Journal of Water Process Engineering 76 (2025) 108108

Contents lists available at ScienceDirect

Journal of Water Process Engineering

journal homepage: www.elsevier.com/locate/jwpe

Machine learning modeling of microplastics removal by coagulation in water and wastewater treatment

Ahmad Hosseinzadeh [a], Farid Amirkhani [a], Nahid Azizi [b], Amir Dashti [a], John L. Zhou [a,*], Ali Altaee [a]

[a] Centre for Green Technology, School of Civil and Environmental Engineering, University of Technology Sydney, NSW 2007, Australia
[b] Khomein University of Medical Sciences, Khomein, Iran

ARTICLE INFO

Editor: Ludovic F. Dumée

Keywords:
Coagulation
Coagulation aid
Machine learning modeling
Microplastics
Water treatment

ABSTRACT

Microplastics (MPs) pose a global concern due to their persistence and potential toxicity. Coagulation is the common treatment technology for removing particles including MPs in water and wastewater. This research aims to address this challenge by developing machine learning models, including Artificial Neural Network (ANN), Least Square Support Vector Machine (LSSVM), Particle Swarm Optimization-Adaptive Neuro-Fuzzy Inference System (PSO-ANFIS), and Radial Basis Function (RBF) to predict the removal efficiency of MPs by coagulation under different conditions. Various input parameters, such as MP and coagulant concentration, solution pH and temperature were considered in these models. Through statistical analyses, the RBF model exhibited the highest accuracy with an R^2 value of 0.96 and R^2 value for ANN, PSO-ANFIS and RBF was 0.91, 0.83 and 0.79, respectively. Sensitivity analysis revealed that water temperature had the most significant negative effect, while coagulant aid showed the most positive effect on the coagulation performance for MP removal. The modeling approach and its findings provide valuable insights for improving the efficiency of MP removal in dynamic water and wastewater treatment processes.

1. Introduction

Water scarcity and pollution are major global challenges, impacting public health, environmental sustainability, and economic development [1]. With population growth and expanding industrial activities, the demand for clean water is rising [2]. However, water bodies are increasingly contaminated by pollutants such as industrial chemicals and pharmaceuticals [3], agricultural runoff [4], and microplastics (MPs) [5]. Plastics are so widespread in daily life that human exposure to MPs is unavoidable. In recent years, studies have found MPs in tap water from different origins such as ground or surface water, and desalinated water [6]. This exposure has expressed worries regarding possible health hazards. As access to safe drinking water is a UN Sustainable Development Goal [7], it's crucial to accurately assess the effect of presence of MPs on health. Additionally, research shows that MPs can enter the food chain, particularly through seafood, further contributing to human exposure [8].

Coagulation is a widely used primary treatment method for removing fine particles from both drinking water and wastewater [9,10]. It has proven effective in removing MPs through mechanisms like surface charge neutralization and bridging [11]. The jar test is commonly used to determine the optimal coagulant dosage for particle removal in drinking water treatment plants, simulating coagulation under controlled conditions. However, it faces significant challenges, including being time-consuming, labor-intensive, and reliant on operator's judgment, which can introduce errors. Additionally, it is impractical for real-time adjustments as it cannot quickly respond to alterations in untreated water characteristics, such as fluctuations in pH, turbidity, or contaminants [12]. To overcome these challenges, advanced methods like nonlinear models and automatic control systems are needed to quickly adjust coagulant dosages based on real-time raw water quality, especially for microplastic removal, while reducing human error [13].

In recent years, ML has garnered much attention, with various algorithms being used to propose accurate models for coagulant dosage rates. Heddam conducted a comparative study applying three ML models to predict coagulant dosage rates at the Boudouaou water treatment plant [14]. Jayaweera et al. noted that modeling coagulant dosage rates for water with elevated turbidity necessitates a greater

* Corresponding author.
E-mail addresses: Hosseinzadeh26@gmail.com (A. Hosseinzadeh), junliang.zhou@uts.edu.au (J.L. Zhou).

https://doi.org/10.1016/j.jwpe.2025.108108
Received 7 April 2025; Received in revised form 27 May 2025; Accepted 4 June 2025
Available online 20 June 2025

number of input variables compared to models applied for low turbidity water. The results indicated that the ELM-RBF model exhibited moderate performance ($R^2 < 0.59$) for high turbidity water, and high performance ($R^2 = 0.88$) for low turbidity water [15]. Wadkar et al. employed CFFNN to model the coagulant dosage rate. The performance of the CFFNN was deemed favorable ($R^2 = 0.88$) [16]. Raj et al. evaluated the elimination of 1-μm carboxylated PS microspheres using the iron-chitosan-based CFS technique. The experimental results were validated by an ANN, yielding an RMSE of 1.0643 and an R^2 of 0.9997 [17]. Liu et al. developed a deep learning (DL) model to estimate the turbidity of effluent from the sedimentation tank using 5761 data points and PAC floc images from a DWTP in East China (2022). The DL model was accurate than traditional ML models, achieving an R^2 of 0.97, compared to R^2 of 0.76 for the traditional model. Incorporating operational data from the sedimentation tank improved prediction accuracy by 79.6 % [18]. Hossain explored the use of hybrid BOA and ML techniques to predict microplastic removal during coagulation. The results showed that the hybrid BOA-BRT model outperformed others, with a high R^2, demonstrating its robustness in forecasting MPs elimination efficiency. In comparison, the hybrid BOA-ANN model had higher MAE value than BOA-BRT [19].

Based on the literature review, coagulation is the dominating process for removing particles such as MPs in water and wastewater. However, the coagulant dosage in water treatment plants is not easily adjusted in response to changes in raw water quality. Coagulation as a dominating separation process for removing fine particles in water and wastewater treatment plants needs innovative modeling by ML methods which can capture non-linear complex relationships in a timely fashion. Recognizing the limitations posed by the availability of measured water parameters, ML methods can effectively leverage raw water quality parameters measured at its inlet. While earlier studies primarily concentrated on the application of standalone ML models, it is noteworthy that, to the best of the author's knowledge, none has explored the potential benefits derived from integrating metaheuristic algorithms with ML models. Due to complicated nature and knowledge of the modeling process of MP removal using coagulants and lack of a comprehensive modeling method for the prediction of this process, the aim of this study is to develop novel ML models for simulating and predicting the removal efficiency of MPs by different coagulants in water and wastewater treatment plants.

For the first time, accurate ML methods were used to predict the removal of MPs as the hazardous emerging contaminants using the various coagulants in different operational situations by varying the MP size, coagulant dose, pH, temperature of the solution and coagulant aid material concentration. ANN, PSO-ANFIS, LSSVM, and RBF by MATLAB software were developed to predict the removal of MPs by various coagulants in water treatment plants. Various input parameters, such as MP and coagulant aid concentration, MP size, pH, temperature, and sedimentation time, were considered in these models. Various graphical and statistical methods were utilized to assess the models performance and accuracy.

2. Model theory

2.1. Least squares support vector machine (LSSVM)

The SVM functions are supervised learning techniques that evaluate sets of data for classification or regression research utilizing correlated learning methods. The solution computations in SVM can be optimized using the LSSVM technique, as proposed by Suykens and Vandewalle [20]. This modification enhances the performance of SVM by incorporating least squares principles, making it particularly effective in certain scenarios where traditional SVM methods may face challenges [21,22].

A support vector machine (SVM) model is a robust technique for supervised learning, originally established through statistical learning theory, and is widely recognized for its effectiveness in data categorization, predictive modeling, and function estimation in high-dimensional feature spaces [23]. In regression contexts, SVMs are typically formulated to address convex quadratic programming problems governed by inequality constraints, aiming to identify support vectors. Nonetheless, the computational burden introduced by constrained optimization algorithms often results in substantial time and memory consumption when handling large-scale datasets. Furthermore, SVMs are susceptible to uncertainties arising from data and process noise [24], and they perform poorly in the presence of non-Gaussian noise and outlier contamination [25].To mitigate these limitations, Suykens et al. proposed a modified framework known as Least Squares Support Vector Machine (LSSVM) [20], which employs equality constraints in place of inequality ones and adopts a sum-of-squares error cost function similar to that used in conventional neural network training [26]. A notable advantage of LSSVM in regression applications is its ability to utilize the entire dataset during model construction [27]. Fig. 1 shows the schematic diagram of CSA-LSSVM model.

2.1.1. Coupled simulated annealing (CSA)

Simulated Annealing (SA) is a probabilistic search algorithm designed to locate global optima in complex combinatorial optimization tasks. Initially introduced by Metropolis et al. [29] and later advanced by Kirkpatrick et al. [30], SA draws its conceptual basis from the metallurgical process of annealing, in which a material is heated to a molten state and then gradually cooled to refine its crystalline structure and minimize imperfections. This metaheuristic has proven highly effective for large-scale datasets, as it systematically explores the search space to approximate the global minimum of an objective function. Coupled Simulated Annealing (CSA), inspired by the concept of coupled local minimizers (CLM), involves multiple SA processes operating concurrently in a coordinated framework. While various global search algorithms have been utilized to solve high-dimensional optimization problems, a common limitation across them is the requirement for extensive evaluations of the objective function to ensure convergence to the global optimum. To mitigate this issue, hybrid strategies combining global and local search techniques are often employed to accelerate convergence, albeit with a reduced likelihood of attaining the true global minimum. To overcome these limitations, Xavier-de-Souza et al. [31] introduced a novel optimization framework based on the interconnection of several SA instances. The fundamental idea behind this coupling mechanism is to enable collaborative exploration by sharing information between solvers, thereby improving robustness and increasing the probability of reaching the global solution.

2.2. Adaptive neuro-fuzzy inference system (ANFIS)

The ANFIS strategy was originally introduced by Jang [32] as a five-layered hybrid algorithm designed for the combining of both fuzzy systems and neural networks. The ANFIS topology can typically be trained using a hybrid learning algorithm that incorporates both fuzzy reasoning and back-propagation techniques. The fundamental explanation of the layers in the ANFIS architecture, involving two input parameters and one output parameter, has been provided before [33,34].

2.3. Multilayer perceptron (MLP) networks

Intelligent systems have the capacity to gain experience, improve their performance, and adapt to changes in the environment [35]. Neural networks offer several advantages, including the ability to handle large volumes of data and the capacity to generalize results. ANNs are parallel distributed systems composed of various artificial neurons. Multi-Layer Perceptron (MLP) networks, a type of neural network, consist of three constituents: input, hidden, and output layers. MLP networks can have one or multiple hidden layers, with each layer containing several neurons. Determining the optimum of neurons number in the hidden layer can be accomplished through either trial and

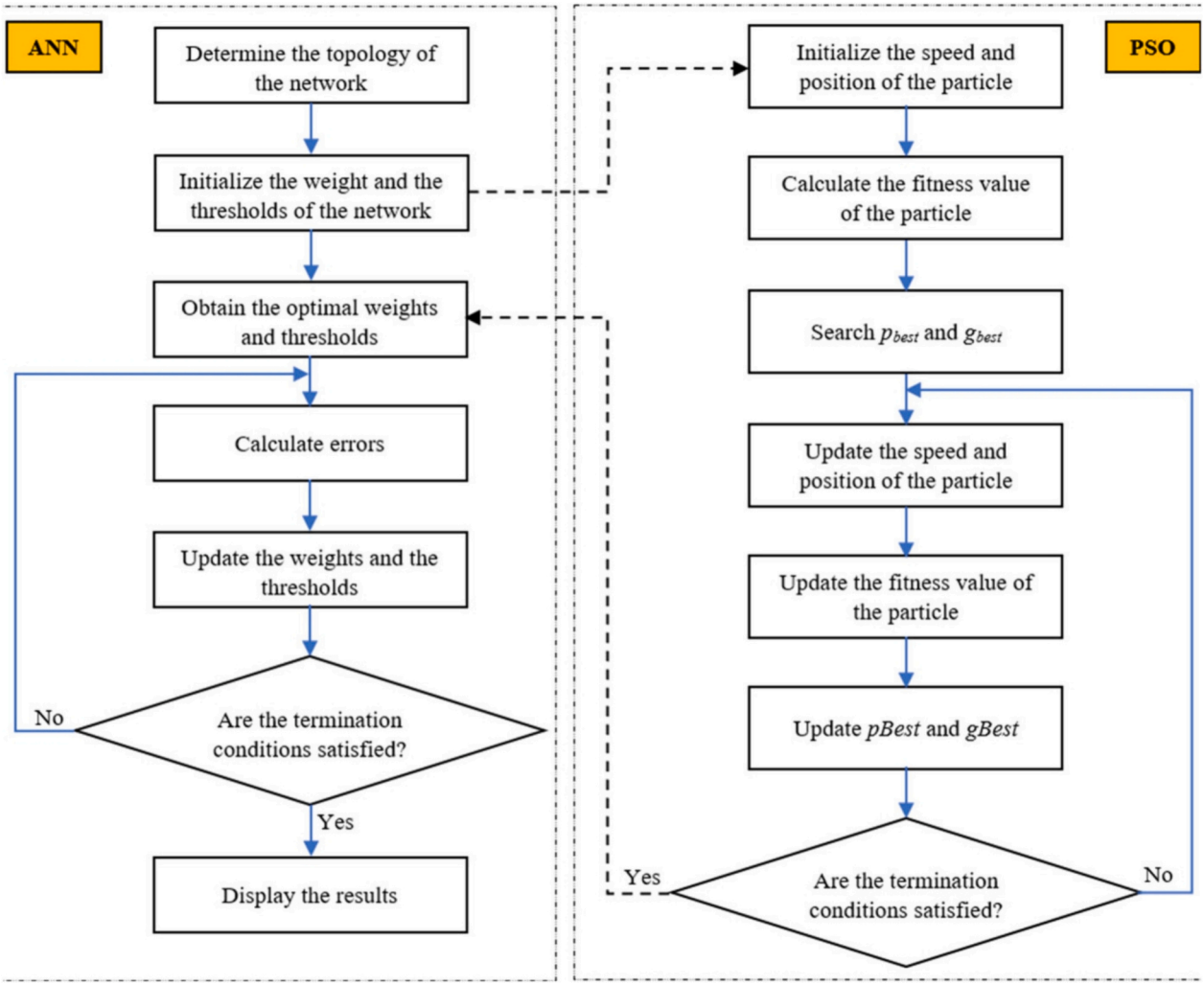

Fig. 1. A layout of a CSA-LSSVM algorithm [28]. Reproduced from ref [28].

error methods or an intelligent approach. The performance of developed systems is often evaluated using mean square error. In these systems, errors are backpropagated through the network, and weights and biases are optimized through iterations called epochs. It is crucial to determine an appropriate number of epochs to avoid undertraining or overtraining the system. Undertraining occurs when the system does not have sufficient time to complete the learning process, while overtraining results in the system memorizing data without actually learning, leading to poor performance in forecasting test datasets [36]. Striking a balance in the number of epochs is essential for optimal performance in predicting test datasets.

2.4. Radial basis function (RBF)

RBF represents a variant of feed-forward neural networks, devised on the premise of localized basis functions and iterative function approximation networks [37,38]. The training of RBF networks exhibits a comparative ease when juxtaposed with MLP networks, attributable to the straightforward and fixed three-layer structural configuration. RBF displays exceptional adaptability to data not encountered during the training phase. The RBF neural network consists of an input, a hidden, and an output layer [39]. Each neuron within the hidden layer incorporates a radial basis function serving as a non-linear activation function. The outputs of these functions are reciprocally correlated with the distance from the neuron's center. Noteworthy model parameters encompass the neurons' centers, the distance scale, and the configuration of the radial function. Leveraging linear optimization techniques, RBF attains a globally optimal solution for the adjustable weights, minimizing the Mean Square Error [39].

3. Experimental data for the coagulation removal of MPs

To investigate the coagulation removal of MPs, various input parameters derived experimentally were considered in the models. These parameters include MP size, coagulant type and dose (mg/L), water pH, water temperature (°C), and type and concentration (mg/L) of coagulant aid material. It is important to note that, given the complexity of coagulants and coagulant aids as shown in Fig. 2, a numerical value is assigned to distinguish them as a specific input. The output of the models is the MP removal efficiency (%). In total, 405 datasets were selected from literature work. Of the total datasets, 75 % were randomly selected for training the models, and 25 % used for testing model performance. Data were selected randomly using MATLAB. Like similar previous studies [40–42], train-test split method was used since it is quick and has lower computational time. No preprocessing was made on the data. Table 1 provides details of the data used in this study, and Table 2 provides a summary statistics of input data. The accuracy of the employed models for estimating actual values are assessed using metrics such as the correlation coefficient (R^2), mean relative error (MRE) and mean absolute error (MAE). These metrics are formulated in Eqs. (1)–(3):

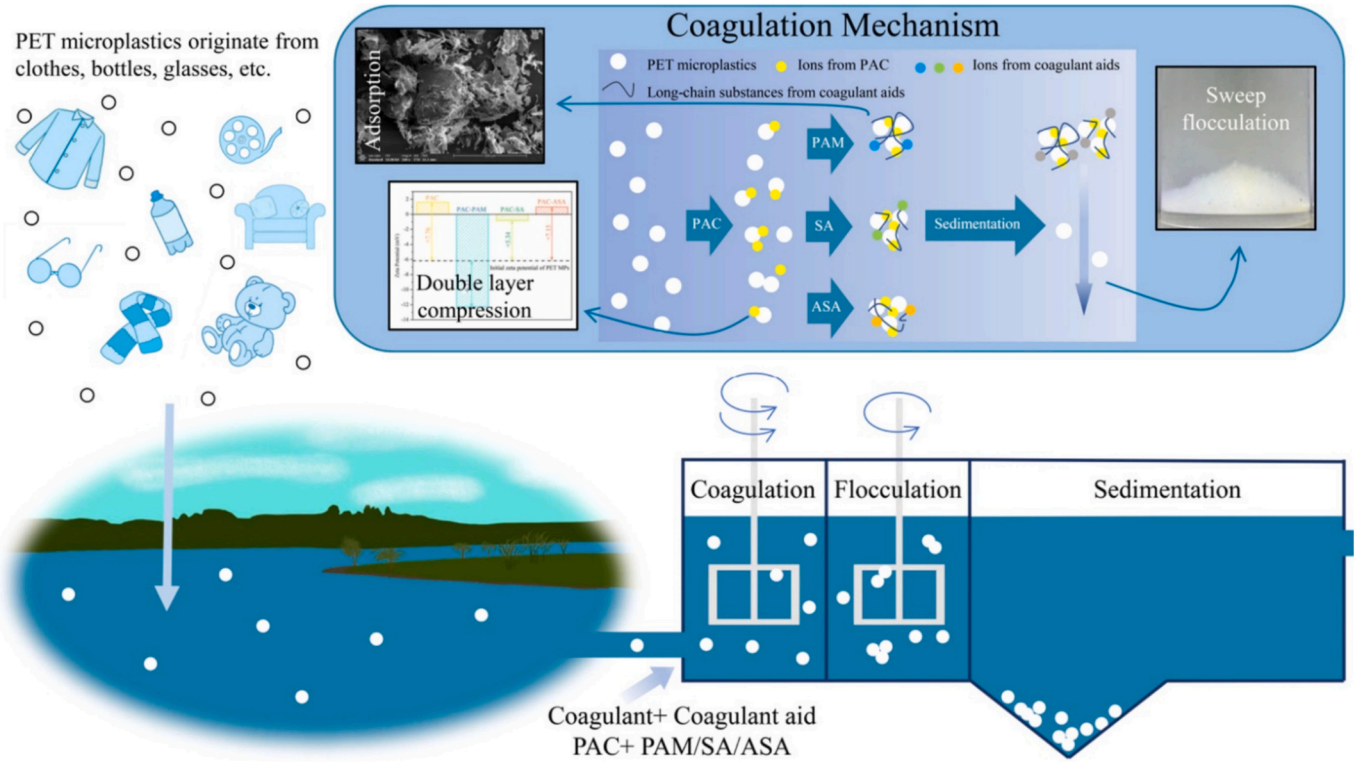

Fig. 2. Experimental workflow for coagulation and MP removal [43]. Reproduced from ref. [43] with permission from Elsevior, License Number: 6045670478746 © Elsevier.

Table 1
Details of actual data employed in this research.

MP size (μm)	MP type	Coagulant	Coagulant dose (mg/L)	Coagulant aid	Coagulant aid dose (mg/L)	pH	T (°C)	Removal efficiency (%)	Dataset	Ref.
0–500	Polyethylene terephthalate	PAC	20–50	Anionic PAM, ASA, SA	0–100	3–9	25	33.14–100	44	[43]
1–6.3	Polystyrene	$FeCl_3$	5.62–199	None	–	6.5–7.3	25	14.50–100	22	[44]
1	Polystyrene	C-577	26.71–46.19	None	–	7.3	25	27.98–63.91	7	[44]
0–5000	Polyethylene	$FeCl_3{\cdot}6H_2O$	27.03–1351.45	None	–	7	25	0.32–13.23	24	[45]
0–5000	Polyethylene	$AlCl_3{\cdot}6H_2O$	24.14–3621.45	Anionic PAM, Cationic PAM	0–15	6–8	25	0.19–64.29	136	[45]
0–5000	Polyethylene	$FeCl_3{\cdot}6H_2O$	27.03–1351.45	Anionic PAM, Cationic PAM	0–15	7	25	0.38–91.24	128	[46]
10–100	Polyethylene	$AlCl_3{\cdot}6H_2O$	10–50	PC sand	0–500	7	20	8.11–100	32	[47]
1	commercial fluorescent polystyrene microspheres	$AlCl_3{\cdot}6H_2O$	250	SDS, Tween	0–10	7	30	76.42–99.18	12	[48]

Table 2
Summary statistics of experimental data.

	Min	Max	Mean	Standard deviation
MPs LB size	0	2000	638.3627	727.8529
MPs UB size	1	5000	1557.6714	1725.0906
Coagulant code	1	5	1.8247	0.9473
Coagulant dose (mg/L)	5.6231	3621.4500	391.0174	570.4163
pH	3	9	6.9780	0.5689
Temp (°C)	20	30	24.7531	1.6315
Coagulant aid code	0	7	1.0988	1.4793
Coagulant aid dose (mg/L)	0	500	16.5556	63.6522

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left[x_i^{predicted} - x_i^{experimental}\right]^2}{\sum_{i=1}^{n}\left[x_i^{experimental} - x^m\right]^2}, x_m = \frac{\sum_{i=1}^{n} x_i^{experimental}}{n} \quad (1)$$

$$MRE = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|x_i^{predicted} - x_i^{experimental}\right|}{x_i^{experimental}} \quad (2)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|x_i^{experimental} - x_i^{predicted}\right| \quad (3)$$

4. Results and discussion

4.1. ML model development

Various ML methodologies, including ANN, PSO-ANFIS, LSSVM and RBF in MATLAB were applied for the computation of MP removal in

coagulant systems. Training data were used to train the models. In particular, an ANFIS model coupled with PSO was employed for the precise prediction of MP removal in the coagulant system like previous studies [40,49]. PSO served the purpose of training the model and determining optimal parameter values within the ANFIS framework. The determination of optimal values for Particle Swarm Optimization (PSO) parameters was accomplished through a trial-and-error approach. Table 3 presents a comprehensive overview of the PSO-ANFIS parameters. Trial and error method was used to find the optimum parameters of PSO algorithm.

Utilizing the coupled simulated annealing Algorithm (CSA) [31,50] for optimization purposes, the optimal values for the LSSVM algorithm, including γ and σ^2, were derived employing the RBF kernel function. Specifically, these values were identified as 42.3863 and 73.4652, respectively. In the context of MLP networks, it was determined that a single hidden layer is sufficient. Since with one hidden layer, we have a simpler model and adequate accuracy and according to previous studies [51–53], the model was developed with one hidden layer. Consequently, our MLP structure was designed with a single hidden layer to reduce computational time.

The default loss function utilized for the ANN model developed is the mean squared error (MSE). A logsig activation function was used as the hidden layer activation function. The optimizer used is the Levenberg-Marquardt algorithm (trainlm), which is MATLAB's standard training function for feedforward neural networks.

The algorithm's validity was assessed by incrementally increasing the number of neurons from 1 to 50 within the determined hidden layer. Notably, the highest efficiency was observed when the MLP comprised 45 neurons in the hidden layer.

The RBF model necessitates the determination of two tuning variables, specifically spread and the maximum number of neurons (MNN), integral to the structure of the RBF model. Optimizing these variables is imperative for enhancing the accuracy of the model. The trial and error algorithm was employed to ascertain the optimal values for these tuning variables. The optimum values for Spread and MNN were identified as 62 and 120, respectively.

4.2. Comparison of different ML models

In this study, we employed four ML models of PSO-ANFIS, RBF, ANN, and LSSVM for predicting MP removal using the coagulant systems. Fig. 3 shows the accuracy of the developed models based on R^2 value. The strengths of the models developed by PSO-ANFIS, RBF, ANN, and LSSVM in predicting MP removal using the coagulant systems were assessed using various statistical indices (Table 4). The findings indicate that RBF model presented the highest accuracy. RBF exhibited higher prediction strengths in terms of R^2 for the total dataset. Regarding MAE and MRE, the RBF model demonstrated fewer errors compared to the other models.

The application of the developed ML algorithms in estimating MP removal using the coagulant system is provided in Table 5. RBF predicted MP removal using different coagulants, with the highest performance with the lowest MRE. The minor disparities in the forecast capabilities of present models applied in alternative uses might be ascribed to the characteristics of the data, like the variety of clotting agents, quantity and kind of inputs, and the form of computational procedures. Distinct computational procedures might produce diverse outcomes in various applications. The prediction accuracy of each model significantly lower when $FeCl_3$-H_2O is used as a coagulant compared to other coagulants. The reason for that can be unseen experimental procedures and situation and maybe the interaction between experimental parameters like pH and temperature compared to other data and $FeCl_3$-H_2O performance due to its hydrolysis, charge neutralization, or floc formation kinetics.

Table 3
Details of the PSO-ANFIS model for predicting MP removal in coagulation systems.

Parameter	Value
Maximum iteration number	100
Maximum particle number	1000
Initial inertia weight (Wmin)	0.7
Inertia weight damping ratio (Wdamp)	0.89
Cognitive acceleration (C1)	1
Social acceleration (C2)	2
Number of fuzzy rules	10

4.3. Effects of input parameters on MP removal efficiency

To comprehend the behavior of MP removal during coagulation, varying dosages of $FeCl_3{\cdot}6H_2O$ and $AlCl_3{\cdot}6H_2O$ were employed. Notably, $AlCl_3{\cdot}6H_2O$ exhibited superior performance in MP removal compared to $FeCl_3{\cdot}6H_2O$. As depicted in Fig. 4a, the removal efficiency initially exhibited an upward trend, followed by stabilization. Conversely, even with a substantial dosage of $FeCl_3{\cdot}6H_2O$, a low removal efficiency for PE was noted [45]. Fig. 4b illustrates the comparison of removal efficiency for small MP particles under different dosages of $FeCl_3{\cdot}6H_2O$ at pH 7.0. The results indicated that smaller PE particle sizes correlated with higher removal efficiency [45]. Polyacrylamide (PAM) is frequently employed to augment coagulation in water treatment due to its outstanding performance [54]. In comparison to cationic PAM, however, the growth rate of removal efficiency was significantly higher in the presence of anionic PAM, as depicted in Fig. 4c [45].

For lower dosage of coagulant (0.5 mM) of $AlCl_3{\cdot}6H_2O$, the removal of MP was hardly changing, even for the tiny particle size. However, the elimination efficiency of PE diminished with raising solution pH with a high coagulant dosage (5 mM $AlCl_3{\cdot}6H_2O$), particularly for the tiny sized particles [45]. The investigation extended to assess MP removal efficiencies at pH 3–9 [43]. As shown in Fig. 4d, pH has a slight impact on the removal performance. In the Poly-Aluminum Chloride (PAC) system, pH 3 has resulted in a slight reduction of removal efficiency. As is known, the hydrolysis reaction involves the amalgamation of Al^{3+} with OH^-, the OH^- concentration serves as the pivotal factor determining pH levels. Consequently, the reduction induced by low pH might be linked to the restraining impact of acidic environments on the hydrolysis of PAC. This, in turn, exerts a detrimental influence on the hydrolysis by-products, leading to a corresponding limitation in double layer compression. Conversely, elevated pH levels can positively influence the efficacy of removal processes [55]. Garvasis et al. [56] discovered that the efficiency of aluminum sulfate nanoparticles remained unaffected when the solution pH varied between 3 and 9. Likewise, Zahrim et al. [57] observed minimal impact of pH on color removal using ferric chloride-APAM, even with pH fluctuations from 6 to 10. These findings suggest that enhanced coagulation remains effective when dealing with water at diverse pH levels. To gain deeper insights into the chemical dynamics and stability of solutions during the flocculation phase, zeta potentials across varying pH levels were measured [58]. Chen et al. [8] noted that the zeta potentials of nanoparticles exhibited minimal variation as pH was increased, consistently maintaining a stable negatively charged milieu, indicating the unaltered electrical characteristics of NPs in response to pH shifts. In addition, the results show that PSO-ANFIS predictions are inferior to the other models, except at pH 5.

4.4. Sensitivity analysis

Sensitivity analyses was conducted to elucidate the impact of each input parameter on the target variable, namely MP removal. The quantitative effect of each parameter was determined using a relevancy factor, as per the previously outlined method. The relevancy factor is constrained within the range of −1 to 1, where higher absolute values indicate a more pronounced effect of the respective parameter. In this context, a positive effect signifies an increase in the target variable as a

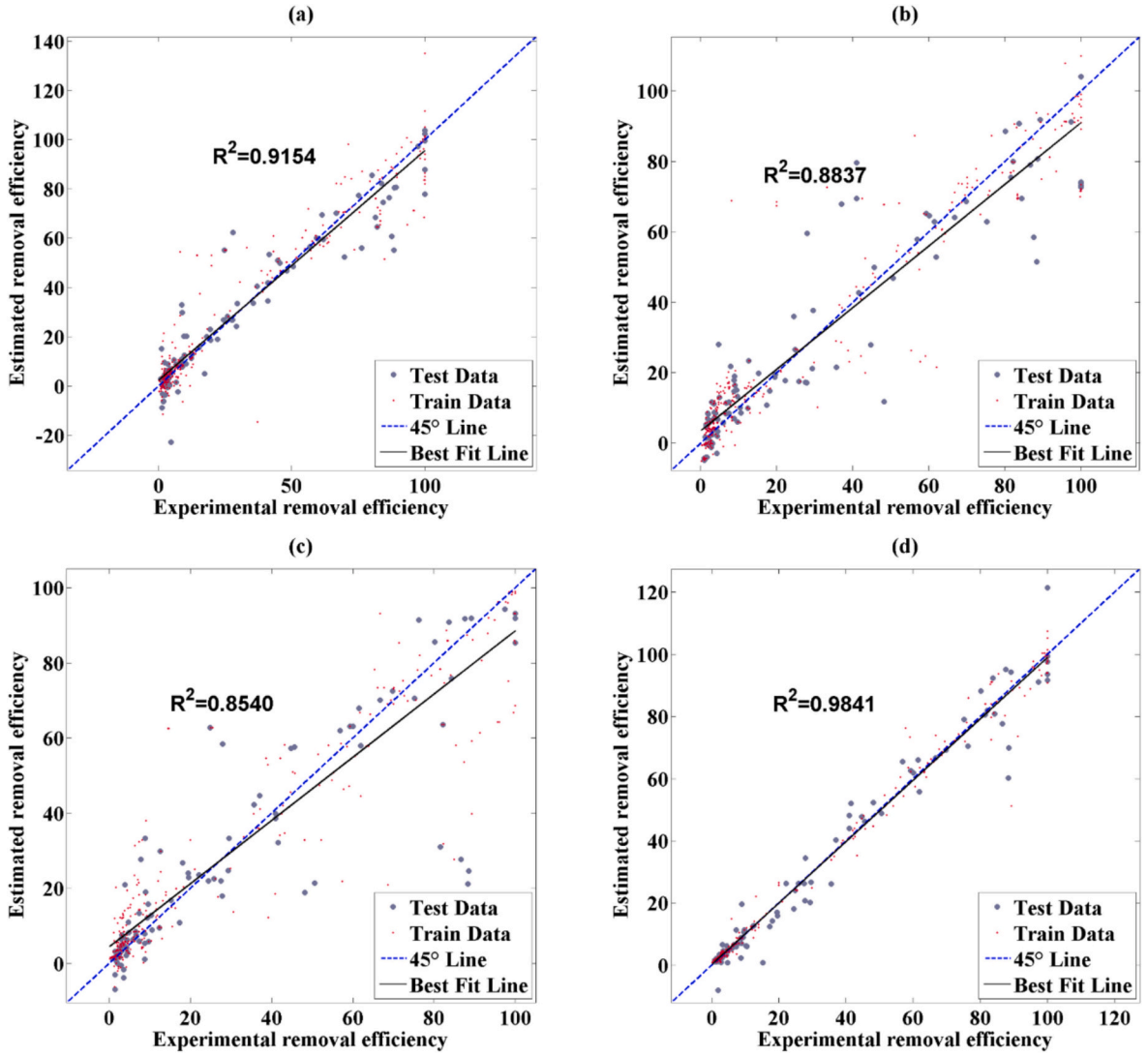

Fig. 3. Comparison of estimated MP removal efficiency by the models of a) ANN, b) PSO-ANFIS, c) LSSVM, and d) RBF in relation to experimentally determined results.

Table 4
Statistical values of the proposed models for the training, test and total datasets.

Parameter	Training dataset	Test dataset	Total dataset
ANN			
R^2	0.9162	0.9153	0.9154
MAE	5.6151	6.0379	5.7205
MRE	0.7399	0.8091	0.7571
PSO-ANFIS			
R^2	0.8982	0.8379	0.8837
MAE	6.3770	8.7538	6.9698
MRE	0.8382	0.8497	0.8411
LSSVM			
R^2	0.8748	0.7936	0.8540
MAE	6.9953	8.3995	7.3455
MRE	0.7304	0.5673	0.6897
RBF			
R^2	0.9887	0.9695	0.9841
MAE	1.5891	3.5847	2.0868
MRE	0.1401	0.2454	0.1664

Table 5
MRE of the proposed models for the prediction of MP removal using different coagulants.

Coagulant	ANN	PSO-ANFIS	LSSVM	RBF
PAC	0.0886	0.1557	0.0653	0.0308
$FeCl_3$	0.6390	0.0996	0.7896	0.0210
C-577	0.2822	0.2556	0.2525	0.0853
$FeCl_3{\cdot}6H_2O$	0.8271	0.9181	0.9933	0.2056
$AlCl_3{\cdot}6H_2O$	0.8943	1.0570	0.5908	0.1873

specific input parameter rises, while a negative effect indicates a decrease in the target variable with an increase in the specific input parameter. The relevancy factor serves as a metric to quantitatively assess the influence of each parameter on the target variable, providing insights into the sensitivity of the system to individual input variations. The details about sensitivity analysis by relevancy factor method and further information can be found in previous studies (Amirkhani et al.; [59]).

Among the main operating parameters, the minimum and maximum size of MPs, coagulant dose, pH and temperature were identified to have a direct impact on the results, meaning that changes in any of these parameters could affect the performance of MP removal during

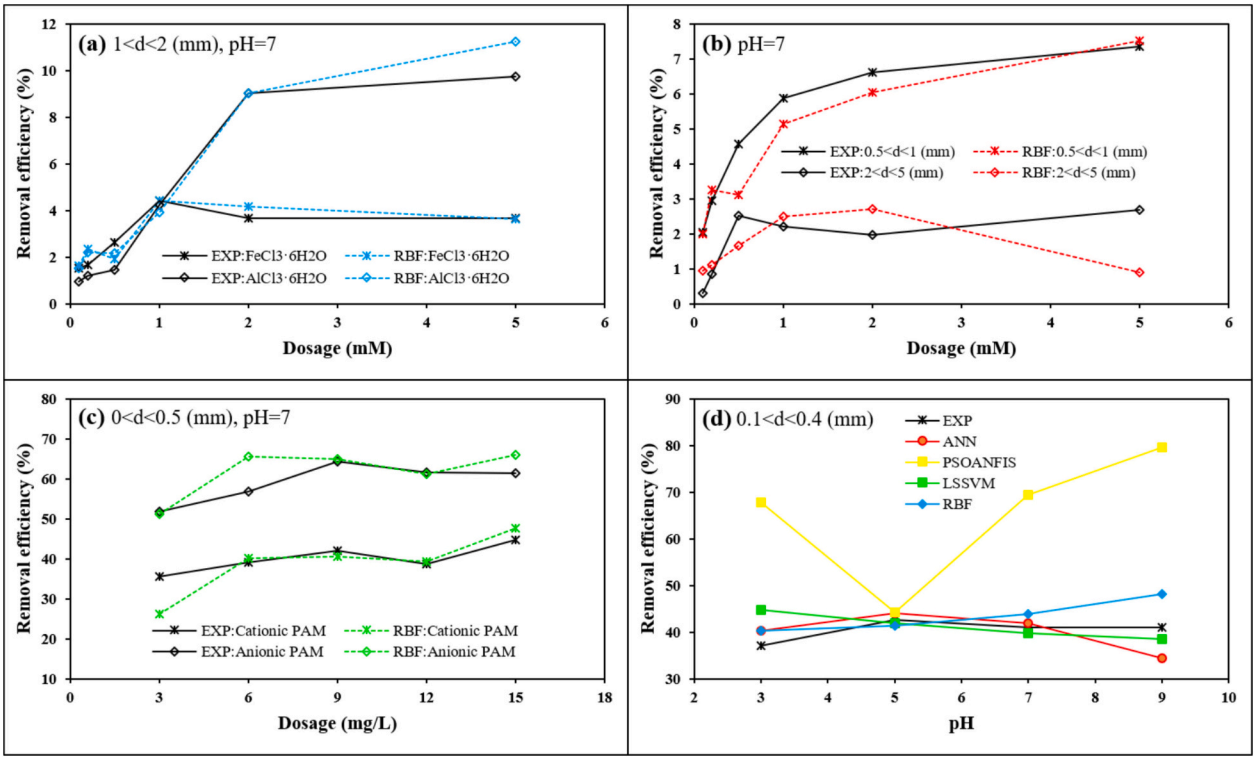

Fig. 4. Comparison of experimental and estimated MPs removal efficiency at room temperature in the presence of a) $FeCl_3.6H_2O$ and $AlCl_3{\cdot}6H_2O$, b) $FeCl_3{\cdot}6H_2O$, c) 5 mM $AlCl_3{\cdot}6H_2O$ with cationic and anionic PAM, and d) 50 mg/L PAC. Experimental data from [43,45].

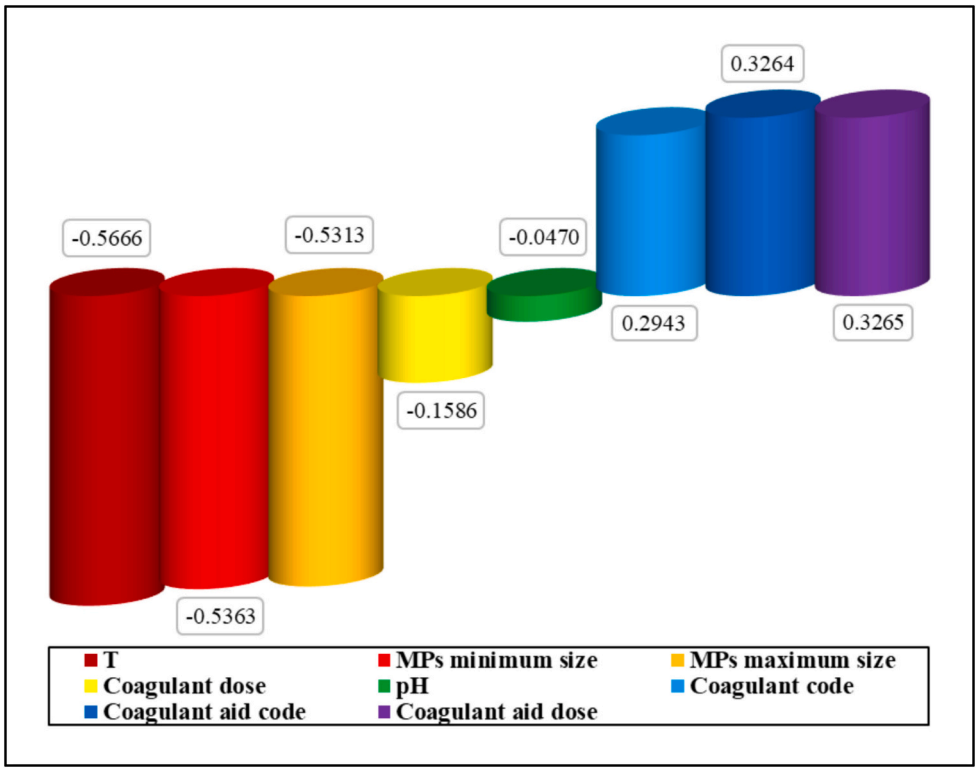

Fig. 5. Sensitivity analysis of the operating parameters in the modeling of MP removal during coagulation process in water and wastewater treatment.

coagulation. The sensitivity analysis results (Fig. 5) highlight that water temperature had the most negative effect on MP removal with a relevancy factor of −0.5666, followed by the minimum and maximum size of MPs (0.5365 and 0.5313, respectively). It was observed that pH and coagulant dose had a negligible impact on the removal efficiency. In addition, the type and dose of coagulant aid presented the most positive

impacts on MP removal efficiency, with a relevance factor of +0.3264 and + 0.3265, respectively.

5. Conclusions

In this study, the MP removal from water and wastewater treatment plants was estimate using ML models. Four ML models, including ANN, PSO-ANFIS, LSSVM and RBF were employed to estimate the removal efficiency of MPs during the coagulation treatment of water and wastewater. Different operating parameters, such as MP minimum and maximum size, coagulant dose, pH, temperature of the solution and coagulant aid were used as input parameters. The developed models, namely ANN, PSO-ANFIS, LSSVM, and RBF demonstrated high accuracy with R^2 values of 0.91, 0.88, 0.85, and 0.98 respectively. According to the sensitivity analysis, water temperature exhibited the most significant adverse impact on MP removal efficiency. The second most crucial factors were the minimum and maximum size of MP, whereas pH and coagulant dose were found to have minimal influence on MP removal efficiency. The type and dose of coagulant aid demonstrated the most favorable effect on enhancing MP removal efficiency.

The study addresses a significant void by offering a comparative assessment of various ML approaches for modeling microplastic removal, a topic that has not been thoroughly explored in previous research. Additionally, the research presents a data-driven modeling framework that allows for quick predictions and modifications of coagulation performance, eliminating the necessity for repeated tests. From a practical perspective, the models created can help water treatment personnel optimize the use of coagulants, boost the effectiveness of MP removal, and decrease operating cost. This directly benefits public health and aligns with the wider objective of reducing plastic pollution in water ecosystems. In summary, the results showcase the potential of ML to improve environmental process management and highlight the significance of combining data analysis with traditional water treatment methods.

CRediT authorship contribution statement

Ahmad Hosseinzadeh: Writing – review & editing, Resources, Investigation, Conceptualization. **Farid Amirkhani:** Writing – review & editing, Writing – original draft, Visualization, Validation, Data curation. **Nahid Azizi:** Data curation. **Amir Dashti:** Writing – original draft, Software, Methodology. **John L. Zhou:** Writing – review & editing, Supervision. **Ali Altaee:** Writing – review & editing.

Declaration of 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.

Acknowledgements

We thank the University of Technology Sydney for a PhD studentship and strategic support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jwpe.2025.108108.

Data availability

Data will be made available on request.

References

[1] J. Chen, Y. Xie, S. Sun, M. Zhang, P. Yan, F. Xu, L. Tang, S. He, Efficient nitrogen removal through coupling biochar with zero-valent iron by different packing modes in bioretention system, Environ. Res. 223 (2023) 115375.
[2] Y. Jurczynski, R. Passos, L.C. Campos, A review of the most concerning chemical contaminants in drinking water for human health, Sustainability 16 (16) (2024) 7107.
[3] P.K. Singh, U. Kumar, I. Kumar, A. Dwivedi, P. Singh, S. Mishra, C.S. Seth, R. K. Sharma, Critical review on toxic contaminants in surface water ecosystem: sources, monitoring, and its impact on human health, Environ. Sci. Pollut. Res. 31 (45) (2024) 56428–56462.
[4] R. Nehru, C.-W. Chen, C.-D. Dong, A review of smart electrochemical devices for pesticide detection in agricultural food and runoff contaminants, Sci. Total Environ. 935 (2024) 173360.
[5] Y. Ma, X. Gu, Y. Zhang, P. Yan, M. Zhang, S. Sun, T. Ren, L. Tang, S. He, Unveiling the microplastic perturbation on surface flow constructed wetlands with macrophytes of different life forms: responses of nitrogen removal and sensory quality, J. Hazard. Mater. 477 (2024) 135283.
[6] I.V. Kirstein, F. Hensel, A. Gomiero, L. Iordachescu, A. Vianello, H.B. Wittgren, J. Vollertsen, Drinking plastics?–quantification and qualification of microplastics in drinking water distribution systems by μFTIR and Py-GCMS, Water Res. 188 (2021) 116519.
[7] W. WHO, Microplastics in drinking-water, World Health Organization, Geneva, 2019.
[8] G. Chen, Q. Feng, J. Wang, Mini-review of microplastics in the atmosphere and their risks to humans, Sci. Total Environ. 703 (2020) 135504.
[9] Y. Duan, S.-Y. Sun, J. Zhao, H. Yuan, Microplastics affect the removal of dye in textile wastewater: adsorption capacity and its effect on coagulation behavior, Sep. Purif. Technol. 359 (2025) 130505.
[10] Y. Mao, Z. Hu, H. Li, H. Zheng, S. Yang, W. Yu, B. Tang, H. Yang, R. He, W. Guo, Recent advances in microplastic removal from drinking water by coagulation: Removal mechanisms and influencing factors, Environ. Pollut. 349 (2024) 123863.
[11] E. Haddad, E.A. Ben-David, M. Habiby, D.L. Angel, A.M. Booth, I. Sabbah, Enhancing nano and microplastics destabilization: synergistic effects of natural mucin and conventional coagulants in water and wastewater treatment, Environ. Technol. Innov. 38 (2025) 104189.
[12] M. Ghasemi, M. Hasani Zonoozi, N. Rezania, M. Saadatpour, Predicting coagulation–flocculation process for turbidity removal from water using graphene oxide: a comparative study on ANN, SVR, ANFIS, and RSM models, Environ. Sci. Pollut. Res. 29 (48) (2022) 72839–72852.
[13] G. Rytel, Floating facilities providing social services–a floating salvation Army refuge in Paris and "school ship" in Vienna, J. Water Land Dev. 45 (2020).
[14] S. Heddam, Water Engineering Modeling and Mathematic Tools, Elsevier, 2021, pp. 475–489.
[15] C. Jayaweera, M. Othman, N. Aziz, Improved predictive capability of coagulation process by extreme learning machine with radial basis function, J. Water Process Eng. 32 (2019) 100977.
[16] D.V. Wadkar, R.S. Karale, M.P. Wagh, Application of cascade feed forward neural network to predict coagulant dose, J. Appl. Water Eng. Res. 10 (2) (2022) 87–100.
[17] S. Raj, B. Mahanty, S. Hait, Coagulative removal of polystyrene microplastics from aqueous matrices using FeCl3-chitosan system: experimental and artificial neural network modeling, J. Hazard. Mater. 468 (2024) 133818.
[18] H. Liu, Y. Chen, X. Pan, J. Zhang, J. Huang, E. Lichtfouse, G. Zhou, H. Ge, Image recognition enhances efficient monitoring of the coagulation-settling in drinking water treatment plants, J. Clean. Prod. 482 (2024) 144251.
[19] S.Z. Hossain, Machine learning approaches for predicting microplastic removal, Arab. J. Sci. Eng. (2025) 1–14.
[20] J.A. Suykens, J. Vandewalle, Least squares support vector machine classifiers, Neural. Process. Lett. 9 (1999) 293–300.
[21] J. Suykens, T. Van Gestel, J. De Brabanter, B. De Moor, J. Vandewalle, Least squares support vector machines, World Scientific Publishing, Singapore, 2002.
[22] H. Wang, D. Hu, Comparison of SVM and LS-SVM for Regression, IEEE, 2005, pp. 279–283.
[23] M. Seyyedattar, S. Zendehboudi, S. Butt, Relative permeability modeling using extra trees, ANFIS, and hybrid LSSVM–CSA methods, Nat. Resour. Res. 31 (1) (2022) 571–600.
[24] H. Xu, V. Hassani, C.G. Soares, Truncated least square support vector machine for parameter estimation of a nonlinear manoeuvring model based on PMM tests, Appl. Ocean Res. 97 (2020) 102076.
[25] X. Lu, W. Liu, C. Zhou, M. Huang, Robust least-squares support vector machine with minimization of mean and variance of modeling error, IEEE Trans. Neural Netw. Learn. Syst. 29 (7) (2017) 2909–2920.
[26] A. Kamari, A.H. Mohammadi, A. Bahadori, S. Zendehboudi, Prediction of air specific heat ratios at elevated pressures using a novel modeling approach, Chem. Eng. Technol. 37 (12) (2014) 2047–2055.
[27] A. Chamkalani, S. Zendehboudi, A. Bahadori, R. Kharrat, R. Chamkalani, L. James, I. Chatzis, Integration of LSSVM technique with PSO to determine asphaltene deposition, J. Pet. Sci. Eng. 124 (2014) 243–253.
[28] A. Azarpour, S. Zendehboudi, N.M.C. Saady, Deterministic models for performance analysis of lignocellulosic biomass torrefaction, ACS Omega 10 (2025) 6470–6501.
[29] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, Equation of state calculations by fast computing machines, J. Chem. Phys. 21 (6) (1953) 1087–1092.
[30] S. Kirkpatrick, C.D. Gelatt Jr., M.P. Vecchi, Optimization by simulated annealing, Science 220 (4598) (1983) 671–680.

[31] S. Xavier-de- Souza, J.A. Suykens, J. Vandewalle, D. Bollé, Coupled simulated annealing, IEEE Trans. Syst. Man Cybern. B Cybern. 40 (2) (2009) 320–335.
[32] J.-S. Jang, ANFIS: adaptive-network-based fuzzy inference system, IEEE Trans. Syst. Man Cybern. 23 (3) (1993) 665–685.
[33] K.H. Lee, First course on fuzzy theory and applications, Springer Science & Business Media, 2004.
[34] M. Seyyedattar, M.M. Ghiasi, S. Zendehboudi, S. Butt, Determination of bubble point pressure and oil formation volume factor: extra trees compared with LSSVM-CSA hybrid and ANFIS models, Fuel 269 (2020) 116834.
[35] R. Santos, M. Rupp, S. Bonzi, A. Fileti, Comparison between multilayer feedforward neural networks and a radial basis function network to detect and locate leaks in pipelines transporting gas, Chem. Eng. Trans. 32 (2013) 1375–1380.
[36] S. Haykin, N. Network, A comprehensive foundation, Neural Netw. 2 (2004) (2004) 41.
[37] K.-L. Du, M. Swamy, Radial basis function networks, in: Neural networks in a softcomputing framework, 2006, pp. 251–294.
[38] M.J. Powell, Radial basis functions for multivariable interpolation: a review, Algorithms for approximation (1987) 143–167.
[39] A. Karkevandi-Talkhooncheh, A. Rostami, A. Hemmati-Sarapardeh, M. Ahmadi, M. M. Husein, B. Dabir, Modeling minimum miscibility pressure during pure and impure CO2 flooding using hybrid of radial basis function neural network and evolutionary techniques, Fuel 220 (2018) 270–282.
[40] A. Dashti, F. Amirkhani, M. Jokar, A. Mohammadi, K.-W. Chau, Insights into the estimation of heavy metals ions sorption from aqueous environment onto natural zeolite, Int. J. Environ. Sci. Technol. 18 (2021) 1773–1784.
[41] A. Dashti, M. Asghari, H. Solymani, M. Rezakazemi, A. Akbari, Modeling of CaCl2 removal by positively charged polysulfone-based nanofiltration membrane using artificial neural network and genetic programming, Desalin. Water Treat. 111 (2018) 57–67.
[42] H. Riasat Harami, A. Dashti, P. Ghahramani Pirsalami, S.K. Bhatia, A. Ismail, P. Goh, Molecular simulation and computational modeling of gas separation through polycarbonate/p-nitroaniline/zeolite 4A mixed matrix membranes, Ind. Eng. Chem. Res. 59 (38) (2020) 16772–16785.
[43] Y. Zhang, G. Zhou, J. Yue, X. Xing, Z. Yang, X. Wang, Q. Wang, J. Zhang, Enhanced removal of polyethylene terephthalate microplastics through polyaluminum chloride coagulation with three typical coagulant aids, Sci. Total Environ. 800 (2021) 149589.
[44] K. Rajala, O. Grönfors, M. Hesampour, A. Mikola, Removal of microplastics from secondary wastewater treatment plant effluent by coagulation/flocculation with iron, aluminum and polyamine-based chemicals, Water Res. 183 (2020) 116045.
[45] B. Ma, W. Xue, C. Hu, H. Liu, J. Qu, L. Li, Characteristics of microplastic removal via coagulation and ultrafiltration during drinking water treatment, Chem. Eng. J. 359 (2019) 159–167.
[46] B. Ma, W. Xue, Y. Ding, C. Hu, H. Liu, J. Qu, Removal characteristics of microplastics by Fe-based coagulants during drinking water treatment, J. Environ. Sci. 78 (2019) 267–275.
[47] N.K. Shahi, M. Maeng, D. Kim, S. Dockko, Removal behavior of microplastics using alum coagulant and its enhancement using polyamine-coated sand, Process. Saf. Environ. Prot. 141 (2020) 9–17.
[48] Y. Xia, X.-M. Xiang, K.-Y. Dong, Y.-Y. Gong, Z.-J. Li, Surfactant stealth effect of microplastics in traditional coagulation process observed via 3-D fluorescence imaging, Sci. Total Environ. 729 (2020) 138783.
[49] A. Dashti, M. Jokar, F. Amirkhani, A.H. Mohammadi, Quantitative structure property relationship schemes for estimation of autoignition temperatures of organic compounds, J. Mol. Liq. 300 (2020) 111797.
[50] J.A. Suykens, J. Vandewalle, B. De Moor, Intelligence and cooperative search by coupled local minimizers, Int. J. Bifurc. Chaos 11 (08) (2001) 2133–2144.
[51] F. Amirkhani, A. Dashti, H. Abedsoltan, A.H. Mohammadi, K.-W. Chau, Towards estimating absorption of major air pollutant gasses in ionic liquids using soft computing methods, J. Taiwan Inst. Chem. Eng. 127 (2021) 109–118.
[52] F. Amirkhani, A. Dashti, H. Abedsoltan, A.H. Mohammadi, A.G. Chofreh, F.A. Goni, J.J. Klemeš, Estimating flashpoints of fuels and chemical compounds using hybrid machine-learning techniques, Fuel 323 (2022) 124292.
[53] A. Dashti, M. Raji, A. Azarafza, A. Baghban, A.H. Mohammadi, M. Asghari, Rigorous prognostication and modeling of gas adsorption on activated carbon and zeolite-5A, J. Environ. Manag. 224 (2018) 58–68.
[54] R. Zhang, S. Yuan, W. Shi, C. Ma, Z. Zhang, X. Bao, B. Zhang, Y. Luo, The impact of anionic polyacrylamide (APAM) on ultrafiltration efficiency in flocculation-ultrafiltration process, Water Sci. Technol. 75 (8) (2017) 1982–1989.
[55] M. Sillanpää, M.C. Ncibi, A. Matilainen, M. Vepsäläinen, Removal of natural organic matter in drinking water treatment by coagulation: a comprehensive review, Chemosphere 190 (2018) 54–71.
[56] J. Garvasis, A.R. Prasad, K. Shamsheera, P. Jaseela, A. Joseph, Efficient removal of Congo red from aqueous solutions using phytogenic aluminum sulfate nano coagulant, Mater. Chem. Phys. 251 (2020) 123040.
[57] A.Y. Zahrim, Z. Dexter, C.G. Joseph, N. Hilal, Effective coagulation-flocculation treatment of highly polluted palm oil mill biogas plant wastewater using dual coagulants: Decolourisation, kinetics and phytotoxicity studies, J. Water Process Eng. 16 (2017) 258–269.
[58] A. Pandey, V.V. Pathak, R. Kothari, P.N. Black, V. Tyagi, Experimental studies on zeta potential of flocculants for harvesting of algae, J. Environ. Manag. 231 (2019) 562–569.
[59] A. Hemmati-Sarapardeh, A. Larestani, N.A. Menad, S. Hajirezaie, Applications of artificial intelligence techniques in the petroleum industry, Gulf Professional Publishing, 2020.