatment of coal gasification wastewater.[1–3]

In recent years, sodium percarbonate (SPC) has emerged as an effective oxidizing agent in water treatment. When dissolved in water, SPC decomposes to produce percarbonate and carbonate ions, the latter further breaking down to form oxygen and hydroxyl radicals (·OH), which enhance the oxidative process. Unlike traditional oxidizing agents, SPC overcomes the limitation of operating within a narrow acidic pH range due to the buffering effect of coexisting carbonate ions. For example, Guo *et al.*[4] investigated the use of SPC in an ozone (O_3)/SPC system to accelerate the degradation of sulfamethoxazole (SMX) in water. Their study showed that after 30 min of treatment, the O_3/SPC system achieved substantial enhancement in eliminating total organic carbon (TOC) and chemical oxygen demand (COD) compared to ozone-only treatment. At the optimal SPC concentration of 0.2 g L^{-1}, the SMX degradation rate improved by 16.4%

[a] *Beijing Key Laboratory of Fuels Cleaning and Advanced Catalytic Emission Reduction Technology, College of New Materials and Chemical Engineering, Beijing Institute of Petrochemical Technology, Beijing 102617, China.*
E-mail: malei@bipt.edu.cn
[b] *Chen Ping Laboratory of TIANS Engineering Technology Group Co., Ltd., Shijiazhuang, Hebei 050000, China*
[c] *Division of Physical Sciences and Engineering, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia.*
E-mail: xiaowei.liu@kaust.edu.sa

relative to the O_3-only system, with the kinetic constant increasing by 1.7-fold.

In parallel, artificial intelligence (AI) technologies have introduced new opportunities for in-depth exploration and intelligent analysis in water environment management, addressing complex problems in water treatment.[5] Traditional methods, *e.g.*, adsorption, membrane separation/filtration, precipitation, flotation, coagulation/flocculation, aerobic and anaerobic processes, *etc.*, are often fitted using physical models. However, the inherent complexity of real-world water treatment reactions makes them difficult to capture accurately with simple physical models. Artificial neural networks (ANN) have emerged as a powerful tool for resolving this issue, exhibiting improved predictive performance across a wide range of complex operational scenarios.[6] By utilizing ANN modelling, the applicability of alternative models can be extended, enabling more effective removal of a variety of environment pollutants.[7] With the computational power of modern central processing units (CPU), ANN have made significant progress in chemical and environmental applications.[5] These advancements range from aiding robots in discovering enhanced photocatalysts[8] to serving as predictive tools in water resource management and environmental toxicology,[9–11] effectively modelling and optimizing pollutant elimination processes to improve treatment efficiency and cost-effectiveness.[12,13] Response surface methodology (RSM) analysis represents a hybrid framework that integrates experimental design, mathematical statistics, and parameter optimization.[14] Its core concept involves approximating implicit functions by constructing explicit polynomial expressions. Through the use of multidimensional quadratic regression equations, RSM quantifies interdependencies between factors and system responses in multifactor tests, effectively addressing multivariable problems.[15–17] Compared with other conventional methods, RSM offers the advantage of requiring fewer experimental runs while maintaining strong interpretability, making it especially suitable in certain AOP optimization studies.

In this paper, we employ the RSM and the ANN models to evaluate the predictive accuracy of TOC elimination during SPC oxidation treatment of *m*-cresol polluted wastewater. SPC was selected as the oxidant, with six key experimental parameters: the initial pH of solution, reaction time, dosage of SPC, reaction temperature, dosage of catalyst and the initial *m*-cresol concentration. These parameters were selected because they represent the core operational factors governing SPC activation, ·OH generation, and pollutant degradation efficiency, while providing a balanced and practical basis for model optimization with strong literature precedent. The removal rate of TOC was taken as the primary performance indicator. An advanced SPC oxidation system was developed to study the treatment of *m*-cresol wastewater. To determine the optimized reaction conditions for wastewater treatment, the prediction accuracies of the two studied models were compared, aiming to improve the system's treatment efficiency.

2. Experimental

2.1. Materials

m-Cresol (C_7H_8O, Beijing Bailingwei), SPC ($Na_2CO_3{\cdot}1.5H_2O_2$, Shanghai Aladdin), hydrogen peroxide (H_2O_2, Tianjin Fuyu), ammonia solution ($NH_3{\cdot}H_2O$, 28–30%, Shanghai Aladdin), tetrabutyl titanate ($C_{16}H_{36}O_4Ti$, Shanghai Meryl Biochemical), and ferric nitrate ($Fe(NO_3)_3{\cdot}9H_2O$, Shanghai Maclin Biochemical) were used as received unless otherwise specified, and all experiments utilized deionized (DI) water with resistivity exceeding 18 MΩ.

2.2. Catalyst preparation

The titanium oxide-supported iron catalyst was prepared as follows: $Fe(NO_3)_3{\cdot}9H_2O$ was dissolved in 100 mL of anhydrous ethanol and subjected to ultrasonic treatment to ensure uniform dispersion. At 45 °C, a specific amount of tetrabutyl titanate was slowly added under continuous stirring, ensuring that the final catalyst composition corresponded to a ~5% Fe loading on TiO_2. The impregnated sample was then cooled to room temperature, centrifuged, and rinsed with ethanol several times. After drying, the sample was calcined in a tube furnace at 500 K and naturally cooled down to yield the FeO_x/TiO_2 catalyst.[18,19]

2.3. Experimental setup and SPC oxidation process

The experimental setup for SPC oxidation in this study is displayed in Fig. 1. The procedure was conducted as follows: 100 mL of simulated wastewater with an adjusted pH was added into the reactor. The agitator was turned on and set to a fixed speed. Once the wastewater temperature reached the target value, a predetermined amount of SPC and catalyst, as specified by the experimental design (Table 1 and S1) was added promptly. After the reaction, the resulting mixture underwent filtration *via* a 0.45 μm membrane to remove catalyst particulates preceding analytical measurements.

2.4. TOC measurement

The TOC level was determined by utilizing a Shimadzu TOC-LCPN analyzer. TOC removal rates were computed as follows:

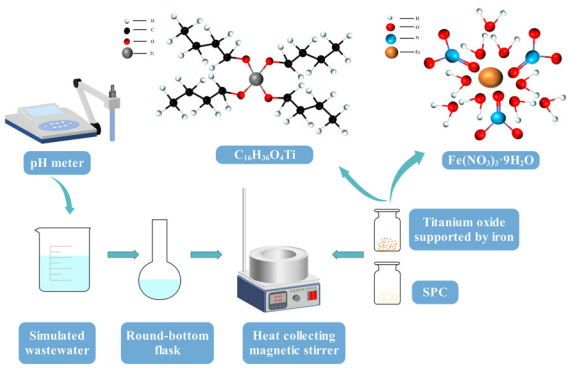

Fig. 1 SPC oxidation experimental setup.

Table 1 Experimental control variables and level coding

Experimental factor	Unit	Symbol	Horizontal coding				
			−1.565	−1	0	1	1.565
Initial pH of solution		P	1.0	3	6.5	10	12
Reaction time	min	t	4.3	10	20	30	35.7
SPC dosage	g L^{-1}	S	1.1	1.5	2.3	3.1	3.6
Reaction temperature	°C	T	14.4	20	30	40	45.7
Catalyst dosage	g L^{-1}	c	8.2	9	10.5	12	12.9
m-Cresol concentration	mg L^{-1}	M	35.9	50	75	100	114.1

$$\mathrm{TOC_{removal}}\,(\%) = \frac{\mathrm{TOC_0} - \mathrm{TOC}_t}{\mathrm{TOC}_t} \times 100\% \quad (1)$$

where $\mathrm{TOC_0}$ is the initial TOC level in the raw *m*-cresol solution, and TOC_t is the TOC level at a given time during the reaction.

2.5. Model design and evaluation

2.5.1. RSM design. To maximize degradation efficiency, optimizing process variables is crucial. In this study, optimization was performed using the statistical RSM model, employing central composite design (CCD) within the Design-Expert software (V8.0.5b, Stat-Ease, Inc., Minneapolis).[20] Six control variables were investigated: the initial pH (P), reaction time (t), SPC dosage (S), reaction temperature (T), catalyst dosage (c), and initial *m*-cresol concentration (M). The removal rate of TOC (%) served as the response factor. The experimental design, structured according to the RSM-CCD framework, and the corresponding horizontal encoded variable levels are detailed in Table 1.

The variables in the experimental matrix were encoded using formula (2):

$$x_i = \frac{\alpha(2X_i - (X_{\max} + X_{\min}))}{X_{\max} - X_{\min}} \quad (2)$$

where x_i represents the factor's coded value; X_i gives its actual value; $X_{\max}$ and $X_{\min}$ correspond to the max and min limits of the actual value of the factor, respectively.

A multivariate regression analysis was performed on the experimental dataset and design matrix associated with CCD, developing an encoded quadratic polynomial model aimed at predicting the response variable values:

$$Y = b_0 + \sum_{i=1}^{5} b_i x_i + \sum_{i=1}^{4}\sum_{i<j}^{5} b_{ij} x_i x_j + \sum_{i=1}^{5} b_{ii} {x_i}^2 + \varepsilon \quad (3)$$

where the predicted response (Y) corresponds to TOC removal efficiency, with b_0 designating the offset term. The coefficients b_i (linear), b_{ii} (quadratic), and b_{ij} (interactive) quantify variable effects, while ε encapsulates the random error. Each independent variable is expressed as x_i in encoded units.

2.5.2. ANN design. While ANN models typically require large datasets to produce accurate outputs, statistically well-distributed experimental data, such as that generated through RSM using design of experiments (DOE), can provide an effective basis for building reliable ANN models.[21] Based on the specific characteristics of this study, a customized ANN model was employed. This model utilized response surface data for the aforementioned six factors in the RSM design to predict the outcomes (Fig. 2). For data partitioning, a common approach involves dividing the dataset into distinct training and test sets to evaluate model efficacy. Typical partitioning ratios often range between 80 : 20 and 90 : 10 training-test ratios, depending on the size and characteristics of the dataset.[22] Also, the implementation of cross-validation protocols during model training enhances reliability. This method, widely recognized as robust for the selection of parameters in AI-based algorithms, builds high-quality networks that divide the dataset into multiple folds. The model is then trained multiple times, with each fold serving as validation in turn. The final performance evaluation is based on the average results across all folds, providing an overall assessment of the model's effectiveness.

2.5.3. Model evaluation. The effectiveness of the proposed models was assessed using three key metrics: the coefficient of correlation (R^2), root mean square error (RMSE), and average absolute relative error (AARE%). The analysis of variance approach (ANOVA) was conducted to evaluate the proportion of predicted variability relative to observational data variance. Within this framework, R^2 served as a statistical measure for polynomial model adequacy, while AARE% quantified the magnitude of divergence between predicted and experimental results. These indicators are described by the following equations:

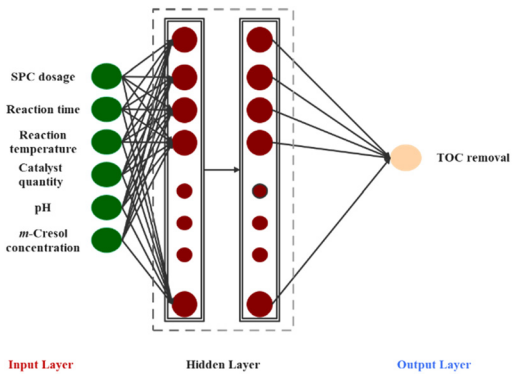

Fig. 2 Schematic for the ANN model in this work.

$$R^2 = 1 - \frac{\sum \left(Q_p - Q_o\right)^2}{\sum \left(\overline{Q_o} - Q_o\right)^2} \quad (4)$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{m}\sum \left(Q_p - Q_o\right)^2} \quad (5)$$

$$\mathrm{AARE}\,(\%) = \frac{1}{m}\sum \frac{\left|Q_p - Q_o\right|}{Q_o} \times 100\% \quad (6)$$

where Q_p denotes the model-predicted quantity and Q_o gives the experimentally observed value with $\overline{Q_o}$ representing their arithmetic mean. Generally, R^2 serves as a quantitative metric bounded between 0 and 1, where values approaching unity ($R^2 \geq 0.7$) demonstrate stronger congruence between theoretical predictions and empirical observations.[23] A smaller RMSE indicates a better model fit,[24] while a diminishing AARE% closer to 0 reflects a reduced discrepancy between model predictions and observed values,[25] implying higher model prediction accuracy.

To compare the model prediction accuracies of ANN against RSM, the following formula is applied:

$$\text{Increased accuracy}\,(\%) = -\frac{\mathrm{RMSE_{ANN}} - \mathrm{RMSE_{RSM}}}{\mathrm{RMSE_{RSM}}} \times 100\% \quad (7)$$

3. Results and discussion

3.1. Catalyst characterization

To enhance SPC oxidation, an appropriate catalyst is required in the system. Based on the literature review and preliminary experiments, our team-developed FeO_x/TiO_2 catalyst demonstrated superior catalytic efficiency in the SPC system and was selected for this study. The catalyst preparation method is described in the Experimental section. Powder X-ray diffraction (PXRD) analysis as depicted in Fig. 3a confirmed the coexistence of Fe_3O_4, rutile TiO_2, and anatase TiO_2, with distinct FeO_x/TiO_2 characteristic peaks indicating successful synthesis. Scanning electron microscopy (SEM) imaging (Fig. 3b) revealed uniformly distributed FeO_x particles, marked by red circles, with an average size of ~15 nm, located on a porous structure that can promote efficient mass transfer during catalysis. Transmission electron microscopy (TEM) imaging (Fig. 3c) identified lattice spacings of 0.16 and 0.35/0.32 nm, attributed to Fe_3O_4 (511) and TiO_2 (101)/(110) planes, respectively, corroborating the catalyst's structural integrity. N_2 adsorption–desorption measurements at 77 K (Fig. S1) revealed the porous

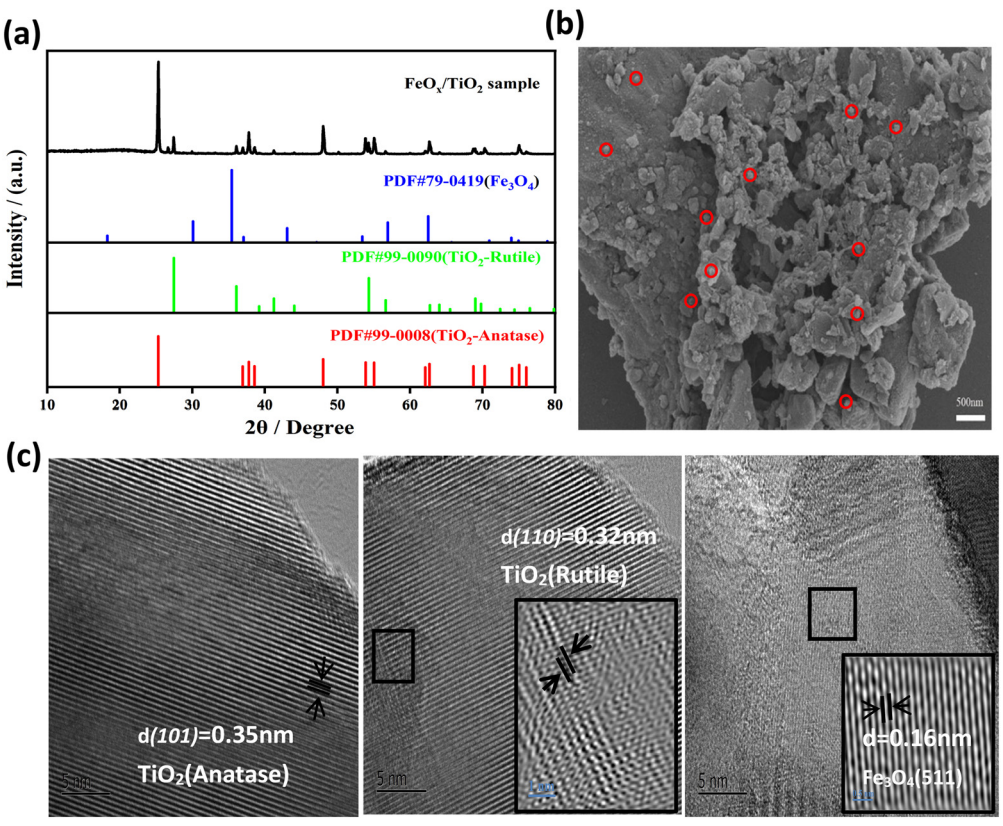

Fig. 3 (a) PXRD curves, (b) SEM image (red circles: FeO_x particles on the support), and (c) TEM images of the FeO_x/TiO_2 catalyst sample (the inset figures are enlarged versions of the rectangular zoomed areas).

structure of the FeO_x/TiO_2 catalyst, showing a specific surface area of 6.0 $m^2\ g^{-1}$ and an average pore size of 32.7 nm.

3.2. RSM results and optimization

According to the RSM-CCD experiment design outlined in Table 1, a total of 86 experimental runs are included to evaluate the six factors across one horizontal scheme (Table S1). The factors are represented as follows: x_1 (initial pH, *P*), x_2 (reaction time, *t*), x_3 (SPC dosage, *S*), x_4 (reaction temperature, *T*), x_5 (catalyst dosage, *c*), and x_6 (initial *m*-cresol concentration, *M*). The response variable, Y_1, represents the experimental results for the TOC removal rate.

ANOVA was carried out to statistically quantify response model variations while validating regression coefficient significance (Table S2). The adeq precision metric, serving as a signal-to-noise ratio indicator, evaluates the model's ability to provide reliable signals relative to background noise. Herein, adeq precision was calculated as 16.62, indicating that the model provides sufficient signals for reliably predicting the TOC removal rate under specific conditions. An *F*-value of 32.8 suggests that the model makes a highly significant contribution to the variance of the interpreted data, with only a 0.01% chance that random fluctuations would lead to an increased *F*-value. However, the *F*-value for the lack-of-fit term is 1.9, implying a 15.5% probability that it is inflated by random fluctuations, suggesting non-significant terms present in the model. Terms with a *P*-value < 0.1 are treated as statistically important, while highly significant factors ($P < 0.0001$) strongly influence the response variable. These findings indicate the need for optimization to refine the model and enhance predictive performance.

Model optimization typically involves excluding non-significant terms identified through ANOVA. In this process, variables with *P*-values greater than 0.05 are iteratively removed, while those with $P < 0.0001$ are regarded as key factors due to their strong influence on the response variable. Additionally, the adeq precision metric, with values greater than 4 considered acceptable, is used as a criterion for evaluating model reliability. In this context, single variables represent the effects of individual factors, two variables capture interactions between factors, and quadratic terms account for the nonlinear effects of individual factors. Streamlining the model by removing non-significant terms improves its focus on key influencing factors and enhances predictive accuracy. As shown in Table 2, the optimization retained significant terms, including single-factor conditions and interaction terms such as *BE*, A^2, B^2, and F^2. Post-optimization, the error between R^2 and $R_{pred}{}^2$ reduced from 35% to 16%, while adeq precision increased from 16.6 to 27. The improved signal-to-noise ratio and *F*-value confirmed the optimization's effectiveness, demonstrating enhanced model reliability and predictive accuracy. Based on the transformation of eqn (2), the final optimized formula for TOC removal is expressed as follows:

Table 2 Analysis of variance and statistical parameters of the optimized RSM

Source of variance	TOC removal rate		
	Coefficient	*F*-Value	*P*-Value
Model		37.3	<0.0001
Cut moment	36.5		
A-*P*	−5.6	73.0	<0.0001
B-*t*	8.6	170.8	<0.0001
C-*M*	−3.0	20.2	<0.0001
D-*S*	−1.6	6.2	0.0153
E-*T*	5.1	59.0	0.0003
F-*c*	2.5	14.1	<0.0001
BE	1.4	4.2	<0.0001
A^2	−4.3	9.5	0.0028
B^2	−3.9	7.8	0.0068
F^2	4.6	11.1	0.0014
Lack of fit	—	1.8	0.1761
R^2	0.8	—	—
$R_{Adj}{}^2$	0.8	—	—
$R_{pred}{}^2$	0.7	—	—
Adeq precision	27.0	—	—

$$TOC_{removal} = 215.73 + 2.94P + 1.99t - 0.12S - 2.05T + 0.23c - 41.52M - 0.01Tt - 0.35\,P^2 - 0.04T^2 + 2.06c^2 \quad (8)$$

Three-dimensional (3D) response surface plots with associated contour plots are generated to elucidate the correlations between independent variables and their corresponding responses.[26,27] Table 2 identifies reaction time (*B*) and reaction temperature (*E*) as key interactive factors influencing the optimized response surface. As shown in Fig. 4a, the curved contour lines highlight the nonlinear interactions between *B* and *E* affecting the TOC removal rate.[28] The response surface within the designated ranges captures the complete trend of these interactions, demonstrating the complex relationship between these two factors. Fig. 4b further illustrates that the TOC removal rate increases progressively with extended reaction times and elevated temperatures. This trend likely results from higher temperatures accelerating the SPC oxidation reaction and longer reaction times enabling more complete interactions between pollutants and oxidants, collectively enhancing degradation efficiency.

Next, the RSM model's predictive performance was further evaluated by comparing experimental and predicted values. Fig. 5a shows a discrepancy between these values, with an AARE% of 15%. Under the predicted optimal conditions – an initial solution pH of 4.4, a reaction time of 25.4 minutes, an SPC dosage of 1.8 g L^{-1}, a reaction temperature of 21.6 °C, a catalyst amount of 11.3 g L^{-1}, and an *m*-cresol concentration of 56.5 mg L^{-1} – the experimental TOC removal rate was merely 38.2%. This value is lower than many experimental test results, and also shows an absolute error of 5.1% compared to the predicted value of 43.3%. These deviations suggest that although the RSM model offers general predictive capability, its accuracy is limited, implying the need for a more robust predictive model.

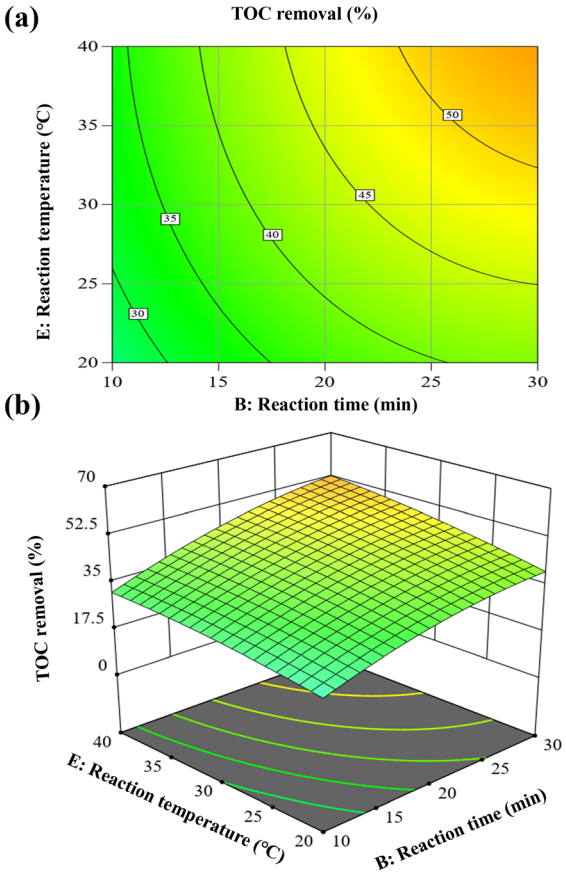

Fig. 4 (a) The contour map and (b) 3D response surface map for factors *B* and *E*.

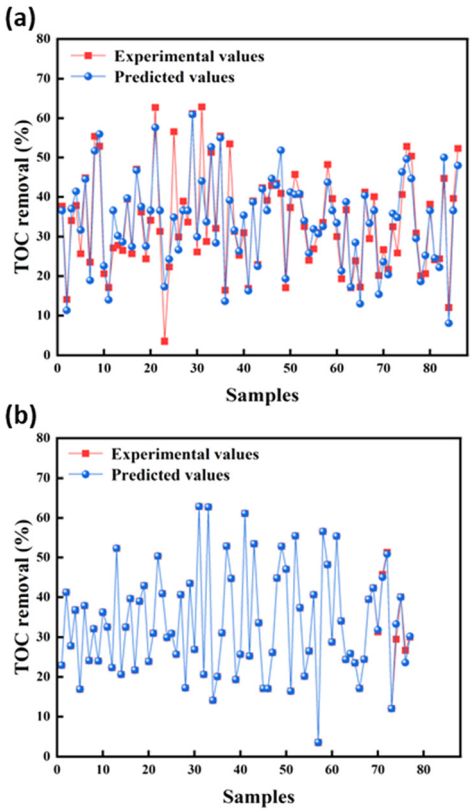

Fig. 5 Comparison of predicted TOC removal rates with experimental results using (a) the RSM model and (b) the ANN model.

3.3. ANN fitting and optimization

The ANN model offers high flexibility and adaptability to various types and complexities of data. It can handle multidimensional input and output data, making it suitable for modeling complex relationships, including those in RSM. ANN models possess the capability to learn patterns and regularities from data automatically, allowing them to predict new outcomes effectively.[21] For this study, the six input variables were selected as the aforementioned. To improve model performance and stability, the datasets were standardized using StandardScaler to achieve zero-mean normalization with unit variance. After removing 9 duplicate data sets from the RSM dataset, 69 data entries were allocated for training while 8 served as validation benchmarks. The ANN architecture incorporated two hidden layers (Fig. 2), with neuron quantity optimization requiring empirical and iterative adjustments. In this case, after trials the first and second hidden layers were each set to 14 neurons, with cross-verification performed using k =10. Additionally, two important parameters – solver and activation function – were optimized to enhance the algorithm performance. The limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm was adopted for minimizing the loss function, while the tanh activation function was applied in the output layer.

After training, the ANN model demonstrates excellent performance, achieving a prediction score of 1 for the training set and 0.98 for the test (Table 3). The AARE% is just 0.4%, highlighting the model's exceptional generalization ability and accuracy. Fig. 5b presents the prediction results for all data in the training and test sets, revealing robust performance across both datasets. Compared with RSM (Fig. 5a), the ANN model clearly shows superior accuracy in predicting TOC removal rates. The model predicted the optimal reaction conditions as follows: an initial solution pH of 2.3, a reaction time of 35.7 min, a SPC dosage of 2.9 g L^{-1}, a reaction temperature of 45.7 °C, a catalyst dosage of 12.9 g L^{-1}, with a pollutant concentration of 75 mg L^{-1}. Under these conditions, the predicted TOC removal rate was 63.1%, whilst the experimental value was 67.8%, leading to a 4.8% absolute error. This high level of predictive accuracy, coupled with 77% enhancement of experimental TOC removal rate than that obtained from RSM, indicates that the current model effectively fits the datasets and captures the relationship between input parameters and TOC removal rate. Although a

Table 3 ANN model evaluation metrics

Data	R^2	RMSE	AARE%
Training set	1	1.78	0.4%
Test set	0.98	0.01	

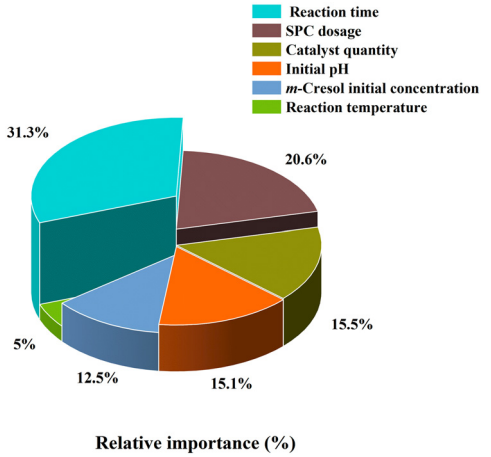

Fig. 6 Relative importance of input parameters on TOC removal.

minor discrepancy still exists between the experimental and predicted value, the difference is within an acceptable range. Experimental values are often subject to errors arising from experimental operation difficulties, equipment accuracy or environmental factors, which may contribute to deviations.

ANN sensitivity analysis was then performed to examine the relative importance of each input parameter in terms of TOC efficacies for *m*-cresol degradation. As shown in Fig. 6, the ranking of variable significance is as follows: reaction time > SPC dosage > catalyst dosage > initial solution pH > initial *m*-cresol concentration > reaction temperature. This order aligns well with the established mechanistic understanding of advanced oxidation processes. As reaction time increases, the substrate has more opportunities to interact with sodium percarbonate and the catalyst, leading to the generation of more ·OH, which enhances pollutant removal.[18] At suitable pH and reaction temperatures, the collision frequency of reactant molecules increases, along with the energy transferred during collisions, making it easier for the reactant molecules to overcome the reaction activation energy, thus boosting TOC removal rates. However, excessive SPC can result in the formation of oxidative by-products, which may negatively affect the TOC removal rate. Similarly, higher initial concentrations of *m*-cresol can inhibit TOC removal, possibly due to competition for reactive radicals or the depletion of oxidizing agents.

3.4. RSM and ANN modeling comparison

Fig. 7 compares the predictive performance between ANN and RSM models. For the training set, the ANN model demonstrated exceptional fitting capabilities (R^2 = 1, RMSE = 1.78), significantly outperforming the RSM model (R^2 = 0.84,

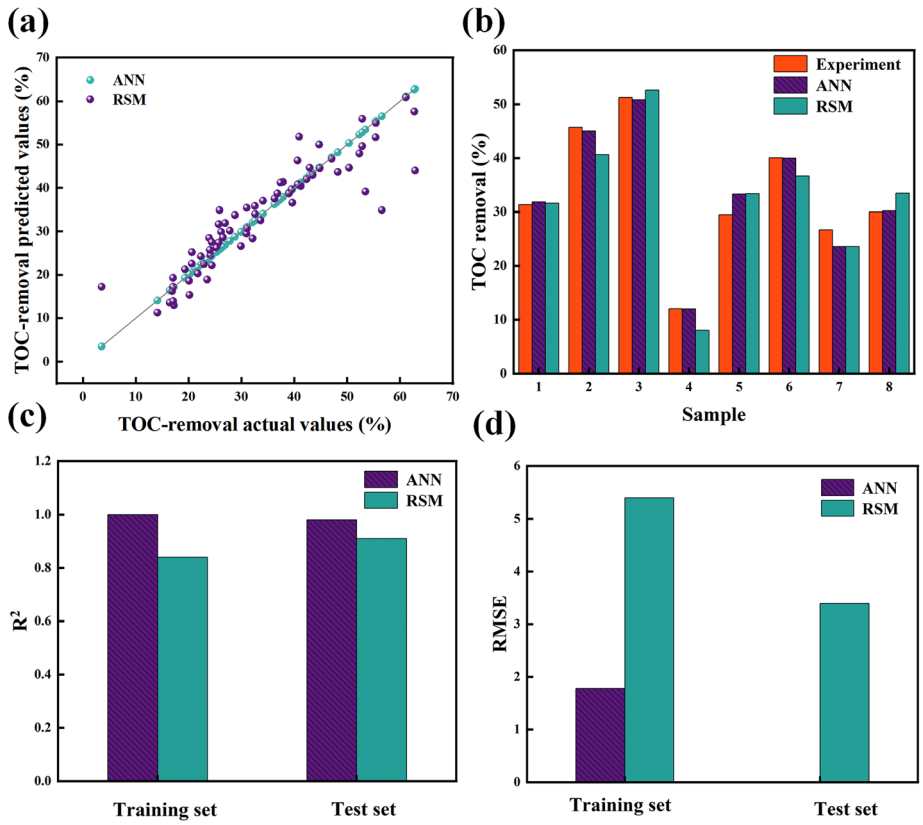

Fig. 7 Prediction of TOC removal rate using RSM and ANN models. (a) Training set, (b) test set, (c) R^2, and (d) RMSE comparisons.

RMSE = 5.4). This superior performance persisted in the test set, where the ANN model achieved near-perfect predictions (R^2 = 0.98, RMSE = 0.01) compared to RSM's respectable but lower performance (R^2 = 0.91, RMSE = 3.39). The approximately 100% improvement (eqn (7)) in prediction accuracy by the ANN model underscores its enhanced capability for forecasting TOC removal rates. These comprehensive results establish the ANN model as the more reliable choice for predicting TOC removal rates, particularly when making predictions on new data.

4. Conclusions

This study explored the treatment of *m*-cresol-contaminated wastewater using SPC as an oxidizer and compared the prediction performance of ANN and RSM models. The ANN model, with a 6-14-14-1 architecture, predicted optimal conditions as an initial solution pH of 2.3, a reaction time of 35.7 min, a SPC dosage of 2.9 g L^{-1}, a reaction temperature of 45.7 °C, a catalyst dosage of 12.9 g L^{-1}, an initial *m*-cresol concentration of 75 mg L^{-1}, achieving a TOC removal rate of 67.8% (predicted: 63.1%). In comparison, the RSM model predicted a TOC removal rate of 43.3% under its optimal conditions, with an experimental value of 38.2%. The ANN model demonstrated higher accuracy, with an RMSE of 0.01 and AARE% of 0.4%, outperforming the RSM model's RMSE of 3.39 and AARE% of 15%.

In conclusion, the ANN model exhibited superior predictive capabilities and reliability, capturing complex nonlinear relationships, achieving a *ca.* 100% improvement in prediction accuracy over the RSM model. Notably, the experimental TOC removal rate is improved by 77% under the guidance of the ANN model compared with RSM! These findings underscore the potential of ANN models to optimize AOP for wastewater treatment and offer a foundation for enhancing the efficiency of SPC-based systems in addressing environmental challenges.

In future work, this framework may be extended to complex wastewater matrices, integrated with global optimization algorithms, and advanced toward hybrid mechanism–data-driven architectures to enhance both adaptability and interpretability.

Author contributions

L. G.: conceptualization, formal analysis, investigation, methodology, software, and writing – original draft. J. Z.: conceptualization, formal analysis, investigation, and methodology. Y. C.: investigation and methodology. J. Z.: investigation and methodology. Z. L.: methodology and supervision. X. C.: methodology and supervision. L. M.: conceptualization, validation, supervision, and funding acquisition. X. L.: methodology, validation, and supervision. All authors participated in discussions and contributed to the review and editing of the manuscript.

Conflicts of interest

There are no conflicts to declare.

Data availability

The data supporting this article have been included as part of the supplementary information (SI).

Supplementary information: the SI mainly includes catalyst characterization conditions, N_2 adsorption–desorption (Fig. S1), response surface coding and corresponding experimental results (Table S1), analysis of variance and statistical parameters of RSM (Table S2), as well as discussions on the principles and methodological considerations underlying the present work. See DOI: **https://doi.org/10.1039/D5EW00689A**.

Acknowledgements

This research was sponsored by the Undergraduate Research Training, and the Shijiazhuang High-Level Science and Technology Innovation and Entrepreneurship Talent Scheme (No. 08202303). Acknowledgments are extended to King Abdullah University of Science and Technology and the Beijing Institute of Petrochemical Technology for their support.

Notes and references

1 F. Fan and X. Liu, M-cresol wastewater treatment in anaerobic fluidized bed microbial fuel cell and electricity generation performance, *Kezaisheng Nengyuan*, 2018, **36**(6), 828–835, DOI: **10.13941/j.cnki.21-1469/tk.2018.06.006**.

2 C. Yu and Z. Mao, Characteristics and kinetics study of m-cresol biodegradation by Bacillus cereus strain SMC, *Huagong Jinzhan*, 2015, **34**(5), 1453–1458, DOI: **10.16085/j.issn.1000-6613.2015.05.044**.

3 P. Liu, S. He, H. Wei, J. Wang and C. Sun, Characterization of α-Fe_2O_3/γ-Al_2O_3 Catalysts for Catalytic Wet Peroxide Oxidation of *m*-Cresol, *Ind. Eng. Chem. Res.*, 2015, **54**, 130–136, DOI: **10.1021/ie5037897**.

4 H. Guo, D. Li, Z. Li, S. Lin, Y. Wang, S. Pan and J. Han, Promoted elimination of antibiotic sulfamethoxazole in water using sodium percarbonate activated by ozone: Mechanism, degradation pathway and toxicity assessment, *Sep. Purif. Technol.*, 2021, **266**, 118543, DOI: **10.1016/j.seppur.2021.118543**.

5 Y. Zhang, X. Gao, K. Smith, G. Inial, S. Liu, L. B. Conil and B. Pan, Integrating water quality and operation into prediction of water production in drinking water treatment plants by genetic algorithm enhanced artificial neural network, *Water Res.*, 2019, **164**, 114888, DOI: **10.1016/j.watres.2019.114888**.

6 J. A. González-Amaya, *et al.*, Artificial neural network-based QSAR model for predicting degradation techniques of pharmaceutical contaminants in water bodies with experimental verification, *Environ. Sci.: Water Res. Technol.*, 2024, **10**(6), 1492–1498, DOI: **10.1039/D4EW00137K**.

7 H. E. Reynel-Ávila, I. A. Aguayo-Villarreal, L. L. Diaz-Muñoz, J. Moreno-Pérez, F. J. Sánchez-Ruiz, C. K. Rojas-Mayorga, D. I. Mendoza-Castillo and A. Bonilla-Petriciolet, A review of the modeling of adsorption of organic and inorganic pollutants from water using artificial neural networks, *Adsorpt. Sci. Technol.*, 2022, 9384871, DOI: **10.1155/2022/9384871**.

8 B. Burger, P. M. Maffettone, V. V. Gusev, C. M. Aitchison, Y. Bai, X. Wang, X. Li, B. M. Alston, B. Li and R. Clowes, A mobile robotic chemist, *Nature*, 2020, **583**(7815), 237–241, DOI: **10.1038/s41586-020-2442-2**.

9 H. Kermet-Said and N. Moulai-Mostefa, Modeling and prediction of COD and turbidity removals from dairy wastewaters by Fenton process using RSM and ANN, *Biomass Convers. Biorefin.*, 2024, **14**, 8419–8431, DOI: **10.1007/s13399-022-03187-5**.

10 I. Pantic, J. Paunovic, J. Cumic, S. Valjarevic, G. A. Petroianu and P. R. Corridon, Artificial neural networks in contemporary toxicology research, *Chem.-Biol. Interact.*, 2023, **369**, 110269, DOI: **10.1016/j.cbi.2022.110269**.

11 P. Kumar, S. Verma, R. Kaur, J. Papac, H. Kušić and U. L. Štangar, Enhanced photo-degradation of N-methyl-2-pyrrolidone (NMP): Influence of matrix components, kinetic study and artificial neural network modelling, *J. Hazard. Mater.*, 2022, **434**, 128807, DOI: **10.1016/j.jhazmat.2022.128807**.

12 M. Fan, J. Hu, R. Cao, W. Ruan and X. Wei, A review on experimental design for pollutants removal in water treatment with the aid of artificial intelligence, *Chemosphere*, 2018, **200**, 330–343, DOI: **10.1016/j.chemosphere.2018.02.111**.

13 J. Zhang, X. Yao, Y. Zhao, R. Li, X. Chen, H. Jin, H. Wei, L. Ma, Z. Mu and X. Liu, Optimizing the UV-fenton degradation of m-cresol wastewater: an experimental and artificial intelligence modeling approach, *Ind. Eng. Chem. Res.*, 2024, **63**, 921–929, DOI: **10.1021/acs.iecr.3c03847**.

14 X. Zhou, H. Sun and Z. Zhang, Application of process optimization of wastewater treatment using response surface methodology, *Huaxue Yanjiu Yu Yingyong*, 2017, **29**, 753–760.

15 R. L. Mason, R. F. Gunst and J. L. Hess, *Statistical design and analysis of experiments: with applications to engineering and science*, John Wiley & Sons, 2003.

16 Y. Wang and C. Wang, The Application of Response Surface Methodology, *Journal of Minzu University of China (Natural Sciences Edition)*, 2005, 236–240.

17 H. Zhang, C. Ye, H. Yang, X. Zhang, F. Yang and G. Liu, Optimization for treatment of oil field wastewater by coagulation sedimentation process using response surface methodology, *Sustainable Environ. Res.*, 2013, **7**, 169–174.

18 S. Qiu, T. Ren, H. Yin, Q. Cheng and S. Ma, Photodegradation properties of titanium oxide nanoparticles supported by trace iron, *Huaxue Yanjiu Yu Yingyong*, 2023, **35**, 2038–2046.

19 A. Khaleel and A. Al-Nayli, Supported and mixed oxide catalysts based on iron and titanium for the oxidative decomposition of chlorobenzene, *Appl. Catal. B*, 2008, **80**, 176–184, DOI: **10.1016/j.apcatb.2007.11.027**.

20 R. Pravina, H. Uthayakumar and A. Sivasamy, Hybrid approach based on response surface methodology and artificial neural networks coupled with genetic algorithm (RSM-GA-ANN) for the Prediction and optimization for the Photodegradation of dye using nano ZnO anchored glass fiber under solar light irradiation, *J. Taiwan Inst. Chem. Eng.*, 2023, **153**, 105248, DOI: **10.1016/j.jtice.2023.105248**.

21 K. M. Desai, S. A. Survase, P. S. Saudagar, S. S. Lele and R. S. Singhal, Comparison of artificial neural network (ANN) and response surface methodology (RSM) in fermentation media optimization: case study of fermentative production of scleroglucan, *Biochem. Eng. J.*, 2008, **41**, 266–273, DOI: **10.1016/j.bej.2008.05.009**.

22 J. Zhang, *Application of Artificial Intelligence in Advanced Oxidation Treatment of Wastewater*, Beijing Institute of Petrochemical Technology, Beijing, 2024.

23 F. X. Yang, Y. Li, X. Li and J. Yuan, The beauty premium of tour guides in the customer decision-making process: An AI-based big data analysis, *Tour. Manag.*, 2022, **93**, 104575, DOI: **10.1016/j.tourman.2022.104575**.

24 H. M. Bui, H. N. Bui, T. M. Le and R. R. Karri, Application of artificial neural networks on water and wastewater prediction: A review, *Soft Computing Techniques in Solid Waste and Wastewater Management*, 2021, pp. 95–109. DOI: **10.1016/B978-0-12-824463-0.00011-2**.

25 Q. Shang, X. Liu, M. Zhang, P. Zhang, Y. Ling, G. Cui, W. Liu, X. Shi, J. Yue and B. Tang, Photocatalytic degradation of ofloxacin antibiotic wastewater using TS-1/C_3N_4 composite photocatalyst: Reaction performance optimisation and estimation of wastewater component synergistic effect by artificial neural network and genetic algorithm, *Chem. Eng. J.*, 2022, **443**, 136354, DOI: **10.1016/j.cej.2022.136354**.

26 S. A. Ahmed and E. S. Gogina, Phenol degradation by electropersulfate process: Statistical modeling using CCD-RSM optimization, *AIP Conf. Proc.*, 2023, **2560**(1), 040007, DOI: **10.1063/5.0127226**.

27 K. Yetilmezsoy, S. Demirel and R. J. Vanderbei, Response surface modeling of Pb (II) removal from aqueous solution by Pistacia vera L.: Box–Behnken experimental design, *J. Hazard. Mater.*, 2009, **171**(1–3), 551–562, DOI: **10.1016/j.jhazmat.2009.06.035**.

28 S. Ahmed, E. Gogina and N. Makisha, Photodegradation of phenol using UVC-activated persulfate and chelated-Fe (III) combined system: RSM modeling, mechanism, effect of coexisting ions, and effect of heavy species, *J. Water Process Eng.*, 2024, **60**, 105104, DOI: **10.1016/j.jwpe.2024.105104**.