



OPEN Optimization and modeling of sulfur removal from liquid fuel using carbon-based adsorbents through synergistic application of RSM and machine learning

Karim Maghfour Sarkarabad, Mohsen Shayanmehr & Ahad Ghaemi✉

This research investigates the adsorption desulfurization of liquid fuels using artificial neural networks (ANN) and response surface methodology (RSM) approaches. The effectiveness of sulfur removal was predicted by analyzing five important factors: temperature, concentration, surface area, fuel/adsorbent, and time. We employed radial basis function (RBF) and multilayer perceptron (MLP) algorithms for ANN modeling. The optimal MLP configuration, utilizing the Levenberg–Marquardt (Trainlm) algorithm, consisted of three hidden layers with 20, 17, and 9 neurons, respectively, while the optimal RBF network contained 43 neurons. The MLP network's determination coefficient (R^2) was 0.98 over 30 epochs, and its mean squared error (MSE) was 0.0028. The RBF network also obtained an R^2 of 0.98 and an MSE of 0.0026 over 40 epochs. A two-factor interaction design served as the basis for the RSM model, which produced an R^2 of 0.91. A comparison of the RSM, MLP, and RBF models, using the average absolute relative deviation, indicated that the ANN models, particularly the RBF model, produced more accurate predictions than the RSM model. The findings show that temperature and concentration were the two most significant factors influencing sulfur removal efficiency. Overall, artificial neural networks outperformed the RSM approach in predicting desulfurization performance, providing a more reliable modeling tool for optimizing the sulfur removal process.

Keywords Adsorptive desulfurization, Activated carbon, Artificial neural network, Response surface methodology

List of symbols

BET	Surface area (m^2/g)
X_i	Fuel/adsorbent ratios (g/g)
T	Temperature (K)
C	Sulfur concentration (ppm)
t	Time (min)
f	Function
i	Number of neurons in the hidden layer
N	Number of neurons
N	Number of data sets for training (–)
R^2	Correlation coefficient
w	Weight factor (–)
w_{ij}	Weight related to each hidden neuron (–)
w_{ik}	Weight connecting the i th neuron to the k th neuron
x	Input variable (–)
x_i	Input examples (attributes)
X	Independent variable of RSM method
$x^{(j)}$	Input of j th layer
$x_k^{(j)}$	Value of the k th neuron in the j th layer

School of Chemical, Petroleum and Gas Engineering, Iran University of Science and Technology, Tehran, Iran.
✉email: aghaemi@iust.ac.ir

y_j	Target output
Z	Packed height (m)
Greek letters	
α	Constants in turbulence models (-)
β_{jk}	Bias weight for neuron j in layer k
γ_i	Output value of the i th neuron
ε_0	Residual factor of RSM method
σ	Width of radial basis function (RBF) kernel (-)
σ_i	Spread of Gaussian function (-)
$\varphi(x)$	Nonlinear activation function
φ_i	Gaussian radial basis function

Abbreviations

2FI	Two-factor interaction
AARD	Average absolute relative deviation
ANN	Artificial neural networks
GDR	Generalized delta rule
MLP	Multilayer perceptron
MSE	Mean square error
RBF	Radial basis function
RSM	Response surface methodology

Terminology

Activation function	The activation function is a mathematical function in between the input feeding the current neuron and its output going to the next layer.
Bias	Bias is a constant which helps the model in a way that can fit best for the given data
Epoch	In the training process, the inputs enter each training step and give outputs that are compared with the target to calculate an error. With this process, weights and biases are calculated and modified in each epoch
Neurons	Neurons are the basic units of the large neural network
Weight	Represents the importance and strengths of the feature/input to the Neurons

The combustion of sulfides in transportation fuel produces toxic sulfur components, prompting strict regulations worldwide to limit their emissions. Within the petroleum industry, organic sulfur compounds like disulfides, mercaptans, and thiophenes contribute to these emissions, impacting the environment, human health, and refinery equipment. Recent mandates from authorities such as the US Environmental Protection Agency and the European Union highlight the urgency in addressing this issue^{1–3}. Over the past 2 decades, researchers have identified several alternative methods to hydrodesulfurization (HDS)^{4,5} for sulfur removal that are effective under mild conditions. The leading technologies among these are^{6,7}: bio-desulfurization (BDS)⁸, adsorptive desulfurization (ADS)^{9–11}, extractive desulfurization (EDS)¹², and oxidative desulfurization (ODS)¹³. The global challenge of removing organic sulfur from liquid fuel prompts the exploration of cost-effective solutions. Adsorptive desulfurization emerges as a promising technology to tackle this issue. With its competitiveness, ADS has garnered considerable interest from both academia and industry^{14,15}. Porous solid sorbents, including activated carbon (AC)^{14,16,17}, zeolites^{9,18} and MOF^{19,20} have gained attention for their cost-effectiveness, low energy consumption, and ease of regeneration. In air pollution control, they are extensively utilized for desulfurization due to their ability to achieve ultra-low sulfur levels and high selectivity for sulfur compounds. Additionally, adsorptive desulfurization with porous adsorbents is gaining traction as an alternative for ultra-deep desulfurization. This method offers advantages such as easy accessibility and renewability of adsorbents, moderate operating conditions, low energy consumption, thermal stability, and desirable textural features^{21,22}. Activated carbon is a highly effective material for removing pollutants due to its extensive microporosity and surface area²³. This makes it a valuable adsorbent for capturing sulfur oxides (SO₂) from various sources. Studies confirm that AC's superior performance stems from its large surface area, functional surface groups, and spacious pores²⁴. Notably, AC can be derived from various renewable precursors like cow dung, rice husk, or wood. However, raw AC may require modification for optimal sulfur removal. Through a process involving impregnation with transition metals (like copper or nickel), AC's efficiency in capturing aromatic sulfur compounds from fuel oils is significantly enhanced⁷. The understanding of adsorption equilibria is facilitated through mathematical formulations known as adsorption isotherms. Additionally, adsorption kinetics models are utilized to describe the temporal evolution of the adsorption process through rate equations^{25,26}. Artificial intelligence (AI) algorithms have demonstrated their efficacy in addressing diverse real-world challenges across engineering and industrial sectors, with notable applications in the petroleum industry²⁷. Given the time-consuming and expensive nature of laboratory experiments, the utilization of predictive mathematical models becomes imperative. Artificial neural networks have emerged as powerful tools in various fields, offering superior accuracy in predicting nonlinear multivariate simulations compared to conventional linear methods. ANN, comprising input, hidden, and output layers, facilitate modeling without prior knowledge of the underlying physical or chemical processes, making them applicable in diverse areas such as adsorption and reaction kinetics²⁸. In the context of desulfurization capacity forecasting, ANN have proven beneficial, offering a viable alternative where traditional insights are lacking. Utilizing experimental data, both multilayer perceptron and radial basis function networks are developed, along with a quadratic model, to establish semi-empirical predictive models. A comparative analysis between ANN and response surface methodology models highlights the predictive capabilities of each approach, with ANN showing promise in optimizing the adsorption process to maximize desulfurization uptake capacity. The versatility of ANN models further extends to facilitating

desulfurization adsorption plant design, underscoring their significance in modern nonlinear statistical modeling^{29,30}. Noora Naif Darwish³¹ conducted a study investigating the adsorptive desulfurization of diesel oil using two commercial powdered activated carbons (PAC1 and PAC2) and one granular activated carbon (GAC). Darwish's research evaluated the impact of adsorbent amount, temperature, and contact time on sulfur removal and the ignition quality of diesel fuel. Adsorption isotherms were analyzed using Langmuir and Freundlich models, with the Freundlich model fitting PAC1 and PAC2 well. Additionally, a feed-forward neural network was used to correlate experimental data, demonstrating accurate predictions of sulfur removal capacities. This study aims to develop advanced neural networks for the adsorption desulfurization of liquid fuels, a critical process in reducing harmful sulfur emissions. We will explore the adsorption desulfurization using both RSM and ANN, with a focus on the MLP and RBF approaches. Key independent factors in this study include surface area, temperature, concentration, time, and fuel/adsorbent. These variables will serve as inputs, while the removal percentage of sulfur will be the response variable for the modeling approaches. The RSM technique will be employed to identify ideal parameters and design a semi empirical model that accurately reflects the impact of these independent variables on the removal percentage. This method is particularly valuable for understanding the interactions between variables and optimizing the desulfurization process under various conditions. Similarly, the primary goal of the artificial neural networks approach is to establish the optimal network architecture and ascertain the most effective network biases and weights. This involves a detailed exploration of the relationship between the removal percentage and independent variables. For the MLP network, this includes determining the ideal size of the hidden layer, selecting the appropriate number of neurons and the most suitable training functions. In the radial basis function network, the focus will be on optimizing the number of neurons, the spread value, and the total count of epochs. In addition to the core modeling objectives, this study will also compare the predictive accuracy and robustness of the RSM and ANN approaches. By leveraging the strengths of both methods, we aim to develop a comprehensive modeling framework that can be applied to various desulfurization scenarios. The insights gained from this study will not only advance the field of fuel desulfurization but also contribute to the broader application of neural network modeling in environmental engineering. Furthermore, this research will address the scalability and practical implementation of the optimized models in industrial settings. By conducting sensitivity analyses and validation with experimental data, we aim to ensure that the developed models are both reliable and applicable to real-world desulfurization processes. The ultimate goal is to enhance the efficiency and effectiveness of adsorption desulfurization, thereby contributing to cleaner fuel production and reduced environmental impact. Mazen Khaled³² conducted a detailed investigation into the adsorption isotherms and kinetics of thiophene and dibenzothiophene (DBT) in hexane, serving as a model diesel fuel. This study compared the adsorption efficiency of thiophene and DBT on multi-walled carbon nanotubes (MWCNT) and graphene oxide (GO) with that of activated carbon, which was used as a benchmark. Results indicated that MWCNT achieved a DBT removal efficiency of 68.8%, almost twice the efficiency of GO, though it was about 25% less efficient than AC. Furthermore, the research demonstrated that these adsorbents were significantly more proficient in removing DBT from the model fuel than thiophene. Olawumi Oluwafolakemi Sadare³³ research centered on the creation and assessment of a dual process combining adsorption and biodesulfurization (AD/BDS) to eliminate sulfur compounds from South African petroleum distillates. The study utilized adsorbents like pomegranate and neem leaf powder, activated carbon, carbon nanotubes (CNTs) and functionalized carbon nanotubes (FCNT) to test their efficiency in removing dibenzothiophene from both model and actual diesel. The effectiveness of adsorption was measured through comprehensive characterization and kinetic analyses. The results demonstrated that the integrated AD/BDS process is highly effective in desulfurizing diesel, ensuring compliance with strict environmental standards. Ibrahim and Aljanabi³⁴ conducted a study on batch adsorption desulfurization of diesel fuel with an initial sulfur content of 580 ppm, utilizing activated carbon for the physical adsorption of refractory sulfur compounds. They examined the effects of various parameters such as temperature, time, ratio of diesel to activated carbon, particle size of activated carbon, mixing speed and initial sulfur concentration on desulfurization efficiency. The process reduced the sulfur concentration to 247 ppm, achieving a 57% desulfurization efficiency. Kinetic analysis revealed that the pseudo-second order model accurately predicted the equilibrium sorption capacity, while the Freundlich isotherm provided the best fit for the adsorption data, indicating physical adsorption. Nunthaprechachan et al.³⁵ investigated the removal of dibenzothiophene from n-octane using activated carbons derived from sewage sludge (S-ACs). They investigated how various activating agents, ratios, carbonization temperatures, and durations impact the properties of S-ACs and their ability to adsorb DBT. The S-AC activated with KOH demonstrated the most effective DBT removal, achieving a rate of 70.6%, which exceeded that of commercially available activated carbon. The adsorption process adhered to the Langmuir isotherm model. Al-Khodor et al.³⁶ studied sulfur removal from crude oil 2.5 wt.% from Iraq's Halfaya Oil Field using activated carbon adsorption. They investigated AC dosage, time, and temperature effects, applying Langmuir, Freundlich, and Temkin isotherm models. Temkin models best fit the data. Thermodynamic assessments indicated spontaneous, endothermic adsorption, reducing 28% desulfurization efficiency crude oil.

The summary of experimental adsorption desulfurization processes are presented in Table 1. The results of the conducted research show that there is a need for recognition of effective factors and the interaction of factors with each other on the efficiency percentage of removing sulfur compounds. Therefore, in this research, two methods of RSM and neural networks models have been used to evaluate and analyze the behavior of the adsorptive desulfurization process. Using the combination of the two methods, the effective factors on the removal percentage of sulfur compounds were determined.

The novelty of this study is uses both artificial neural networks and response surface methodology to simulate and optimize the removal of sulfur from liquid fuels using carbon-based adsorbents. In this study, ANN models and RSM are thoroughly compared using both radial basis function (RBF) and multilayer perceptron (MLP) algorithms, with ANNs exhibiting better prediction accuracy. In addition to outperforming RSM, the

Researcher	Surface area (m ² /m ³)	Fuel/adsorbent (g/g)	Temperature (K)	Sulfur concentration (ppm)	Time (min)	Removal (%)	Refs.
Mazen Khaled	217.24, 882.24	21.97–359.32	303	1.77–251.00	8.30–241.00	1.28–96.70	³²
Sadare	225.05, 245.49, 1154.87	9.41–66.23	298–308	174.08–1513.78	9.71–180.00	9.41–83.80	³³
Ibrahim and Aljanabi	930.10	1.66–8.32	313	281.00–581.00	9.91–211.00	13.12–98.79	³⁴
Nunthaprechachan et al.	896.00	140.60	401–701	100.00	29.88–119.40	8.91–80.60	³⁵
Al-Khodor et al.	519.40	32.74–158.93	303	580.00	90.00–120.00	3.48–27.90	³⁶

Table 1. Summary of experimental works on adsorption desulfurization processes.

Author	Adsorbent	Variable	Response	ANN type	Hidden layer	Training function	R ²	Refs.
Maghfour Sarkarabad et al.	Carbon-based	Temperature, concentration, surface area, fuel/adsorbent, time	Removal (%)	RSM-ANN	3 layers	Levenberg–Marquardt	0.98	This study
Shokouhi et al.	<i>N</i> -methylpyrrolidone (NMP)	Temperature, pressure	H ₂ S solubility	ANN-GA	Not specified	Genetic algorithm	0.94	³⁷
Niknam Shahrak et al.	Porous MOF materials	Surface area, pore volume, pore diameter	Water vapor adsorption	ANN-ANFIS	Multiple layers	Back propagation	0.98	³⁹
Esfandiyari et al.	[BMIM][AlCl ₄] (ionic liquid)	Volume ratio, extraction temperature, extraction time	Thiophene removal	GA-ANFIS-PSO	Multi-layered	Particle swarm optimization	0.95	⁴⁰
Esfandiyari et al.	Co(III)/Al ₂ O ₃ catalyst	Pressure, temperature, CO/H ₂ ratio, time	CH ₄ , CO ₂ , CO composition	MLP-ANN	2 layers	Levenberg–Marquardt	0.97	⁴¹
Amiri Delouei et al.	Nanofluid + ultrasonic vibrations	Fluid inlet temperature, volume fraction, flow rate	Nusselt number, heat transfer rate	ANN-ANFIS-PSO	3 layers	Back propagation with PSO	0.94	³⁸
Nobakht et al.	AC/Fe ₂ (MoO ₄) ₃	pH, stirring rate, temperature, dosage, contact time	Dye adsorption efficiency	ANN	Single layer	Back propagation	0.93	⁴²

Table 2. Overview of studies on adsorption and prediction models.

ANN models improved the determination coefficient (R²) and reduced prediction errors (MSE) to provide higher accuracy when compared to conventional experimental techniques. Finding important variables like temperature and concentration also helps to improve the optimization process and provides a more reliable and effective option for process improvement than traditional experimental methods. Mohammad Shokouhi et al. used artificial intelligence techniques to create a neural network model optimized using a genetic algorithm (GA) to forecast the solubility of hydrogen sulfide (H₂S) in *N*-methylpyrrolidone (NMP). The model was configured using experimental data, with solubility as the output and temperature and pressure as inputs. Using GA, neural network parameters such as learning rates, momentum, and neuron count were changed to optimize the model's architecture design. Using the Peng–Robinson–Stryjek–Vera (PRSV) cubic equation of state, the predictions of the GA-ANN model were evaluated using experimental and thermodynamic data. The GA-ANN model showed remarkable accuracy and was a dependable and effective method of predicting H₂S solubility in NMP, at least in comparison to the PRSV model, which had an absolute relative deviation (ARD%) of 1.91%³⁷. Morteza Esfandiyari et al. used particle swarm optimization (PSO) to develop an adaptive neuro-fuzzy inference system (ANFIS) and artificial neural networks (ANNs) for predicting important performance parameters in a double-pipe counter-flow heat exchanger. The hot flow rates (113–257 l/h), inlet fluid temperatures (40–60 °C), ultrasonic excitation power levels (0 and 60 W), and nanoparticle volume percentages (0–0.8%) were all varied in their studies. The Nusselt number, number of transfer units (NTU), heat transfer rate, and system effectiveness were all predicted using the models. With correlation coefficients above 94.84%, their results showed that the ANN-PSO and ANFIS-PSO models were both quite accurate³⁸ (Table 2).

Process description

Adsorption desulfurization is a pivotal technique in liquid fuel treatment, known for its effective removal of sulfur contaminants under relatively gentle operational parameters. Adsorption desulfurization operates under mild conditions, exhibits high selectivity towards thiophenic compounds, and achieves superior desulfurization efficiency. Adsorption, fundamentally involves a process of mass transfer where molecules adhere to a surface through intermolecular forces. This phenomenon plays a crucial role in the effectiveness of adsorption-based desulfurization methods³⁶. The adsorption quantity is determined using the equation provided below:

$$q_e = \frac{V(C_i - C_e)}{M} \quad (1)$$

C_i represents the initial concentration and C_e denotes the equilibrium concentration. V stands for the volume, while M denotes the mass. Adsorption stands as a promising method for desulfurizing liquid fuels, offering a selective approach to remove sulfur compounds. The core principle involves extracting sulfur contaminants, constituting less than 1% of the fuel mass, through targeted adsorption. This method effectively purifies the fuel, preserving 99% of its sulfur-free composition⁴³. Adsorption desulfurization involves the utilization of an adsorbent to extract sulfur compounds from liquid fuels. This method optimizes temperature and pressure

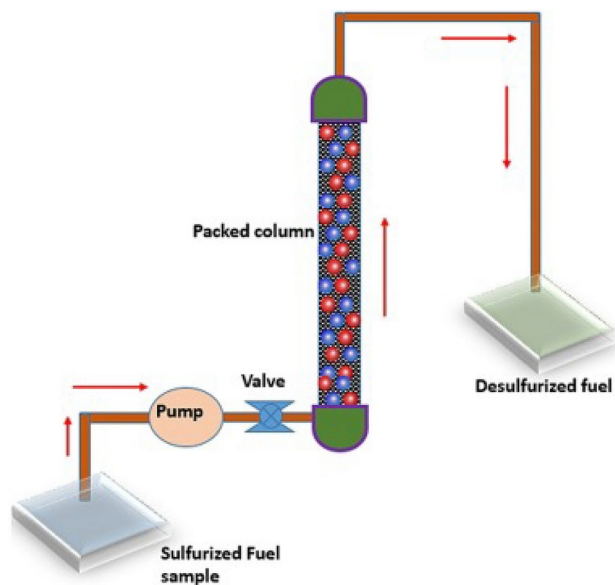


Fig. 1. Flowchart diagrams of an experimental setup for a fixed-bed system in adsorptive desulfurization¹⁷.

Factor	Type	Name	Units	Minimum	Maximum	Coded low	Coded high	Mean	Std. Dev
A	Input	BET	m ² /g	217.24	1154.87	-1 ↔ 217.24	+1 ↔ 1154.87	680.30	415.93
B	Input	Fuel/adsorbent	g/g	1.66	343.59	-1 ↔ 1.66	+1 ↔ 359.32	46.24	52.25
C	Input	Temp	K	298.00	701.00	-1 ↔ 298.00	+1 ↔ 799.00	308.39	43.31
D	Input	Sulfur concentration	ppm	1.77	1513.78	-1 ↔ 1.77	+1 ↔ 1513.78	1053.99	615.59
E	Input	Time	min	9.55	241.00	-1 ↔ 8.30	+1 ↔ 241.00	74.79	50.63
R1	Output	Removal efficiency	%	1.28	98.7939	-	-	44.73	26.02

Table 3. The information and specifics of the data utilized in adsorptive desulfurization.

conditions to ensure effective interaction between the adsorbent and sulfur-containing fuel. During the desulfurization process, functional groups within the adsorbent form chemical bonds with sulfur compounds in the liquid, creating covalent bonds or π - π complexes⁴⁴.

Adsorption desulfurization factors

The adsorbent's specific surface area and pore structure are pivotal in determining its capacity to capture sulfur. Temperature and adsorption time between the adsorbent and liquid fuel are critical parameters affecting adsorption. Proper optimization of these conditions is essential for achieving maximum sulfur removal efficiency. The type and concentration of sulfur compounds in the liquid fuel directly impact the kinetics of adsorption and overall desulfurization performance. There are three kinds of factors including adsorbent characteristics, operating conditions, and fuel properties impact on adsorption capacity of desulfurization adsorbents.

ADS can achieve profound sulfur removal, effectively reducing sulfur content to very low levels (Fig. 1). ADS applies to a wide array of liquid fuels, encompassing petroleum-based fuels, biofuels, and synthetic fuels. ADS processes are generally regarded as environmentally friendly, generating minimal waste and emissions. ADS systems can be tailored to function under diverse conditions, encompassing both continuous and batch processes^{44,45}.

Data collection

In this study, the dataset employed for training, testing, and validation was from earlier empirical research, yielding a substantial compilation of 274 data points across five separate references³²⁻³⁶. The experimental data, detailed in Table 3, pertain to the adsorption desulfurization of liquid fuel and encompass a variety of characteristics. The integration of this comprehensive dataset facilitates robust model training and validation, thereby enhancing the reliability and accuracy of the predictive models.

Response surface methodology

Response surface methodology is a statistical method employed to explore the relationships between input parameters and sulfur removal efficiency using preprocessed data. RSM, rooted in experimental design and analysis, aims to identify the optimal conditions for various processes. This methodology offers numerous advantages, including high efficiency, low cost, robust optimization and data analysis capabilities, and broad

applicability. For researchers and engineers, RSM is an invaluable tool for optimizing process parameters, thereby enhancing production efficiency and product quality while simultaneously reducing costs and boosting competitiveness⁴⁶. RSM is characterized by its visual, nonlinear, and multivariable optimization capabilities, leveraging multiple regression methods to address complex nonlinear and multivariable problems. It also provides visual analysis results that facilitate a clearer understanding of the data⁴⁷. By empirically analyzing the effects of input variables, RSM aids in system optimization. The three fundamental processes in RSM are the systematic design of experiments, regression modeling approach, and optimization technique. Implementing RSM minimizes the total number of experiments required and allows for monitoring the synergistic influence of each independent variable throughout the process⁴⁸. The two-factor interaction (2FI) model is a tool employed in response surface methodology to examine the linear effects of two factors and their interactions. In this model, it is assumed that the factors X_1 and X_2 have a direct linear impact on the response variable Y , while the interaction between these two factors is also taken into account. The regression equation for the 2FI model can be expressed as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{12} X_1 X_2 + \varepsilon \quad (2)$$

In this equation, β_0 represents the intercept, indicating the value of the response Y when both X_1 and X_2 are zero and β_1 is the regression coefficient associated with factor X_1 , reflecting its linear effect on the response Y and β_2 is the regression coefficient for factor X_2 , representing its linear impact on the response Y and β_{12} denotes the interaction coefficient between X_1 and X_2 , showing how the combination of these two factors influences the response and ε accounts for the random error within the model^{49–51}. The design of experiments is a systematic approach that incorporates various input variables into a process to gain insights and establish optimal input, output relationships. This approach provides multiple benefits, including identifying crucial factors impacting the process, reducing costs, developing process models, and developing process models. By examining how independent variables interact and impact the dependent variable, experimental design can develop a semi-empirical model that precisely pinpoints influential factors and optimizes the process, all while reducing costs. Response surface methodology is a statistical technique utilized in experimental design to forecast the relationship between independent and dependent variables through a surface model. This approach includes a variety of mathematical and statistical models aimed at improving, optimizing, and analyzing different processes by investigating the links between input variables and output responses⁵². Table 2 shows the input and output parameters of the process.

Artificial neural network

The first study on artificial neural networks traces back to 1943, when McCulloch and Pitts⁵³ devised a rudimentary neural network computational model. In 1949, Hebb⁵⁴ proposed learning principles for neural networks. Since then, ANN has undergone rapid development and has found extensive application across diverse domains⁵⁵. Artificial neural networks offer an alternative approach to tackling complex formulation problems by adeptly managing intricate data and establishing mathematical links between input and output variables, all without necessitating a comprehensive theoretical understanding of the underlying phenomenon. Moreover, ANN possess predictive capabilities, enabling them to forecast outcomes based on the input data provided⁵⁶. Artificial neural networks are experiencing widespread application across numerous scientific and engineering domains. In chemical engineering, ANN finds utility in various areas such as process dynamics, process modeling, and the optimization of industrial chemical processes⁵⁷. The ANN is a computational model that takes inspiration from the human brain, consisting of interconnected nodes called neurons. Typically, an ANN is structured with an input layer, one or more hidden layers, and an output layer. The input layer corresponds to the predictor variables in regression analysis. Within the hidden layers, complex computations occur via weighted connections. The output layer finalizes the computation, producing the network's outputs⁵⁸. The essence of an artificial neural network lies in its representation of a process through mathematical modeling, forged not through conventional mass and energy balances, but rather through empirical development. Mimicking the intricate network of neurons in the human brain, an ANN comprises interconnected processing elements or nodes, organized into layers. These layers are linked through interconnections, forming a complex web of information processing. Illustrated in Fig. 2, the fundamental structure of a single neuron or node within an ANN model encompasses inputs, an activation function, and a solitary output. The connections between nodes are quantified by calculated

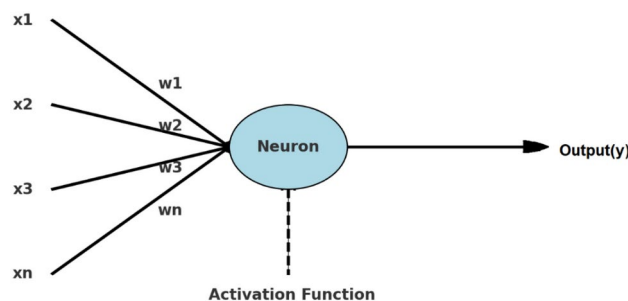


Fig. 2. A schematic depiction of a node within a neural network.

values known as weights, which denote the “strength” of inter-neuronal connections, ultimately influencing the output, denoted by y ⁵⁹.

Within the hidden layer of an ANN, each neuron processes weighted inputs and biases from every neuron in the previous layer, as shown in Eq. (3):

$$Z_i = \left(\sum_{k=1}^{N_{j-1}} X_k^{j-1} W_{k,i} - b_k \right) \quad (3)$$

where X_k^{j-1} signifies the input originating from the k -th node within the j -th layer, $W_{k,j}$ stands for the weight of the connection linking node k with all nodes in the preceding layers, while b_i represents the bias associated with the node. N_{j-1} denotes the total count of nodes within layer $j-1$.

Following the summation, an activation function is applied to compute the output of each node, which is computed as $Y_i = f(Z_i)$. Among the various activation functions, the sigmoid function is the most frequently employed, which is defined as follows⁵⁹:

$$f(Z_i) = \frac{1}{1 + e^{Z_i}} \quad (4)$$

The mean square error function is also utilized to compute the network output error. This can be represented by the following equation:

$$MSE = \frac{1}{N} \sum_{i=1}^N (Y_i - T_i)^2 \quad (5)$$

where Y_i and T_i represent the network output and the target value, respectively. N represents the quantity of experimental data points used in the investigation. The MSE is instrumental in determining the adjustments to the network's weights and biases. This process continues, with the MSE being recalculated iteratively, until it falls below a predefined threshold or another stopping condition is met⁶⁰.

In this study, the absolute value of the relative error, AARE, is used to compare neural networks with the RSM model and the performance of the resultant network is demonstrated through R^2 and MSE. Furthermore, the efficacy of the generated network was assessed via the computation of R^2 , MSE, and AARD metrics.

$$R^2 = \frac{\sum_{i=1}^n (Y_{predicted} - Y_{actual})^2}{\sum_{i=1}^n (Y_{predicted} - Y_{mean})^2} \quad (6)$$

$$AARD (\%) = \frac{100}{n} \sum_{i=1}^n \left| \frac{Y_{predicted} - Y_{actual}}{Y_{predicted}} \right| \quad (7)$$

In this context, $Y_{predicted}$ denotes the anticipated value of Y derived from both ANN and RSM, whereas Y_{actual} signifies the empirical value of Y . Y_{mean} corresponds to the average value of Y ⁶¹.

Multilayer perceptron neural networks

The perceptron algorithm, pioneered by Rosenblatt in the late 1950s, is widely acknowledged as a prevalent and extensively utilized model in supervised machine learning⁶². The MLP is a widely recognized type of ANN and is considered a reliable technique for addressing modeling tasks across various fields⁶³. The term “ANN model” refers to the specific arrangement and interconnections of neurons within the network. Over time, numerous ANN models have been proposed, each tailored to different applications and advancements in the field. The MLP is a specific case of the feedforward neural network (FNN), characterized by fully connected neurons between each preceding and succeeding layer. When the number of hidden layers increases, the MLP is also referred to as a deep neural network (DNN). In the training phase, the connection weights are updated in a backward manner based on the error magnitude, leading to the alternative name “feedforward backpropagation networks” (FFBN). These terms are often used interchangeably in the literature⁶⁴. A typical MLP is composed of three layers: the input layer, one or more hidden layers, and the output layer. The hidden layer can comprise one or more layers, each containing specific processing units known as neurons. The size of the input and output datasets dictates the number of neurons present in these layers. Various methods are employed to evaluate the optimal number of neurons and hidden layers, including the trial-and-error approach and more sophisticated intelligent methods⁶⁵. The MLP's capability to approximate any nonlinear function is attributed to its ability to tune weights and activation functions effectively⁶⁶. One of the most commonly used methods for training MLPs is the backpropagation algorithm, which is based on the error-correction learning rule. This algorithm adjusts the model parameters in a manner that minimizes the error function, thereby reducing the overall model error. The backpropagation training process can be summarized in two main steps:

- Forward pass: The input vector is fed into the multilayer network, propagating through the hidden layers until it reaches the output layer. The output layer generates a response vector, which represents the MLP's final output.

- Backward pass: During this phase, the MLP's parameters are updated and refined. This involves adjusting the weights of the neurons in the hidden layers according to the error-correction rule. The adjustments aim to minimize the discrepancy between the predicted outputs of the neural network and the actual target values.

Through these iterative processes, the MLP continuously improves its performance, making it a powerful tool for modeling complex, nonlinear relationships in data⁶⁷. The performance of the MLP model can fluctuate due to the initial weights assigned to the input parameters, which is considered a significant drawback. To overcome this issue, the model is executed multiple times, and the most accurate iteration is chosen as the final version⁶⁸. Equation (7) illustrates the output of any perceptron⁶⁹.

$$\gamma_i(x^{(j)}) = \varphi\left(\sum_{k=1}^N w_{ik}x_k^{(j)} + b_i\right) \quad (8)$$

Here, $\varphi(x)$ denotes non-linear activation function, γ_i represents the output of the i th neuron, $x^{(j)}$ denotes the input for the j th layer, $x_k^{(j)}$ signifies the value of the k th neuron within the j th layer, w_{ik} stands for the weight between the i th and k th neuron, and b_i corresponds to the bias value linked to the i th neuron⁷⁰.

Figure 3 depicts the structure of an MLP featuring three hidden layers, with each neuron in one layer is fully connected to every neuron in the preceding layer. The main goal during training is to reduce the mean squared error. This is achieved by creating an error signal from the activation function, which is then propagated backward through the layers to update the weights. Adjustments that effectively reduce the overall MSE are prioritized.

Radial basis function (RBF)

In the late 1980s, Moody and Darken introduced the RBF neural network, renowned for its self-learning abilities, efficient storage, rapid search speed, and quick computation time⁷¹. Radial basis function neural networks (RBF-NNs) utilize radial basis functions as activation functions, distinguishing them as a specific type of artificial neural network. These networks are particularly known for their ability to approximate functions and are widely used in pattern recognition, classification, and interpolation problems. The architecture of an RBF-NN typically consists of three layers: the input layer, a hidden layer with a non-linear RBF activation function, and a linear output layer. Radial basis function neural networks are a subset of artificial neural networks characterized by their three-layer structure: the input layer, the hidden layer, and the output layer. Notably, RBF-NNs have a single hidden layer that plays a critical role in their operation. Within this hidden layer, there are units known as neurons or hidden units, each of which utilizes a radial basis function as its activation function. The performance of RBF-NNs is significantly influenced by certain parameters known as weight terms. These weights are pivotal to the network's ability to learn and generalize from the data. The values of these weight terms are optimized using a method called the gradient descent approach. This optimization technique iteratively adjusts the weights to minimize the error between the network's predictions and the actual target values, thereby enhancing the network's accuracy and performance⁶⁵. The radial basis function neural network is a type of feedforward neural network renowned for its outstanding performance. This network excels at approximating non-linear

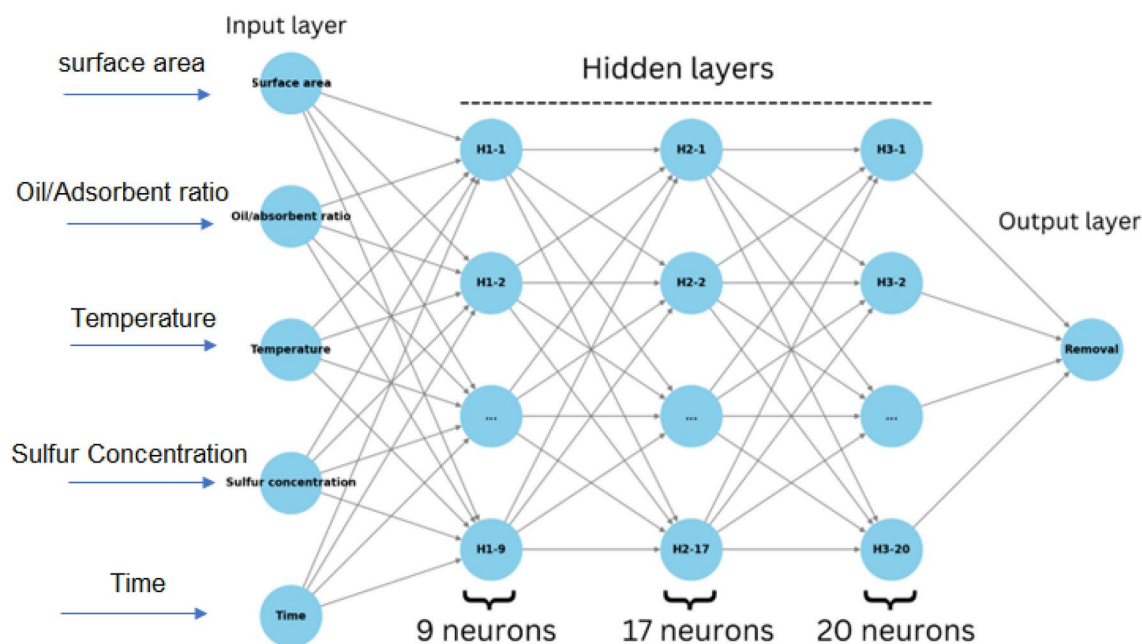


Fig. 3. Schematic diagram of MLP network with three hidden layers.

functions with a high degree of accuracy and achieves global approximation without encountering local minima. Additionally, due to its compact topological structure, the RBF network boasts a rapid learning speed, making it highly efficient and effective for various applications⁷². While numerous activation functions exist for radial-based neurons, the Gaussian function stands out as the most widely utilized⁷³. It is mathematically represented as Eq. (8)⁷⁴:

$$\varphi_i (\|x - c_i\| * b) = \exp \left(-\frac{1}{2\sigma_i^2} (\|x - c_i\| * b)^2 \right) \quad (9)$$

Here, x signifies the input, φ_i represents the output, while c_i and σ denote the center and spread parameters of the Gaussian function, respectively. Moreover, the variable b stands for the bias term. To derive the network's output (y), a linear combination of the activation function and the weight vector w of the output layer is computed.

$$y = \sum_{i=0}^n \varphi_i w_i \quad (10)$$

In this context, w_i represents the combined weighted value attributed to the i th basis function. Figure 4 showcases the configuration of the RBF network employed in this investigation.

ANN model design

As outlined in the Fig. 5 algorithm, the initial step involved consolidating experimental data. Input variables such as surface area, fuel/adsorbent ratio, temperature, sulfur concentration, and time were integrated, with removal percentage serving as the output. Subsequent to data integration, the second phase entailed normalizing both input and output data. The artificial neural network model was trained on 70% of the complete dataset, adjusting parameters such as weights, biases, and thresholds to improve the model's effectiveness. Validation of the network was conducted using a dedicated 15% subset, while testing utilized another 15%. Evaluating the trained model's accuracy involved employing statistical measures like R^2 value and MSE, contrasting predictions against observed data. Determining the best configuration for the multilayer perceptron model involved experimenting with different parameters, including hidden layer structures and training algorithms, to achieve optimal predictive performance. In contrast, the process of determining the ideal neuron count for the RBF network often involves iterative experimentation, initially employing a surplus of neurons and gradually reducing their number until reaching minimal MSE levels. Training algorithms cease operation upon achieving optimized error levels during input testing.

Results and discussion

Both MLP and RBF models demonstrated satisfactory performance in predicting sulfur removal efficiency. Performance evaluation metrics, such as mean squared error and coefficient of determination (R^2), were employed to assess model accuracy. The RBF network exhibited superior performance compared to the MLP network, achieving a higher R^2 value and lower MSE, indicating a more accurate prediction of sulfur removal efficiency. Sensitivity analysis was conducted to investigate the influence of each input parameter on sulfur

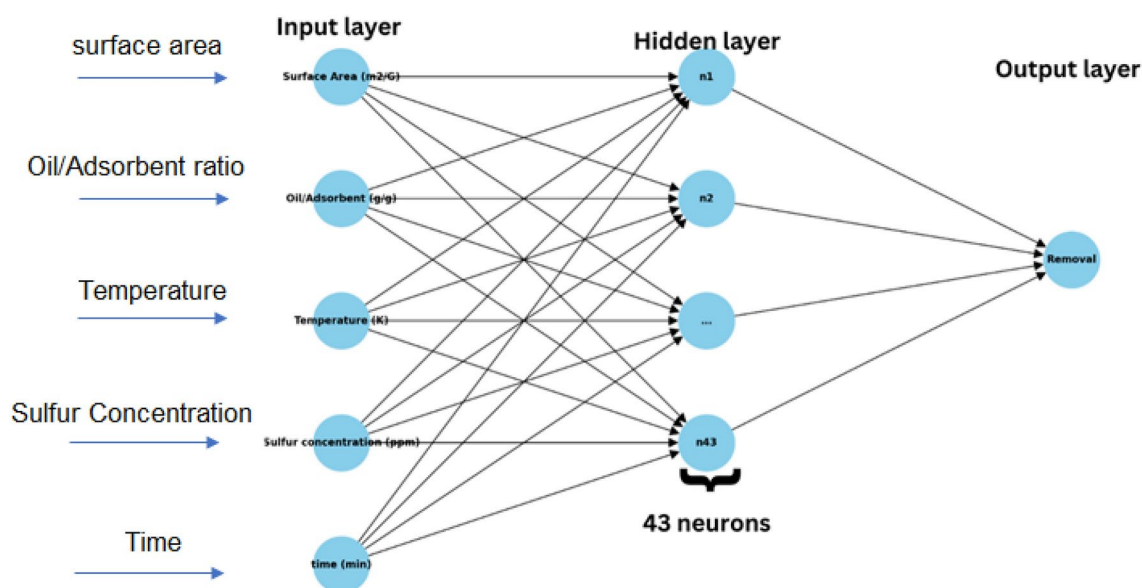


Fig. 4. Architecture of a typical RBF network model.

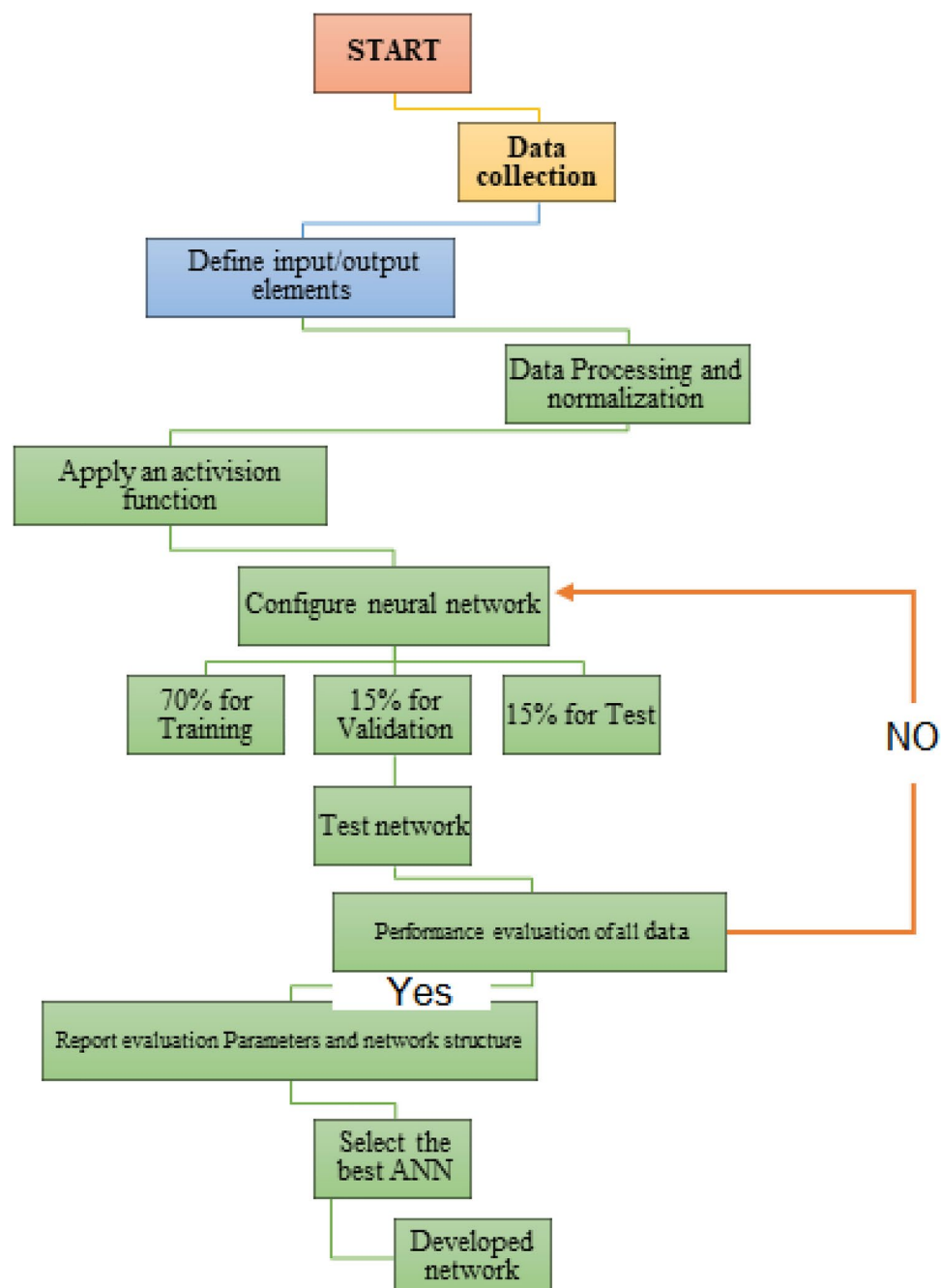


Fig. 5. Flow chart of ANN model.

removal. The analysis revealed that temperature and fuel-to-adsorbent ratio were the most significant factors affecting sulfur removal efficiency.

RSM results

Analysis of variance (ANOVA)

The results of the ANOVA analysis are presented in Table 4, showing the significance of the model with the F-value and the likelihood associated with the ANOVA test with the P value⁷⁵. Factors in the model are deemed significant if their P values are below 0.05, while P values exceeding 0.1 suggest that the factors are not significant²⁹. The analysis revealed that the model exhibited a significant F-value of 194.32. Among the independent variables, only parameter A had a value slightly above 0.05, indicating that nearly all parameters are significant. Term C demonstrated the highest F-value and the lowest $p > F$ value, signifying that temperature has the greatest influence on the removal percentage, followed by concentration and the fuel-to-adsorbent ratio. The model also showed a strong signal with a ratio of 52.5912. Table 5 presents the static parameters of the fit for the responses based on 274 observations.

Source	Sum of squares	df	Mean Square	F-value	p value
Model	1.698E+05	15	11,320.01	194.32	<0.0001
A-BET	196.98	1	196.98	3.38	0.0671
B-fuel/adsorbent	21,332.80	1	21,332.80	366.19	<0.0001
C-Temp	23,998.75	1	23,998.75	411.96	<0.0001
D-Sulfur concentration	22,932.82	1	22,932.82	393.66	<0.0001
E-Time	446.32	1	446.32	7.66	0.0061
AB	502.01	1	502.01	8.62	0.0036
AC	205.60	1	205.60	3.53	0.0614
AD	237.74	1	237.74	4.08	0.0444
AE	4.05	1	4.05	0.0695	0.7922
BC	20,737.88	1	20,737.88	355.98	<0.0001
BD	4903.18	1	4903.18	84.17	<0.0001
BE	3940.31	1	3940.31	67.64	<0.0001
CD	24,324.23	1	24,324.23	417.54	<0.0001
CE	7.49	1	7.49	0.1286	0.7201
DE	8379.37	1	8379.37	143.84	<0.0001
Residual	15,029.92	258	58.26		
Lack of fit	13,669.05	247	55.34	0.4473	0.9867
Pure error	1360.86	11	123.71		
Cor total	1.848E+05	273			

Table 4. ANOVA results.

Fit statics	Removal
R ²	0.9187
Adjusted R ²	0.9140
Predicted R ²	0.9013
Adeq precision	52.5912

Table 5. The statistical parameters were utilized to model the responses in the 2FI model.

Correlations

The 2FI equation presented in Eq. (10) was used to model the experimental results, demonstrating the influence and interaction between the variables.

$$\begin{aligned}
 \text{Removal} = & -1883.55111 + 0.221732 \times A + 10.81465 \times B + 6.24666 \times C + 1.52885 \times D \\
 & + -0.346199 \times E + -0.000141 \times AB + -0.000621 \times AC + -8.28035e - 06 \times AD \\
 & + -5.88762e - 06 \times AE + -0.035731 \times BC + -0.000366 \times BD + 0.002749 \times BE \\
 & + -0.004976 \times CD + -0.000143 \times CE + 0.000293 \times DE
 \end{aligned} \quad (11)$$

Analysis of parameter interdependencies using correlation matrix

The Fig. 6 illustrates a correlation matrix that displays the impact of various parameters on each other. Positive values indicate a positive correlation between two parameters, while negative values indicate a negative correlation. The intensity of the colors represents the strength of the correlation; darker colors signify stronger correlations.

Perturbation plots

The perturbation plot facilitates a comprehensive comparison of the impacts of all five process parameters on the removal percentage at the center point. By employing a single plot, the influence of each parameter is distinctly observable. Figure 7 shows the perturbation plot for Removal percentage. It has been well-established that the removal percentage increases significantly as the sulfur concentration rises, while it significantly decreases with higher fuel/adsorbent ratios and elevated temperatures. In contrast, surface area and contact time have relatively moderate effects.

Effect of adsorptive desulfurization parameters on removal percentage

In this study, 3D response surface plots were generated using Design-Expert version 13.0 to illustrate the interactions between important independent variables and their impact on desulfurization efficiency. Figure 8a–e provide a detailed examination of these interactions. Figure 8a shows the removal percentage as a function of the fuel/adsorbent ratio and desulfurization time. The plot indicates that the removal percentage increases

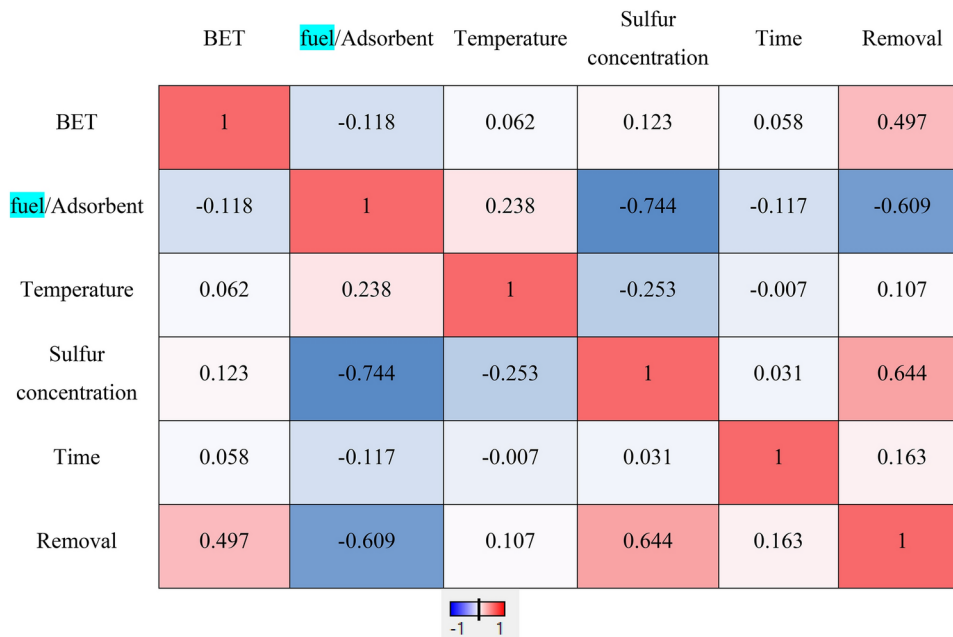


Fig. 6. Correlation matrix illustrating the interdependencies among various parameters.

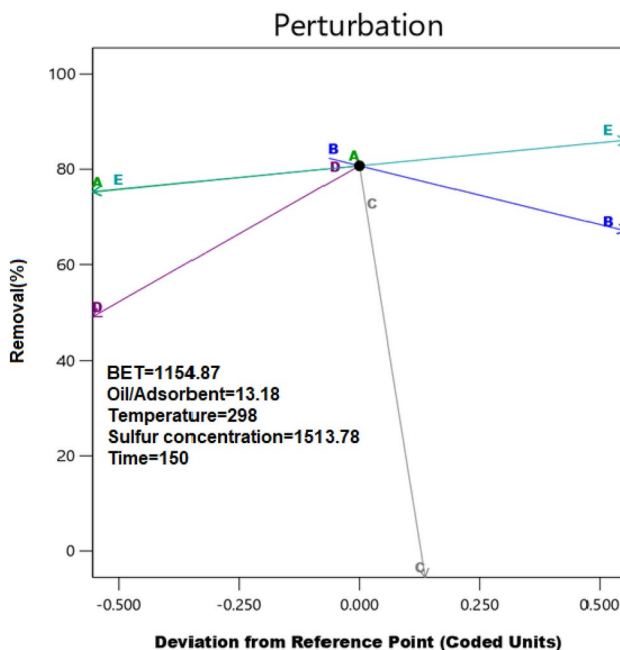


Fig. 7. Perturbation plots for removal percentage.

with an increase in desulfurization time and a decrease in the fuel/adsorbent ratio. This suggests that longer processing times and higher adsorbent concentrations are beneficial for enhancing the desulfurization efficiency. Figure 8b illustrates the removal percentage as a function of sulfur concentration and desulfurization time. The removal percentage is observed to increase with both increasing sulfur concentration and desulfurization time. This indicates that higher initial sulfur concentrations and extended treatment durations can lead to more effective sulfur removal. Figure 8c depicts the removal percentage as a function of sulfur concentration and BET surface area during desulfurization. The plot shows that the removal percentage increases with an increase in both BET surface area and sulfur concentration. This suggests that a larger surface area available for adsorption, along with higher sulfur concentrations, improves the efficiency of sulfur removal. Figure 8d shows the removal percentage as a function of the fuel/adsorbent ratio and sulfur concentration. The plot indicates that the removal percentage increases with increasing sulfur concentration and decreasing fuel/adsorbent ratio. This implies

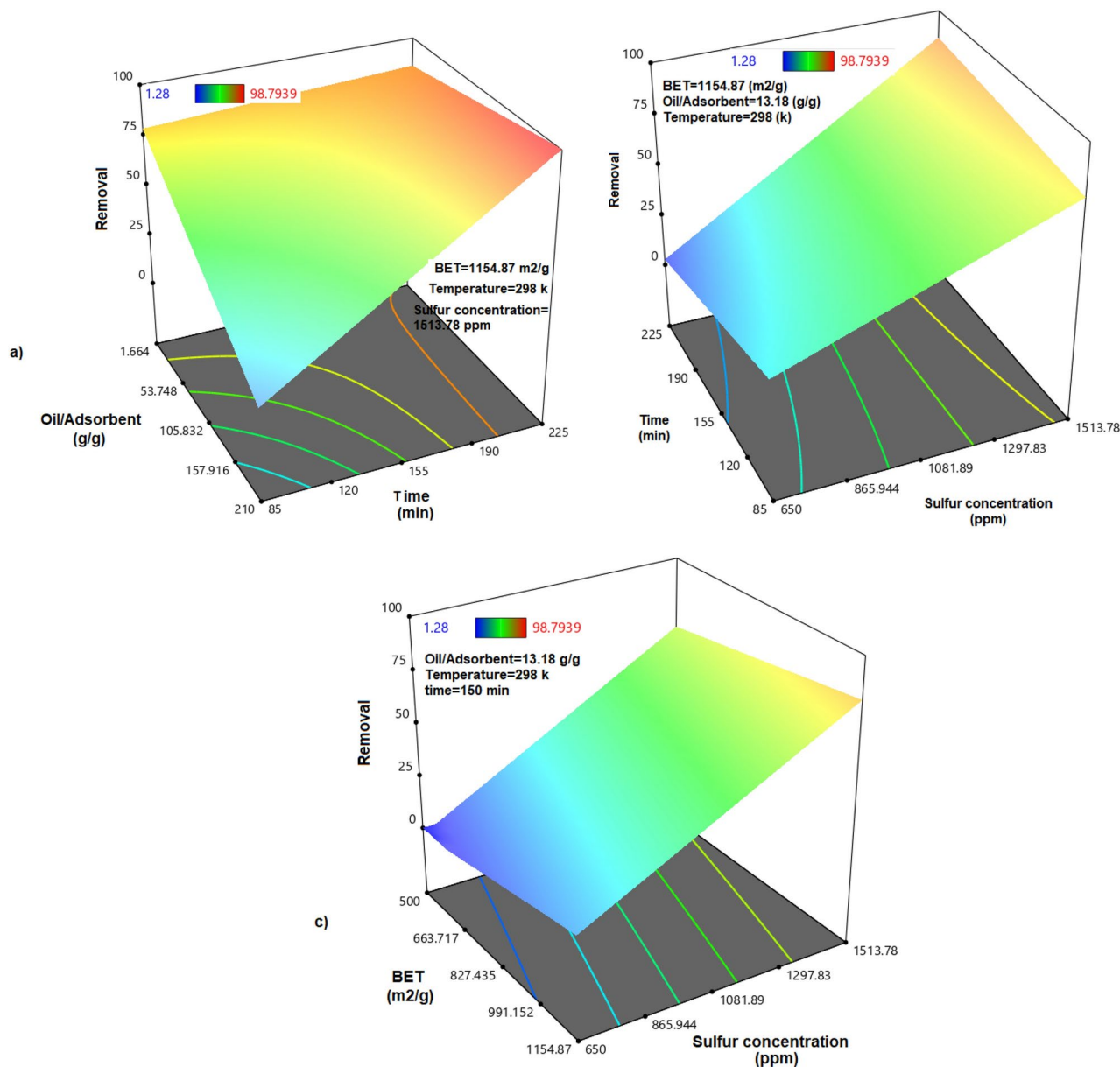


Fig. 8. (a) Fuel/adsorbent ratios and time on removal percentage; (b) time and sulfur concentration on removal percentage; (c) surface area and sulfur concentration on removal percentage; (d) sulfur concentration and fuel/adsorbent ratios on removal percentage; and (e) fuel/adsorbent ratios and surface area on removal percentage.

that higher sulfur concentrations and greater amounts of adsorbent relative to oil enhance the desulfurization process. Figure 8e illustrates the removal percentage as a function of the fuel/adsorbent ratio and BET surface area. The plot demonstrates that the removal percentage increases with an increase in BET surface area and a decrease in the fuel/adsorbent ratio. This suggests that maximizing the adsorbent's surface area and optimizing the fuel/adsorbent ratio are crucial for achieving higher desulfurization efficiency.

These observations provide valuable insights into the optimization of desulfurization processes, highlighting the importance of key parameters such as desulfurization time, sulfur concentration, fuel/adsorbent ratio, and surface area (BET) in enhancing removal efficiency.

Results of ANN

Optimization and prediction

In order to achieve the best results on the test data, the architecture of the multilayer perceptron network was carefully crafted. Various factors were meticulously taken into account, including the quantity of layers and neurons within each layer, the activation function utilized for each layer, the number of epochs, and the training function. In the same way, the RBF network underwent optimization through fine-tuning its characteristics,

Model	R ²	MSE
MLP	0.98	0.0028
RBF	0.98	0.0026

Table 6. Analytical standards for evaluating various models.

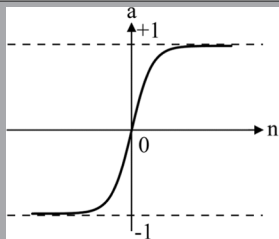
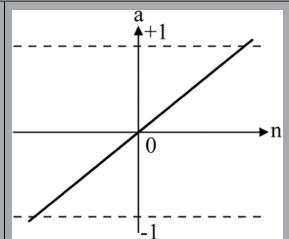
Graph		
Mathematical algorithm function	$a = \text{tansig}(n) = \frac{2}{1+e^{-2n}} - 1$	$a = \text{purelin}(n) = n$

Table 7. The Tansig and Purelin activation functions were used, with n represents the input signal and a_n represents the output.

Backpropagation algorithm	Function	Number of neurons	R ²	MSE train	MSE test	Epoch
Levenberg–Marquardt	Trainlm	46	0.980	0.0028	0.0336	32
Bayesian regularization	Trainbr	23	0.975	0.0034	0.0415	48
Scaled conjugate gradient	Trainscg	46	0.910	0.0152	0.0983	23

Table 8. The outcome of training algorithms for Backpropagation in Multilayer Perceptron.

such as adjusting the number of neurons, epochs, and the training function. In order to avoid overtraining and make reliable predictions, it is essential to reduce the network's runtime and assess the model's accuracy using data that was not used in the training phase. Overtraining occurs when the network fits the training data effectively but struggles to perform well on new data, leading to low accuracy and lack of generalizability. Hence, it is crucial to validate the network's performance by comparing predicted values with test data. As shown in the preceding equations, the models' performance was assessed using analytical criteria, specifically the MAE, MSE, RMSE, and R². The results are shown in Table 6.

Three different training algorithms were used to determine the best algorithm for the ANN model: Levenberg–Marquardt (trainlm)⁴⁶, Bayesian Regularization (trainbr)⁴⁷, and Scaled Conjugate Gradient (trainscg)⁴⁸. Various activation functions were also explored in this study. The sigmoid function (tansig) was selected for the neurons in the hidden layers, and the linear function (purelin) was chosen for the output layer, as detailed in Table 7.

Various performance metrics, such as MSE, R squared R², and the number of epochs, were assessed to identify the optimal neural network configuration as shown in Table 8. The MLP model was tested with neuron counts ranging from 1 to 46. The data clearly shows that the trainlm algorithm performs better than all other configurations in terms of all performance metrics. Therefore, the trainlm algorithm was chosen as the preferred approach for training the network.

The study observed that the mean squared error for the multilayer perceptron network increased with the addition of more hidden layers and neurons. However, beyond two hidden layers, the increase in MSE was not significant. Consequently, the number of hidden layers was capped at three to avoid unnecessary computational slowdowns. For a comparative analysis with the MLP network, the parameters of the Radial Basis Function network needed optimization. The optimization of the spread parameter was carried out using network data, as depicted in Fig. 9. The Radial Basis Function network, producing two outcomes, employed the average MSE to identify the optimal spread. Furthermore, Fig. 10 illustrates that increasing the number of neurons led to a further decrease in MSE and improved the results. The Gaussian function-based RBF model achieved optimal performance with 43 neurons. Figure 11a and b demonstrate that the multilayer perceptron and radial basis function networks achieved the most favorable MSE validation outcomes following 32 and 40 epochs, respectively. The optimal MSE value for the MLP network's removal percentage was 0.0028, while for the RBF network, it was 0.0026. Figure 12 demonstrates that the regression values R² for training, testing, and the overall selected structure were nearly identical for both MLP and RBF networks, with R²=0.980 for MLP and R²=0.981 for RBF. The optimal MLP network configuration consisted of three hidden layers with 9, 17, and 20 neurons, respectively. This configuration was found to be the most effective in achieving low MSE and high regression accuracy, underscoring the importance of careful network architecture design in optimizing desulfurization processes.

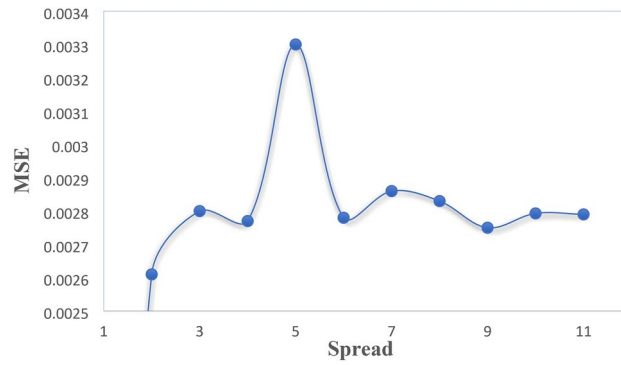


Fig. 9. The relationship between the spread and MSE in RBF neural network.

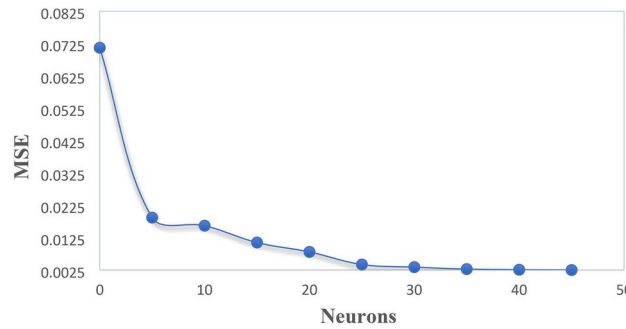


Fig. 10. The relationship between the number of neurons and mean squared error in RBF neural network.

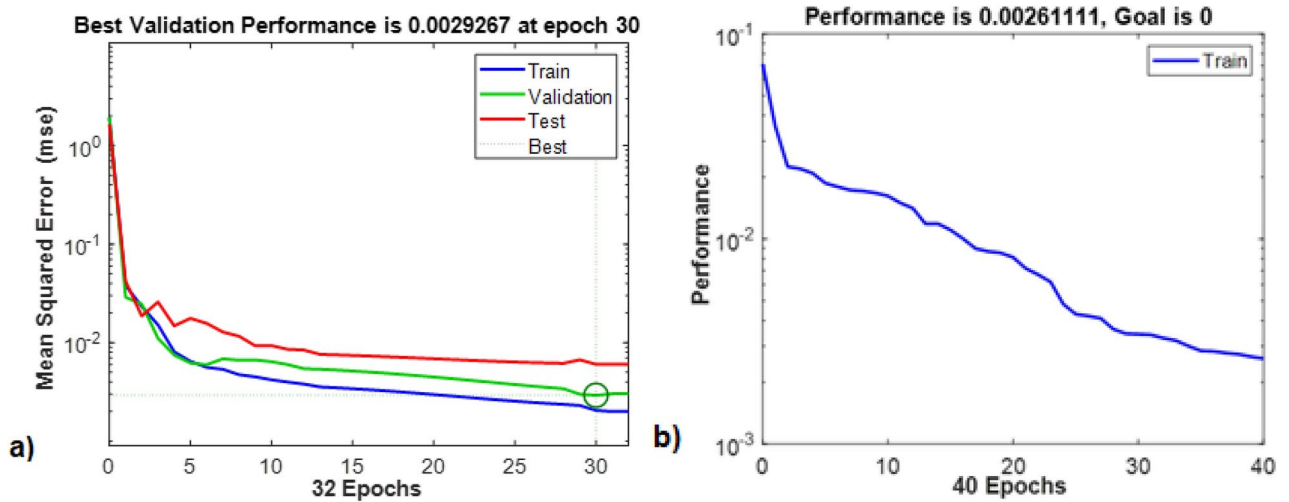


Fig. 11. The MSE for the MLP and RBF models is validated to assess the network performance for removal percentage.

3D response surfaces

The response surface of the ANN model is depicted in the three-dimensional plots in Fig. 13a–c. In these plots, three factors are varied while keeping the other two constant to emphasize the interaction effects between the factors. The results from the response surface methodology indicate that similar behaviors, trends, and magnitudes are observed for the parameters influencing the response.

Comparison of results between the RSM and ANN

In order to assess the efficiency of the RSM model and the trained neural networks, a random set of experimental data was selected and inputted into both the RSM model and the MLP and RBF networks for comparison. The

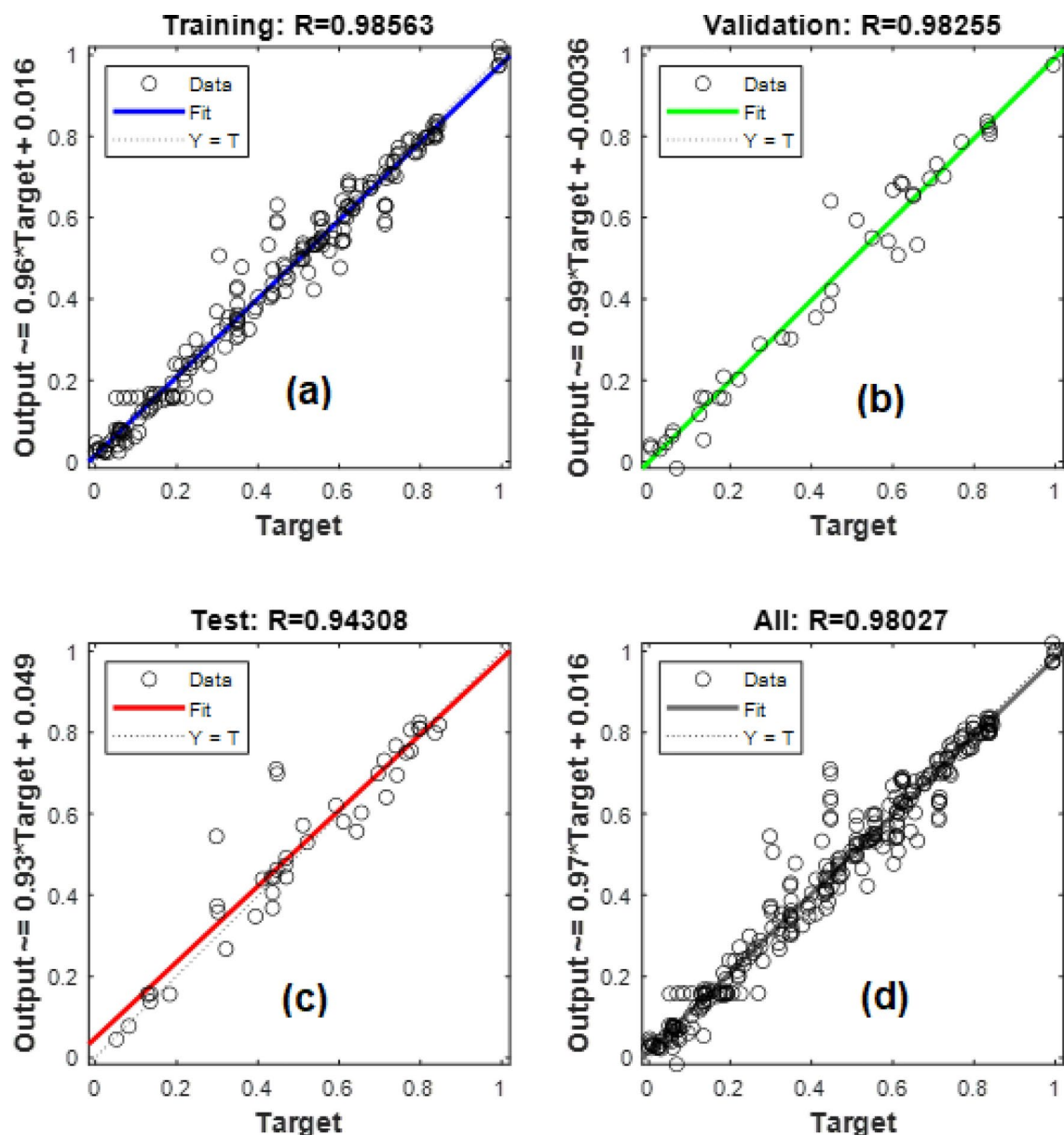


Fig. 12. MLP network regression status of (a) training data, (b) validation data, (c) testing data, and (d) all data.

accuracy of these models was evaluated by comparing their predicted response values with the experimental data and calculating the average absolute relative deviation (%AARD). The results of this comparison are shown in Table 9. The neural networks demonstrated lower %AARD values compared to the RSM model, suggesting better predictive accuracy. In addition, the RBF network was more accurate than the MLP network in predicting the output parameter.

Conclusion

The growing need for process analysis and optimization, coupled with the enhanced availability of statistical software and computing power, has significantly contributed to the widespread adoption of response surface methodology and artificial neural network modeling tools. This study centers on adsorption desulfurization. Key parameters such as temperature, concentration, surface area, fuel-to-adsorbent ratio, and time were utilized to predict removal efficiency through both response surface methodology and artificial neural network modeling. The response surface methodology approach enabled a detailed examination of the interactions among these factors. Predictions from both modeling techniques were closely aligned with the experimental data. Moreover, response surface methodology was utilized to create two correlations for the response with the 2FI model. ANOVA results indicated that temperature and concentration were the most influential factors affecting removal efficiency. In the context of ANN modeling, both MLP and RBF algorithms were assessed,

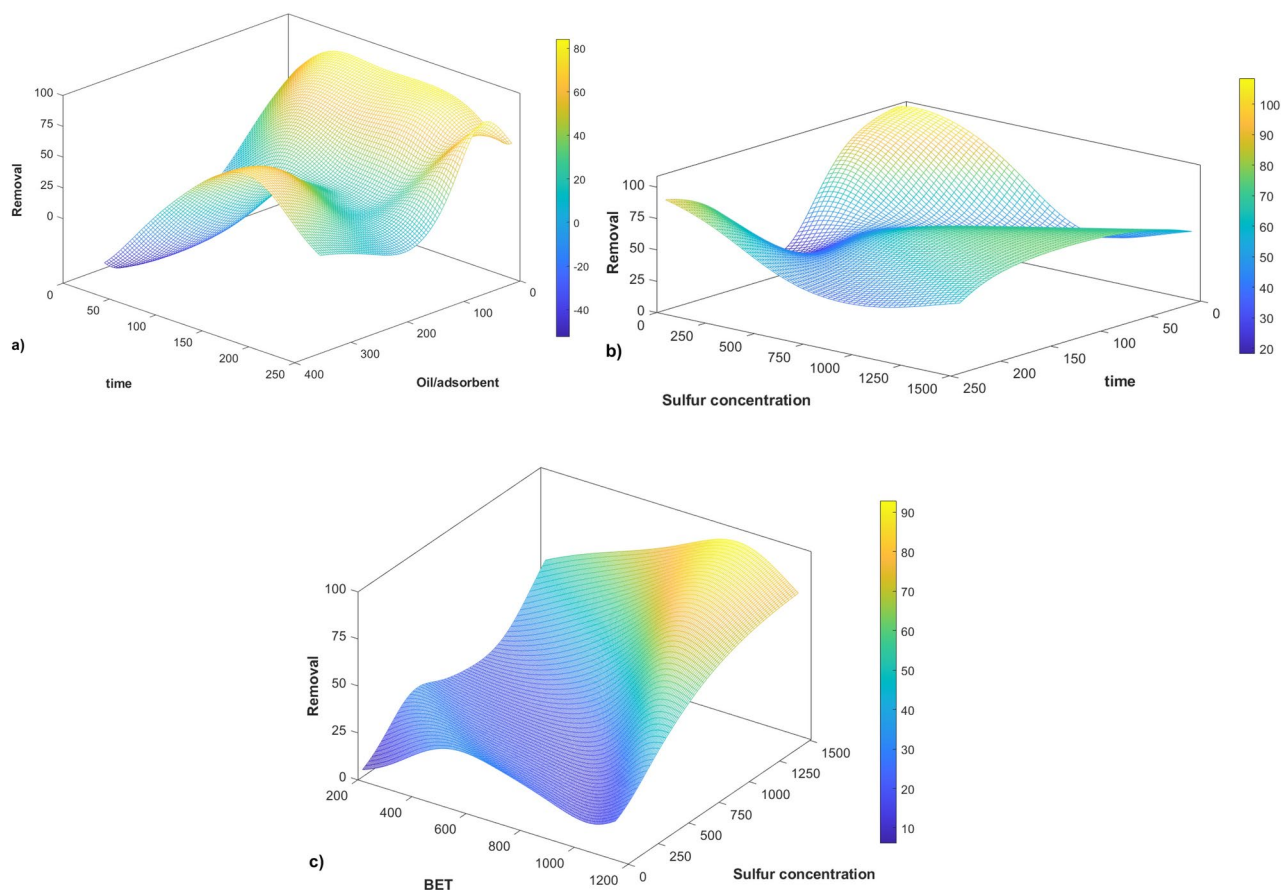


Fig. 13. 3D response surface plots using an artificial neural network with a multilayer perceptron model for (a) fuel/adsorbent ratios and time on removal percentage, (b) time and sulfur concentration on removal percentage, (c) surface area and sulfur concentration on removal percentage.

Run	Surface area (m ² /g)	Fuel/adsorbent (g/g)	T (K)	Sulfur concentration (ppm)	Time (min)	RSM		MLP		RBF	
						Y _{RSM}	%AARD	Y _{MLP}	%AARD	Y _{RBF}	%AARD
1	225.05	13.18	298	1513.78	39.8	51.06	9.65	54.5746	13.7554	55.6221	14.9791
2	225.05	13.18	298	1513.78	90.2	55.58	10.74	59.5313	7.9581	62.0272	6.1955
3	245.49	13.18	298	1513.78	49.7	52.40	34.52	49.8430	30.3920	51.8799	33.6773
4	245.49	13.18	308	1513.78	50	33.26	3.01	33.7536	2.3254	32.5108	4.0808
5	1154.87	13.18	308	1513.78	89.8	50.90	3.88	49.3266	5.3201	53.1421	1.8640
6	1154.87	21.9667	298	1513.78	90	73.13	2.51	78.9710	1.2799	78.7533	1.1385
7	882.24	32.6238	303	250	60	20.92	10.37	26.9702	1.0799	26.5051	0.1990
8	882.24	109.833	303	50.5	60	17.82	11.47	18.1012	12.4181	16.5827	7.1818
9	930.1	4.992	313	580	180	34.17	1.59	35.2982	0.0025	36.0969	1.1288
10	519.4	40.925	303	800	90	20.50	4.43	24.4471	4.3269	18.0130	9.9711

Table 9. Comparison of results from several case studies using ANN and RSM.

with minimal differences observed between them. The trainlm algorithm stood out as the most effective among the MLP algorithms, utilizing the tansig function in with three hidden layers (20, 17, and 9 neurons) and the purelin function in the output layer. The MLP network demonstrated optimal performance with an MSE of 0.0028 for removal percentage over 30 epochs, whereas the Radial Basis Function network achieved a superior MSE of 0.0026 for removal percentage over 40 epochs. The ANN and RSM models exhibited average R² values of 0.981 and 0.919, respectively. ANN models are very useful for optimizing complicated processes like adsorption desulfurization because they provide significant advantages over experimental methods in terms of prediction speed, cost efficiency, and adaptability to a wide range of situations. All things considered, ANN outperformed RSM in terms of efficiency, offering a dependable and potent instrument for next developments in sulfur removal optimization. This study not only shows how effective ANN is in adsorption desulfurization but also shows

how it may improve predictive accuracy, cut down on the amount of testing required, and expedite process enhancements in environmental applications.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Received: 6 September 2024; Accepted: 28 January 2025

Published online: 05 February 2025

References

- Liu, J. et al. Constructing light-insensitive Ag (I) sites by in situ integration of nanoheaters for efficient adsorptive desulfurization. *Chem. Eng. J.* **484**, 149717 (2024).
- Khalse, N. M. & De, M. Adsorptive desulfurization of thiophenic sulfur compounds using nitrogen modified graphene. *Sep. Purif. Technol.* **331**, 125693 (2024).
- Hamad, K. I. et al. Efficient immobilization of acids into activated carbon for high durability and continuous desulfurization of diesel fuel. *Energy Sci. Eng.* **11**(10), 3662–3679 (2023).
- Shafiq, I. et al. Recent developments in alumina supported hydrodesulfurization catalysts for the production of sulfur-free refinery products: A technical review. *Catal. Rev.* **64**(1), 1–86 (2022).
- Wang, E. et al. Recent advances in the unsupported catalysts for the hydrodesulfurization of fuel. *Fuel Process. Technol.* **235**, 107386 (2022).
- Saha, B., Vedachalam, S. & Dalai, A. K. Review on recent advances in adsorptive desulfurization. *Fuel Process. Technol.* **214**, 106685 (2021).
- Yaseen, M. et al. Fabrication of Zn and Mn loaded activated carbon derived from corn cobs for the adsorptive desulfurization of model and real fuel oils. *Fuel* **284**, 119102 (2021).
- Mamuad, R. Y. & Choi, A. E. S. Biodesulfurization processes for the removal of sulfur from diesel oil: A perspective report. *Energies* **16**(6), 2738 (2023).
- Sokol, H. J., Lee, K. X. & Valla, J. A. The mechanism of adsorptive desulfurization on Cu and Ce exchanged Y zeolite using density functional theory. *Comput. Mater. Sci.* **235**, 112813 (2024).
- Qasim, A. & Alwan, H. H. Adsorptive desulfurization of Iraqi light naphtha using calcite and modified calcite. *Iraqi J. Chem. Pet. Eng.* **25**(1), 83–93 (2024).
- Wen, Z. et al. La doped silica aerogel as selective adsorbent for the desulfurization of model fuel. *Colloids Surf. A Physicochem. Eng. Asp.* **687**, 133493 (2024).
- Solouki, A. & Chaouki, J. Extractive desulfurization of crude petroleum oil and liquid fuels using trihexyl tetradecyl phosphonium bis (2-ethylhexyl) phosphate ionic liquid. *Chem. Prod. Process Model.* **19**, 263–273 (2024).
- Aslam, S. et al. Facile fabrication of confined spaces with molybdenum atoms for fast oxidative desulfurization of fuel oil. *Sep. Purif. Technol.* **344**, 127301 (2024).
- Ali, I. S. et al. Adsorptive desulfurization of model and real fuel via wire-, rod-, and flower-like Fe₃O₄@ MnO₂@ activated carbon made from palm kernel shells as newly designed magnetic nanoadsorbents. *Fuel* **340**, 127523 (2023).
- Chen, K. et al. Evaluation of adsorptive desulfurization performance and economic applicability comparison of activated carbons prepared from various carbon sources. *RSC Adv.* **10**(66), 40329–40340 (2020).
- Khalid, R. Y. & Fadhil, A. B. Adsorptive removal of dibenzothiophene from model fuel over activated carbon developed by KOH activation of pinecone: Equilibrium and kinetic studies. *Asia Pac. J. Chem. Eng.* **19**(1), e2976 (2024).
- Azeez, M. O. & Ganiyu, S. A. Review of biomass derived-activated carbon for production of clean fuels by adsorptive desulfurization: Insights into processes, modifications, properties, and performances. *Arab. J. Chem.* **16**, 105182 (2023).
- Huang, E. et al. Cu (I)-faujasite-type zeolites with hierarchical structure for adsorption desulfurization. *Microporous Mesoporous Mater.* **348**, 112397 (2023).
- Zhou, T. et al. Preparation of Ti-MOFs for efficient adsorptive desulfurization: Synthesis, characterization, and adsorption mechanisms. *Fuel* **339**, 127396 (2023).
- Ghassa, M. et al. Developing new MOF-based hybrid materials for efficient adsorptive desulfurization of a refractory sulfur compound. *Energy Fuels* **37**(18), 13769–13784 (2023).
- Sun, Z. et al. Regeneration performance of activated carbon for desulfurization. *Appl. Sci.* **10**(17), 6107 (2020).
- Azeez, M. O. et al. Synergistic effect of nitrogen and molybdenum on activated carbon matrix for selective adsorptive desulfurization: Insights into surface chemistry modification. *Arab. J. Chem.* **15**(1), 103454 (2022).
- Spessato, L. et al. KOH-super activated carbon from biomass waste: Insights into the paracetamol adsorption mechanism and thermal regeneration cycles. *J. Hazard. Mater.* **371**, 499–505 (2019).
- 贾冯馨 et al. Preparation and SO₂ adsorption behavior of coconut shell-based activated carbon via microwave-assisted oxidant activation. *中国炼油与石油化工* **20**(1), 67 (2018).
- Azizian, S. & Eris, S. Adsorption isotherms and kinetics. In *Interface Science and Technology* 445–509 (Elsevier, 2021).
- Babatunde, K. A. et al. Adsorption of gases on heterogeneous shale surfaces: A review. *J. Pet. Sci. Eng.* **208**, 109466 (2022).
- Al-Jamimi, H. A., BinMakhashen, G. M. & Saleh, T. A. Multiobjectives optimization in petroleum refinery catalytic desulfurization using machine learning approach. *Fuel* **322**, 124088 (2022).
- Mashhadimoslem, H. et al. Development of predictive models for activated carbon synthesis from different biomass for CO₂ adsorption using artificial neural networks. *Ind. Eng. Chem. Res.* **60**(38), 13950–13966 (2021).
- Moradi, M. R., Ramezanipour Penchah, H. & Ghaemi, A. CO₂ capture by benzene-based hypercrosslinked polymer adsorbent: Artificial neural network and response surface methodology. *Can. J. Chem. Eng.* **101**(10), 5621–5642 (2023).
- Khoshrافتar, Z. & Ghaemi, A. Preparation of activated carbon from *Entada africana* Guill. & Perr for CO₂ capture: Artificial neural network and isotherm modeling. *J. Chem. Pet. Eng.* **56**(1), 165–180 (2022).
- Darwish, N. N. *Adsorption Study of Desulfurization of Diesel Oil Using Activated Charcoal* (2015).
- Khaled, M. Adsorption performance of multiwall carbon nanotubes and graphene oxide for removal of thiophene and dibenzothiophene from model diesel fuel. *Res. Chem. Intermed.* **41**, 9817–9833 (2015).
- Sadare, O. O. *Development and Evaluation of Adsorption Coupling Biodesulfurization (AD/BDS) Process for the Desulphurization of South African Petroleum Distillates* (University of the Witwatersrand, Faculty of Engineering and the Built Environment, 2018).
- Ibrahim, N. K. & Aljanabi, S. K. Desulfurization and kinetic study of diesel fuel by batch adsorption on activated carbon. *Eng. Technol. J.* **33**(8), 16 (2015).
- Nunthaprechachan, T., Pengpanich, S. & Hunsom, M. Adsorptive desulfurization of dibenzothiophene by sewage sludge-derived activated carbon. *Chem. Eng. J.* **228**, 263–271 (2013).
- Al-Khodior, A., Yusra, A. & Albayati, T. M. Adsorption desulfurization of actual heavy crude oil using activated carbon. *Eng. Technol. J.* **38**(10A), 1441–1453 (2020).

37. Shokouhi, M. et al. Thermodynamical and artificial intelligence approaches of H₂S solubility in N-methylpyrrolidone. *Chem. Phys. Lett.* **707**, 22–30 (2018).
38. Esfandyari, M., Delouei, A. A. & Jalai, A. Optimization of ultrasonic-excited double-pipe heat exchanger with machine learning and PSO. *Int. Commun. Heat Mass Transf.* **147**, 106985 (2023).
39. Niknam Shahrak, M., Esfandyari, M. & Karimi, M. Efficient prediction of water vapor adsorption capacity in porous metal-organic framework materials: ANN and ANFIS modeling. *J. Iran. Chem. Soc.* **16**(1), 11–20 (2019).
40. Esfandyari, M., Esfandyari, M. & Jafari, D. Prediction of thiophene removal from diesel using [BMIM][AlCl₄] in EDS process: GA-ANFIS and PSO-ANFIS modeling. *Pet. Sci. Technol.* **36**(16), 1305–1311 (2018).
41. Esfandyari, M., Amiri, M. & Salooki, M. K. Neural network prediction of the Fischer–Tropsch synthesis of natural gas with Co(III)/Al₂O₃ catalyst. *Chem. Eng. Res. Bull.* **17**(1), 25–33 (2015).
42. Nobakht, A., Jafari, D. & Esfandyari, M. New insights on the adsorption of phenol red dyes from synthetic wastewater using activated carbon/Fe₃(MoO₄)₃. *Environ. Monit. Assess.* **195**(5), 574 (2023).
43. Muzic, M. et al. Study of diesel fuel desulfurization by adsorption. *Chem. Eng. Res. Des.* **88**(4), 487–495 (2010).
44. Omar, R. A. & Verma, N. Review of adsorptive desulfurization of liquid fuels and regeneration attempts. *Ind. Eng. Chem. Res.* **61**(25), 8595–8606 (2022).
45. Hernández-Maldonado, A. J. & Yang, R. T. Desulfurization of transportation fuels by adsorption. *Catal. Rev.* **46**(2), 111–150 (2004).
46. Zhang, Y. & Xu, H. Optimization of desulfurization process via choline phosphotungstate coupled with persulfate using response surface methodology. *Catalysts* **14**(5), 326 (2024).
47. Tang, J. et al. Optimization studies on biological desulfurization of sulfide ore using response surface methodology. *Minerals* **11**(6), 583 (2021).
48. Majid, M. F. et al. Liquid polymer eutectic mixture for integrated extractive-oxidative desulfurization of fuel oil: An optimization study via response surface methodology. *Processes* **8**(7), 848 (2020).
49. Aslan, N. & Cebeci, Y. Application of Box–Behnken design and response surface methodology for modeling of some Turkish coals. *Fuel* **86**(1–2), 90–97 (2007).
50. Nariman, N. A. et al. Optimum design of flexural strength and stiffness for reinforced concrete beams using machine learning. *Appl. Sci.* **11**(18), 8762 (2021).
51. Pereira, L. M. S., Milan, T. M. & Tapia-Blácido, D. R. Using response surface methodology (RSM) to optimize 2G bioethanol production: A review. *Biomass Bioenergy* **151**, 106166 (2021).
52. Zafari, P. & Ghaemi, A. Modeling and optimization of CO₂ capture into mixed MEA-PZ amine solutions using machine learning based on ANN and RSM models. *Results Eng.* **19**, 101279 (2023).
53. McCulloch, W. S. & Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* **5**, 115–133 (1943).
54. Donald, H. *The Organization of Behavior* (Wiley, 1952).
55. Fath, A. H., Madanifar, F. & Abbasi, M. Implementation of multilayer perceptron (MLP) and radial basis function (RBF) neural networks to predict solution gas-oil ratio of crude oil systems. *Petroleum* **6**(1), 80–91 (2020).
56. Valera, V. Y., Martins, T. D. & Codolo, M. C. Experimental evaluation and neural networks modeling of removal efficiency and volumetric mass transfer coefficient for gas desulfurization in spray tower. *Chem. Eng. Sci.* **285**, 119568 (2024).
57. Gheisari, M. *Modeling of Oxidative Desulfurization Process by Artificial Neural Network* (EasyChair, 2022).
58. Sadrara, M. & Khorrami, M. K. Principal component analysis–multivariate adaptive regression splines (PCA-MARS) and back propagation-artificial neural network (BP-ANN) methods for predicting the efficiency of oxidative desulfurization systems using ATR-FTIR spectroscopy. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **300**, 122944 (2023).
59. Arce-Medina, E. & Paz-Paredes, J. I. Artificial neural network modeling techniques applied to the hydrodesulfurization process. *Math. Comput. Model.* **49**(1–2), 207–214 (2009).
60. Wu, S. et al. Prediction of endpoint sulfur content in KR desulfurization based on the hybrid algorithm combining artificial neural network with SAPSO. *IEEE Access* **8**, 33778–33791 (2020).
61. Mansouri, K., Bahmanzadegan, F. & Ghaemi, A. Evaluation of hydrogen production via steam reforming and partial oxidation of dimethyl ether using response surface methodology and artificial neural network. *Sci. Rep.* **14**(1), 15570 (2024).
62. Ghaemi, A., Dehnavi, M. K. & Khoshraftar, Z. Exploring artificial neural network approach and RSM modeling in the prediction of CO₂ capture using carbon molecular sieves. *Case Stud. Chem. Environ. Eng.* **7**, 100310 (2023).
63. Zhu, X. et al. Application of machine learning methods for estimating and comparing the sulfur dioxide absorption capacity of a variety of deep eutectic solvents. *J. Clean. Prod.* **363**, 132465 (2022).
64. Bilgiç, G. et al. Recent advances in artificial neural network research for modeling hydrogen production processes. *Int. J. Hydrog. Energy* **48**, 18947–18977 (2023).
65. Khosravi-Nikou, M. et al. A robust method to predict equilibrium and kinetics of sulfur and nitrogen compounds adsorption from liquid fuel on mesoporous material. *Iran. J. Oil Gas Sci. Technol.* **9**(2), 93–118 (2020).
66. Qikun, M. Machine learning-based optimization for catalytic sulfur removal: Computational modeling and analysis of fuel purification for reduction of environmental impacts. *Case Stud. Therm. Eng.* **53**, 103835 (2024).
67. Xu, R. & Yang, X. Machine learning optimization for catalytic desulfurization of petroleum: Multi-layered perceptron, multi task Lasso, and Gaussian process regression models. *J. Mol. Liq.* **400**, 124508 (2024).
68. Khoshraftar, Z. & Ghaemi, A. Prediction of CO₂ solubility in water at high pressure and temperature via deep learning and response surface methodology. *Case Stud. Chem. Environ. Eng.* **7**, 100338 (2023).
69. Satapathy, S. K. et al. *EEG Brain Signal Classification for Epileptic Seizure Disorder Detection* (Academic Press, 2019).
70. Naderi, K. et al. Modeling based on machine learning to investigate flue gas desulfurization performance by calcium silicate absorbent in a sand bed reactor. *Sci. Rep.* **14**(1), 954 (2024).
71. Naderi, K., Foroughi, A. & Ghaemi, A. Analysis of hydraulic performance in a structured packing column for air/water system: RSM and ANN modeling. *Chem. Eng. Process. Process Intensif.* **193**, 109521 (2023).
72. Hong, H. et al. Radial basis function artificial neural network (RBF ANN) as well as the hybrid method of RBF ANN and grey relational analysis able to well predict trihalomethanes levels in tap water. *J. Hydrol.* **591**, 125574 (2020).
73. Mashhadimoslem, H., Ghaemi, A. & Palacios, A. Analysis of deep learning neural network combined with experiments to develop predictive models for a propane vertical jet fire. *Heliyon* **6**(11), e05511 (2020).
74. Pashaei, H., Mashhadimoslem, H. & Ghaemi, A. Modeling and optimization of CO₂ mass transfer flux into Pz-KOH-CO₂ system using RSM and ANN. *Sci. Rep.* **13**(1), 4011 (2023).
75. Ghaemi, A. et al. Hydrodynamic behavior of standard liquid-liquid systems in Oldshue-Rushton extraction column; RSM and ANN modeling. *Chem. Eng. Process. Process Intensif.* **168**, 108559 (2021).

Author contributions

K.M.S.: Conceptualization, methodology, conceived and designed the experiments, validation, formal analysis, investigation, resources, writing—original draft, writing—review and editing. M.S.: Conceptualization, software, conceived and designed the experiments, validation, formal analysis, investigation, resources, data curation. A.G.: Conceptualization, methodology, software, conceived and designed the experiments, validation,

formal analysis, investigation, resources, data curation, writing—original draft, writing—review and editing, supervision visualization, project administration, supervision, funding acquisition.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to A.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025