RSC Advances

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Check for updates

Cite this: *RSC Adv.*, 2026, **16**, 1993

Efficacy of advanced Fenton-photo systems for the degradation of petroleum hydrocarbons using complex neural networks

Zainab Ashkanani, *[a] Rabi Mohtar,*[b] Salah Al-Enezi,[c] Faten Khalil,[d] Muthanna Al-Momin,[e] Xingmao Ma, [f] Patricia K. Smith,[a] Salvatore Calabrese,[a] Meshal Abdullah[g] and Najeeb Aladwani[h]

Environmental contamination from petroleum hydrocarbons poses significant challenges, particularly in regions affected by oil industry activities and conflict-related environmental disasters. This study investigated AI-optimized photo-Fenton systems for remediation of oil-contaminated desert soils through three key objectives: (1) systematically evaluating multiple catalysts, particularly iron(III) sulfate pentahydrate, and EDTA as a chelating agent for TPH removal, determining optimal conditions for specific reduction thresholds; (2) assessing four advanced photo-Fenton systems incorporating EDTA and $Fe_2(SO_4)_3\cdot 5H_2O$ across multiple performance metrics, using statistical validation and machine learning to identify key performance factors; and (3) employing neural network modeling and hierarchical clustering to evaluate predictive accuracy, identify influential factors, and discover natural PAH classifications based on degradation patterns. Soil samples from Kuwait's Great Burgan Field were treated using various combinations of $Fe_2(SO_4)_3\cdot 5H_2O$ (5–10 g L^{-1}), Fe_2O_3 (5–35 g L^{-1}), EDTA (15–20 g L^{-1}), and H_2SO_4 (5–25 mL L^{-1}). $Fe_2(SO_4)_3\cdot 5H_2O$ demonstrated superior catalytic activity, achieving 99% TPH removal at 10 g L^{-1}, compared to Fe_2O_3's 92% at 35 g L^{-1}. Advanced photo-Fenton system 3, combining both catalysts, showed exceptional performance with >98% removal of high-molecular-weight PAHs. The synergistic effect is proposed to arise from enhanced radical generation through both hydroxyl ($^{\bullet}OH$) and sulfate radical ($SO_4^{\bullet-}$) pathways, based on an established mechanistic precedent for Fe(III)-sulfate photochemistry. Neural network models successfully predicted PAH removal with $R^2 > 0.91$, while hierarchical clustering revealed distinct contaminant groupings. Treatment efficiency was primarily governed by EDTA concentration and Fe^{2+}/H_2O_2 ratio, with degradation mechanisms varying based on PAH structure. This AI-driven optimization provides an efficient framework for soil remediation in petroleum-contaminated desert environments where traditional methods face significant challenges.

Received 25th September 2025
Accepted 15th December 2025

DOI: 10.1039/d5ra07305g

rsc.li/rsc-advances

1. Introduction

Crude oil (petroleum) is a complex mixture composed primarily of hydrocarbons and small amounts of nitrogen, sulfur, and oxygen. Petroleum hydrocarbons include various chemicals such as gasoline, diesel, and fuel oil.[1,2] The main components of crude oil, collectively referred to as total petroleum hydrocarbons (TPH), encompass both aliphatic and aromatic compounds. Within the aromatic group, polycyclic aromatic hydrocarbons (PAHs) are of particular concern due to their toxic, genotoxic, mutagenic, and carcinogenic properties. Regulatory agencies worldwide focus on identifying and analyzing highly toxic PAHs to evaluate their prevalence and persistence in the environment.[3–7] The United States Environmental Protection Agency (US EPA) designated sixteen PAHs as priority pollutants: naphthalene (NAP), acenaphthylene (ACY), acenaphthene (ACE), fluorene (FLU), phenanthrene (PHEN), anthracene (ANTH), fluoranthene (FLTH), pyrene (PYR), benzo[*a*]anthracene (B[*a*]A), chrysene (CHRY), benzo[*b*]fluoranthene (B[*b*]F), benzo[*k*]fluoranthene (B[*k*]F), benzo[*a*]pyrene (B[*a*]P), benzo[*g*,*h*,*i*]perylene (B[*ghi*]P), indeno[1,2,3-*c*,*d*]pyrene (IND), and dibenz[*a*,*h*]anthracene (D[*ah*]A). These are categorized based on molecular weight and substrate properties into low

---

[a] *Department of Biological and Agricultural Engineering, Texas A&M University, Texas, USA. E-mail: Ashkanani@tamu.edu*
[b] *Department of Biological and Agricultural Engineering and Zachry Department of Civil and Environmental Engineering, Energy Institute, Texas A&M University, USA. E-mail: Mohtar@tamu.edu*
[c] *Petroleum Research Center (PRC), Kuwait Institute for Scientific Research (KISR), Ahmadi, Kuwait*
[d] *Environmental Life Science-Resource Center, Kuwait Institute for Scientific Research (KISR), Al-Shuwaikh, Kuwait*
[e] *Soil Remediation Group – Kuwait Oil Company, Ahmadi, Kuwait*
[f] *Department of Civil & Environmental Engineering, Texas A&M University, Texas, USA*
[g] *Department of Geography, Sultan Qaboos University, Muscat, Oman*
[h] *Department of Earth and Environmental Sciences, Kuwait University, Kuwait*

molecular weight (LMW) PAHs (two to three rings), intermediate molecular weight (IMW) PAHs (four rings), and high molecular weight (HMW) PAHs (five or more rings).[8,9] The molecular weight of PAHs significantly influences their behavior and fate in soil systems and their resistance to conventional remediation approaches.[10,11]

Advanced oxidation processes (AOPs), introduced in the 1980s, are effective in remediating water contaminants using powerful, short-lived, oxidizing agents like hydroxyl radicals ($^{\bullet}$OH).[12,13] These processes involve chemical treatments to decompose organic and inorganic materials in various environmental contexts.[13,14] Among the widely studied AOPs, Fenton reaction generates hydroxyl radicals essential for degrading organic contaminants, including petroleum hydrocarbons and heavy metals.[13,15–17] Fenton-based AOPs (*i.e.* Fenton, Fenton-like, electro-Fenton, and photo-Fenton systems) are appreciated for the high removal efficiency of recalcitrant organic contaminants under different operational and environmental conditions.[13] In these processes, hydrogen peroxide (H_2O_2) is activated by ferrous or ferric iron (Fe^{2+}/Fe^{3+}) under acidic conditions to produce hydroxyl radicals.[18,19] These radicals degrade organic matter by abstracting hydrogen atoms from organic carbons or cleaving carbon–carbon bonds and ultimately converting contaminants into carbon dioxide (CO_2) and water (H_2O).[20,21] The photo-Fenton process, achieved by irradiating Fenton reactions with UV light, further enhances production of hydroxyl radicals: cost-effective and environmentally friendly AOPs extensively used in soil treatments.[21–24] Despite the proven effectiveness of traditional Fenton processes, significant challenges remain in their application to petroleum-contaminated soils, particularly regarding catalyst stability, pH sensitivity, and removal efficiency for recalcitrant high-molecular-weight PAHs. Conventional iron catalysts often suffer from rapid deactivation, limited reusability, and poor performance under near-neutral pH conditions typical of natural soil environments.[25,26] These limitations highlight the need for advanced catalyst systems with enhanced stability, broader pH operating ranges, and superior oxidative capacity.

Artificial Intelligence (AI) and Machine Learning (ML) offer potential improvements in efficiency and effectiveness of environmental remediation as technologies that provide time- and cost-saving solutions to address the complexities of contaminated soils. Their versatility and applicability have been demonstrated across various fields.[27–31] Specifically, ML aids in establishing nonlinear relationships between bio-indicators and the physicochemical properties of contaminated sites, crucial for developing robust models that predict soil contamination levels and assess phytotoxicity with high accuracy.[32] Techniques such as ensemble learning are effective in modeling soil and image data, enhancing the estimation of heavy metal distribution and integrating environmental variables to enable precise predictions about soil reclamation processes.[27,30,31,33]

Deep learning, a subset of ML, offers a framework for supervised learning tasks that involve complex functions. By layering multiple levels of artificial neurons, deep networks can represent intricate patterns, making them suitable for tasks like pattern recognition, classification, and prediction.[28,34–36] Artificial neural networks (ANNs), which mimic neural processing in the human brain, are recognized as robust nonlinear estimators and have been validated in various domains. Their application in environmental science extends to analyzing soil contaminants. Recent studies show their effectiveness in predicting concentrations of toxic compounds like TPH and PAHs in contaminated soils.[37] ML methods are increasingly used to detect and quantify heavy metals and petroleum hydrocarbons, providing detailed insight into the physicochemical parameters of soils.[38]

Recent studies have demonstrated the potential of artificial neural networks (ANNs) for optimizing Fenton and photo-Fenton processes in water treatment applications Elmolla *et al.* achieved $R^2 = 0.997$ in predicting antibiotic degradation *via* Fenton oxidation.[39] Sarı *et al.* applied ANN and NARX-ANN models to optimize photo-Fenton treatment of textile wastewater, achieving 94% COD removal.[40] Deep reinforcement learning has been employed for automated parameter optimization in photo-Fenton disinfection,[41] while recent reviews have highlighted the growing integration of ML techniques in advanced oxidation processes.[42,43]

However, these applications have predominantly focused on wastewater treatment with operational parameters as model inputs. Critical gaps remain in: (1) ML-driven optimization of photo-Fenton systems for soil remediation, where soil matrix interactions introduce additional complexity; (2) integration of molecular descriptors to predict compound-specific degradation based on chemical structure; and (3) systematic evaluation of novel catalyst systems such as $Fe_2(SO_4)_3 \cdot 5H_2O$. This study addresses these gaps by developing a comprehensive neural network framework that combines molecular descriptors with treatment parameters to predict PAH-specific removal efficiency in contaminated soils. The complex interactions between multiple catalysts, chelating agents, and target contaminants present an ideal scenario for advanced computational modeling approaches that can identify non-obvious relationships and optimization pathways beyond traditional experimental design.[44,45]

Comprehensive site characterization is essential for effective remediation design, particularly in heterogeneous contaminated environments. Electrical resistivity methods have emerged as powerful non-invasive tools for delineating subsurface hydrocarbon contamination, providing critical spatial information that complements point-based chemical sampling.[46,47] The high electrical resistivity of petroleum hydrocarbons relative to water-saturated soils creates distinct geophysical signatures that enable three-dimensional mapping of contamination plumes.[48,49] Recent advances integrate resistivity surveys with machine learning algorithms to predict contamination severity and optimize sampling strategies, demonstrating the synergy between geophysical characterization and computational modeling.[50,51] For weathered hydrocarbon sites, resistivity variations also reflect biodegradation processes, as microbial activity produces conductive metabolites that alter bulk soil conductivity.[52,53] This multi-scale characterization approach combines surface observations, subsurface geophysics, and advanced computational analysis. The integrated methodology provides comprehensive understanding necessary for designing targeted remediation strategies in complex contaminated environments.

The Gulf War (1990–1991) resulted in one of the most severe environmental disasters in history, leading to a massive hydrocarbon spill in Kuwait. It is estimated that approximately 6–8 million barrels of crude oil were released into terrestrial and marine ecosystems causing extensive ecological damage.[54,55] Despite receiving $460 million in compensation for environmental restoration, Kuwait's oil fields remain contaminated with residual oil lakes and tarcrete formations.[54,55] Reports from the Kuwait Environmental Public Authority (KEPA) and the Kuwait Oil Company (KOC) have documented significant alterations in soil texture, loss of local wildlife, and deep penetration of oil into soil strata reaching freshwater aquifers.[54,56] Studies suggest that the consequences of this oil-induced pollution may persist for decades and pose long-term environmental challenges.[55,57,58] In response, the United Nations Compensation Commission (UNCC) allocated $3 billion to Kuwait to initiate a remediation project aimed at cleansing oil-contaminated soil.[59] Despite these efforts, clean-up operations remain financially burdensome and logistically complex, especially when applied to extensive areas of contaminated soil. Some of the employed remediation strategies have served as temporary measures and highlight the need for more effective, sustainable solutions.[60]

The goal of this research is to develop and validate an optimized, AI-enhanced photo-Fenton system for the efficient, cost-effective remediation of petroleum hydrocarbon-contaminated soils, particularly those impacted by oil industry activities and conflicts. It aims to provide a widely applicable solution to address persistent environmental contamination caused by petroleum hydrocarbons.

The central aim is to conduct a thorough investigation into the efficacy of various advanced Fenton-photo system configurations, utilize multiple catalysts and chelating agents to enhance the oxidation process, and employ deep learning and other statistical techniques to analyze the results. By combining experimental work with advanced computational methods, the study seeks to identify the optimal treatment parameters that maximize contaminant removal and understand the underlying mechanisms. The research is structured around the following key objectives:

(1) Systematically evaluate and compare the effectiveness of multiple catalysts, particularly iron(III) sulfate pentahydrate, and EDTA as a chelating agent in Fenton and modified Fenton processes for TPH removal.

This objective focuses on determining the optimal conditions (concentration, reaction time) for each catalyst and chelating agent to achieve specific TPH reduction thresholds, with particular emphasis on investigating iron(III) sulfate pentahydrate, a catalyst previously employed in wastewater treatment, in the novel context of petroleum-contaminated soil remediation, and its strategic combination with Fe_2O_3 for enhanced synergistic performance.

(2) Assess and compare four advanced photo-Fenton systems incorporating EDTA and iron(III) sulfate pentahydrate for remediating oil-contaminated soil.

This objective focuses on evaluating multiple performance metrics including TPH removal efficiency and PAH degradation capabilities, using statistical validation and machine learning approaches to identify key performance factors, and establishing optimal operating conditions for field applications.

(3) Employ neural network modeling and hierarchical clustering analysis to gain deeper insight into PAH removal across the four treatment systems.

This objective seeks to evaluate predictive model accuracy, identify influential factors affecting removal efficiency, determine optimal operational parameters, discover natural groupings of PAHs based on treatment responses, and provide comprehensive comparative evaluation of the treatment systems for specific PAH profiles in contaminated environments.

This work represents a significant advancement in materials chemistry applied to environmental remediation through the novel integration of established iron catalysts ($Fe_2(SO_4)_3\cdot 5H_2O$ and Fe_2O_3) with AI-based optimization frameworks. While Fe(III) sulfate has been previously utilized in modified Fenton systems for wastewater treatment, its strategic combination with Fe_2O_3 coupled with neural network modeling for petroleum hydrocarbon-contaminated soil remediation has not been systematically investigated. This dual-catalyst approach creates synergistic reactive interfaces that overcome the limitations of traditional single-catalyst Fenton systems while enabling efficient oxidation under near-neutral pH conditions. By elucidating the structure–activity relationships governing photo-Fenton PAH degradation and developing predictive computational models, this research contributes fundamental insights to the field of advanced photocatalytic oxidation while addressing a critical environmental challenge.

2. Materials and methods

2.1 Overall methodology

We investigated the remediation of oil-contaminated soil from the Great Burgan Field in Kuwait using both Fenton-like and advanced photo-Fenton oxidation processes. A multi-faceted approach was employed, combining field sampling, laboratory experiments, analytical chemistry, statistical analysis, and computational modeling.

Soil samples were collected using a stratified random sampling approach, with triplicate samples taken from contaminated and uncontaminated (control) sites. Samples were prepared following standard protocols, including air-drying, homogenization, and sieving. Soil physicochemical properties (pH, alkalinity, specific conductance, moisture content, total organic matter, and total organic carbon) were characterized.

Fenton-like experiments assessed the individual impacts of Fe_2SO_4, Fe_2O_3, H_2SO_4, and EDTA on contaminant removal, using predetermined dosages of H_2O_2 and Fe^{2+}. Four advanced photo-Fenton systems were developed, combining H_2O_2, $Fe_2(SO_4)_3\cdot 5H_2O$, Fe_2O_3, EDTA, Fe^{2+}, and H_2SO_4 in various configurations and concentrations. Experiments were conducted in soil slurries (1 : 100 soil-to-water ratio) under controlled conditions, with a full-spectrum halogen lamp providing irradiation. Treatment durations ranged from 5 to 30 minutes.

Total Petroleum Hydrocarbons (TPH) were analyzed using a modified USEPA Method 3540C with Soxhlet extraction and GC-FID. Polycyclic Aromatic Hydrocarbons (PAHs) were analyzed using the same extraction method, with additional cleanup steps, followed by GC-MS in selected ion monitoring mode. Rigorous quality control measures, including method blanks, matrix spikes, and certified reference materials, were implemented throughout.

Statistical analysis was performed using R (version 4.2.0). Repeated measures ANOVA, mixed-effects modeling, and principal component analysis (PCA) were used to evaluate treatment effects and degradation patterns. Power analysis and effect size calculations were performed.

A comprehensive computational framework (Python 3.9.12) was developed to analyze PAH degradation, including: calculating 14 molecular descriptors (using RDKit), performing PCA, and developing neural network models to predict removal efficiency. Feature importance analysis (Random Forest Regressor) and visualization techniques (violin plots, 3D surface plots) were used to interpret the model results. A complementary machine learning approach used molecular fingerprints (ECFP4, MACCS keys, Atom Pair fingerprints) and a separate neural network architecture. Hierarchical clustering and network graphs were used to analyze molecular similarities. All computational methods were rigorously validated.

2.2 Site description and sample collection

Soil samples were collected from the heavily impacted Great Burgan Field in Kuwait, which experienced severe oil contamination during the 1990–1991 Gulf War. Initial characterization revealed TPH concentrations ranging from 129 726–86 653 mg kg^{-1} and total PAH concentrations between 500 and 800 μg kg^{-1}. Control samples obtained from an uncontaminated proximate area contained less than 50 mg kg^{-1} TPH and PAH concentrations below the detection limit of 5 μg kg^{-1}. A stratified random sampling approach was implemented to account for spatial variability, with triplicate samples collected from three distinct locations within each designated stratum using a stainless-steel soil corer (5 cm diameter) at depths of 10–15 cm. This sampling depth was selected based on previous studies indicating peak concentrations of weathered petroleum hydrocarbons in this horizon. Detailed site characteristics, sampling coordinates, and soil properties are provided in SI (Table S1 and Fig. S1).

2.3 Geophysical site characterization

Electrical resistivity surveys were conducted to complement surface sampling and provide comprehensive subsurface contamination assessment. Vertical Electrical Sounding (VES) measurements at 15 locations across the study area employed Wenner array configuration with electrode spacings ranging from 1 to 10 m. The survey revealed extreme heterogeneity in resistivity values (1.51–464.72 Ω m), reflecting complex subsurface contamination patterns from the 1991 oil spill. Three distinct contamination signatures were identified: heavy surface contamination, moderate contamination with moisture influence, and low contamination or degraded zones. Complete geophysical characterization methodology and results are detailed in SI Section S6.

2.4 Sample preparation and characterization

Collected samples were stored in pre-cleaned amber glass containers and transported at 4 °C. Samples were air-dried, homogenized, sieved through 2 mm mesh, and stored at −20 °C until analysis. Characterization revealed pH 7.8, organic matter content 4.5%, and particle size distribution of 60% sand, 25% silt, and 15% clay. Detailed preparation protocols are provided in SI Section S7.

2.5 Fenton-like and advanced Fenton-photo systems

This study investigated both Fenton-like and advanced photo-Fenton systems for remediation of oil-contaminated soil (Table 1). Fenton-like experiments were conducted using 8 g soil samples, while advanced photo-Fenton experiments utilized 20 g samples. In the Fenton-like experiments, the impact of varying concentrations of Fe_2SO_4, Fe_2O_3, H_2SO_4, and EDTA on contaminant removal was individually assessed. These experiments were performed with predetermined doses of H_2O_2 (0.349 M and 0.419 M) and Fe^{2+} (0, 0.01 M, 0.02 M, and 0.03 M) (Table 2).

Four distinct advanced photo-Fenton systems were developed to explore synergistic effects of different reagent combinations (detailed in Table 1). System 1 consisted of H_2O_2, $Fe_2(SO_4)_3 \cdot 5H_2O$, and EDTA. System 2 combined H_2O_2, $Fe_2(SO_4)_3 \cdot 5H_2O$, EDTA, Fe^{2+}, and H_2SO_4. System 3 was formulated with H_2O_2, $Fe_2(SO_4)_3 \cdot 5H_2O$, Fe_2O_3, EDTA, Fe^{2+}, and H_2SO_4. System 4 utilized H_2O_2, Fe_2O_3, EDTA, and Fe^{2+}. These advanced systems were tested under varying concentrations of H_2O_2 (0, 0.349 M, 0.419 M), Fe^{2+} (0, 0.01, 0.02, and 0.03 M), and EDTA (15 and 20 g L^{-1}).

All experiments were conducted under controlled laboratory conditions using soil slurries prepared with a soil-to-water ratio of 1 : 100 (w/v). The reaction pH, maintained at 6.5 using dilute NaOH or H_2SO_4, based on preliminary studies that identified this pH as optimal for Fenton reaction efficiency. A full-spectrum 500 watt halogen lamp (Halco ProLume H500Q/CL), positioned 30 cm above the reaction vessels, provided consistent irradiation at an intensity of 1000 μW cm^{-2}. Treatment durations of 5, 10, 15, and 30 minutes were evaluated. Reactions were terminated at the designated time points by the addition of sodium thiosulfate, effectively quenching the Fenton process. All reagents used in this study, including $Fe_2(SO_4)_3 \cdot 5H_2O$, Fe_2O_3, H_2O_2, EDTA, and $FeSO_4 \cdot 7H_2O$, are commercially available at industrial scale from multiple suppliers, facilitating potential scale-up and field implementation.

2.6 Analytical methods

TPH analysis employed modified USEPA Method 3540C with Soxhlet extraction and GC-FID analysis. PAH analysis utilized the same extraction method with additional cleanup steps,

Table 1 Overview of experimental parameters for various Fenton-like systems

Fenton-like systems	Soil (g)	pH	Time (minutes)	$Fe_2(SO_4)_3 \cdot 5H_2O$ (g L^{-1})	Fe_2O_3 (g L^{-1})	EDTA (g L^{-1})	H_2SO_4 (mL L^{-1})	Fe^{2+} (M)	H_2O_2 (M)
Iron(III) sulfate pentahydrate	8	6.5	10	5	0	0	0	0	
				10				0.01	
								0.02	
								0.03	
Iron(III) oxide				0	5	0	0	0	
					15			0.01	
					25			0.02	
					35			0.03	
Ethylenediaminetetraacetic acid (EDTA)				0	0	15	0	0	
								0.01	
						20		0.02	
								0.03	
Sulfuric acid				0	0	0	5	0	
							10	0.01	
							15	0.02	0
							25	0.03	0.349
Advanced photo-Fenton system 1	20		30	10	0	15	0	0	0.419
						20			
Advanced photo-Fenton system 2				10	0	15	25	0	
								0.01	
						20		0.02	
								0.03	
Advanced photo-Fenton system 3				10	15	15	25	0	
								0.01	
						20		0.02	
								0.03	
Advanced photo-Fenton system 4				0	15	15	0	0	
								0.01	
						20		0.02	
								0.03	

followed by GC-MS in selected ion monitoring mode. Method detection limits were 5 mg kg^{-1} (TPH) and 0.05–0.1 μg kg^{-1} (individual PAHs). Complete analytical protocols, instrument parameters, and quality assurance procedures are detailed in SI Section S8.

2.7 Quality control and statistical analysis

Quality control measures included method blanks, matrix spikes, certified reference materials with recovery studies (80–120% acceptable range), and precision measurements (≤20%

Table 2 Optimal treatment parameters for different TPH reduction levels across various catalyst systems

	Low			Medium			High		
	(70% reduction)			(80% reduction)			(90% reduction)		
Treatment	Time to reduction (minutes)	Optimal H_2O_2 (M)	Optimal Fe^{2+} (M)	Time to reduction (minutes)	Optimal H_2O_2 (M)	Optimal Fe^{2+} (M)	Time to reduction (minutes)	Optimal H_2O_2 (M)	Optimal Fe^{2+} (M)
EDTA (15 g L^{-1})	10	0.349	0	10	0.349	0.01	30	0.349	0.01
EDTA (20 g L^{-1})	30	0.349	0	30	0.419	0.03	10	0.419	0.02
Iron(III) oxide (5 g L^{-1})	10	0.349	0	10	0.349	0.01	30	0.349	0
Iron(III) oxide (15 g L^{-1})	30	0.419	0.01	30	0.349	0.01	10	0.419	0.02
Iron(III) oxide (25 g L^{-1})	10	0.349	0.01	30	0.419	0.02	10	0.419	0.03
Iron(III) oxide (35 g L^{-1})	30	0.349	0.03	10	0.349	0	30	0.349	0
Iron(III) sulfate pentahydrate (5 g L^{-1})	10	0.419	0.02	10	0.349	0	30	0.419	0.02
Iron(III) sulfate pentahydrate (10 g L^{-1})	30	0.349	0.03	30	0.349	0.01	10	0.419	0.02
Sulfuric acid (5 mL L^{-1})	30	0.349	0.03	10	0.419	0	30	0.349	0
Sulfuric acid (10 mL L^{-1})	30	0.419	0.03	—	—	—	10	0.419	0.02
Sulfuric acid (15 mL L^{-1})	—	—	—	—	—	—	10	0.349	0
Sulfuric acid (25 mL L^{-1})	—	—	—	—	—	—	10	0.419	0

RSD). Statistical analysis employed R (version 4.2.0) with repeated measures ANOVA, mixed-effects modeling, and PCA. Power analysis validated sample size adequacy ($\beta = 0.80$). Full quality control procedures and statistical methodologies are provided in SI Section S9.

2.8 Neural network modeling

2.8.1 Advanced computational framework for PAH analysis and degradation assessment. A comprehensive computational framework was developed to analyze PAH degradation patterns, integrate molecular descriptor analysis, machine learning, and statistical visualization. Fourteen molecular descriptors were calculated for each PAH using RDKit, including molecular weight, $\log P$, hydrogen bond acceptors, and various topological indices. These were validated against experimental values with maximum acceptable deviation of 5%. These descriptors were selected based on their established mechanistic relationships to oxidative degradation processes. Analysis revealed that $X\,\text{Log}\,P$ governs the transition between reaction-controlled kinetics (time-dependent, $X\,\text{Log}\,P < 5.2$) and catalyst-limited kinetics (Fe concentration-dependent, $X\,\text{Log}\,P > 5.8$), reflecting the role of hydrophobicity in controlling aqueous-phase availability. Molecular dimensions (width, length) correlate with optimal treatment system selection, with extended molecules (length > 13 Å) requiring dual-catalyst System 3. The complexity index encodes ring fusion patterns that determine reactive site accessibility for radical attack. Detailed descriptor–kinetics correlations and mechanistic interpretations are provided in SI Section S17 and Table S3.

The dataset comprised PAH removal efficiency measurements for 16 EPA priority PAHs across four treatment systems, with varying concentrations of EDTA (15, 20 g L^{-1}), H_2O_2 (0.349, 0.419 M), Fe^{2+} (0, 0.01, 0.02, 0.03 M), and treatment times (0, 10, 30 minutes), totaling 7488 data points (triplicate measurements). The 14 molecular descriptors were combined with 4 treatment parameters as input features (18 total), with PAH removal efficiency (%) as the output variable. Data preprocessing included outlier detection using Mahalanobis distance ($p < 0.001$) and StandardScaler normalization. Detailed dataset composition and preprocessing steps are provided in SI Section S10.

A neural network model with four hidden layers was developed to predict PAH removal efficiency, using molecular descriptors and treatment parameters as inputs. The model achieved mean absolute error < 0.1 for normalized descriptor predictions through nested cross-validation. Feature importance analysis was conducted using a complementary Random Forest Regressor, a standard approach for interpreting black-box models, with results validated through permutation importance (30 repetitions).

A complementary machine learning approach utilized molecular fingerprints (ECFP4, MACCS keys, Atom Pair fingerprints) with a six-layer neural network architecture. The dataset was partitioned using stratified approach (90% training/validation, 10% testing) with five-fold cross-validation. Detailed neural network architecture, training procedures, and validation methods are provided in SI Section S10.

To ensure reproducibility and transparency, all codes necessary for reproducing the results are publicly available at https://github.com/SAAAHco/PhotoFenton-PAH-Remediation-ML.

3. Results

3.1 Subsurface contamination architecture

Electrical resistivity surveys at 15 VES stations revealed extreme heterogeneity in subsurface contamination, with resistivity values ranging from 1.51 to 464.72 Ω m. Three distinct contamination signatures were identified: heavy surface contamination (>60 Ω m at shallow depths), moderate contamination with moisture influence (10–80 Ω m), and low contamination or degraded zones (<20 Ω m). The spatial distribution and subsurface architecture of the contamination were visualized through resistivity cross-sections (SI Fig. S2 and S3) and horizontal slice maps (SI Fig. S4). The resistivity data provided critical insight to understanding variable photo-Fenton performance across the study area, with high resistivity zones requiring higher catalyst concentrations and moderate resistivity zones showing optimal treatment conditions. Complete geophysical characterization results, including spatial distribution analysis and treatment implications, are provided in SI Section S12 with Table S2.

3.2 Comparative performance of multiple photocatalysts and EDTA chelation in Fenton-based TPH removal

This study investigated the influence of various catalysts and chelating agents in Fenton and modified Fenton processes for TPH removal from oil-contaminated soil, with particular emphasis on iron(III) sulfate pentahydrate ($Fe_2(SO_4)_3\cdot5H_2O$) as a novel catalyst addition. The experimental matrix encompassed EDTA chelation, iron(III) catalysis, acid-enhanced oxidation, and systematic evaluation of $Fe_2(SO_4)_3\cdot5H_2O$'s catalytic performance, with removal efficiency assessed across three reduction thresholds (70%, 80%, and 90% TPH removal) (Table 2).

The chelating agent EDTA demonstrated complex coordination chemistry at different concentrations. At 15 g L^{-1}, EDTA achieved 70% and 80% reduction within 10 minutes using 0.349 M H_2O_2. The higher EDTA concentration (20 g L^{-1}) exhibited different reaction kinetics, achieving 90% removal in 10 minutes with 0.419 M H_2O_2 and 0.02 M Fe^{2+}. Iron(III) oxide demonstrated concentration-dependent catalytic activity, with optimal performance at 15 g L^{-1} achieving 90% reduction in 10 minutes. Iron(III) sulfate pentahydrate's performance revealed insights into slurry-phase Fenton chemistry, with 10 g L^{-1} concentration achieving 90% reduction in 10 minutes with 0.419 M H_2O_2. Sulfuric acid's role extended beyond pH modification, with superior performance at higher concentrations (15–25 mL L^{-1}) achieving 90% reduction in 10 minutes. The observation that 25 mL L^{-1} required only H_2O_2 without additional Fe^{2+} suggests acid-catalyzed H_2O_2 decomposition as a significant pathway.

The results demonstrate that optimal performance depends on achieving the right balance between radical generation, catalyst availability, and mass transfer processes. Higher catalyst

Table 3 Comparative analysis of advanced photo-Fenton systems for PAH degradation in contaminated soil

System	Optimal conditions	TPH removal (%)	PAH degradation characteristics	Statistical significance	PCA results
System 1: H_2O_2 + $Fe_2(SO_4)_3\cdot 5H_2O$ + EDTA	H_2O_2: 0.419 M EDTA: 20 g L^{-1} $Fe_2(SO_4)_3\cdot 5H_2O$: 25 g L^{-1} pH: 6.5	99	Rapid degradation of low MW PAHs >95% removal of 4–6 ring PAHs at 30 minutes Rebound effect observed	$p < 0.05$ for H_2O_2 $\varphi = 0.55$ $R^2 = 0.902$	PC1: 72.3% (time) PC2: 18.7% (H_2O_2)
System 2: H_2O_2 + $Fe_2(SO_4)_3\cdot 5H_2O$ + Fe^{2+} + H_2SO_4 + EDTA	H_2O_2: 0.419 M EDTA: 15 g L^{-1} Fe^{2+}: 0.02 M H_2SO_4: 25 g L^{-1}	99	Enhanced degradation of 5–6 ring PAHs >97% removal of benzo[*a*]pyrene Improved selectivity	$F(15, 30) = 5.55$, $p < 0.001$ $\varphi = 0.61$ $R^2 = 0.928$	PC1: 68.5% (time) PC2: 22.3% (H_2O_2 + Fe^{2+})
System 3: H_2O_2 + $Fe_2(SO_4)_3\cdot 5H_2O$ + Fe_2O_3 + Fe^{2+} + H_2SO_4 + EDTA	H_2O_2: 0.419 M EDTA: 15 g L^{-1} Fe_2O_3: 15 g L^{-1} Fe^{2+}: 0.02 M	99	Best performance for high MW PAHs >98% removal of recalcitrant PAHs Accelerated degradation in the first 10 minutes	$F(15, 30) = 5.55$, $p < 0.001$ $\varphi = 0.63$ $R^2 = 0.923$	PC1: 75.6% (time) PC2: 15.8% (catalysts)
System 4: H_2O_2 + Fe_2O_3 + Fe^{2+} + H_2SO_4 + EDTA	H_2O_2: 0.419 M EDTA: 15 g L^{-1} Fe^{2+}: 0.02 M Fe_2O_3: 15 g L^{-1}	92	>95% removal of 2–3 ring PAHs in 10 minutes Slower degradation of high MW PAHs Uniform EDTA effect across PAHs	$F(15, 30) = 5.55$, $p < 0.001$ $\varphi = 0.54$ $R^2 = 0.896$	PC1: 70.2% (time + Fe^{2+}) PC2: 19.5% (EDTA)

concentrations generally enhanced removal rates through increased active site availability. Detailed mechanistic discussions and extended analysis are provided in SI Section S13.

3.3 Performance evaluation of advanced photo-Fenton systems for TPH and PAH degradation

A comprehensive evaluation of four advanced photo-Fenton systems was conducted utilizing EDTA as a core chelating agent across all configurations, with iron(III) sulfate pentahydrate incorporated in three systems (Tables 3 and 4).

System 1 achieved 99% TPH removal within 10 minutes under optimal conditions (0.419 M H_2O_2, 20 g L^{-1} EDTA, 25 g L^{-1} $Fe_2(SO_4)_3\cdot 5H_2O$), with excellent removal of low molecular weight PAHs and >95% removal of 4–6 ring PAHs after 30 minutes. Statistical analysis revealed significant H_2O_2 concentration effects ($p < 0.05$, $R^2 = 0.902$). Principal component analysis identified treatment time as the primary factor (PC1: 72.3%) and H_2O_2 concentration as secondary (PC2: 18.7%).

System 2 achieved 99% TPH removal in 30 minutes, demonstrating better high molecular weight PAH degradation, particularly >97% removal of benzo[*a*]pyrene. The system exhibited high process stability and superior model performance ($R^2 = 0.928$). Statistical analysis showed significant treatment effects ($F(15, 30) = 5.55$, $p < 0.001$). PCA revealed treatment time (PC1: 68.5%) and combined H_2O_2/Fe^{2+} effects (PC2: 22.3%) as primary performance drivers (SI Fig. S4).

System 3 achieved 99% TPH removal in 10 minutes with the best performance for high MW PAHs (>98% removal). This system demonstrated strong model performance ($R^2 = 0.923$) and robust temporal dependence ($\varphi = 0.63$). Statistical analysis confirmed significant treatment effects ($F(15, 30) = 5.55$, $p < 0.001$). PCA indicated treatment time dominance (PC1: 75.6%) with catalyst interactions as secondary factors (PC2: 15.8%) (SI Fig. S7). A minimal rebound effect was observed in some treatment conditions between 10 and 30 minutes (Fig. S5 and S8), attributed to oxidized intermediate formation as discussed in SI Section S5. The relatively small magnitude of this effect in the dual-catalyst systems supports sustained radical generation that promotes complete mineralization rather than intermediate accumulation.

Table 4 Comparative performance metrics of modified photo-Fenton systems

Performance indicator	System 1	System 2	System 3	System 4
Maximum TPH removal (%)	99	99	99	92
Time to max removal (minutes)	10	30	10	30
Optimal EDTA concentration (g L^{-1})	20	15	15	15
High MW PAH removal	Good	Better	Best	Moderate
Low MW PAH removal	Excellent	Excellent	Excellent	Excellent
Process stability	Moderate	High	High	Moderate
Model performance (R^2)	0.902	0.928	0.923	0.896

Table 5 Comparative evaluation of PAH treatment systems with performance metrics and compound clustering analysis

System	Model performance	Feature importance (top 5)	Removal rates	PAH clustering analysis
System 1	MSE: 0.0342 ± 0.0028	(1) Benzo[*b*]fluoranthene (B[*b*]F)	EDTA (15 g L^{-1}): 92.1 ± 2.8% – optimized radical generation and stability	**Group 1 (known carcinogens)**
	R^2: 0.912 ± 0.024 Convergence: 800 epochs	(2) Pyrene (PYR) (3) Treatment time	EDTA (20 g L^{-1}): 88.4 ± 3.1% H_2O_2 (0.349 M): 93.2 ± 2.5% – optimized peroxide stability	Compounds: NAP–CHRY Similarity: 0.85
	Silhouette width: 0.72	(4) EDTA concentration (5) H_2O_2 concentration	H_2O_2 (0.419 M): 87.6 ± 3.2% Treatment time: increased efficiency over time	Silhouette score: 0.72 **Mechanism** • Shared carcinogenicity • Similar chemical properties **Group 2A (probable carcinogens)** Compounds: B[*a*]A and D[*ah*]A Similarity: clustered together Silhouette score: 0.72 **Mechanism** • Higher probability of carcinogenicity • Similar genotoxicity • Similar inflammatory potential **Non-carcinogenic** Compounds: ACY Similarity: distinct clustering Silhouette score: 0.72 **Mechanism** • Divergent treatment response • Unique degradation pattern
System 2	MSE: 0.0315 ± 0.0024 R^2: 0.934 ± 0.021 Convergence: 750 epochs	(1) Anthracene (ANTH) (2) Pyrene (PYR) (3) Fluorene (FLU)	EDTA (15 g L^{-1}): 92.3 ± 3.2% EDTA (20 g L^{-1}): 89.1 ± 2.9% Fe^{2+} (0.030 M): maximum efficiency achieved	**Unique cluster** Compounds: B[*a*]P Similarity: 0.81
	Silhouette width: 0.75	(4) Benzo[*a*]pyrene (B[*a*]P)	H_2O_2 (0.349 M): 90.5 ± 2.8% H_2O_2 (0.419 M): 94.2 ± 2.4%	Silhouette score: 0.75 **Mechanism** • Distinct degradation pattern • Molecular fingerprint correlation $r > 0.82$ **Group 2A** Compounds: B[*a*]A–D[*ah*]A Similarity: 0.78 Silhouette score: 0.75 **Mechanism** • Similar removal mechanisms • Moderate efficiency **Mixed group** Compounds: ACY–ACE–FLU–ANTH Similarity: clustered together Silhouette score: 0.75 **Mechanism** • Interconnected degradation pathways • Uncertain carcinogenic risks

Table 5 (*Contd.*)

System	Model performance	Feature importance (top 5)	Removal rates	PAH clustering analysis
System 3	MSE: 0.0298 ± 0.0021 R^2: 0.945 ± 0.018 Convergence: 800 epochs Silhouette width: 0.76	(1) Pyrene (PYR) (2) Acenaphthylene (ACY) (3) Fe^{2+} concentration (4) H_2O_2 concentration	EDTA (15 g L^{-1}): 94.2 ± 2.1% EDTA (20 g L^{-1}): 96.4 ± 1.8% Fe^{2+} (0.030 M): 95.1 ± 2.0% H_2O_2 (0.349 M): 92.8 ± 2.4% H_2O_2 (0.419 M): 95.1 ± 2.0%	**Distinct cluster** Compounds: B[*a*]A–ANTH Similarity: 0.83 Silhouette score: 0.76 **Mechanism** • Enhanced removal under high EDTA • Molecular fingerprint correlation $r > 0.88$ **Linked compounds** Compounds: NAP–CHRY, B[*k*]F–IND Similarity: 0.79 Silhouette score: 0.76 **Mechanism** • Consistent degradation patterns • Strong correlation with operational parameters **Mixed group** Compounds: FLU, PYR, B[*ghi*]P Similarity: grouped together Silhouette score: 0.76 **Mechanism** • Similar degradation patterns • Concentration-dependent efficiency
System 4	MSE: 0.0325 ± 0.0023 R^2: 0.928 ± 0.019 Convergence: 800 epochs Silhouette width: 0.74	(1) Pyrene (PYR) (2) Acenaphthylene (ACY) (3) Fe^{2+} concentration (4) H_2O_2 concentration	EDTA (15 g L^{-1}): 90.5 ± 2.6% EDTA (20 g L^{-1}): 93.2 ± 2.2% Fe^{2+} (0.030 M): 91.8 ± 2.4% H_2O_2 (0.349 M): 89.8 ± 2.7% H_2O_2 (0.419 M): 91.8 ± 2.4%	**Group 2A** Compounds: B[*a*]A–D[*ah*]A Similarity: 0.81 Silhouette score: 0.74 **Mechanism** • Highest removal efficiency • Requires high oxidant concentration • Molecular fingerprint correlation $r > 0.84$ **Mixed group** Compounds: CHRY–IND Similarity: 0.77 Silhouette score: 0.74 **Mechanism** • Moderate removal rate • EDTA-dependent efficiency **Uncertain status** Compounds: FLU, PHEN, PYR, B[*ghi*]P Similarity: clustered together Silhouette score: 0.74 **Mechanism** • Ambiguous carcinogenic status • Similar chemical behaviors

System 4 achieved 92% TPH removal in 30 minutes, showing excellent efficiency for low MW PAHs but moderate performance for high MW PAHs ($R^2 = 0.896$). Statistical analysis showed significant treatment effects ($F(15, 30) = 5.55$, $p < 0.001$). PCA identified combined time/Fe^{2+} effects (PC1: 70.2%) and EDTA concentration (PC2: 19.5%) as key performance factors.

Comparative analysis revealed that three systems achieved equivalent maximum TPH removal (99%), with Systems 1 and 3 demonstrating superior treatment speed. System 3's exceptional performance can be attributed to synergistic interaction between homogeneous $Fe_2(SO_4)_3 \cdot 5H_2O$ and heterogeneous Fe_2O_3 catalysts. Extended PAH-specific degradation analysis and detailed statistical results are provided in SI Section S14.

3.4 Neural network modeling and hierarchical clustering analysis of PAH removal systems

Neural network modeling coupled with hierarchical clustering analysis revealed significant correlations between treatment parameters, PAH characteristics, and removal efficiencies (Table 5) (SI Fig. S5, S6, S8, and S9). System 3 demonstrated the highest predictive accuracy (MSE: 0.0298 ± 0.0021, R^2: 0.945 ± 0.018), followed by System 2 (R^2: 0.934 ± 0.021), System 4 (R^2: 0.928 ± 0.019), and System 1 (R^2: 0.912 ± 0.024).

Feature importance analysis highlighted pyrene as consistently influential across all systems. Treatment parameters, particularly EDTA and H_2O_2 concentrations, were consistently among the top five features. System 3 demonstrated superior performance at both EDTA concentrations tested, achieving 94.2 ± 2.1% removal at 15 g L^{-1} and 96.4 ± 1.8% at 20 g L^{-1}. The superior performance of System 3 can be attributed to synergistic interactions between the dual catalysts operating through complementary mechanisms. The soluble Fe^{3+} from $Fe_2(SO_4)_3 \cdot 5H_2O$ enables rapid homogeneous Fenton chemistry, while Fe_2O_3 provides a stable heterogeneous surface for sustained H_2O_2 activation at near-neutral pH conditions where homogeneous catalysts lose activity.[61] This dual-pathway approach ensures continuous radical generation through both solution-phase and surface-mediated mechanisms. Furthermore, EDTA chelation maintains iron solubility and forms photoactive Fe(III)–EDTA complexes that undergo rapid photoreduction under UV irradiation, accelerating the rate-limiting Fe^{3+}/Fe^{2+} cycle.[21] The sulfate moiety may additionally contribute sulfate radicals ($SO_4^{\bullet-}$, $E^\circ = 2.5–3.1$ V), which exhibit higher selectivity for electron-rich aromatic compounds like PAHs compared to $^\bullet OH$ alone.[62] This synergistic enhancement between the dual iron catalysts aligns with recent advances in iron-promoted oxidation chemistry, where cooperative interactions between iron and complementary elements have been demonstrated to significantly improve catalytic efficiency for PAH degradation, including anthracene oxidation.[63]

Hierarchical clustering analysis revealed distinct patterns in PAH behavior across treatment systems. System 1 demonstrated clear clustering based on carcinogenic potential, with known carcinogens (NAP–CHRY) grouping together with high similarity (0.85). Probable carcinogens (Group 2A: B[*a*]A and D[*ah*]A), referenced by studies, formed a distinct cluster. System 2 identified B[*a*]P as forming its own unique cluster with similarity index of 0.81, while ACY–ACE–FLU–ANTH compounds clustered together, showing interconnected degradation pathways and uncertain carcinogenic risks (SI Fig. S6). System 3 showed the most refined clustering with B[*a*]A–ANTH forming a distinct group with similarity of 0.83, NAP–CHRY and B[*k*]F–IND compounds showing similarity of 0.79, while FLU, PYR, and B[*ghi*]P clustered together. System 4 revealed B[*a*]A–D[*ah*]A clustering with similarity of 0.81, CHRY–IND forming a distinct cluster with similarity of 0.77, while FLU, PHEN, PYR, and B[*ghi*]P clustered together, exhibiting ambiguous carcinogenic status and similar chemical behaviors as supported by reference risks (SI Fig. S9). The observed clustering patterns correlate strongly with molecular topology and electronic properties rather than simply molecular weight or ring number.

System 3 emerged as the most effective treatment approach, combining highest predictive accuracy with well-defined clustering and superior removal efficiencies. These results demonstrate that treatment system optimization must consider both operational parameters and PAH-specific characteristics with clearer correlation analysis. Complete molecular fingerprint analysis, detailed clustering results, and extended statistical evaluations are provided in SI Section S15.

4. Conclusion

This comprehensive investigation demonstrates the effectiveness of iron(III) sulfate pentahydrate ($Fe_2(SO_4)_3 \cdot 5H_2O$) in advanced Fenton and photo-Fenton systems for oil-contaminated soil remediation, with particular emphasis on its novel synergistic combination with Fe_2O_3 and the application of AI-based optimization for treatment parameter identification. Through rigorous mechanistic analysis, multivariate statistical evaluation, and advanced machine learning techniques, we uncovered important relationships between catalyst configurations and remediation performance that advance the field's theoretical understanding and practical applications.

Our research demonstrates that the multi-catalyst photo-Fenton system (System 3) provides significant improvements in remediation technology, achieving 99% TPH removal within 10 minutes while simultaneously delivering superior degradation of recalcitrant high molecular weight PAHs (>98% removal efficiency). This notable performance is supported by strong statistical reliability metrics (MSE: 0.0298 ± 0.0021, R^2: 0.945 ± 0.018, silhouette width: 0.76), validating its effectiveness across diverse contamination scenarios. Synergistic combination of $Fe_2(SO_4)_3 \cdot 5H_2O$ and Fe_2O_3 catalysts creates a multi-phase reaction system that overcomes the limitations of single-catalyst approaches, offering enhanced performance under environmentally relevant conditions (pH 6.5) and minimizing the formation of iron sludge waste.

Neural network modeling revealed valuable insights into structure–activity relationships, with System 3 demonstrating well-defined PAH clustering patterns (B[*a*]A–ANTH forming a distinct group with similarity of 0.83) and robust molecular fingerprint correlations (r > 0.88). These findings establish quantifiable connections between molecular structure and degradation mechanisms in $Fe_2(SO_4)_3 \cdot 5H_2O$-catalyzed systems, providing useful biomarkers for monitoring remediation effectiveness and optimizing treatment parameters.

Principal component analysis identified treatment time as the primary performance driver across all systems (PC1: 68.5–75.6%), while feature importance analysis highlighted pyrene as a consistently influential compound with significant predictive value. The optimal balance of catalysts in System 3 created synergistic radical generation pathways that maintained effective oxidative capacity throughout the treatment process, addressing limitations of single-catalyst approaches for complex contaminant mixtures. From a materials chemistry perspective, this study demonstrates that rationally designed catalyst systems can dramatically enhance treatment efficiency through strategic manipulation of surface properties, redox potentials, and coordinated electronic environments. The integration of $Fe_2(SO_4)_3\cdot 5H_2O$ with Fe_2O_3 in an AI-optimized framework represents a significant advancement in applying established Fenton catalysts to soil remediation challenges. While Fe(III) sulfate has been utilized in wastewater treatment, its combination with heterogeneous iron oxide catalysts and neural network-guided optimization provides optimal catalytic activity through complementary reaction pathways, offering significant practical advantages including cost-effectiveness, environmental compatibility, and operational stability under near-neutral conditions.

This research contributes to the mechanistic understanding of catalyst-driven remediation processes and provides a foundation for designing improved treatment technologies. The enhanced performance of the $Fe_2(SO_4)_3\cdot 5H_2O$-based systems, particularly when combined with complementary catalysts, represents a valuable advancement in addressing persistent organic pollutants in soil environments. These findings have important implications for improving remediation efficiency, reducing treatment times, and enhancing the degradation of carcinogenic compounds in environmental restoration efforts.

This study has certain limitations that warrant acknowledgment. The proposed sulfate radical involvement was inferred from literature precedent rather than directly verified through EPR spectroscopy or radical quenching experiments. Additionally, while sludge formation was observed during treatment, quantitative assessment of Fe leaching and catalyst reusability were not performed. Future research should employ spin-trapping methods and selective scavenger studies to quantify the relative contributions of $^{\bullet}OH$ and $SO_4^{\bullet-}$ to PAH degradation, alongside comprehensive evaluation of Fe leaching kinetics and catalyst recovery for environmental safety assessment. Scale-up validation, techno-economic analysis, and life cycle assessment will be essential to evaluate practical implementation potential for large-scale environmental restoration projects. Although this study focused on petroleum hydrocarbons in desert soils, Fenton-based processes have demonstrated effectiveness for diverse organic contaminants including pesticides, polychlorinated biphenyls (PCBs), and chlorinated solvents across various soil matrices.[14,64] The fundamental chemistry and machine learning framework developed here are therefore potentially transferable to other contamination scenarios, provided site-specific calibration is performed. The use of commercially available, low-cost reagents further supports economic feasibility for field-scale applications. The promising results achieved in this study offer new pathways for developing more effective, sustainable solutions to address petroleum hydrocarbon contamination in soil environments worldwide.

Author contributions

Zainab Ashkanani: conceptualization, data curation, formal analysis, investigation, methodology, resources, software, supervision, validation, visualization, writing – original draft, writing – review & editing. Rabi Mohtar: conceptualization, methodology, project administration, resources, supervision, validation, writing – original draft, writing – review & editing. Salah Al-Enezi: funding acquisition, methodology, resources, supervision, writing – review & editing. Faten Khalil: formal analysis, funding acquisition, resources, supervision. Muthanna Al-Momin: project administration, resources, supervision. Xingmao Ma: validation, writing – review & editing. Patricia K. Smith: supervision, validation, writing – review & editing. Salvatore Calabrese: supervision, validation, writing – review & editing. Meshal Abdullah: validation, writing – review & editing. Najeeb Aladwani: formal analysis, investigation, methodology, resources, validation, visualization, writing – review & editing.

Conflicts of interest

The authors declare no conflict of interest.

Data availability

The code for neural network modeling and PAH degradation analysis can be found at https://github.com/SAAAHco/PhotoFenton-PAH-Remediation-ML. The repository includes all Python scripts for molecular descriptor calculations, neural network architectures, hierarchical clustering analysis, and visualization tools employed in this study. Additional experimental data, including TPH and PAH concentration measurements, treatment parameters, and statistical analyses have been included as part of the supplementary information (SI). Supplementary information: extended theoretical frameworks, detailed methodological protocols, and additional experimental results supporting the main manuscript. Specifically, it includes: (1) theoretical framework (Sections S1–S5) providing extended discussion on advanced oxidation processes, catalytic mechanisms of $Fe_2(SO_4)_3\cdot 5H_2O$ and Fe_2O_3, the role of EDTA chelation, and the rebound effect in Fenton-like systems; (2) materials and methods (Sections S6–S11) detailing geophysical site characterization protocols, sample preparation procedures, complete analytical methods for TPH and PAH analysis (GC-FID and GC-MS), quality control procedures, and comprehensive neural network architecture specifications including molecular descriptor calculations and machine learning validation procedures; (3) results (Sections S12–S17) presenting electrical resistivity survey data revealing subsurface contamination architecture, extended catalyst performance analysis, PAH-specific degradation kinetics for all four treatment systems, and molecular fingerprint correlations. Supporting tables include soil sample characteristics (Table S1), electrical resistivity measurements at 15 VES locations (Table S2), and comprehensive molecular descriptor analysis (Table S3). Supporting figures include site location maps (Fig. S1), resistivity cross-sections and depth-slice maps (Fig. S2–S4), and complete analysis results for advanced Fenton-photo systems 2–4 including TPH removal efficiency, PCA plots, feature importance analysis, violin plots, and hierarchical clustering dendrograms (Fig. S5–S10). See DOI: https://doi.org/10.1039/d5ra07305g.

Acknowledgements

This research was supported by the Scientific Research – Petroleum Research Center (KISR-PRC), the National Unit for Environmental Research and Services at Kuwait University, and the Kuwait Environmental Remediation Program (KERP) at Kuwait Oil Company (KOC). We extend our gratitude to Prof. Abdulazim Marafi, Director of Operations at KISR-PRC and to Ronia Jose and Fousiya Majeed from KISR for their invaluable collaboration in providing essential resources and laboratory space necessary for our experiments. We wish to thank Nada Al-Qallaf, Team Leader at KOC, for her crucial assistance in facilitating access to the Greater Burgan Oil Field and supporting the sample collection process. Our collaboration with the National Unit for Environmental Research and Services at the Research Sector of Kuwait University significantly enriched this study. We acknowledge the substantial contributions and access to laboratory facilities provided by Prof. Abdullah Alhendal, Alia Awan, Akhila Thomas, Melba Kishor, Milu Jose, and Alberto Orayan. Their support greatly enhanced our research capabilities and overall study outcomes.

References

1 S. Kuppusamy, *et al.*, An overview of total petroleum hydrocarbons, *Total petroleum hydrocarbons*, 2020, pp. 1–27.
2 M. Čolnik, *et al.*, Hydrothermal decomposition of polyethylene waste to hydrocarbons rich oil, *J. Supercrit. Fluids*, 2021, **169**, 105136.
3 D. Brion and É. Pelletier, Modelling PAHs adsorption and sequestration in freshwater and marine sediments, *Chemosphere*, 2005, **61**(6), 867–876.
4 N. Raza, *et al.*, Aromatic hydrocarbons in air, water, and soil: Sampling and pretreatment techniques, *TrAC, Trends Anal. Chem.*, 2018, **103**, 56–73.
5 A. B. Patel, *et al.*, Polycyclic Aromatic Hydrocarbons: Sources, Toxicity and Remediation Approaches, *Front. Microbiol.*, 2020, **11**, 2675.
6 Y. Liu, *et al.*, Distribution and sources of polycyclic aromatic hydrocarbons in surface sediments of rivers and an estuary in Shanghai, China, *Environ. Pollut.*, 2008, **154**(2), 298–305.
7 S. L. Simonich and R. A. Hites, Organic pollutant accumulation in vegetation, *Environ. Sci. Technol.*, 1995, **29**(12), 2905–2914.
8 R. Xu and J. P. Obbard, Biodegradation of polycyclic aromatic hydrocarbons in oil-contaminated beach sediments treated with nutrient amendments, *J. Environ. Qual.*, 2004, **33**(3), 861–867.
9 H. I. Abdel-Shafy and M. S. Mansour, A review on polycyclic aromatic hydrocarbons: source, environmental impact, effect on human health and remediation, *Egypt. J. Pet.*, 2016, **25**(1), 107–123.
10 D. M. Brown, *et al.*, Heavy hydrocarbon fate and transport in the environment, *Q. J. Eng. Geol. Hydrogeol.*, 2017, **50**(3), 333–346.
11 P. Logeshwaran, *et al.*, Petroleum hydrocarbons (PH) in groundwater aquifers: An overview of environmental fate, toxicity, microbial degradation and risk-based remediation approaches, *Environ. Technol. Innovation*, 2018, **10**, 175–193.
12 W. H. Glaze, J.-W. Kang and D. H. Chapin, The chemistry of water treatment processes involving ozone, hydrogen peroxide and ultraviolet radiation, *Ozone: Sci. Eng.*, 1987, **9**(4), 335–352.
13 Z. Ashkanani, *et al.*, AI-assisted systematic review on remediation of contaminated soils with PAHs and heavy metals, *J. Hazard. Mater.*, 2024, 133813.
14 M. Cheng, *et al.*, Hydroxyl radicals based advanced oxidation processes (AOPs) for remediation of soils contaminated with organic compounds: a review, *Chem. Eng. J.*, 2016, **284**, 582–598.
15 J. J. Pignatello and S. L. Nason, Importance of soil properties and processes on bioavailability of organic compounds, *Bioavailability of Organic Chemicals in Soil and Sediment*, 2020, pp. 7–41.
16 R. Wuana, P. Adie and N. A. Immaculate, *Seasonal Variation in Bioavailability of Some Toxic Metals in Waste Dump Soils of Makurdi*, North-Central Nigeria, 2012.
17 C. Min, *et al.*, Hydroxyl radicals based advanced oxidation processes (AOPs) for remediation of soils contaminated with organic compounds: A review, *Chem. Eng. J.*, 2016, **284**, 582–598.
18 X. Chen, D. Hulbert and P. B. Shepson, Measurement of the organic nitrate yield from OH reaction with isoprene, *J. Geophys. Res.:Atmos.*, 1998, **103**(D19), 25563–25568.
19 E. Veignie, *et al.*, Fenton degradation assisted by cyclodextrins of a high molecular weight polycyclic aromatic hydrocarbon benzo [a] pyrene, *J. Hazard. Mater.*, 2009, **168**(2–3), 1296–1301.
20 S. Gan and H. K. Ng, Inorganic chelated modified-Fenton treatment of polycyclic aromatic hydrocarbon (PAH)-contaminated soils, *Chem. Eng. J.*, 2012, **180**, 1–8.
21 J. J. Pignatello, E. Oliveros and A. MacKay, Advanced oxidation processes for organic contaminant destruction based on the Fenton reaction and related chemistry, *Crit. Rev. Environ. Sci. Technol.*, 2006, **36**(1), 1–84.
22 L. Clarizia, *et al.*, Homogeneous photo-Fenton processes at near neutral pH: a review, *Appl. Catal., B*, 2017, **209**, 358–371.
23 D. Huguenot, *et al.*, Combination of surfactant enhanced soil washing and electro-Fenton process for the treatment of soils contaminated by petroleum hydrocarbons, *J. Environ. Manage.*, 2015, **153**, 40–47.
24 S. Matavos-Aramyan and M. Moussavi, Advances in Fenton and Fenton based oxidation processes for industrial effluent contaminants control-a review, *Int. J. Environ. Sci. Nat. Resour.*, 2017, **2**(4), 1–18.
25 M. Pereira, L. Oliveira and E. Murad, Iron oxide catalysts: Fenton and Fentonlike reactions–a review, *Clay Miner.*, 2012, **47**(3), 285–302.
26 F.-X. Wang, *et al.*, Fe-Cu bimetal metal-organic framework for efficient decontamination via Fenton-like process: Synthesis, performance and mechanism, *J. Colloid Interface Sci.*, 2023, **649**, 384–393.

27 R. Roohi, *et al.*, Application of artificial neural network model for the identification the effect of municipal waste compost and biochar on phytoremediation of contaminated soils, *J. Geochem. Explor.*, 2020, **208**, 106399.

28 K. Gautam, *et al.*, A review on control and abatement of soil pollution by heavy metals: Emphasis on artificial intelligence in recovery of contaminated soil, *Environ. Res.*, 2023, **225**, 115592.

29 C. Lopes, *et al.*, Application of rhamnolipid surfactant for remediation of toxic metals of long-and short-term contamination sites, *Int. J. Environ. Sci. Technol.*, 2021, **18**, 575–588.

30 Y. Yang, *et al.*, Performance prediction model for desalination plants using modified grey wolf optimizer based artificial neural network approach, *Desalin. Water Treat.*, 2024, 100411.

31 Y. Chu, *et al.*, Artificial neural network prediction models of heavy metal polluted soil resistivity, *Eur. J. Environ. Civ. Eng.*, 2021, **25**(9), 1570–1590.

32 K. Padma and K. Don, Artificial neural network applications in analysis of forensic science, *Cyber Security and Digital Forensics*, 2022, pp. 59–72.

33 D. Lakshmi, *et al.*, Artificial intelligence (AI) applications in adsorption of heavy metals using modified biochar, *Sci. Total Environ.*, 2021, **801**, 149623.

34 H. Guo, *et al.*, An ensemble deep neural network approach for predicting TOC concentration in lakes along the middle-lower reaches of Yangtze River, *J. Intell. Fuzzy Syst.*, 2022, **42**(3), 1455–1482.

35 T. Matsui, *et al.*, A natural language processing model for supporting sustainable development goals: translating semantics, visualizing nexus, and connecting stakeholders, *Sustainability Sci.*, 2022, **17**(3), 969–985.

36 T. S. Anggraini, *et al.*, Machine learning-based global air quality index development using remote sensing and ground-based stations, *Environ. Adv.*, 2024, **15**, 100456.

37 R. Olawoyin, Application of backpropagation artificial neural network prediction model for the PAH bioremediation of polluted soil, *Chemosphere*, 2016, **161**, 145–150.

38 B. Anifowose and F. Anifowose, Artificial Intelligence and Machine Learning in environmental impact prediction for soil pollution management–Case for EIA Process, *Environ. Adv.*, 2024, 100554.

39 E. S. Elmolla, M. Chaudhuri and M. M. Eltoukhy, The use of artificial neural network (ANN) for modeling of COD removal from antibiotic aqueous solution by the Fenton process, *J. Hazard. Mater.*, 2010, **179**(1–3), 127–134.

40 B. Sarı, *et al.*, The Utilization and Modeling of Photo-Fenton Process as a Single Unit in Textile Wastewater Treatment, *Clean:Soil, Air, Water*, 2023, **51**(1), 2100328.

41 S. Hernández-García, *et al.*, Deep reinforcement learning for automated search of model parameters: photo-fenton wastewater disinfection case study, *Neural Comput. Appl.*, 2023, **35**(2), 1379–1394.

42 D. Palma, *et al.*, Artificial neural networks applied to photo-Fenton process: An innovative approach to wastewater treatment, *Water Sci. Eng.*, 2025, **18**(3), 324–334.

43 S. Pritsis, *et al.*, The role of hyetograph shape and designer subjectivity in the design of an urban drainage system, *Water Sci. Technol.*, 2024, **90**(3), 920–934.

44 Q. Yuan, *et al.*, Machine Learning-Assisted Catalysts for Advanced Oxidation Processes: Progress, Challenges, and Prospects, *Catalysts*, 2025, **15**(3), 282.

45 H. Yang, *et al.*, Predicting heavy metal adsorption on soil with machine learning and mapping global distribution of soil adsorption capacities, *Environ. Sci. Technol.*, 2021, **55**(20), 14316–14328.

46 E. A. Atekwana and E. A. Atekwana, Geophysical signatures of microbial activity at hydrocarbon contaminated sites: a review, *Surv. Geophys.*, 2010, **31**(2), 247–283.

47 A. Revil, *et al.*, Understanding biogeobatteries: Where geophysics meets microbiology, *J. Geophys. Res.:Biogeosci.*, 2010, **115**(G1), G00G02.

48 W. A. Sauck, A model for the resistivity structure of LNAPL plumes and their environs in sandy sediments, *J. Appl. Geophys.*, 2000, **44**(2–3), 151–165.

49 D. D. Werkema Jr, *et al.*, Investigating the geoelectrical response of hydrocarbon contamination undergoing biodegradation, *Geophys. Res. Lett.*, 2003, **30**(12), 1647.

50 X. Kang, *et al.*, Hydrogeophysical characterization of nonstationary DNAPL source zones by integrating a convolutional variational autoencoder and ensemble smoother, *Water Resour. Res.*, 2021, **57**(2), e2020WR028538.

51 J. Heenan, *et al.*, Electrical resistivity imaging for long-term autonomous monitoring of hydrocarbon degradation: Lessons from the Deepwater Horizon oil spill, *Geophysics*, 2015, **80**(1), B1–B11.

52 E. A. Atekwana, W. A. Sauck and D. D. Werkema Jr, Investigations of geoelectrical signatures at a hydrocarbon contaminated site, *J. Appl. Geophys.*, 2000, **44**(2–3), 167–180.

53 F. M. Mewafy, *et al.*, Evidence that bio-metallic mineral precipitation enhances the complex conductivity response at a hydrocarbon contaminated site, *J. Appl. Geophys.*, 2013, **98**, 113–123.

54 M. Al-Sarawi, M. Massoud and F. Al-Abdali, Preliminary assessment of oil contamination levels in soils contaminated with oil lakes in the greater Burgan oil fields, Kuwait, *Water, Air, Soil Pollut.*, 1998, **106**(3–4), 493–504.

55 K. Hirschmann, *The Kuwaiti Oil Fires*, Infobase Publishing, 2009.

56 UNCC, *Report and Recommendations Made by the Panel of Commissioners Concerning the Third Instalment of "F4": Claims*, United Nations Compensation Commission Governing Council Kuwait, UNCC, Kuwait, 2002.

57 G. Brown and S. Porembski, Flora and vegetational aspects of miniature dunes in a sand-depleted Haloxylon salicornicum community in the Kuwait desert, *Flora*, 1998, **193**(2), 133–140.

58 G. Brown and S. Porembski, Phytogenic hillocks and blow-outs as' safe sites' for plants in an oil-contaminated area of northern Kuwait, *Environ. Conserv.*, 2000, 242–249.

59 Commission UNC, *Follow-up Programme for Environmental Awards*, State of Kuwait, 2013.

60 S. S. Al-Ateeqi, *Phytoremediation of Oil-Polluted Desert Soil in Kuwait Using Native Plant Species*, University of Glasgow, 2014.

61 J. He, *et al.*, Interfacial mechanisms of heterogeneous Fenton reactions catalyzed by iron-based materials: A review, *J. Environ. Sci.*, 2016, **39**, 97–109.

62 L. W. Matzek and K. E. Carter, Activated persulfate for organic chemical degradation: A review, *Chemosphere*, 2016, **151**, 178–188.

63 X. Li, *et al.*, A concise synthesis of Se/Fe materials for catalytic oxidation reactions of anthracene and polyene, *Chin. Chem. Lett.*, 2025, **36**(3), 110036.

64 M. Usman, *et al.*, Fenton oxidation for soil remediation: A critical review of observations in historically contaminated soils, *J. Hazard. Mater.*, 2022, **424**(Pt D), 127670.