

RESEARCH

Open Access



Volatilization of benzene on soil surface under different factors: evaluation and modeling

Qian Wang^{1,2,3}, Jianmin Bian^{2,3*} and Dongmei Ruan^{2,3}

Abstract

The volatilization of volatile organic compounds following a leakage event is a crucial mechanism that influences their migration and transformation in the soil. It is noteworthy that this process is intricately shaped by soil properties and environmental factors, exhibiting highly complex nonlinear relationships. However, there is currently no reliable mathematical model to predict the nonlinear relationship. To address this gap, the study conducted dynamic experiments considering various factors, including particle size, organic matter content, temperature, wind speed and moisture content. The volatilization rate (k), an important parameter in volatilization kinetics reflecting the speed of volatilization, was calculated by first-order kinetic principle. Finally, an innovative approach was introduced using a Back Propagation Neural Network (BPNN) model for prediction. The findings indicate that wind speed exerts the most significant impact on the volatilization rate of benzene among the examined factors. The application of BPNN demonstrates the model's accuracy in simulating benzene volatilization rates under diverse conditions. The results of K-fold cross-validation alleviate concerns of potential over-prediction, affirming the reliability of the constructed model. This research introduces a novel methodology for predicting volatilization parameters in real-world scenarios.

Highlights

: Five factors including particle size, organic matter content, temperature, wind speed and moisture content were considered.

: Wind speed has the most significant impact on the volatilization of benzene among five factors.

: A Back Propagation Neural Network model was successfully implemented to predict the volatilization rates of benzene.

Keywords Volatilization of benzene, Soil properties, Environmental influential factors, Back Propagation, K-fold cross-validation

*Correspondence:

Jianmin Bian
bianjm@jlu.edu.cn

¹ School of Environmental Engineering, Xuzhou University of Technology, Xuzhou 221018, China

² Key Laboratory of Groundwater Resources and Environment (Ministry of Education), Jilin University, Changchun 130021, China

³ College of New Energy and Environment, Jilin University, Changchun 130021, China

1 Introduction

In various industries, such as chemical manufacturing, metal smelting, and petroleum processing, the inadvertent release of volatile organic compounds (VOCs) poses significant threats to air, soil, and groundwater [1–3]. Due to their high vapor pressure, low dissolution, and high coefficient of diffusion, the volatilization process emerges as a crucial mechanism determining



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/>.

the fate of VOCs [4, 5]. Vapor retardation through volatilization can decelerate migration and reduce the total contaminant mass in soils, thereby influencing the distribution characteristics of contaminants in the unsaturated zone and groundwater. Consequently, understanding the volatilization process of VOCs on the soil surface holds significant scientific importance in elucidating environmental pollution mechanisms and aiding in the remediation of petrochemical-contaminated fields.

Previous research has demonstrated that the volatilization is influenced by the type of VOCs [6]. Specifically, for different BTEX compounds, the volatilization rate increases with the saturation vapor pressure (10, 3.8, 1.26 and 0.43 kPa for benzene, toluene, xylene, and ethylbenzene, respectively) [7]. Additionally, soil properties and environmental conditions [4, 8–10] are crucial factors influencing the volatilization. For example, Lu et al. [4] investigated the volatilization of toluene in soil through batch experiments under varying moisture content conditions. The results indicated that an increase in moisture content promoted the volatilization of toluene under low moisture conditions, while under high moisture conditions, an increase in moisture content reduced the volatilization of toluene. Indoor simulation experiments conducted by [10] indicated that the increasing wind speed accelerated the volatilization speed of diesel oil. The study of Zhang et al. [9] showed the three most important factors influencing the soil vapor extraction remediation efficiency were the length of time, pollutant type, and temperature.

From above researches, it has been found that there is a non-linear relationship between volatilization and influencing factors, introducing challenges in predicting the volatilization parameters under varying conditions. From 1980s, the artificial neural network (ANN) has garnered widespread attention for its exceptional capabilities in addressing complex nonlinear problems across various engineering disciplines [9, 11–16]. For instance, Qiao et al. [15] demonstrated the effectiveness of the generalized regression neural network with K-fold cross validation (K-CV) in accurately predicting the mechanical properties of hypereutectoid steels; Ding et al. [16] optimized the water quality index assessment model using the combined weights based on machine learning and game theory; Wang et al. [14] successfully employed the back propagation neural network (BPNN) to estimate the benzene adsorption mass on three soils considering five factors. While ANN has been proven to perform well in simulating the nonlinear problem associated with different factors, within our search scope, no relevant research has explored the application of ANN on the volatilization of VOCs.

This study focused on volatilization characteristics of benzene on three types of soils: silt loam, loam, and sandy loam. Comparative experiments were conducted to analyze the volatilization characteristics of benzene under varying conditions involving temperature, wind speed, and media moisture content in three soils. Subsequently, the volatilization models were employed to fit experimental data, determining benzene volatilization rates. Then the impact of different factors on the volatilization process was analyzed, identifying the primary controlling factor in the volatilization process. Furthermore, the BPNN model was applied to predict the volatilization rates on different factors. Also the K-CV was applied to validate the BPNN model reliability. These findings contribute to a deeper understanding of the migration behavior of VOCs on soils. Moreover, the enhanced prediction accuracy of volatilization parameters provides valuable insights for the development of actual multiphase and multiscale models.

2 Materials and methods

2.1 Study area

The study area is situated in a chemical industrial park in northeast China, as shown in Fig. S1. This region is a significant industrial base with over 240 chemical enterprises specializing in new chemical materials, petrochemicals, and fine chemicals. The climate in the study area is classified as the continental monsoon climate of the northern temperate zone. The average annual precipitation is 882 mm, the average annual evaporation is 871 mm, the average temperature is 5.8 °C, and the average annual wind speed is 2.7 m s⁻¹. During the preliminary investigation, a significant explosion accident was identified within the chemical park, resulting in a severe exceedance of benzene homologues' concentrations in groundwater. Notably, benzene, with a concentration surpassing the III standard limit of Chinese Standards for groundwater quality (10 µg L⁻¹) by 7368 times, emerged as a major concern. Due to its small molecular weight, considerable water solubility, and high vapor density, benzene was chosen as the representative in this study. Unpolluted soil samples were collected from three distinct types of soils in the study area, as detailed in Table S1. The testing methodologies were in accordance with those outlined in articles authored by the researcher [14]. The organic matter content (OMC) was determined using potassium dichromate oxidation with external heating. Particle size was measured with a laser particle size analyzer (Bettersize 2000, Dandong Baite Instrument Co., China).

2.2 Experimental apparatus and procedures

The volatilization of organic pollutants can be investigated by the concentration measurement method, which monitors concentration changes of VOCs in soil columns or air by chromatography [17, 18] or portable volatile gas detectors [4]. Another method, weight measurement method [10], is considered a simple and effective experimental technique, involves using various containers such as glass dishes, aluminum boxes, iron cans, and soil columns filled with distilled water and media to simulate water and soil surfaces, respectively. Subsequently, VOCs are introduced into the containers, and the weight of the pollutants is recorded at different time intervals to calculate the losses due to volatilization.

(1) Experimental apparatus

The static volatilization experimental apparatus, as depicted in Fig. 1, comprises three main components. The first component, serving as the core of the experimental setup, includes a portable refrigerator, a thermometer, and an ion fan. The refrigerator regulates the temperature, monitored by the embedded thermometer, while the ion fan controls wind speed. The ion fan is chosen for its small volume, light weight, and easy of placement. The ion fan is placed on a grid support above the soil, allowing the wind to blow directly onto the soil surface. The second component features a weight monitoring device utilizing a precision analytical balance with an accuracy of 0.0001 g. This device is employed to measure losses due to volatilization. The third component is a ventilation chamber that extracts the air inside the cabinet, treats it appropriately, and then disperses it into the atmosphere outside. All the

experiments were conducted in the ventilation chamber to prevent benzene polluting the surrounding air.

(2) Experimental apparatus and procedures

Twenty g sterilized soil were carefully placed into an aluminum box with a height of 3 cm with a diameter of 5 cm. According to the dry density of the soils, the height of silty loam, loam and sandy loam in aluminum box were 0.91, 0.79 and 0.88 cm, respectively. Also, 2 mL benzene was injected into top of soil, with a parallel sample without the addition of benzene to avoid the potential confounding factors. The initial mass of the apparatus was measured, and subsequent mass measurements were conducted at regular intervals (5, 10, 20, 30, 40, 60, 80, 100, and 120 min). The experiment continued until the mass difference between consecutive measurements fell within a range of 0.05 g. The volatilization quantity of organic compounds at different time intervals was calculated. Then various volatilization models were applied to fit the experimental data and the volatilization rates were determined.

Three primary influencing factors of temperature, wind speed, and moisture content were selected to explore the dynamic characteristics of volatilization under different conditions. Controlled temperatures of 10, 15, 20, 25, and 30 °C were achieved using a refrigerator for precise temperature regulation. An ion fan was utilized to adjust wind speed, with airflow rates at 2.6, 3.1, and 3.6 m³ min⁻¹. For moisture content, 20 g sterilized soil, initially dried, were spread evenly on a plastic wrap. Moisture levels of 10, 20, and 30% were achieved by uniformly adding 2, 4, and 6 mL of ultra-pure water, respectively. The soil samples were enclosed in plastic wrap and allowed to stand for 12 h to ensure uniform

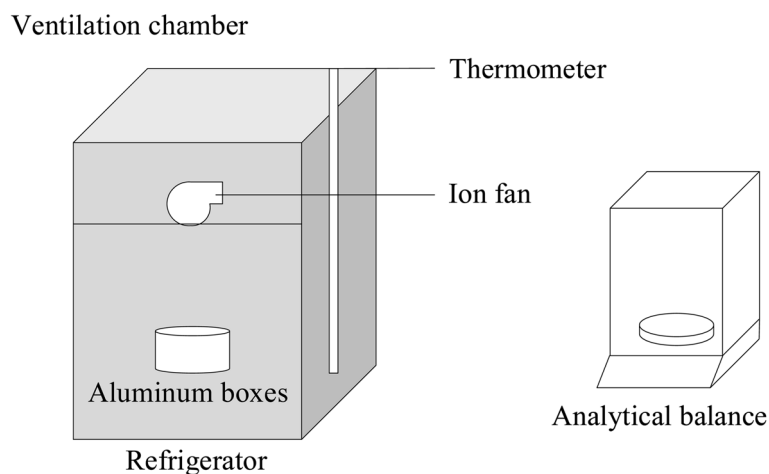


Fig. 1 Experimental apparatus for volatilization

water distribution. All experiments were repeated three times, with the average value being recorded.

2.3 Prediction of volatilization rates under different influential factors

The BPNN algorithm, serving as one typical ANN method, possesses the excellent mechanism of error feedback optimization, thus exhibits outstanding capabilities in mapping multidimensional functions and adeptly handles complex regression problems [19, 20]. The BPNN network structure comprises three layers: input, hidden, and output [21]. Notably, the hidden layer is particularly effective in addressing nonlinear problems. Feature data from the dataset enters the network through the input layer and propagates to the hidden layer. The hidden layer, functioning as a neural network's black box, consolidates various function capabilities for handling complex problems. The output layer receives the processed data from the hidden layer and adjusts its internal parameters (weights and biases) to minimize the difference between the predicted and actual output.

A three-layer BPNN model was developed using Python. The input layer includes temperature, wind speed, moisture content, median particle size (D50), and organic matter content. Notably, the study emphasizes the influence of media on volatility by considering median particle size (D50) and OMC as key descriptors. The volatilization rates were served as the output. The dataset of 33 samples were randomly divided into three parts: 70% allocated for training, 15% for validation samples, and 15% for testing purposes.

Given the limited dataset size in the volatilization experiment, the potential risk of overfitting on training data exists when employing the BPNN model [22]. Overfitting occurs when the model fits the training data well but encounters challenges in generalizing to data outside the training set. To address this concern, cross-validation is employed as a strategy to assess and validate the performance of the constructed BPNN model, thereby mitigating the risk of overfitting [23].

K-CV is a popular resampling method in cross-validation [24]. The process involves three steps. Firstly, the dataset D is randomly partitioned into m folds of approximately equal size. Then, one fold is designated for testing, while the remaining $m-1$ folds are donated for training. This process is repeated m times in total. Finally, the cross-validation error is calculated as the average of the mean squared error (MSE) across all iterations. In the context of this study, the value of m is set to 4, taking into account the dimensions of the dataset and the availability of computational resources.

To evaluate the performance of the machine learning models, the correlation coefficient (R) and MSE were employed as assessment metrics for the BPNN model.

$$R = \frac{\sum_{i=1}^n (O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})^2 \sum_{i=1}^n (P_i - \bar{P})^2}} \quad (1)$$

$$MSE = \frac{\sum_{i=1}^n (P_i - O_i)^2}{n - 1} \quad (2)$$

where n is the total number of samples; O_i denotes the i th experimental value of the volatilization rate; P_i represent the i th modeling value of the volatilization rate by BPNN; \bar{O} and \bar{P} represent the average values of the experimental and simulated volatilization rates for the total samples, respectively.

3 Results and discussion

3.1 Effects of soil properties on volatilization

The volatilization kinetics results for benzene on silty loam, loam and sandy loam are presented in Fig. 2. As depicted in the curves, the residual mass of benzene on the three soils decreases over time, and the declining rate decreases with time until reaching a stable state. Under static water conditions, the volatilization loss of benzene is directly proportional to time [7]. In porous media, during the initial stages of volatilization, the volatilization loss was approximately proportional to time, similar to the trend in static water condition. However, as time progressed, benzene infiltrated into the soil, and the volatilization process was influenced by media factors such as dispersion, surface adsorption, and gas diffusion channels. Therefore, the examination of volatilization loss in relation to time reveals a non-linear correlation.

Various equations have been employed in previous research to express this non-linear correlation, including first-order kinetic principle [7], logarithmic principle [6], and parabolic equation [10]. These equations were employed to fit the volatilization data. Among these models, the first-order kinetic model (Eq. (3)) provided the best fit to the data, as illustrated in Fig. 2. The fitting parameters are detailed in Table 1.

$$m_t = m_0 e^{-kt} \quad (3)$$

where m_0 is the initial mass of pollutant; m_t is the residual mass at t time; and k is the rate constant for benzene volatilization.

Among the three soils, the stable times of silty loam, loam and sandy loam are 200, 160 and 100 min, respectively. The volatilization rate of silty loam, loam, and sandy loam are 0.21, 0.22 and 0.26 min^{-1} . It is found that the volatilization rate decreases with a decrease in the

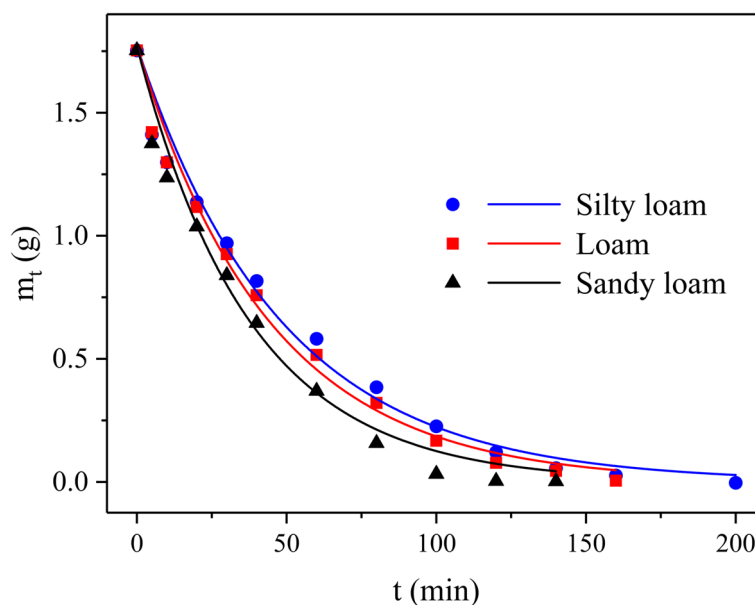


Fig. 2 The volatilization kinetics results for benzene on silty loam, loam and sandy loam incorporating experimental data points and the corresponding model-fitted curve

Table 1 The parameters of the kinetic models to fit the volatilization data

Soils	k (min^{-1})	R^2
Silty loam	0.021	0.982
Loam	0.022	0.986
Sandy loam	0.026	0.980

particle size. The conclusion is consistent with the result reported by Tong et al. [7]. As the particle size decreases, the soil tends to have larger specific surface areas, smaller porosity, and higher clay content. The results in increased benzene entrapment and a reduction in effective diffusion channels for gas, thus hindering volatilization.

3.2 Effects of temperature on volatilization

The curves depicting the residual mass of benzene in three soils over time under varying temperature conditions are shown in Fig. 3. Meanwhile, the fitting results of the corresponding kinetic models are presented in Table S2. Analysis of Fig. 3 reveals that the volatilization rate of benzene in soil increases with the rising temperatures. This phenomenon can be explained by two reasons. Firstly, the saturation vapor pressure of benzene is increased with the rise in temperature [25], making it easier for benzene to volatilize into the air. Secondly, elevated temperature reduces the viscosity of the non-aqueous phase [26], thereby enhancing the mobility of benzene in the soil.

As the temperature escalated from 10 to 30 °C, the volatilization equilibrium time of silty loam, loam and sandy loam decreased from 240 to 80 min, 200 to 80 min and 140 to 60 min, respectively. The volatilization rate of silty loam, loam, and sandy loam increased from 0.015 to 0.037 min^{-1} , 0.015 to 0.037 min^{-1} , and 0.017 to 0.049 min^{-1} , respectively. The variations in volatilization equilibrium time for silty loam, loam, and sandy loam are 160, 120, and 80 min, respectively. Silty loam has a greater duration than loam and sandy loam. The corresponding increases in volatilization rates are 0.022, 0.022, and 0.032 min^{-1} , with sandy loam exhibiting a higher increment than loam and sandy loam. This suggests that temperature exerts a more significant impact on soils with larger particle sizes.

3.3 Effects of wind speed on volatilization

In Fig. 4, a comparison between scenarios with no wind and those with wind (at three wind speeds: 2.6, 3.1, and 3.6 $\text{m}^3 \text{min}^{-1}$) reveals significant variations in the volatilization loss on soils. Meanwhile, the fitting results of the corresponding kinetic models on different wind speeds are presented in Table S3. The presence of wind significantly increases the volatilization rate of benzene compared to the scenario without wind, as depicted in Fig. 4. According to boundary-layer theory, volatile organic compounds traverse the air boundary layer from soil surface to the turbulent edge, then rapidly leave the soil surface with the turbulent flow. The increase in wind speed enhances the volatile of pollutant molecules from the

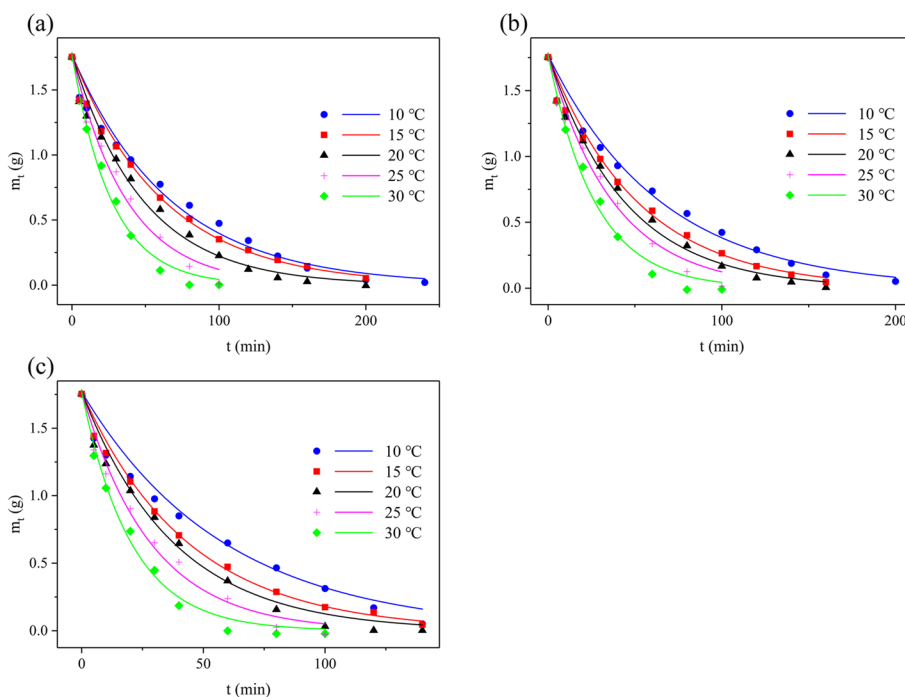


Fig. 3 The curves of the residual mass of benzene with time on different temperatures in three soils of (a) silty loam, (b) loam and (c) sandy loam

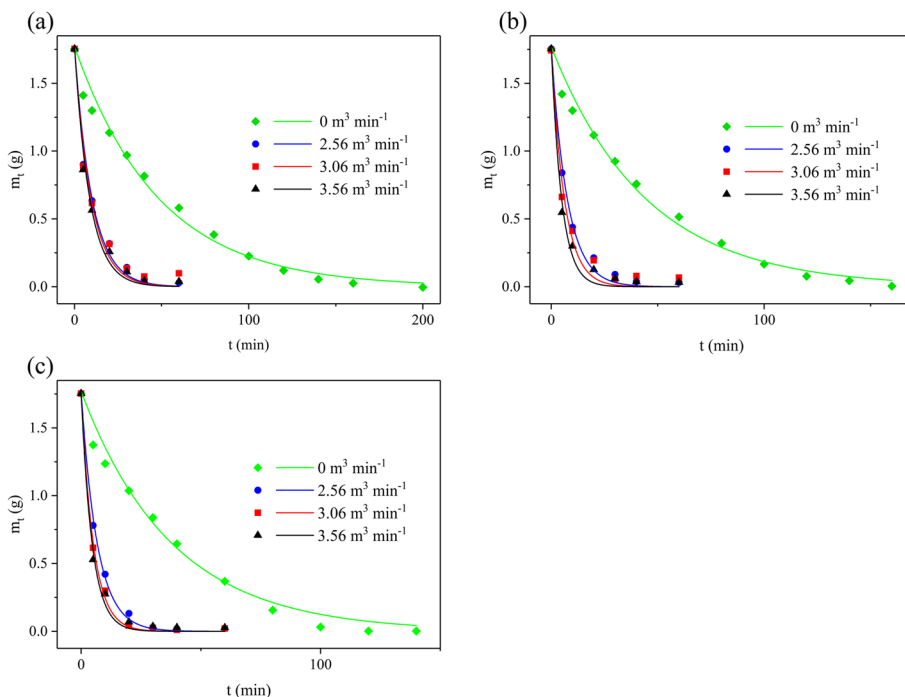


Fig. 4 The curves of the residual mass of benzene with time on different wind speeds in three soils of (a) silty loam, (b) loam and (c) sandy loam

boundary layer to the air, while not significantly affecting the kinetic energy of organic molecules on the solid surface [10]. The volatilization at the solid surface quickly

reaches equilibrium in wind speed of $2.6 \text{ m}^3 \text{ min}^{-1}$. Consequently, further increases in wind speed result in minimal changes in the volatilization rate of pollutants.

Moreover, under the two scenarios for three soils, there is a substantial difference in pollutant loss in the early-stage, while the trends converge in the later stage. In the early stages of volatilization, benzene have not yet infiltrated into soil, the rate of air flow determines the thickness of the boundary layer. With higher wind speeds, there is a thinner boundary layer and an increased concentration gradient within the boundary layer, enhancing diffusion capacity and consequently increasing the volatilization rate.

In the later stages of volatilization, benzene infiltrates into the soil, and the volatilization above the soil surface is weakened, and thus the volatilization rate is minimally controlled by wind speed.

For silty loam, loam, and sandy loam, the volatilization equilibrium time decreases from 200 to 30 min, 160 to 30 min, and 100 to 20 min, respectively. The volatilization rate increases from 0.019 to 0.11 min⁻¹, 0.022 to 0.21 min⁻¹, and 0.026 to 0.22 min⁻¹. The variations in volatilization equilibrium time for silty loam, loam, and sandy loam are 170, 130, and 80 min, respectively, and the corresponding increases in volatilization rates are 0.09, 0.19, and 0.19 min⁻¹. It is indicated that wind speed exerts a more significant impact on soils with larger particle sizes.

3.4 Effects of moisture content on volatilization

In experiments under varying moisture conditions for benzene volatilization, a blank control group was established to calculate the potential mass losses due to water evaporation. The equation for calculating the mass loss of benzene is as follows: where m_t is the residual mass of pollutant at t time; m_{tk} is the loss caused by water evaporation; m_{t1} is the measured mass of pollutant at t time.

$$m_t = m_{t1} + m_{tk} \tag{4}$$

Calculating the actual volatilization of benzene under different moisture conditions, the curves depicting the variation of volatilization with moisture content are shown in Fig. 5. The corresponding kinetics fitting outcomes are presented in Table S4. The comparison in Fig. 5 on different moisture content reveals that the volatilization rate of benzene in all soils exhibits an increasing trend with the increase of moisture content. The volatilization process of benzene in the subsurface layer of soil can be categorized into pre-infiltration diffusion and post-infiltration diffusion. The greater the benzene retained on the soils, the slower its volatilization. On one hand, the increase in soil moisture content result in a reduction of effective porosity in the soil, impeding the infiltration of benzene into the soil interstices, promoting pre-infiltration diffusion. On the other hand, an increase

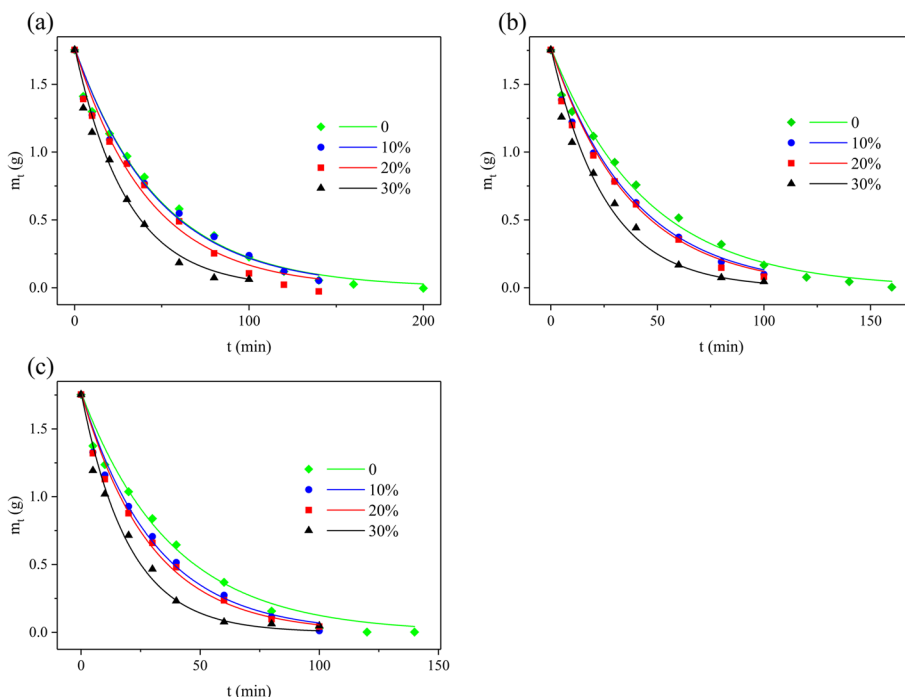


Fig. 5 The curves of the residual mass of benzene with time on different moisture contents in three soils of (a) silty loam, (b) loam and (c) sandy loam

in moisture content reduces the soil’s adsorption capacity [4], facilitating the volatilization of benzene into the air.

Under the moisture content of 30%, for silty loam, loam and sandy loam, the volatilization equilibrium time decreases from 200 to 80 min, 160 to 80 min, and 100 to 60 min, respectively. The volatilization rate increases from 0.021 to 0.033 min⁻¹, 0.022 to 0.040 min⁻¹, and 0.026 to 0.050 min⁻¹. The variations in volatilization equilibrium time for silty loam, loam, and sandy loam are 120, 80, and 40 min, respectively. The corresponding increases in volatilization rates for silty loam, loam, and sandy loam are 0.012, 0.018, and 0.024 min⁻¹, respectively. It is indicated that moisture content exerts a more significant impact on soils with larger particle sizes.

3.5 Correlation analysis and prediction of the volatilization rate under different key factors

The experimental results regarding the factors influencing the volatilization of benzene include temperature, wind speed, moisture content and soil properties. To characterize how variations in these parameters contribute to the overall volatilization, correlations between various influencing factors and the volatilization rate of benzene were computed. Zero-order correlation coefficients and partial correlation coefficients [27] were

obtained by Pearson Correlation Analysis as shown in Table S5. Observably, across the five factors, the volatilization rate is strongly related with wind speed ($r > 0.9$), while other factors and the volatilization rate is non-correlated.

According to the prediction result by BPNN shown in Fig. 6 and the evaluation result shown in Table 2, the BPNN model demonstrates excellent capability in simulating the volatilization rate of benzene in soil. The correlation coefficient of testing is reported as 0.981, and its RMSE is calculated at 0.011.

The results of K-CV by BPNN are shown in Fig. S2 and Table 2. As can be seen, the predictive result of testing with the average R of 0.963 and MSE of 4.054e-4 indicates that there is a good possibility to estimate the volatilization rate of benzene in different influential factors by BPNN. K-CV was used in this study to provide a robust assessment of the BPNN model’s performance by considering multiple training and testing sets, thus enhancing its ability to generalize beyond the original dataset.

The application of BPNN in this paper has shown promise in predicting volatility rates. However, some limitations persist with the BPNN model. Overfitting remains a potential concern despite the application of K-CV to mitigate its risk. Additionally, the interactive

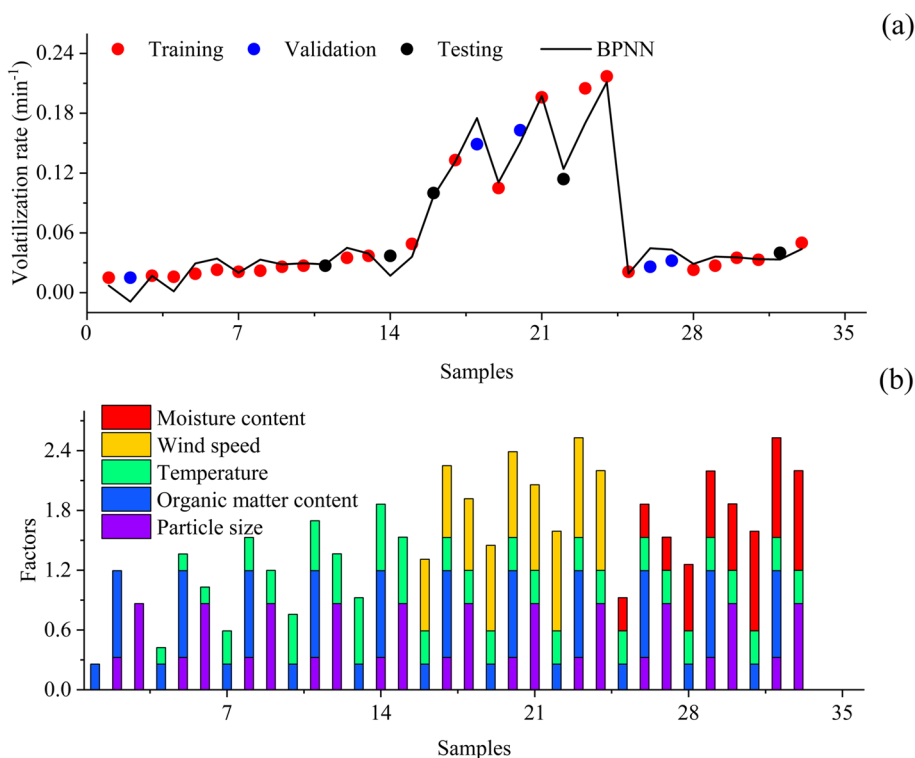


Fig. 6 The prediction result of the volatilization rates of benzene (a) under different influencing parameters by BPNN and the stacked bar chart of the proportion of influencing factors (b), this proportion were obtained by normalizing the values of each influencing factor between 0 and 1

Table 2 The statistical indicators of prediction result of BP Artificial Neutral Network and K-fold Cross Validation

BP Artificial Neutral Network						
Indicator	Training		Validation		Testing	
	R	MSE	R	RMSE	R	MSE
Value	0.988	1.026e-4	0.963	3.757e-4	0.981	1.125e-4
K-fold Cross Validation						
Indicator	Training		Testing			
	R	MSE	R	MSE	R	MSE
K-fold-1	0.990	8.581e-5	0.977	1.620e-4		
K-fold-2	0.983	1.423e-4	0.959	2.017e-4		
K-fold-3	0.981	1.101e-4	0.959	8.452e-4		
K-fold-4	0.983	1.367e-4	0.947	4.125e-4		
Average value	0.984	1.187e-4	0.960	4.054e-4		

effects among the various influential factors affect the accuracy of predictions of volatility rates, highlighting the need for consideration in future studies.

4 Conclusions

Comparative volatilization experiments were conducted under different influential factors of soil properties, temperature, wind speed, and moisture content. Our systematic analysis revealed the impact of these factors on the volatilization and provided insights into the associated mechanisms. The results indicated that an increase in particle size, temperature, and moisture content corresponds to an elevated volatilization rate of benzene. The presence of wind significantly enhances the volatilization rate, with higher wind speeds having a relatively minor impact.

To delve into the complex non-linear relationships between influencing factors and the volatilization coefficients, Pearson correlation analysis was employed. Among all the factors, wind speed emerged as the most influential on the volatilization. Furthermore, we utilized the BPNN as a powerful tool to simulate the volatilization rates on different influential factors, and the fitting results demonstrated a good performance. Finally, to address potential overfitting concerns due to the limited dataset, K-fold cross-validation was conducted, affirming the reliability of the BPNN model.

In conclusion, this study provides a comprehensive understanding of the factors influencing the volatilization of benzene, and the application of advanced modeling techniques of BPNN enhances the accuracy and reliability of predictions in scenarios with limited data.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s42834-024-00225-x>.

Supplementary Material 1.

Authors' contributions

All authors contributed to the study conception and design. Qian Wang: Data curation, methodology and writing—original draft; Jianmin Bian: funding acquisition, writing—review & editing; Dongmei Ruan: Data curation. All authors read and approved the final manuscript.

Funding

This research reported here was funded by the National Key R&D Program of China 'The formation mechanism and multiphase distribution characteristics of controlling organic pollution in chemical industry sites' (2018YFC1800400).

Availability of data and materials

The authors confirm that the data supporting the findings of this study are available within the article.

Declarations

Competing interests

The authors have no relevant financial or non-financial interests to disclose.

Received: 1 April 2024 Accepted: 29 July 2024

Published online: 06 August 2024

References

- Martin-Olmedo P, Sanchez-Cantalejo C, Ancona C, Ranzi A, Bauleo L, Fletcher T, et al. Industrial contaminated sites and health: results of a European survey. *Epidemiol Prev.* 2019;43:238–48.
- Rajasekhar B, Nambi IM, Govindarajan SK. Human health risk assessment for exposure to BTEXN in an urban aquifer using deterministic and probabilistic methods: A case study of Chennai city, India. *Environ Pollut.* 2020;265:114814.
- Wang W, Yan Y, Fang H, Li J, Zha S, Wu T. Volatile organic compound emissions from typical industries: Implications for the importance of oxygenated volatile organic compounds. *Atmos Pollut Res.* 2023;14:101640.

4. Lu L, Liu Z, Li S, Bai M, Fan Z. Volatile organic compound (VOC) emissions from soil and soil under consideration of co-existence of pore water. *Environ Sci Pollut R*. 2021;28:68920–9.
5. Sookhak Lari K, Davis GB, Rayner JL, Bastow TP. Advective and diffusive gas phase transport in vadose zones: Importance for defining vapour risks and natural source zone depletion of petroleum hydrocarbons. *Water Res*. 2024;255:121455.
6. Bani-Hani EH, Al-Khatib F, Khanafer K. Analysis of variants within different models of diesel oil volatilization process using various soil materials. *Spec Top Rev Porous Media* 2015;6:261–6.
7. Tong L, Zheng XL, Li M, Hu ZF. Volatilization behavior of BTEX on different underlying materials. *Environ Sci*. 2008;29:2058–62 [In Chinese]. https://www.hjxx.ac.cn/hjxx/ch/reader/view_abstract.aspx?file_no=20080750&st=alljournals (Accessed 1 Jul 2024).
8. Davardoost F, Kahforoushan D. Evaluation and investigation of the effects of ventilation layout, rate, and room temperature on pollution dispersion across a laboratory indoor environment. *Environ Sci Pollut R*. 2019;26:5410–21.
9. Zhang S, Zhao J, Zhu L. Predicting removal efficiency of organic pollutants by soil vapor extraction based on an optimized machine learning method. *Sci Total Environ*. 2024;927:172438.
10. Ma Y, Zheng X, Anderson SH, Lu J, Feng X. Diesel oil volatilization processes affected by selected porous media. *Chemosphere*. 2014;99:192–8.
11. Wang Y, Chen J, Tang W, Xia D, Liang Y, Li X. Modeling adsorption of organic pollutants onto single-walled carbon nanotubes with theoretical molecular descriptors using MLR and SVM algorithms. *Chemosphere*. 2019;214:79–84.
12. Rahmanifard H, Plaksina T. Application of artificial intelligence techniques in the petroleum industry: a review. *Artif Intell Rev*. 2019;52:2295–318.
13. Belhaj AF, Elraies KA, Alnarabiji MS, Abdul Kareem FA, Shuhli JA, Mahmood SM, et al. Experimental investigation, binary modelling and artificial neural network prediction of surfactant adsorption for enhanced oil recovery application. *Chem Eng J*. 2021;406:127081.
14. Wang Q, Bian J, Ruan D, Zhang C. Adsorption of benzene on soils under different influential factors: an experimental investigation, importance order and prediction using artificial neural network. *J Environ Manage*. 2022;306:114467.
15. Qiao L, Wang Z, Zhu J. Application of improved GRNN model to predict interlamellar spacing and mechanical properties of hypereutectoid steel. *Mat Sci Eng A*. 2020;792:139845.
16. Ding F, Zhang W, Cao S, Hao S, Chen L, Xie X, et al. Optimization of water quality index models using machine learning approaches. *Water Res*. 2023;243:120337.
17. Kleinknecht SM, Class H, Braun J. Experimental study on retardation of a heavy NAPL vapor in partially saturated porous media. *Hydrol Earth Syst Sci*. 2017;21:1381–96.
18. Ma Y, Shi Y, Hou D, Zhang X, Chen J, Wang Z, et al. Treatability of volatile chlorinated hydrocarbon-contaminated soils of different textures along a vertical profile by mechanical soil aeration: A laboratory test. *J Environ Sci-China*. 2017;54:328–35.
19. Lei T, Guo X, Sun X, Ma J, Zhang S, Zhang Y. Prediction of soil urea conversion and quantification of the importance degrees of influencing factors through a new combinatorial model based on cluster method and artificial neural network. *Chemosphere*. 2018;199:676–83.
20. Perez-Sanchez B, Fontenla-Romero O, Guijarro-Berdinas B. A review of adaptive online learning for artificial neural networks. *Artif Intell Rev*. 2018;49:281–99.
21. Bishop CM. *Neural Networks for Pattern Recognition*. Oxford: Oxford University Press; 1995.
22. Chu J, Liu X, Zhang Z, Zhang Y, He M. A novel method overcoming overfitting of artificial neural network for accurate prediction: Application on thermophysical property of natural gas. *Case Stud Therm Eng*. 2021;28:101406.
23. Lei J. Cross-validation with confidence. *J Am Stat Assoc*. 2020;115:1978–97.
24. Wong TT, Yeh PY. Reliable accuracy estimates from k-fold cross validation. *IEEE T Knowl Data En*. 2020;32:1586–94.
25. Mofidi A, Asilian H, Jafari AJ. Vapors generation of volatile organic compounds (VOCs) from pure liquids. *Iran Occup Health*. 2014;11:75–87 [In Persian]. <https://api.semanticscholar.org/CorpusID:227013962> (Accessed 16 Jul 2024).
26. Philippe N, Davarzani H, Colombano S, Dierick M, Klein PY, Marcoux M. Experimental study of thermally enhanced recovery of high-viscosity DNAPL in saturated porous media under non-isothermal conditions. *J Contam Hydrol*. 2021;243:103861.
27. Kleinbaum DG. *Applied Regression Analysis and Other Multivariable Methods*. Scituate: Duxbury Press; 1978.

Publisher's Note

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