ACS OMEGA

Open Access

http://pubs.acs.org/journal/acsodf

Article

Development of the Quantitative Property−Consequence Relationship Model for Prediction of Hydrogen Leakage and Dispersion Using Response Surface Method and Artificial Neural Network Approaches

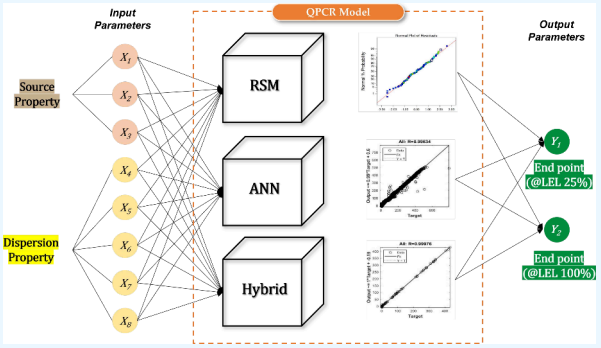
Junseo Lee, Sehyeon Oh, Seongheon Baek, Changbock Chung, and Byungchol Ma*

Cite This: *ACS Omega* 2024, 9, 40857−40869

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Hydrogen, an environmentally friendly and highly regarded future energy source, can form flammable vapor clouds upon leakage, which may transition into explosion. Predicting the dispersion behavior of hydrogen is crucial for preventing such incidents. This study aims to develop a quantitative property−consequence relationship (QPCR) model using the response surface method (RSM) and artificial neural network (ANN) to swiftly and accurately predict dispersion behavior. Initially, 8 variables were defined from source and dispersion models, constructing a data set through 6,561 PHAST simulations. Subsequently, the RSM-BBD (Box-Behnken design) and ANN-BPNN (Backpropagation neural network) models were developed, alongside a hybrid model incorporating BPNN after excluding four low-influence variables based on analysis of variance (ANOVA). All models achieved an R^2 value exceeding 0.99. The hybrid model notably reduced computational costs by 97% compared to ANN-BPNN and exhibited lower mean square error (MSE). These results introduce a cost-effective approach for high-accuracy QPCR modeling and highlight the viability of diverse statistical methods.



INTRODUCTION

Hydrogen is gaining prominence as a next-generation energy solution, as it can serve as a renewable energy source capable of replacing fossil fuels without posing a threat to the environment, while also addressing energy scarcity issues.[1,2] Hydrogen serves not only as a fuel for internal combustion engines in applications such as aircraft propulsion, hydrogen vehicles, and hydrogen refueling stations, but also plays a crucial role in hydrogen fuel cells, where it converts the chemical energy of hydrogen into mechanical energy. Additionally, hydrogen finds utility in various conventional petrochemical processes, including ammonia production and chemical processing.[3,4]

Despite its potential, hydrogen's high flammability poses risks of chemical accidents, including fires and explosions, even with minimum ignition energy.[1,2,5] For example, if the thermal management system of an electric vehicle's lithium-ion battery fails, hydrogen gas as an off-gas can be discharged into the environment, dispersing and potentially causing fires and explosions.[6−8] One notable hydrogen-related accident occurred in June 2019 at the Uno-X hydrogen refueling station in Norway. The incident involved a vapor cloud explosion triggered by progressive hydrogen leakage due to a plug sealing error. The leaking hydrogen spread and accumulated in the surrounding environment, resulting in a vapor cloud that ignited and caused the explosion. This incident resulted in two injuries and temporary closure of the hydrogen refueling station, which caused economic losses.[9] Such accidents demonstrate that hydrogen leaks can rapidly form vapor clouds and escalate into major incidents, including fires and explosions, because of potential ignition sources.[10] Therefore, in order to avoid fire and explosion accidents caused by leaked hydrogen, efforts must be made to quickly and properly understand the dispersion properties of unignited hydrogen emitted in the hydrogen industry.

The leakage and dispersion of chemical substances, including hydrogen, can be evaluated and predicted by understanding the relationship between various input parameters such as leakage hole diameter, operating pressure, wind speed, and concentration of chemicals at specific times and

Received: June 23, 2024
Revised: August 29, 2024
Accepted: September 5, 2024
Published: September 19, 2024

locations, which serve as output variables. Such relationships can be analyzed using the source model and dispersion model.[11,12] The source model evaluates the quantity or rate of chemical leakage and plays a central role in assessing the size of vapor clouds or the potential for fires.[11,12] The dispersion model assesses the process of mixing and spreading in the atmosphere following a release. The dispersion model is classified into: Empirical methods, Computational Fluid Dynamics (CFD) approach and Unified Dispersion Models (UDM).[13]

Empirical and CFD methodologies may produce drastically varied findings depending on the specialist's competence. CFD approaches, in particular, have limitations due to significant computational costs depending on hardware performance, software employed, and user expertise.[14] In contrast, a well-known UDM, PHAST, uses complex models to numerically study the behavior of chemical compounds by solving time-dependent ordinary differential equations. This program addresses the limits of inaccurate empirical models by using enhanced models and provides the advantage of evaluating chemical dispersion behavior quickly and at a lower computational cost than CFD approaches. Furthermore, PHAST has been proven reliable by numerous prior researches, including Wiltox et al.[15−17]

Recently, there has been a trend to combine machine learning algorithms such as Artificial Neural Networks (ANN) with CFD or UDM to predict the dispersion of chemicals.[18−21] The quantitative property-consequence relationship (QPCR) model, originally proposed by Jiao et al., is a well-known machine learning-based dispersion prediction model.[21] QPCR serves as a rapid assessment and prediction tool for consequences such as dispersion by connecting leakage conditions and chemical properties to dispersion consequences. It introduces the concept of "Property descriptor," encompassing source properties, physical properties, criteria properties, and more, allowing for rapid evaluation and prediction of consequences. Jiao et al. developed a QPCR using the Extreme Gradient Boosting algorithm and PHAST to analyze the dispersion of combustible gases.[21] Jiao et al. also built QPCR models for predicting hazardous dispersion using algorithms like XGBoost and Random Forests.[22] In another study, they created a QPCR model to assess the influence of harmful gas dispersion on mortality rates.[13]

Besides the QPCR models, Li et al. used the Back Propagation Neural Network (BPNN) to calculate the maximum horizontal downwind distance utilizing variables from the source model (leakage flow rate, leakage hole diameter, and hydrogen mixing ratio).[23] BPNN is an algorithm for predicting challenging nonlinear issues. It consists of three layers: input, hidden, and output. It modifies the weights and biases of the neurons using forward propagation and back-propagation to minimize the loss function (such as mean squared error), optimizing the network for accurate predictions. However, while these researches created prediction models for generic chemicals that focused on source and physical parameters, they ignored dispersion properties such ambient temperature, wind speed, and release height, which are crucial variables in dispersion models.

Meanwhile, the Response Surface Method (RSM) serves as a useful statistical tool for analyzing the impact of input variables on output variables.[24,25] RSM, in particular, has the advantage of producing accurate results with a limited number of experimental outcomes, since it can construct response surfaces with a small number of experiments or trials, removing the need for lengthy research.[25] Furthermore, Analysis of variance (ANOVA) in RSM uses p-values to analyze the influence of input variables on the regression model, allowing for variable sensitivity analysis.[26] For the reason for these benefits, RSM is adopted in many studies to assess the dispersion consequences of chemicals.

Shi et al. suggested a strategy for predicting vapor cloud size that takes into account variables from both the source model (leakage flow rate and direction) and the dispersion model (wind speed and direction), using RSM and FLACS CFD.[27] Mousavi and Parvini combined RSM and PHAST to simulate the effects of variables from the source model (operating temperature, operating pressure, leakage hole diameter, release height) and the dispersion model (atmospheric temperature, relative humidity, surface roughness, wind speed, atmospheric pressure).[28] However, Shi et al.'s study, which uses CFD results, has limits in terms of simulation costs and computing procedures for input variables, limiting the ability to estimate results quickly. Furthermore, Mousavi and Parvini's study focused entirely on the sensitivity of each variable to the results, which is a weakness.

In this study, we created three types of QPCR models to predict hydrogen leakage and dispersion, considering both the source and dispersion models via RSM and ANN. To begin, scenarios of hydrogen leakage in high-pressure facilities were chosen, with three variables from the source model and five variables from the dispersion model serving as property descriptors. We then used PHAST to calculate the downwind distances for hydrogen at Lower Explosive Limit (LEL) 100% and 25%. We utilized PHAST to simulate 6,561 situations, resulting in 13,122 data points. We subsequently developed three QPCR models: one using RSM, one using ANN, and a hybrid model combining RSM and ANN. Finally, we validated the three QPCR models using the number of data points needed for model building, the determination coefficient (R^2), and mean squared error (MSE). This study's findings illustrate the development of a more accurate and rapid QPCR model, as well as the potential of diverse statistical approaches in QPCR model development.

METHODOLOGY

In this study, the process of developing the QPCR differed from the procedure outlined by Jiao et al.,[29] as it incorporated not only the RSM-based QPCR model but also a hybrid QPCR model that integrates RSM and ANN methods (Figure 1). This section is dedicated to the development and evaluation of QPCR models utilizing statistical methods such as ANN to predict hydrogen leakage and dispersion. Instead of extensively discussing conventional machine learning models, the focus remains on detailing the key steps outlined in Figure 1.

In this study, the models being developed are designed with careful consideration of uncertainties. When modeling complex systems, both aleatoric and epistemic uncertainties need to be addressed with caution.[30,31] Aleatoric uncertainty, also referred to as data uncertainty, represents inherent errors embedded in the data. While aleatoric uncertainty cannot be entirely eliminated, this study aims to mitigate intrinsic noise (e.g., variations in measurement environments) that may occur during the data acquisition process by constructing a data set based on simulation results. The second type, epistemic uncertainty, arises from a lack of knowledge about the values of

https://doi.org/10.1021/acsomega.4c05841
ACS Omega 2024, 9, 40857−40869

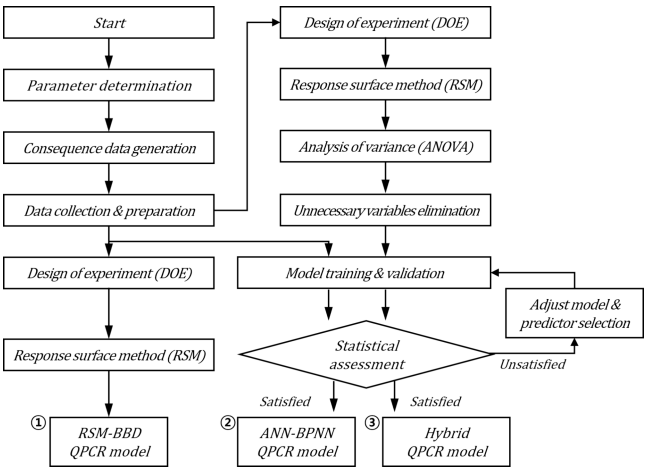


Figure 1. Flowchart for developing the QPCR model.

model parameters. This uncertainty can be reduced by obtaining a sufficient amount of data or conducting a thorough sensitivity analysis.[32] In this study, the ANN-BPNN QPCR model addresses epistemic uncertainty by securing an adequate amount of data, while the RSM-BBD QPCR model and the Hybrid QPCR model mitigate this uncertainty through sensitivity analysis using ANOVA.

Data Generation and Collection. The data generated by integrated models such as PHAST, validated across various leakage scenarios, are suitable for QPCR model development.[29] Selecting variables for the source and dispersion models is essential to run PHAST, preceded by the selection of accident scenarios. In this study, we opted for a scenario involving continuous hydrogen leakage in plume form from high-pressure vessels. When a hydrogen vessel leaks, the PHAST results and output variables are set to determine the distance at which hydrogen is released into the atmosphere and forms a vapor cloud, specifically the downwind distance at 100% of the LEL and a more conservative downwind distance at 25% of the LEL.

Because it is not possible to simulate all conceivable scenarios, many assumptions and restrictions were clearly established in order to carry out source and dispersion modeling. First, variations in the chemical's storage volume and leakage rate over time are not addressed. Instead, the model depicts the size of the highest vapor cloud in steady state following leakage. Second, the leakage and wind directions are fixed. In this study, the leaking direction is set to be horizontal to the ground, and the wind direction is aligned with the leakage direction, focusing primarily on the situation in which the vapor cloud develops most widely. Finally, the default values of PHAST for solar radiation and surface temperature are employed. Although fluctuations in solar radiation due to day and night or cloudiness might impact the atmospheric temperature gradient, these circumstances are too complex to be considered thoroughly, and the values can vary greatly by area, making it difficult to set a universal standard. Therefore, this study uses the default values of PHAST to account for typical conditions. As a result, this study makes use of PHAST's default settings to adjust for common situations. These procedures reduce the complex source/dispersion modeling process while still producing data that is appropriate for the study's purpose.

According to previous studies, for evaluating the results of PHAST, there are thirty-six adjustable input variables. Among these, 13 variables classified as Class A (operating pressure, operating temperature, leakage hole diameter, liquid height inside the vessel, atmospheric temperature, atmospheric pressure, relative humidity, leak duration, release height, leak angle, surface roughness length, wind speed, and atmospheric specific heat) are dominant in influencing the consequences.[33] Among these, operating pressure, operating temperature, leakage hole diameter, liquid height inside the vessel, and leak duration are variables related to the source model, while the remaining eight variables are associated with the dispersion model.

Source Model. The source model is a model that assesses the mechanical turbulence, which has a dominant influence during the initial stages of leakage, to calculate the initial leakage rate of chemicals. In this study, the source model for high-pressure vessels is referenced, utilizing the method described in the literature.[34] The source model calculates the leakage rate of gas using the following Equation 1.

$$Q_m = C_0 A P_0 \sqrt{\frac{2g_c M}{R_g T_0}\frac{\gamma}{\gamma-1}\left[\left(\frac{P_a}{P_0}\right)^{(2-\gamma)} - \left(\frac{P_a}{P_0}\right)^{\gamma+1/\gamma}\right]} \quad (1)$$

Where, Q_m represents the leakage rate [kg/s], C_0 denotes the leakage coefficient [-], A stands for leakage area [m^2], P_0 signifies the operating pressure [Pa], γ represents the specific heat ratio (C_P/C_V), P_a indicates the atmospheric pressure [Pa], T_0 represents the operating temperature [K], R_g is the gas constant [8.314 m^3·Pa/kg-mole·K], g_c denotes the gravitational accelerator [9.8m/s^2], M represents the molecular weight of the gas [kg/kg-mole].

To model hydrogen leakage from the high-pressure vessel, the key variables considered are the operating pressure (P_0) of hydrogen, the operating temperature (T_0) of hydrogen, and the leakage hole diameter (D_0). Since hydrogen is stored in a gaseous state inside the vessel, the liquid height inside the vessel is excluded. Additionally, since the results are based on reaching a steady state during continuous leakage, no separate setting for leakage duration is provided. Consequently, these three variables (P_0, T_0, D_0) have been selected as source properties.

Dispersion Model. When hydrogen leakage occurs from the vessel, the momentum that dominated the initial leakage decreases as the distance from the leakage source increases, and beyond a certain distance, it is influenced by atmospheric conditions such as wind speed.[33] To interpret this, various dispersion models, such as the Gaussian-Gifford model, are used to analyze the dispersion of gases.[33] The Gaussian-Gifford model combines the interpretation of dispersion processes based on Gaussian normal distribution with correlation equations for dispersion coefficients derived from neutral buoyancy gas experimental results. It analyzes the concentration of chemicals moving in the x, y, and z directions.[35] The Gaussian-Gifford equation is as follows Equation 2.

$$C(x,y,z) = \frac{Q_m}{2\pi\sigma_x\sigma_u u}\times\exp\left[-\frac{u}{2}\left(\frac{y}{\sigma_y}\right)^2\right]\times\left(\exp\left[-\frac{1}{2}\left(\frac{z-Z_h}{\sigma_z}\right)^2\right] + \exp\left[-\frac{1}{2}\left(\frac{z+Z_h}{\sigma_z}\right)^2\right]\right) \quad (2)$$

https://doi.org/10.1021/acsomega.4c05841
ACS Omega 2024, 9, 40857−40869

Here, $C(x,y,z)$ represents the concentration of the chemicals in each direction [ppm], σ is the dispersion coefficient proposed by Gifford[36] [-], u is the wind speed [m/s] and Z_h is the release height [m].

In response to the limitations of the Gaussian-Gifford model, numerous studies have sought to enhance dispersion models. Prior studies have emphasized the influence of atmospheric conditions and surface characteristics on dispersion consequences, particularly in determining the vertical mixing level between chemicals and the atmosphere.[2,37−41] For this reason, factors like atmospheric temperature (T_a), relative humidity (H_a), and surface roughness length (Z_0) have been added as input variables in dispersion properties. However, in the scenario of this study, the influence of operating pressure during leakage in high-pressure vessels is predominant, so atmospheric pressure is disregarded, and the atmospheric specific heat is set to the default value in PHAST. Consequently, 5 variables (T_a, H_a, Z_h, Z_0, u) have been selected as dispersion properties.

Descriptors Definition. For the development of the QPCR model, eight different property descriptors have been chosen as shown in Table 1, categorizing each property descriptor into three source properties and five dispersion properties. The criteria property chosen is the downwind distance to reach the LEL at 100% and 25%. Detailed classification and specifications are provided in Table 1. Each property descriptor corresponds to input variables in the PHAST simulation.

Table 1. Property Descriptors for QPCR Model

Nomenclature	Descriptor	Property Type	Unit	Range
T_0	Operating Temperature	Source	°C	−40∼65
P_0	Operating pressure	Source	barg	1∼700
D_0	Leakage hole diameter	Source	mm	0.178∼102
u	Wind speed	Dispersion	m/s	1∼6
Z_h	Release height	Dispersion	m	−0∼10
Z_0	Surface roughness length	Dispersion	m	0.005∼3
T_a	Atmospheric temperature	Dispersion	°C	−20∼40
H_a	Relative humidity	Dispersion	%	20∼100
D_{LEL}	Downwind distance of LEL fraction	Criteria	%	25,100

T_0 and P_0 are specified to cover the typical operating temperature and pressure in the hydrogen industry, including hydrogen refueling stations.[42,43] D_0 is configured to accommodate leakage hole diameter ranging from 0.178 mm, as recommended by IEC 60079−10−1, to the standard large hole size of 102 mm specified in API 581, ensuring adaptability.[44] Parameters such as u,T_a, and H_a are based on general atmospheric conditions, while Z_h refers to the height of facilities used in the hydrogen industry. Lastly, Z_0 encompasses various terrains from mud flats to center of city.

RSM-BBD. RSM is a mathematical-statistical method for determining correlations between different parameters and responses.[25,27] In other terms, it is a statistical method to assess the extent to which many variables (X_1, X_2, ..., X_k) influence an output variable (Y), by creating a mathematical model of the input-output relationship on a first- or second-order response surface. Equation 3 depicts the link between the input variable (X) and the output variable (Y). And Equation 4 employs the quadratic form to optimize the relationship and examine the impact of each variable.

$$Y = f(X_1, X_2, ..., X_k) \tag{3}$$

$$Y = \beta_0 + \sum_{i=1}^{k} \beta_i X_i + \sum_{i=1}^{k} \beta_{ij} X_i X_j \tag{4}$$

In Equation 4, Y represents the output variable, X represents the input variable, and, β denotes the regression coefficients.

Furthermore, RSM excels in producing valid results with a small number of experimental outcomes based on the design of experiment (DOE) methodology.[25] In this study, a Box-Behnken Design (BBD) model, a sort of fractional factorial design within DOE, was utilized to assess the impact of each variable on the output. To do this, the eight variables selected as descriptors were designed with three levels each, as illustrated in Table 2. The BBD model has the advantage of effectively analyzing interactions between input and output variables with a minimal number of experiments.[27]

Table 2. Range of Various Descriptors in Scenario

				Level of factor		
Nomenclature	Code	Descriptor	Unit	−1	0	1
T_0	X_1	Operating temperature	°C	−40	0	65
P_0	X_2	Operating pressure	barg	1	350	700
D_0	X_3	Leakage hole diameter	mm	0.178	25	102
u	X_4	Wind speed	m/s	1	3	6
Z_h	X_5	Release height	m	0	5	10
Z_0	X_6	Surface roughness length	m	0.005	1	3
T_a	X_7	Atmospheric temperature	°C	−20	10	40
H_a	X_8	Relative humidity	%	20	60	100

Moreover, RSM enables for a sensitivity analysis of each input variable to the output variable, i.e., the impact of variables, using ANOVA. In ANOVA results, the F-value is the ratio of the mean square to the sum of squares for each factor, with a higher value indicating a stronger influence of that variable on the output variable.[45] Another indicator, the p-value, represents the probability of observing extreme results given that the null hypothesis is true, and is used to evaluate the model's dependability and the influence of input variables Based on the results of Fisher's statistical test, a p-value of less than 0.05 suggests that the variable has a significant influence on the output variable, but a value greater than 0.1 implies insignificance.[28,46]

ANN-BPNN. BPNN is an ANN algorithm that adjusts network weights and biases in the direction of the most rapid drop in the objective function. BPNN is widely used in dispersion model predictions due to its outstanding training speed and prediction capabilities.[47,48] BPNN, which includes input, hidden, and output layers, allows manual adjustment of the number of neurons in hidden layer. While tuning the number of the neurons might improve accuracy and convergence speed, poor synchronization can lead to overfitting, resulting in major errors in result accuracy.[49]

https://doi.org/10.1021/acsomega.4c05841
ACS Omega 2024, 9, 40857−40869

In this study, the input layer was built with the eight variables from Table 2, while the output layer was defined with two variables: the maximum downwind distances to LEL 100% and 25%. Additionally, the number of neurons in hidden layer was chosen to be twice as large as the input layer size.[19] The training process was carried out using the MATLAB neural network toolbox, with the Levenberg−Marquardt regularization approach serving as the optimization algorithm. The data set was divided into three sets: training (70%), validation (15%), and testing (15%). If the accuracy did not improve by the sixth epoch on the validation set, the process was set to terminate early, with a maximum of 1,000 epochs allowed unless an early stop happened. The most widely used sigmoid function was adopted as the activation function. The sigmoid function has the virtue of being continuously differentiable everywhere.[50]

Hybrid Model Based on RSM and ANN. The key stage in developing a QPCR model is finding appropriate descriptors as input variables.[13] In other words, if we can roughly estimate the influence of each variable on the outcome before constructing the data set for model development, we can improve the accuracy of the results while minimizing computational costs. To accomplish this, an ANOVA was employed to examine the influence of eight factors on the results. ANOVA evaluates the effect of factors in the model-building process by examining the F-value and p-value calculated using the RSM approach. The data set was reconstructed after analyzing the variables' sensitivity and removing those with little influence. Then, a BPNN algorithm was used to create a hybrid QPCR model.This strategy decreases the number of variables, lowering the computational cost of data set building while using just influential factors, resulting in a highly successful QPCR model.

Model Evaluation. In this study, R^2 (Equation 5) and MSE (Equation 6) were chosen as standard metrics to compare and analyze the RSM, ANN, and hybrid QPCR models. Additionally, the data set size needed for model building was chosen as a criterion.

$$R^2 = \frac{\sum_{i=1}^{n} (\hat{y}_i - \overline{y})^2}{\sum_{i=1}^{n} (y_i - \overline{y})^2} \tag{5}$$

$$\text{MSE} = \frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n} \tag{6}$$

Here, $\hat{y}_i$ represents the predicted value, y_i represents the target value, $\overline{y}$ represents the mean value of the target value, and n denotes the number of data points in the data set.

R^2 measures how effectively the independent variable predicts variation in the dependent variable, with higher values indicating better model fit. MSE is calculated as the average squared difference between predicted and actual values, with smaller values indicating higher model accuracy.

RESULTS AND DISCUSSION

RSM-BBD QPCR Model. Using a BBD-based experimental methodology, 120 simulations were performed for each QPCR model, resulting in the development of QPCR models that predict downwind distances to LEL 100% and LEL 25% based on 240 data points. Each model follows a quadratic form as stated in Equation 7, and the proposed LEL 100% RSM-BBD QPCR model and LEL 25% RSM-BBD QPCR models are shown below.

$$D_{LEL} = f(x) = \frac{1}{2} x^T A x + b^T x + c \tag{7}$$

Where, D_{LEL} represents the downwind distance according to hydrogen concentration, $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{nxn}$, $b \in \mathbb{R}^n$ and $c \in \mathbb{R}$ are defined as

$$x = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \\ X_6 \\ X_7 \\ X_8 \end{bmatrix}$$

$$A_{LEL100\%} = \begin{bmatrix} -0.00805 & -0.01059 & 0.007664 & 0.0194 & -0.04592 & 0.091717 & 0.041924 & -0.02602 \\ -0.01059 & -1.83865 & 0.186508 & 0.010482 & 0.037712 & 0.022158 & -0.03139 & -0.04906 \\ 0.007664 & 0.186508 & -3.09394 & -0.04044 & 0.290955 & 0.027677 & 0.005917 & -0.09572 \\ 0.0194 & 0.010482 & -0.04044 & -0.03981 & -0.28194 & -0.069 & -0.01321 & 0.016512 \\ -0.04592 & 0.037712 & 0.290955 & -0.28194 & -0.07019 & 0.226196 & -0.02551 & -0.03912 \\ 0.091717 & 0.022158 & 0.027677 & -0.069 & 0.226196 & -0.00296 & 0.038093 & 0.018991 \\ 0.041924 & -0.03139 & 0.005917 & -0.01321 & -0.02551 & 0.038093 & -0.06609 & 0.018912 \\ -0.02602 & -0.04906 & -0.09572 & 0.016512 & -0.03912 & 0.018991 & 0.018912 & -0.13505 \end{bmatrix}$$

https://doi.org/10.1021/acsomega.4c05841
ACS Omega 2024, 9, 40857−40869

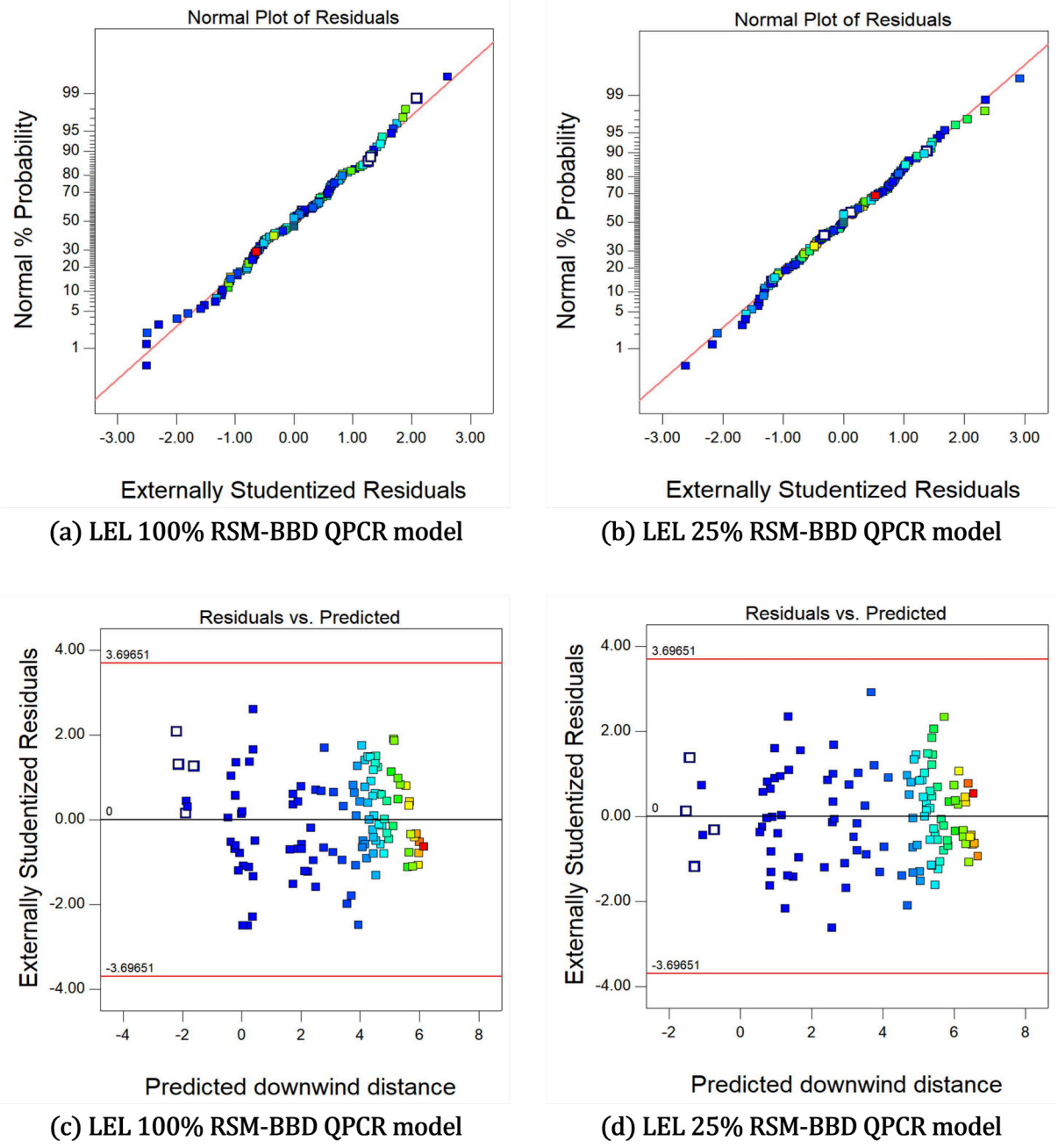
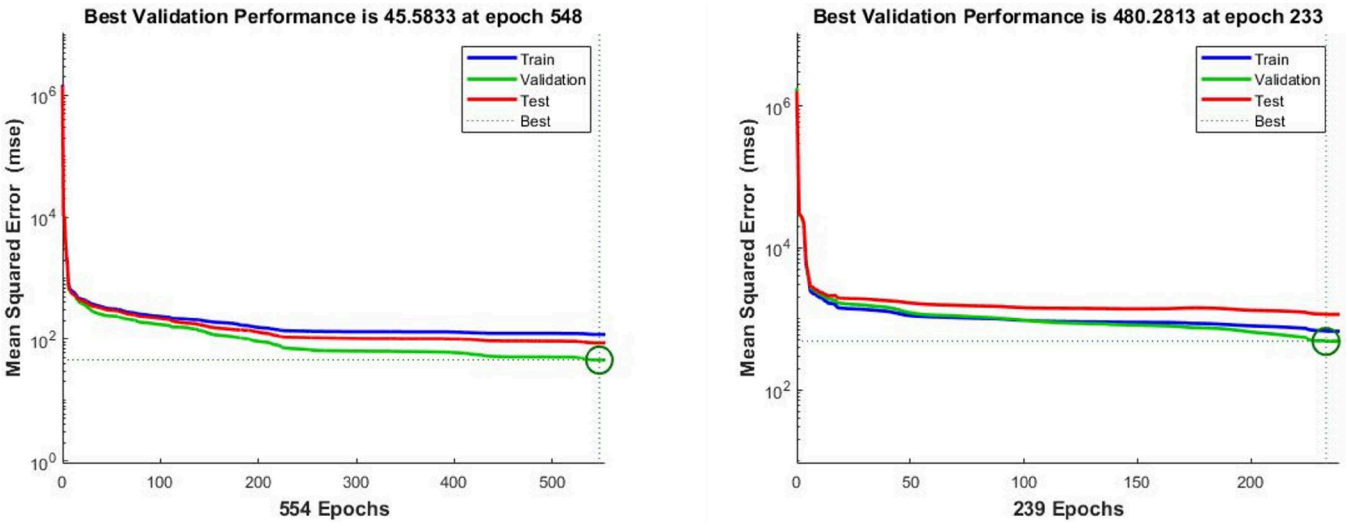


Figure 2. RSM-BBD value of QPCR model; (a) externally studentized residuals vs normal plot of LEL 100% RSM-BBD QPCR model, (b) externally studentized residuals vs normal plot of LEL 25% RSM-BBD QPCR model, (c) predicted vs externally studentized of LEL 100% RSM-BBD QPCR model, (d) predicted vs externally studentized of LEL 25% RSM-BBD QPCR model.

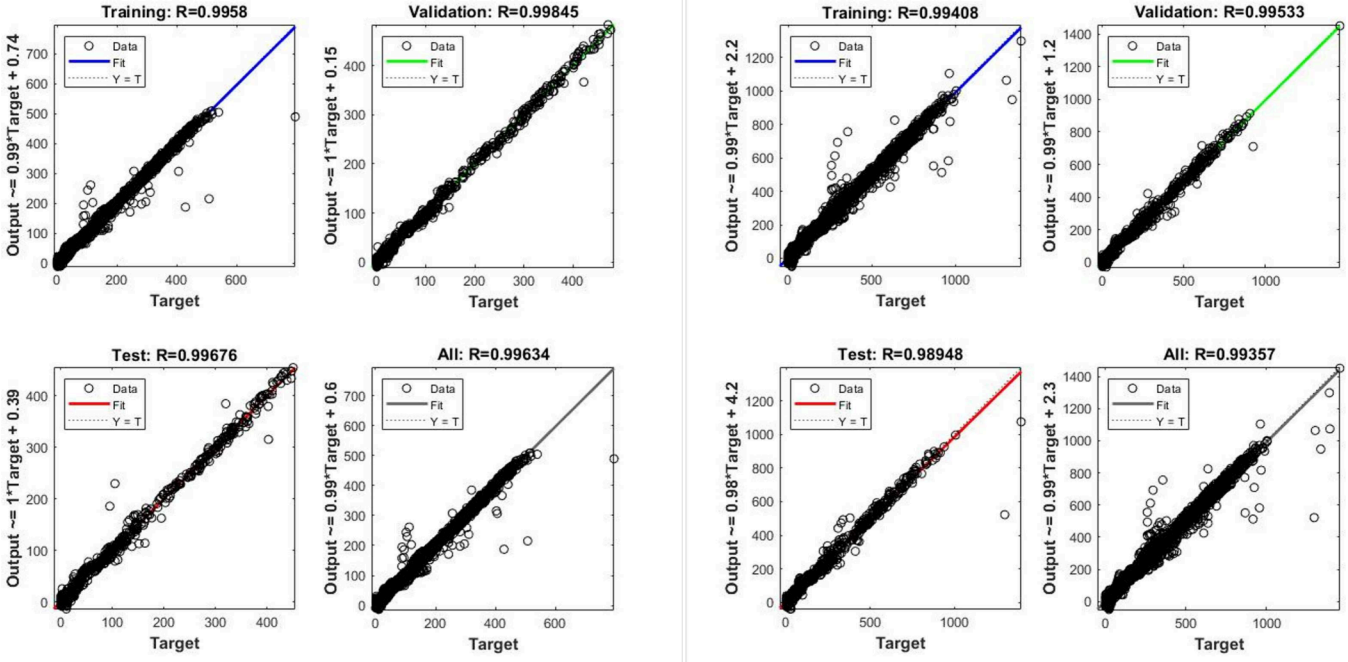
$$A_{LEL25\%} = \begin{bmatrix} -0.00661 & -0.02337 & -0.01734 & 0.008504 & -0.00827 & 0.096024 & 0.024178 & -0.05205 \\ -0.02337 & -2.01285 & 0.053684 & 0.054496 & 0.061372 & 0.072173 & -0.04543 & -0.04187 \\ -0.01734 & 0.053684 & -3.15689 & 0.074646 & 0.225314 & 0.162255 & -0.01782 & -0.09971 \\ 0.008504 & 0.054496 & 0.074646 & -0.01696 & -0.24462 & -0.04676 & 0.004752 & 0.016292 \\ -0.00827 & 0.061372 & 0.225314 & -0.24462 & 0.035535 & 0.133999 & -0.03851 & -0.01013 \\ 0.096024 & 0.072173 & 0.162255 & -0.04676 & 0.133999 & 0.118613 & 0.034141 & 0.024357 \\ 0.024178 & -0.04543 & -0.01782 & 0.004752 & -0.03851 & 0.034141 & -0.11406 & -0.02553 \\ -0.05205 & -0.04187 & -0.09971 & 0.016292 & -0.01013 & 0.024357 & -0.02553 & -0.18793 \end{bmatrix}$$

https://doi.org/10.1021/acsomega.4c05841
ACS Omega 2024, 9, 40857−40869



(a) LEL 100% ANN-BPNN QPCR model (b) LEL 25% ANN-BPNN QPCR model

Figure 3. Model validation performance for downwind distance: (a) LEL 100% ANN-BPNN QPCR model; (b) LEL 25% ANN-BPNN QPCR model.



(a) LEL 100% ANN-BPNN QPCR model (b) LEL 25% ANN-BPNN QPCR model

Figure 4. Regression model fitting, validation, and prediction accuracy performance for downwind distance: (a) LEL 100% ANN-BPNN QPCR model; (b) LEL 25% ANN-BPNN QPCR model. In each subfigure, the top left figure in blue is for training fitting, top right figure in green is for validation fitting, bottom left figure in red is for testing fitting, and bottom right figure in gray is the model fitting performance for all data sets.

https://doi.org/10.1021/acsomega.4c05841
ACS Omega 2024, 9, 40857−40869

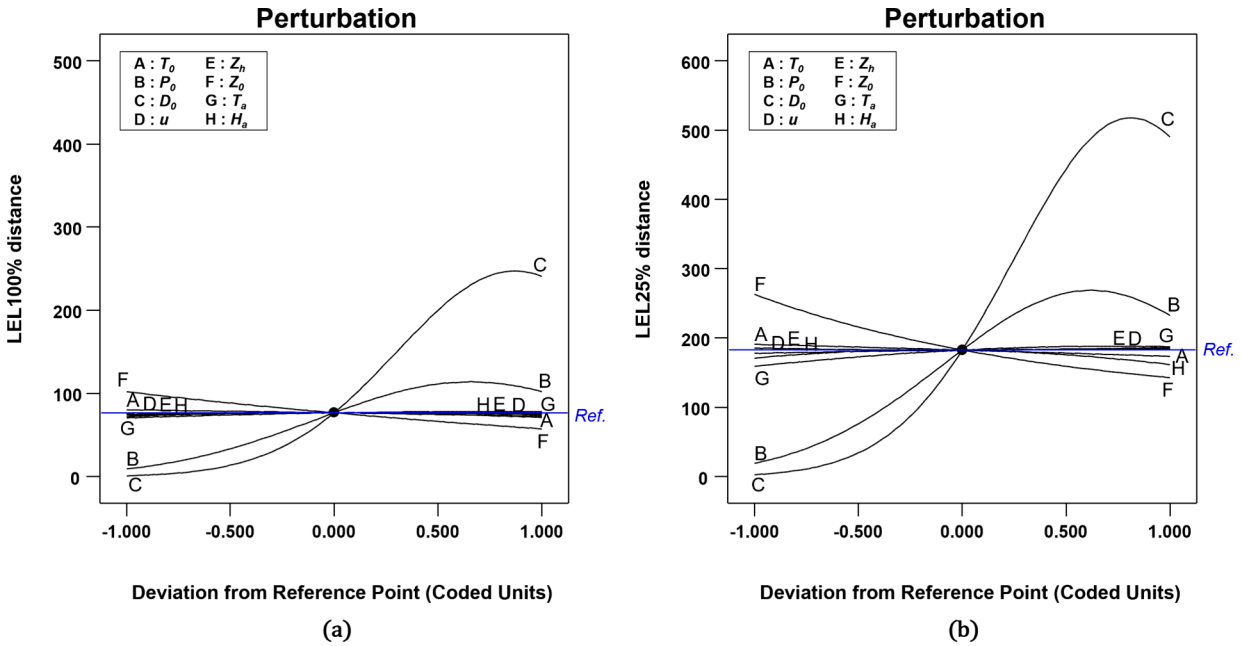


Figure 5. Perturbation plots for evaluating variable importance: (a) LEL 100% RSM-BBD QPCR model; (b) LEL 25% RSM-BBD QPCR model.

Table 3. ANOVA Results of QPCR Models Based on RSM-BBD

QPCR Model	Source	Code	Sum of squares	df	Mean Square	F value	P-value
LEL 100%	**Model**		**587.36**	**44**	**13.35**	**187.43**	**<0.0001**
	T_0	X_1	0.14	1	0.14	1.93	0.1692
	P_0	X_2	81.41	1	81.41	1143.09	<0.0001
	D_0	X_3	405.50	1	405.50	5693.56	<0.0001
	u	X_4	<0.0001	1	<0.0001	<0.0001	0.9225
	Z_h	X_5	<0.0001	1	<0.0001	0.016	0.9007
	Z_0	X_6	4.60	1	4.60	64.55	<0.0001
	T_a	X_7	0.14	1	0.14	2.03	0.1586
	H_a	X_8	<0.0001	1	<0.0001	0.060	0.8069
LEL 25%	**Model**		**562.80**	**44**	**12.79**	**233.61**	**<0.0001**
	T_0	X_1	0.13	1	0.13	2.29	0.1342
	P_0	X_2	87.09	1	87.09	1590.66	<0.0001
	D_0	X_3	368.36	1	368.36	6727.73	<0.0001
	u	X_4	0.019	1	0.019	0.35	0.5583
	Z_h	X_5	<0.0001	1	<0.0001	<0.0001	0.9413
	Z_0	X_6	5.19	1	5.19	94.84	<0.0001
	T_a	X_7	0.38	1	0.38	6.88	0.0105
	H_a	X_8	0.045	1	0.045	0.82	0.3695

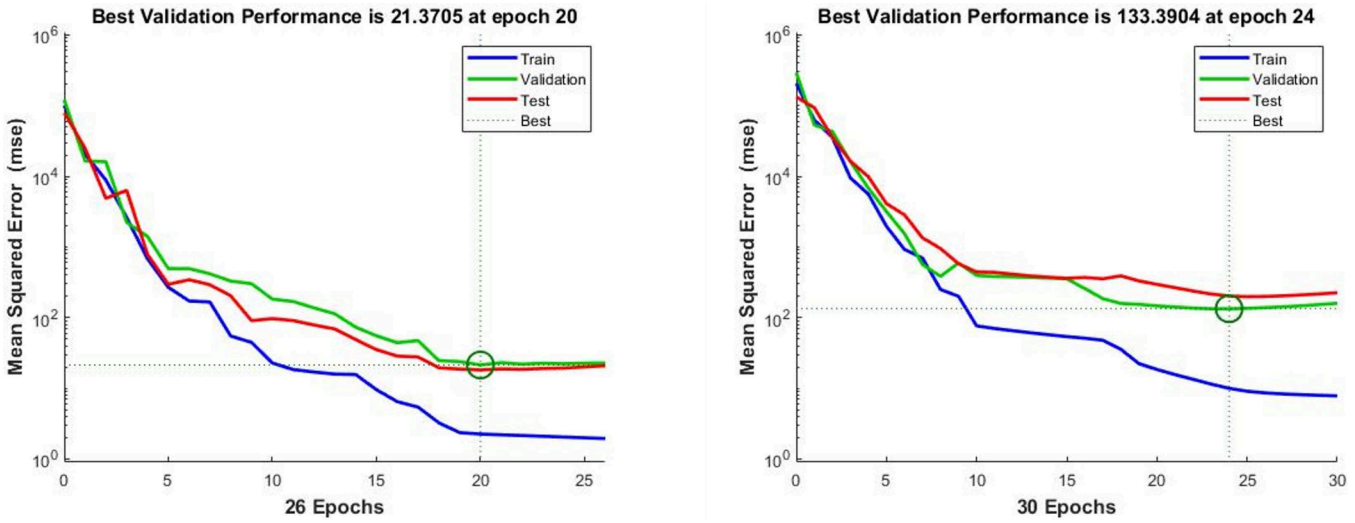
$$b_{\mathrm{LEL100\%}} = \begin{bmatrix} -0.0495 \\ 1.205726 \\ 2.690922 \\ 0.003483 \\ 0.004463 \\ -0.28653 \\ 0.050784 \\ -0.00875 \end{bmatrix}, \quad b_{\mathrm{LEL25\%}} = \begin{bmatrix} -0.04734 \\ 1.247092 \\ 2.564742 \\ 0.018387 \\ 0.00231 \\ -0.30451 \\ 0.082022 \\ -0.02823 \end{bmatrix}$$

$$c_{LEL100\%} = 4.304741,\ c_{LEL25\%} = 5.180752$$

Each of these parameters (X_1, X_2, ·, X_8) refers to coded descriptors, which are listed in Table 2. The aforementioned equations predict the response at given levels of each factor, and the influence of changes in these variables on the outcome may be proven when other factors are held constant, as indicated by regression coefficients. Diagnostics plots are another tool for evaluating the validity of the model.[51] Figure 2(a) and (b) show the normal probability distribution of residuals for the eight descriptors and the QPCR model, respectively, while (c) and (d) illustrate the distribution of data in relation to a straight-line normal distribution. The residuals conform to a linear pattern indicative of a normal distribution, with predicted values closely resembling actual values for all responses. Specifically, all values fall randomly inside the residual limits (−3.69651 to 3.69651), emphasizing the distribution's unpredictability. These findings support the use of the two QPCR models as prediction tools for assessing hydrogen leakage and dispersion.

ANN-BPNN QPCR Model. Each QPCR model utilized 6,561 (3^8) data points to construct its neural network. Among these, 4,593 data points (70%) were randomly selected for the

https://doi.org/10.1021/acsomega.4c05841
ACS Omega 2024, 9, 40857−40869



(a) LEL 100% hybrid QPCR model (b) LEL 25% hybrid QPCR model

Figure 6. Model validation performance for downwind distance: (a) LEL 100% Hybrid QPCR model; (b) LEL Hybrid 25% QPCR model.

training set, 984 data points (15%) for the validation set, and another 984 data points (15%) for the testing set. Performance plots for the developed QPCR models are presented in Figure 3 to visualize their MSE. As previously stated, the number of neurons in hidden layer was increased to 16, double the amount of input variables. As seen in Figure 3, the LEL 100% ANN-BPNN QPCR model (a) exhibits an MSE value of 45.5833 at 548 epochs, while the LEL 25% ANN-BPNN QPCR model (b) shows an MSE value of 480.2813 at 233 epochs. Both models demonstrate similar trends in their training, validation, and testing curves.

Figure 4 displays the models' training, validation, testing, and combined data sets. The overall predictive performance of the two models (LEL 100% ANN-BPNN QPCR model and LEL 25% ANN-BPNN QPCR model) can be considered excellent as the data points closely follow the output vs target fitting line. However, it is observed that some data points are sporadically positioned away from the fitting line. Nevertheless, the two models had R^2 values more than 0.99, indicating acceptable accuracy.

Hybrid QPCR Model. Based on the RSM-BBD results, an ANOVA was performed to determine the sensitivity of each variable. Figure 5 shows perturbation plots for each variable. The perturbation plots illustrate how input parameters affect output variables based on ANOVA results. These charts show the influence of each independent variable on the output variables, while the other input parameters remain constant at their baseline values. A steeper slope suggests more sensitivity. The research determined that D_0, P_0,and Z_0 had the largest influence on the output variables, in that order. Following these, T_a appeared to have a considerable impact; nonetheless, in Figure 5 (a), the difference in influence between T_a and the other four factors is minor.

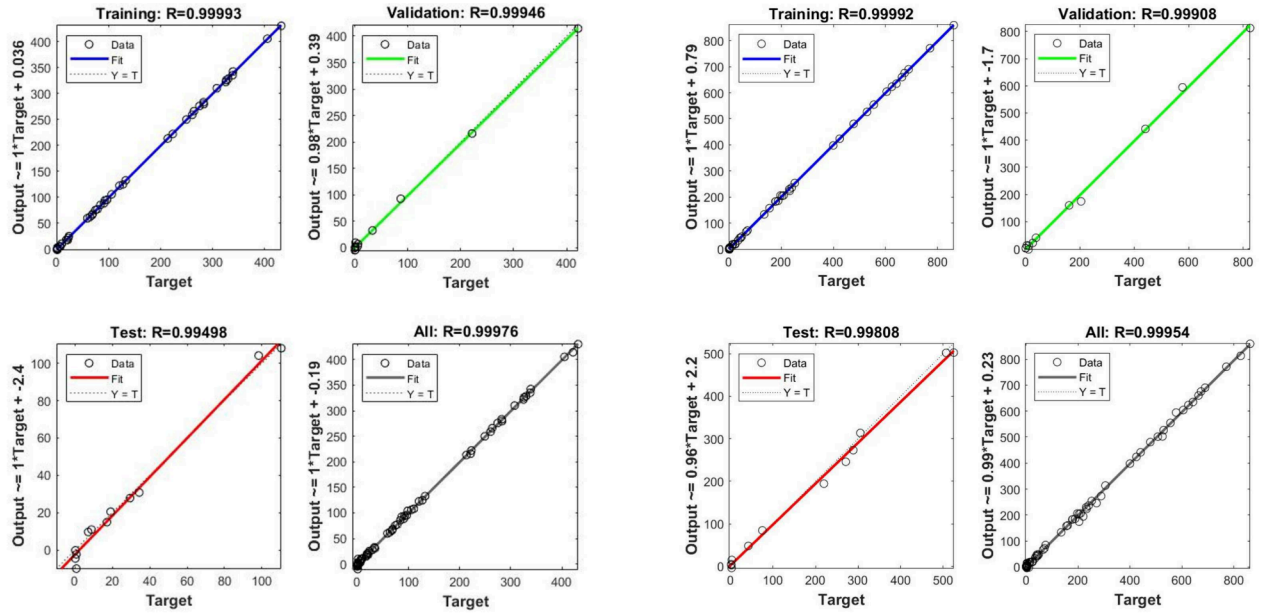
To conduct a more precise study, quantitative p-values for each variable were calculated and the findings are given in Table 3. A p-value of less than 0.05 indicates a substantial influence on the model's outputs.

In the LEL 100% QPCR model, variables P_0, D_0, and Z_0 have high F-values and p-values less than 0.0001, demonstrating their relevance. These findings are congruent with those shown in Figure 5.

P_0 and D_0 are the major factors defining the initial release momentum, which suppresses the buoyancy effect that occurs throughout the process of dispersion into the atmosphere, hence playing an important role in determining the hydrogen dispersion distance. Z_0 is a factor that influences the mixing level of the atmospheric layer by changing hydrogen's horizontal dispersion momentum into vertical momentum, lowering dispersion distance. The increase in T_a is attributable to changes in buoyancy and air viscosity, which can improve mixing with the atmosphere owing to the increased velocity of hydrogen molecules.[37] This can lead in a decrease in hydrogen dispersion. Notably, T_a had a p-value of 0.05 for predicting LEL 25%, demonstrating a high sensitivity in monitoring lower concentrations and thus was picked as a crucial variable.

The variables u, Z_h, H_a, and T_0 have p-values above 0.05, suggesting negligible effect on the results. In this study, u denotes the constant wind speed blowing in the same direction as the leak. Although a strong wind with a speed more than 10m/s has the greatest influence, at ordinary wind speeds of 2m/s, it behaves similarly to no wind circumstances.[52] Furthermore, the same wind direction as the leak direction can be ignored in the hydrogen dispersion process.[52] The temperature differential between the ground and the atmosphere, as well as the stratification in air temperature with height, impact the hydrogen dispersion distance according to Z_h.[53] But, because this study did not account for the atmospheric temperature gradient induced by solar radiation, the sensitivity of the output variable to Z_h is extremely low. Increased H_a may affect dispersion by diminishing hydrogen buoyancy. However, in the context of hydrogen dispersion, H_a is a significant component in vertical dispersion but has little effect on horizontal dispersion distance, which is the key variable in our study.[37,54] Lastly, T_0, the source property, affects the hydrogen leakage rates. Notwithstanding, given the conditions specified in this study, the impacts of P_0 and D_0 are so overwhelming that the T_0's significance is small. As a result, T_0 has a very minor influence on this study.[28]

https://doi.org/10.1021/acsomega.4c05841
ACS Omega 2024, 9, 40857−40869



(a) LEL 100% hybrid QPCR model (b) LEL 25% hybrid QPCR model

Figure 7. Regression model fitting, validation, and prediction accuracy performance for downwind distance: (a) LEL 100% hybrid QPCR model; (b) LEL 25% hybrid QPCR model. In each subfigure, the top left figure in blue is for training fitting, top right figure in green is for validation fitting, bottom left figure in red is for testing fitting, and bottom right figure in gray is the model fitting performance for all data sets.

Overall, ANOVA was conducted to evaluate the sensitivity of each variable, revealing that P_0, D_0, Z_0, and T_a exert dominant effects on the downwind distance of hydrogen. Based on this, a hybrid QPCR model considering only these four variables was developed, while the remaining four variables were set to their default values at Level 0. After modifying the data set with the four selected variables, the ANN technique was applied to develop the hybrid QPCR model, following the same approach as the ANN-BPNN QPCR model. The hidden layer consisted of eight neurons, twice the number of input variables. Out of 81 (3^4) data sets, 57 data points (70%) were allocated to the training set, 12 data points (15%) to the validation set, and another 12 data points (15%) to the test set.

Figures 6 and 7 evaluate the performance of the hybrid QPCR models. Figure 6 plots the MSE of the hybrid models. At 20 epochs, the MSE of the LEL 100% hybrid QPCR model was 21.3705, and at 24 epochs, it was 133.3904. Both models' training, validation, and testing curves followed a similar pattern.

In Figure 7, the distribution of 81 data points around the output vs target fitting line is illustrated. The two hybrid models are distributed across the fitting line, and the R^2 values of both models are close to 1.

This approach yields significant cost savings in simulations. While conducting 6,561 simulations would be necessary for a full factorial design of the original eight variables, excluding less sensitive variables through ANOVA reduces the number to 81 simulations. This reduced data set allows for the development of a hybrid QPCR model capable of predicting hydrogen's downwind distance, and the model's accuracy has been verified to be excellent. In essence, it is expected that by paying only 3% of the expense of a full factorial design, precise results will be attained.

The study compared and analyzed the RSM-BBD QPCR model, ANN-BPNN QPCR model, and hybrid QPCR model based on R^2, MSE, and data set size (Table 4).

Table 4. ANOVA Results of QPCR Models Based on RSM-BBD

QPCR Model		Parameter		
		R^2	MSE	Data points
LEL 100%	RSM-BBD	0.9910	480.8503	120
	ANN-BPNN	0.9963	104.1621	6,561
	Hybrid	0.9998	7.5075	201
LEL 25%	RSM-BBD	0.9928	1436.973	120
	ANN-BPNN	0.9936	715.5987	6,561
	Hybrid	0.9995	57.2931	201

From a statistical standpoint, when checking the predictive performance of each model, it is clear that the hybrid QPCR model, which combines RSM and ANN, has higher R^2 values than the RSM-BBD and ANN-BPNN models. Also, the MSE of the hybrid QPCR model is significantly smaller than the other values, indicating that it is the most accurate among the three models. Plus, the amount of data points required for the hybrid QPCR model was just 3% of that employed in the original full factorial design. This implies that by removing factors with little effects on the findings and focusing on key variables, reasonably good accuracy may be attained with a small data set.

The ANN-BPNN QPCR models benefit from the wide range of meteorological or operating conditions that may be examined using a full factorial design. Unlike the RSM-BBD QPCR model, which uses quadratic form to generalize data, the ANN-BPNN models predict data in a nonlinear fashion, leading to higher accuracy. On the other hand, the RSM-BBD QPCR model offers the advantage of developing models with a

https://doi.org/10.1021/acsomega.4c05841
ACS Omega 2024, 9, 40857−40869

R^2 values greater than 0.99 using the fewest number of data points, hence reducing simulation costs. In conclusion, all three models are deemed sophisticated and demonstrate sufficient performance in predicting the behavior of hydrogen leakage and dispersion.

Table 5 indicates a comparison between the Hybrid model developed in this study and the existing QPCR model from previous research. The comparison model was chosen as the QPCR model,[21] which predicts the dispersion of 41 flammable chemicals, including hydrogen. The QPCR model to be compared predicts the maximum downwind distance of the LEL, which is the same output variable used in this study.

Table 5. Comparison Results of Existing QPCR[21] and Hybrid QPCR Models

Categories		Units	QPCR model	
			Existing	Hybrid
Target chemicals	Chemicals quantity	–	41	1 (H_2)
Source property	Material quantity	m^3	10–4,000	Abundance
	Operating temperature	°C	−50–1,000	−40–65
	Operating pressure	barg	0.069–69	1–700
	Leakage hole diameter	mm	50.8–1,270	0.178–102
Dispersion property	Atmospheric temperature	°C	20	−20–40
	Wind speed	m/s	Unknown	1–6
	Release height	m	Unknown	0–10
	Surface roughness length	m	Unknown	0.005–3
	Relative humidity	%	Unknown	20–100
Physical property	LEL	%	0–100	–
	Vapor density (air = 1)	–	0.068–4.42	–
Data quantity		ea	19,579	201
R^2		–	0.9956	0.9998
MSE		–	0.0066	7.507

The existing QPCR model, constructed from previous research, successfully predicts the downwind distance of the LEL for a variety of flammable chemicals. This model, however, does not consider dispersion properties, limiting its application in real-world circumstances with a variety of atmospheric conditions. In contrast, the hybrid QPCR model developed in this work focuses solely on hydrogen. The development of this hydrogen-specific model allows for a thorough examination of dispersion properties, which were not completely accounted for in earlier studies, while not considering physical properties covered in prior studies.

Also, the hybrid model's design procedure, which involves variable exclusion using ANOVA, allows for the development of a predictive model that reflects a wide range of atmospheric conditions. This method enables the model to attain extremely high accuracy using a much less data set.

CONCLUSION

This work built three QPCR models capable of predicting hydrogen leakage and dispersion: the RSM-BBD QPCR model, the ANN-BPNN QPCR model, and the hybrid QPCR model. These models predict hydrogen leakage from generic high-pressure containers while including all of the essential variables that define the source and dispersion models. The three QPCR models have R^2 values over 0.99, suggesting their ability to reliably predict hydrogen leakage and dispersion.

Although the models developed in this research have the drawback of being particular to hydrogen, they successfully incorporate atmospheric dispersion properties that had previously been disregarded. Additionally, the development approach described in this paper shows how to design a model that accurately predicts hydrogen dispersion while modeling numerous accident scenarios at a low computational cost.

This approach has the benefit of being applicable to various chemicals. Once the model is built, it enables for the acquisition of consequence data suited to the individual characteristics of a facility, without the need for costly simulations or the employment of expert engineers. This is a key strength of the study. The developed QPCR model may be used into process safety management (PSM) systems in facilities that handle hazardous materials. One example of this integration is the application of the QPCR model as a quantitative risk assessment tool. This allows for direct quantitative assessment of leakage and dispersion results from hypothetical incidents, bypassing the necessity for additional modeling procedures.

Additionally, there is a need to make bold decisions, but it can be challenging for experts to develop plans, such as emergency response strategies in the chemical industry, where considering uncertainties is crucial for risk assessment. In such cases, the QPCR model can be utilized as a key component of expert decision-making by providing causal relationships and quantitative predictions based on various variables.

Moreover, statistical approaches such as RSM and ANN played an important part in the creation of the QPCR model. It means that different statistical approaches, such as machine learning algorithms, may be used to develop QPCR models adequately. In the future, building QPCR models with recurrent neural network such as Long Short-Term Memory (LSTM) algorithm might allow for the consideration of changes in leakage rates over time. The approach based on machine learning algorithms is intended to improve the model's robustness and broaden its possible uses.

AUTHOR INFORMATION

Corresponding Author

Byungchol Ma − *School of Chemical Engineering, Chonnam National University, 61186 Gwangju, Republic of Korea;* orcid.org/0000-0003-3158-6940; Phone: +82 62 530 1815; Email: anjeon@jnu.ac.kr; Fax: +82 62 530 1814

Authors

Junseo Lee − *School of Chemical Engineering, Chonnam National University, 61186 Gwangju, Republic of Korea;* orcid.org/0000-0002-9325-5445

Sehyeon Oh − *School of Chemical Engineering, Chonnam National University, 61186 Gwangju, Republic of Korea*

Seongheon Baek − *School of Chemical Engineering, Chonnam National University, 61186 Gwangju, Republic of Korea*

Changbock Chung − *SD Plus Corporation, 61186 Gwangju, Republic of Korea*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.4c05841

https://doi.org/10.1021/acsomega.4c05841
ACS Omega 2024, 9, 40857−40869

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was supported by "Regional Innovation Strategy (RIS)" through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (MOE) (2021RIS-002). The authors thankfully acknowledge the National Research Foundation of Korea (NRF) funded by the Ministry of Education (MOE), for providing adequate funds to procure equipment and consumables.

■ REFERENCES

(1) Norazahar, N.; Ambikabathy, T. M.; Kasmani, R. M.; Ahmad, A.; Jalil, A. A.; Abdullah, T. A. T.; Kamaroddin, M. F. A. Hydrogen application and its safety: An overview of public perceptions and acceptance in Malaysia. *Process Safety and Environmental Protection* **2023**, *180*, 686−698.

(2) Yang, N.; Deng, J.; Wang, C.; Bai, Z.; Qu, J. High pressure hydrogen leakage diffusion: Research progress. *Int. J. Hydrogen Energy* **2024**, *50*, 1029−1046.

(3) Abohamzeh, E.; Salehi, F.; Sheikholeslami, M.; Abbassi, R.; Khan, F. Review of hydrogen safety during storage, transmission, and applications processes. *Journal of Loss Prevention in the Process Industries* **2021**, *72*, 104569.

(4) Park, B.; Kim, Y.; Paik, S.; Kang, C. Numerical and experimental analysis of jet release and jet flame length for qualitative risk analysis at hydrogen refueling station. *Process Safety and Environmental Protection* **2021**, *155*, 145−154.

(5) Cao, W.; Li, W.; Yu, S.; Zhang, Y.; Shu, C. M.; Liu, Y.; Luo, J.; Bu, L.; Tan, Y. Explosion venting hazards of temperature effects and pressure characteristics for premixed hydrogen-air mixtures in a spherical container. *Fuel* **2021**, *290*, 120034.

(6) Jia, Z.; Wang, S.; Qin, P.; Li, C.; Song, L.; Cheng, Z.; Jin, K.; Sun, J.; Wang, Q. Comparative investigation of the thermal runaway and gas venting behaviors of large-format LiFePO4 batteries caused by overcharging and overheating. *Journal of Energy Storage* **2023**, *61*, 106791.

(7) Yang, X.; Wang, H.; Li, M.; Li, Y.; Li, C.; Zhang, Y.; Chen, S.; Shen, H.; Qian, F.; Feng, X.; Ouyang, M. Experimental Study on Thermal Runaway Behavior of Lithium-Ion Battery and Analysis of Combustible Limit of Gas Production. *Batteries* **2022**, *8* (11), 250.

(8) Zhang, Y.; Wang, H.; Li, W.; Li, C. Quantitative identification of emissions from abused prismatic Ni-rich lithium-ion batteries. *eTransportation* **2019**, *2*, 100031.

(9) Guo, L.; Su, J.; Wang, Z.; Shi, J.; Guan, X.; Cao, W.; Ou, Z. Hydrogen safety: An obstacle that must be overcome on the road towards future hydrogen economy. *Int. J. Hydrogen Energy* **2024**, *51*, 1055−1078.

(10) Ma, H.; Wang, C.; Li, Y.; Li, W.; Ding, J.; Zhao, W. Experimental study on the effects of ignition position and inhomogeneous concentration on vented hydrogen deflagrations in a 7 m3 chamber. *Process Safety and Environmental Protection* **2024**, *182*, 579−594.

(11) Pula, R.; Khan, F. I.; Veitch, B.; Amyotte, P. R. A grid based approach for fire and explosion consequence analysis. *Process Safety and Environmental Protection* **2006**, *84*, 79−91.

(12) Al-shanini, A.; Ahmad, A.; Khan, F. Accident modelling and analysis in process industries. *Journal of Loss Prevention in the Process Industries* **2014**, *32*, 319.

(13) Jiao, Z.; Zhang, Z.; Jung, S.; Wang, Q. Machine learning based quantitative consequence prediction models for toxic dispersion casualty. *Journal of Loss Prevention in the Process Industries* **2023**, *81*, 104952.

(14) Li, J.; Li, Q.; Hao, H.; Li, L. Prediction of BLEVE blast loading using CFD and artificial neural network. *Process Safety and Environmental Protection* **2021**, *149*, 711−723.

(15) Witlox, H. W.; Fernandez, M.; Harper, M.; Stene, J. Modelling and validation of atmospheric expansion and near-field dispersion for pressurised vapour or two-phase releases. *Journal of Loss Prevention in the Process Industries* **2017**, *48*, 331−344.

(16) Witlox, H. W.; Fernandez, M.; Harper, M.; Oke, A.; Stene, J.; Xu, Y. Verification and validation of Phast consequence models for accidental releases of toxic or flammable chemicals to the atmosphere. *Journal of Loss Prevention in the Process Industries* **2018**, *55*, 457−470.

(17) Ganci, F.; Carpignano, A.; Mattei, N.; Carcassi, M. N. Hydrogen release and atmospheric dispersion: Experimental studies and comparison with parametric simulations. *Int. J. Hydrogen Energy* **2011**, *36*, 2445−2454.

(18) Qiu, S.; Chen, B.; Wang, R.; Zhu, Z.; Wang, Y.; Qiu, X. Atmospheric dispersion prediction and source estimation of hazardous gas using artificial neural network, particle swarm optimization and expectation maximization. *Atmos. Environ.* **2018**, *178*, 158−163.

(19) Wang, B.; Chen, B.; Zhao, J. The real-time estimation of hazardous gas dispersion by the integration of gas detectors, neural network and gas dispersion models. *Journal of Hazardous Materials* **2015**, *300*, 433−442.

(20) Sun, Y.; Wang, J.; Zhu, W.; Yuan, S.; Hong, Y.; Mannan, M. S.; Wilhite, B. Development of Consequent Models for Three Categories of Fire through Artificial Neural Networks. *Ind. Eng. Chem. Res.* **2020**, *59*, 464−474.

(21) Jiao, Z.; Sun, Y.; Hong, Y.; Parker, T.; Hu, P.; Mannan, M. S.; Wang, Q. Development of flammable dispersion quantitative property-consequence relationship models using extreme gradient boosting. *Ind. Eng. Chem. Res.* **2020**, *59*, 15109−15118.

(22) Jiao, Z.; Ji, C.; Sun, Y.; Hong, Y.; Wang, Q. Deep learning based quantitative property-consequence relationship (QPCR) models for toxic dispersion prediction. *Process Safety and Environmental Protection* **2021**, *152*, 352−360.

(23) Li, Y.; Wang, Z.; Shang, Z. Analysis and prediction of hydrogen-blended natural gas diffusion from various pipeline leakage sources based on CFD and ANN approach. *Int. J. Hydrogen Energy* **2024**, *53*, 535−549.

(24) Chen, W. H.; Chang, C. M.; Mutuku, J. K.; Lam, S. S.; Lee, W. J. Analysis of microparticle deposition in the human lung by taguchi method and response surface methodology. *Environmental Research* **2021**, *197*, 110975.

(25) Chen, W. H.; Carrera Uribe, M.; Kwon, E. E.; Lin, K. Y. A.; Park, Y. K.; Ding, L.; Saw, L. H. A comprehensive review of thermoelectric generation optimization by statistical approach: Taguchi method, analysis of variance (ANOVA), and response surface methodology (RSM). *Renewable and Sustainable Energy Reviews* **2022**, *169*, 112917.

(26) Oh, S.; Lee, J.; Ma, B. Methodology for optimally designing hydrogen refueling station barriers using RSM and ANN: Considering explosion and jet fire. *Int. J. Hydrogen Energy* **2024**, *80*, 234−248.

(27) Shi, J.; Li, J.; Zhu, Y.; Hao, H.; Chen, G.; Xie, B. A simplified statistic-based procedure for gas dispersion prediction of fixed offshore platform. *Process Safety and Environmental Protection* **2018**, *114*, 48−63.

(28) Mousavi, J.; Parvini, M. A sensitivity analysis of parameters affecting the hydrogen release and dispersion using ANOVA method. *Int. J. Hydrogen Energy* **2016**, *41*, 5188−5201.

(29) Jiao, Z.; Hu, P.; Xu, H.; Wang, Q. Machine Learning and Deep Learning in Chemical Health and Safety: A Systematic Review of Techniques and Applications. *ACS Chem. Health Saf.* **2020**, *27*, 316.

(30) Yazdi, M.; Zarei, E.; Adumene, S.; Abbassi, R.; Rahnamayiezekavat, P. In *Methods to Assess and Manage Process Safety in Digitalized Process System*; Khan, F., Pasman, H., Yang, M., Eds.; Methods in Chemical Process Safety; Elsevier: 2022; Vol. *6*; pp 389−416.

(31) Kendall, A.; Gal, Y. What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision? *arXiv*, 2017; https://arxiv.org/abs/1703.04977.

https://doi.org/10.1021/acsomega.4c05841
ACS Omega 2024, 9, 40857−40869

(32) Bjarnadottir, S., Li, Y., Stewart, M. G. In *Climate Adaptation Engineering*; Bastidas-Arteaga, E., Stewar, M. G., Eds.; Butterworth-Heinemann: 2019; pp 271−299.
(33) Pandya, N.; Gabas, N.; Marsden, E. Sensitivity analysis of Phast's atmospheric dispersion model for three toxic materials (nitric oxide, ammonia, chlorine). *Journal of Loss Prevention in the Process Industries* **2012**, *25*, 20−32.
(34) Crowl, D. A.; Louvar, J. F. *Chemical process safety: fundamentals with applications*; Pearson Education: 2001.
(35) Díaz-Ovalle, C.; Vázquez-Román, R.; Sam Mannan, M. An approach to solve the facility layout problem based on the worst-case scenario. *Journal of Loss Prevention in the Process Industries* **2010**, *23*, 385−392.
(36) Gifford, F. A. Use of routine meteorological observations for estimating atmospheric dispersion. *Nucl. Safety* **1961**, *2*, 47−51.
(37) Shu, Z.; Liang, W.; Zheng, X.; Lei, G.; Cao, P.; Dai, W.; Qian, H. Dispersion characteristics of hydrogen leakage: Comparing the prediction model with the experiment. *Energy* **2021**, *236*, 121420.
(38) Shu, Z.; Lei, G.; Liang, W.; Dai, W.; Lu, F.; Zheng, X.; Qian, H. Experimental investigation of hydrogen dispersion characteristics with liquid helium spills in moist air. *Process Safety and Environmental Protection* **2022**, *162*, 923−931.
(39) Le, S. T.; Nguyen, T. N.; Linforth, S.; Ngo, T. D. Safety investigation of hydrogen energy storage systems using quantitative risk assessment. *Int. J. Hydrogen Energy* **2023**, *48*, 2861−2875.
(40) Hanna, S. Meteorological data recommendations for input to dispersion models applied to Jack Rabbit II trials. *Atmos. Environ.* **2020**, *235*, 117516.
(41) Nair, S. R. R.; Ogbeifun, N. O.; Wen, J. Consequence assessment considerations for toxic natural gas dispersion modeling. *Journal of Loss Prevention in the Process Industries* **2022**, *78*, 104792.
(42) Genovese, M.; Fragiacomo, P. Hydrogen refueling station: Overview of the technological status and research enhancement. *Journal of Energy Storage* **2023**, *61*, 106758.
(43) Genovese, M.; Cigolotti, V.; Jannelli, E.; Fragiacomo, P. Hydrogen Refueling Process: Theory, Modeling, and In-Force Applications. *Energies* **2023**, *16*, 2890.
(44) *API RP 754 - Process Safety Performance Indicators for the Refining and Petrochemical Industries. ANSI/API Recommended Practice*; 2016.
(45) Hossain, M. S.; Rahim, N. A.; Aman, M. M.; Selvaraj, J. Application of ANOVA method to study solar energy for hydrogen production. *Int. J. Hydrogen Energy* **2019**, *44*, 14571−14579.
(46) Naderi, K.; Foroughi, A.; Ghaemi, A. Analysis of hydraulic performance in a structured packing column for air/water system: RSM and ANN modeling. *Chemical Engineering and Processing - Process Intensification* **2023**, *193*, 109521.
(47) Wang, R.; Chen, B.; Qiu, S.; Zhu, Z.; Wang, Y.; Wang, Y.; Qiu, X. Comparison of machine learning models for hazardous gas dispersion prediction in field cases. *International Journal of Environmental Research and Public Health* **2018**, *15*, 1450.
(48) Wang, Y.; Chen, B.; Zhu, Z.; Wang, R.; Chen, F.; Zhao, Y.; Zhang, L. A hybrid strategy on combining different optimization algorithms for hazardous gas source term estimation in field cases. *Process Safety and Environmental Protection* **2020**, *138*, 27−38.
(49) Chen, L.; Wu, T.; Wang, Z.; Lin, X.; Cai, Y. A novel hybrid BPNN model based on adaptive evolutionary Artificial Bee Colony Algorithm for water quality index prediction. *Ecological Indicators* **2023**, *146*, 109882.
(50) Olivier, E.; Eldridge, R. B. Prediction of the trayed distillation column mass-transfer performance by neural networks. *Ind. Eng. Chem. Res.* **2002**, *41*, 3436−3446.
(51) Pashaei, H.; Mashhadimoslem, H.; Ghaemi, A. Modeling and optimization of CO2 mass transfer flux into Pz-KOH-CO2 system using RSM and ANN. *Sci. Rep.* **2023**, *13*, 4011.
(52) Li, X.-j.; Xu, Y.-x.; Li, X.; Jin, Z.-j.; Qian, J.-y. Effect of wind condition on unintended hydrogen release in a hydrogen refueling station. *Int. J. Hydrogen Energy* **2021**, *46*, 5537−5547.
(53) Sun, R.; Pu, L.; Yu, H.; Dai, M.; Li, Y. Modeling the diffusion of flammable hydrogen cloud under different liquid hydrogen leakage conditions in a hydrogen refueling station. *Int. J. Hydrogen Energy* **2022**, *47*, 25849−25863.
(54) Liu, Y.; Wei, J.; Lei, G.; Wang, T.; Lan, Y.; Chen, H.; Jin, T. Modeling the development of hydrogen vapor cloud considering the presence of air humidity. *Int. J. Hydrogen Energy* **2019**, *44*, 2059−2068.

https://doi.org/10.1021/acsomega.4c05841
ACS Omega 2024, 9, 40857−40869