



Original Paper

Hydrate phase equilibria in oil-containing fluid: Experimental investigation and machine learning prediction based stacked ensemble



Bing-Yue Han^a, Qing-Yang Yuan^a, Jie Wang^a, Peng-Cheng Li^b, Zheng Ling^a, Lei Yang^a, Yu Liu^a, Yong-Chen Song^a, Lun-Xiang Zhang^{a,*}

^a Key Laboratory of Ocean Energy Utilization and Energy Conservation of Ministry of Education, Dalian University of Technology, Dalian, 116024, Liaoning, China

^b CNOOC Research Institute Ltd, Beijing, 100028, China

ARTICLE INFO

Article history:

Received 4 April 2025

Received in revised form

30 September 2025

Accepted 16 November 2025

Available online 21 November 2025

Edited by Teng Zhu and Min Li

Keywords:

Hydrate phase equilibrium

Flow assurance

Complex oil composition

Model prediction

Inhibitors

Machine learning

ABSTRACT

During oil and gas development and transportation, the presence of oil and complex component in natural gas significantly impacts the hydrate phase equilibrium. However, the specific characteristics of hydrate phase equilibrium in natural gas systems containing oil have not been clearly defined. To address this gap, this study systematically investigated the hydrate phase equilibrium of two natural gas components with oil contents of 40 vol%, 50 vol%, and 60 vol% under pressures ranging from 2 to 8 MPa, based on the actual components from the South China Sea oil and gas fields. Experimental results showed that the oil phase shifted the hydrate phase equilibrium curve towards lower temperatures and higher pressures, with the most significant suppression at 50 vol% oil content. Furthermore, the oil phase had a stronger suppressive effect on hydrates in natural gas containing heavier hydrocarbons (e.g., butane and pentane), with these effects becoming more pronounced at higher pressures. The study indicated that neglecting the oil phase or using simplified natural gas components for hydrate phase equilibrium predictions and inhibitor dosage calculations could have led to overuse of inhibitors, increasing production costs. Based on the experimental data, this study proposed a stacked ensemble model (SELM-DCC) combining multiple POD-RBF base learners and decision coefficients for predicting hydrate phase equilibrium in oil-containing systems. Compared to six other prediction methods, the SELM-DCC model demonstrated lower maximum relative error, enhanced stability, and superior denoising performance. The findings provide practical insights for the study of hydrate phase equilibrium in oil-containing systems and offer a solid framework for model development. Additionally, the proposed model serves as an accurate and versatile predictive tool for the intelligent regulation of inhibitors, offering significant engineering application value.

© 2025 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Natural gas hydrates are ice-like crystals formed from water molecules and gas molecules under low-temperature and high-pressure conditions (Hassanpouryouzband et al., 2020; Chen et al., 2023; Sun et al., 2024). Natural gas hydrates possess a high

energy density, which makes them highly promising for energy supply (Hassanpouryouzband et al., 2020; Yang et al., 2024). Additionally, hydrate-based technologies have been proposed and developed as effective solutions for gas storage (Zhang et al., 2017; Park et al., 2023), wastewater treatment (Geo Lim et al., 2023; Sun et al., 2025), and carbon capture, utilization, and storage (CCUS) (Wang et al., 2025b; Wu et al., 2025). However, natural gas hydrate formation and blockage in oil and gas pipelines are considered some of the most challenging issues in the modern petroleum industry (Sloan, 2003, 2005, 2010; Khan et al., 2023; Liu et al., 2025). Unintended hydrate formation in wellbores and transport

* Corresponding author.

E-mail address: lunxiangzhang@dlut.edu.cn (L.-X. Zhang).

Peer review under the responsibility of China University of Petroleum (Beijing).

pipelines can significantly reduce production, increase the risk of accidents, and potentially cause severe environmental impacts (Liu et al., 2018, 2021; Shi et al., 2021; Wang et al., 2024). Therefore, flow assurance for hydrate blockages has received extensive attention in both scientific research and engineering applications (Lang et al., 2024; Meng et al., 2024; Zhong et al., 2024; Wang et al., 2025a).

The current strategies to mitigate hydrate blockages mainly include real-time monitoring of temperature and pressure (Chee et al., 2014; Wang et al., 2022; Zhang et al., 2022), early detection via wave techniques (Liu et al., 2023; Meng et al., 2024), and the injection of chemical inhibitors (Semenov et al., 2021; Gong et al., 2022; Li et al., 2022b). However, the successful implementation of these methods relies heavily on a thorough understanding of hydrate formation characteristics and accurate predictions of hydrate phase equilibrium. Predicting hydrate formation remains challenging, especially in multiphase systems containing liquid hydrocarbons. In oil and gas development and transportation, the liquid hydrocarbons (oil) present in natural gas, along with their complex component (e.g., methane, ethane, propane, and other hydrocarbons), significantly affect hydrate phase equilibrium (Seo et al., 2009; Dec 2012; Kwon et al., 2014; Shi et al., 2020). While most existing studies focus on simple gas systems with salts or inhibitors, limited attention has been given to liquid hydrocarbon-containing systems (Cha et al., 2016; Khan et al., 2018; Kou and Li, 2019; Sun et al., 2023). Therefore, it is essential to systematically measure and analyze the impact of natural gas component and oil content on hydrate phase equilibrium through experimental methods.

However, traditional methods of data analysis alone are insufficient to fully uncover the phase equilibrium patterns of hydrates under the combined influence of multiple factors (Garapati and Anderson, 2014; Du et al., 2019; Jeong et al., 2022). With the rapid advancement of artificial intelligence and big data technologies, machine learning approaches have shown tremendous potential in handling complex nonlinear problems. In recent years, an increasing number of studies have focused on using machine learning techniques to predict conditions for gas hydrate formation (Hosseini and Leonenko, 2023, 2024; Alavi and Sharifzadeh, 2024; Hosseini et al., 2025). Mesbah et al. developed a model based on the Least Squares Support Vector Machine (LSSVM) algorithm, utilizing 279 experimental data points to forecast a wide range of natural gas hydrate formation temperatures. This model achieved high accuracy with a coefficient of determination (R^2) of 0.9918 (Mesbah et al., 2017). Amar employed a Gene Expression Programming (GEP) model to simulate natural gas hydrate formation temperatures, utilizing a database comprising 279 experimental data points. The established GEP model accurately estimated hydrate formation temperatures with an average absolute relative error (AARE) of 0.1397% (Nait Amar, 2021). However, the precision of these methods relies on the volume of data used during algorithm model construction, and acquiring a satisfactory volume of data through experiments and engineering entails significant time costs. Additionally, reliance on a singular model can reduce prediction stability and increase the likelihood of outlier predictions.

This work systematically measured and discussed the effects of different oil contents and gas components on hydrate phase equilibrium using a self-built high-pressure experimental system. To ensure the reliability of the results and minimize experimental errors, each test condition was repeated at least three times, with a total of nearly 200 experiments conducted. To further enrich the phase equilibrium dataset for complex oil-containing hydrate systems and apply these findings to phase equilibrium predictions in practical engineering, experimental data were used to train

multiple basic predictive models and learners. In designing these base learners, a POD-RBF model based on a reduced-order model was employed to mitigate the dependence on data volume and noise contamination while enhancing the computational efficiency of individual base learners. These base learners were then combined using a stacked ensemble approach to form an ensemble learning model, which incorporated a decision coefficient to further improve its accuracy. This stacked ensemble learning model enriches and extends the phase equilibrium data of complex systems containing oil phases to some extent, thereby providing a new theoretical basis and technical support for the safety assurance of oil and gas transportation pipelines.

2. Experimental methods and materials

2.1. Materials

Two natural gas mixtures used in this work were supplied by Air Liquide (Dalian Special Gas Co., Ltd., China). And the natural gas components and molar ratio for these two gas mixtures shown in Tables 1 and 2 were determined with reference to the actual working conditions of the oil and gas fields in the South China Sea. The two liquid phases used to form hydrate were deionized water and mineral oil. The deionized water with a resistivity of 18.2 M Ω -cm was prepared using an Aqua-plore 2S system (Aqua-pro International Co., LLC, USA). The mineral oil with a specific gravity of 0.822 at 15 °C and a kinematic viscosity of 5 Cst at 40 °C was purchased from a commercial company (Kunshan Siemo Lubrication Technology Co., Ltd., China).

2.2. Experimental apparatus

The schematic illustration of the experimental setup for hydrate phase equilibria is shown in Fig. 1. The major experimental system was a high-pressure cylindrical metal reactor cell with two circular 25 mm diameter sapphire viewing windows. The reactor could work safely under the maximum pressure of 10 MPa, and the volume of it was the 230 mL.

The thermocouples and pressure sensors were applied to detect the temperature and pressure data of the formation and dissociation of natural gas hydrates in real time, and these data were transmitted and stored in the data collection system simultaneously for further analysis. In addition, the accuracy of thermocouples and pressure sensors was 0.01 K and 0.025 MPa respectively. The target temperature was controlled by a constant-temperature water bath (FP51, Temperature Control Company, Germany; the device was operated at 0–1200 rpm and –30 °C–100 °C, with a temperature control accuracy of 0.01 °C), which contained a mixture of commercial ethylene glycol and water for cooling.

Table 1
The natural gas components and molar ratio for Component 1.

Natural gas components	Molar ratio, %
Methane	86.9
Ethane	5.10
Propane	2.04
i-Butane	0.54
n-Butane	0.57
i-Pentane	0.28
Carbon dioxide	3.73
Nitrogen	0.84

Table 2
The natural gas components and molar ratio for Component 2.

Natural gas components	Molar ratio, %
Methane	88.29
Ethane	5.10
Propane	2.04
Carbon dioxide	3.73
Nitrogen	0.84

2.3. Experimental procedure

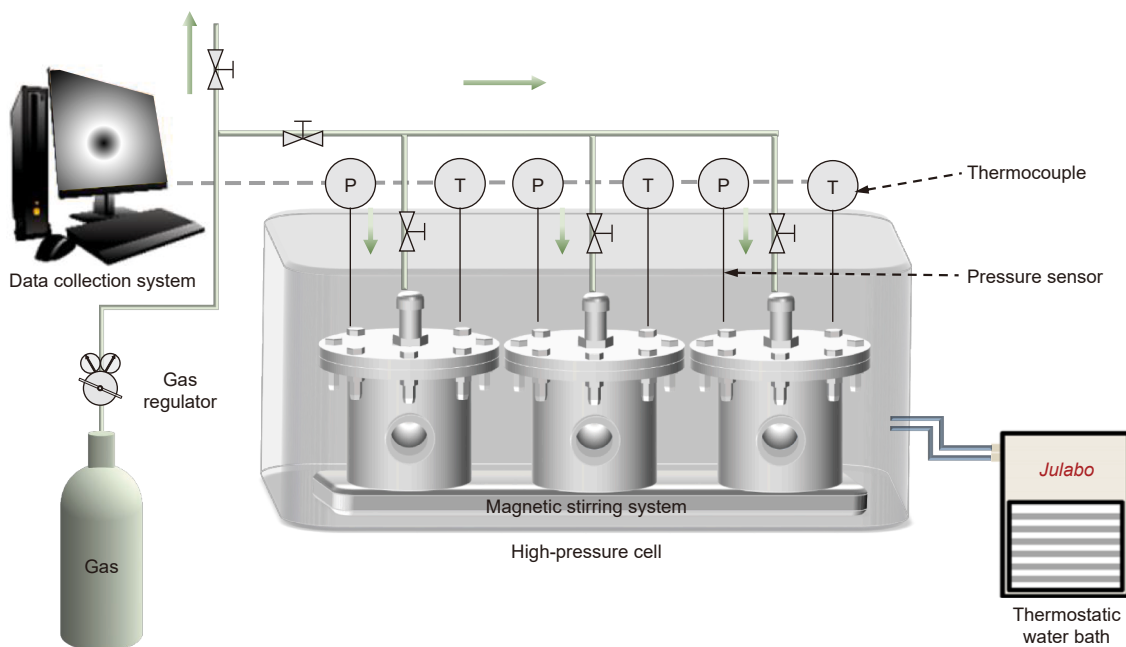


Fig. 1. Schematic illustration of the experimental system for hydrate phase equilibria.

The high pressure reaction cell was cleaned with deionized water at least three times before each experiment. Then the dry reaction cell was filled with 130 mL of the oil-water mixture and sealed. The system was flushed with experimental gas three times to remove air and impurities. To ensure a sufficient mixing of the gas-liquid phase, the rotor in the reaction cell was rotated at 250 rpm. At ambient temperature (about 298 K), the experimental gas was slowly injected until the pressure reached the target pressure (2–8 MPa). The experimental system was kept at ambient temperature for 2 h to make the temperature and pressure in the reaction cell stabilize before starting the experiment.

The phase equilibrium points of natural gas hydrate in an oil-water mixture were measured via an isochoric method, that is, keeping the internal volume of the reaction cell constant while varying the other parameters to form or dissociate hydrates (Li et al., 2022a; Men et al., 2022). The variations of pressure and temperature were recorded in the entire process of the experiment. This credible method has been applied and proven in a mass of previous research (Li et al., 2020b; Park et al., 2022). The specific steps of the experiment were as follows: (i) cooling from the initial temperature to 275.16 K at a rate of 2 K/h, (ii) the system was maintained at 275.16 K for 10 h, during which hydrate formation was observed (determined by pressure drop and temperature rise) and the system eventually reaches a steady state (almost no change in pressure of the system), (iii) the system was heated at a rate of 1 K/h until the temperature was within 2 K of the

prospective phase equilibrium point (predicted by CMSEm), and (iv) the system was slowly heated at 0.1 K/h to dissociate hydrates and return to its initial state. The whole process was shown in Fig. 2. The intersection of the cooling curve and heating curve was the phase equilibrium point of hydrate (Sun et al., 2015). It should be noted that sometimes hydrates formed during the cooling period, as the working conditions were still under the hydrate stable region, and the heating curve still intersected with the expansion/contraction line of the gas after hydrates completely dissociated.

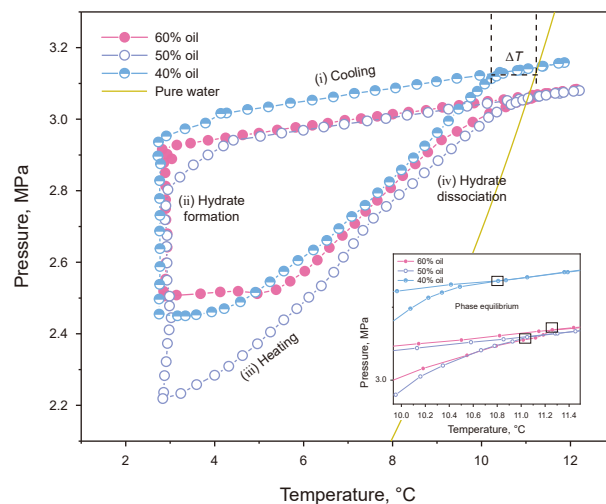


Fig. 2. A typical temperature-pressure curve for determining the hydrate phase equilibrium point in an oil-water system using the isochoric method. The yellow solid line indicates the hydrate equilibrium curve in the pure water system. ΔT denotes the phase equilibrium temperature depression of the hydrate in oil-containing systems relative to that in pure water at the same pressure.

3. Modeling method

The POD-RBF method is a radial basis function neural network based on a reduced order model. This method can combine the feature extraction ability of the reduced order model with the nonlinear fitting ability of the RBF neural network, ensuring high prediction accuracy and generalization ability while miniaturizing the neural network model. At the same time, the POD reduction model itself has a certain degree of noise reduction ability, which can weaken the impact of strong noise interference in the data on the predictive ability of the machine learning model. A single machine learning model may not only exhibit underfitting in cases of insufficient training but also often exhibit poor generalization ability. In this paper, we will construct an ensemble learning model that includes multiple POD-RBF methods as base learners to compensate for the shortcomings in stability, prediction accuracy, and generalization ability of a single machine learning model.

3.1. Proper orthogonal decomposition method

The Proper orthogonal decomposition (POD) method provides a set of orthogonal bases containing the main feature information of the described system. The linear combination of bases can approximate the system, thereby transforming a high-dimensional problem into a low dimensional problem. The “snapshot POD method” proposed by Sirovich (1987) was usually used to reduce the dimensionality of a database. The specific process was as follows.

- (1) Construct a snapshot matrix ensemble \mathbf{D} consisting of M parameter combinations and N moments of data sets. Assuming that the system parameter was denoted by SP (system parameter), and the time variable was denoted by t , where the i -th way of combining the system parameter, the snapshot data column at the q -th moment was denoted as $D(SP_i, t_q)$. Then the matrix form of the sample that collects all the locations of the measurement points can be expressed as a collection of multiple snapshot data columns, as shown in Eq. (1):

$$\mathbf{D} = \left[D_{(SP_1,t)}, D_{(SP_2,t)}, D_{(SP_3,t)} \dots D_{(SP_M,t)} \right] \quad (1)$$

In the above equation t represents all the moments, i.e., $t = (t_1, t_2, t_3, \dots, t_N)$, and SP_i ($i = 1, 2, \dots, M$) denotes the i -th combination of system parameters.

- (2) The rows of the sample matrix form are normalized and the sample matrix \mathbf{D} was normalized to the form $\hat{\mathbf{D}}$.
- (3) Calculate the correlation matrix \mathbf{R} of the normalized sample matrix and find its eigenvalues and eigenvectors as shown in Eqs. (2) and (3):

$$\mathbf{R} = \frac{1}{N} \hat{\mathbf{D}}^T \hat{\mathbf{D}} \quad (2)$$

$$\mathbf{R}\mathbf{A} = \lambda\mathbf{A} \quad (3)$$

where λ was the eigenvalue matrix, λ_j was the j -th eigenvalue on the diagonal in the eigenvalue matrix, \mathbf{A} was the eigenvector matrix, and A_j was the j -th column of the eigenvector matrix. Eqs. (4) and (5) allow the calculation of each order POD orthogonal base ϕ_j and its corresponding modal coefficient $a_j(t)$:

$$\phi_j = \frac{1}{\sqrt{\lambda_j}} \hat{\mathbf{D}} A_j \quad (4)$$

$$a_j(t) = \frac{\phi_j^T \cdot \hat{\mathbf{D}}}{\phi_j^T \cdot \phi_j} \quad (5)$$

The energy captured by the first n POD modes accounts for the energy E of the full-order modes as shown in Eq. (6), where k was the highest order modal order:

$$E = \frac{\sum_{j=1}^n \lambda_j}{\sum_{j=1}^k \lambda_j} \quad (6)$$

The field data at any moment can be reconstructed from the time average of the field data and a set of modal coefficients and orthogonal bases, i.e.,

$$\hat{\mathbf{D}}_{(SP_i,t_q)} = \sum_{j=1}^N \phi_j \cdot a_j(t) \quad (7)$$

3.2. POD-RBF method and basic functions

The Radial Basis Function (RBF) neural network is a feedforward neural network consisting of three layers, characterized by unidirectional transmission of information (Wang and Jin, 2019). This model is known for its minimal hyper parameters and its capability to learn, organize, and adapt. It is particularly effective in addressing challenges such as nonlinear continuous function approximation in high-dimensional spaces and data prediction. In this study, the Gaussian function was chosen as the basis function of the neural network, facilitating the transformation from the input layer to the hidden layer. Its mathematical form was represented by Eq. (8):

$$H = \exp\left(-\frac{|x - C|^2}{2\sigma^2}\right) \quad (8)$$

In the Radial Basis Function (RBF) neural network, H represents the Gaussian function used to process the transformation from the input layer to the hidden layer, while $|\cdot|$ denotes the Euclidean norm. C represents the center vector of the Gaussian function, and σ indicates the standard deviation of the Gaussian function. This study utilized the RBF neural network to establish the relationship between the modal coefficients obtained from Proper Orthogonal Decomposition (POD) and the operating conditions. The process of the POD-RBF method is shown in Fig. 3.

3.3. Stacked ensemble learning model

Stacked ensemble learning models not only combine multiple machine learning models but also ensure that the predictive ability and stability of the ensemble learning stack model are higher than any one of the base learners included, that is, producing a $1 + 1 > 2$ effect. The ensemble learning model used in this article is shown in Fig. 4. Firstly, the n -fold cross validation algorithm (Zhao et al., 2019) was used to split the dataset into n training sets and n validation sets; then, using n training and validation sets, combined with the POD-RBF method described in

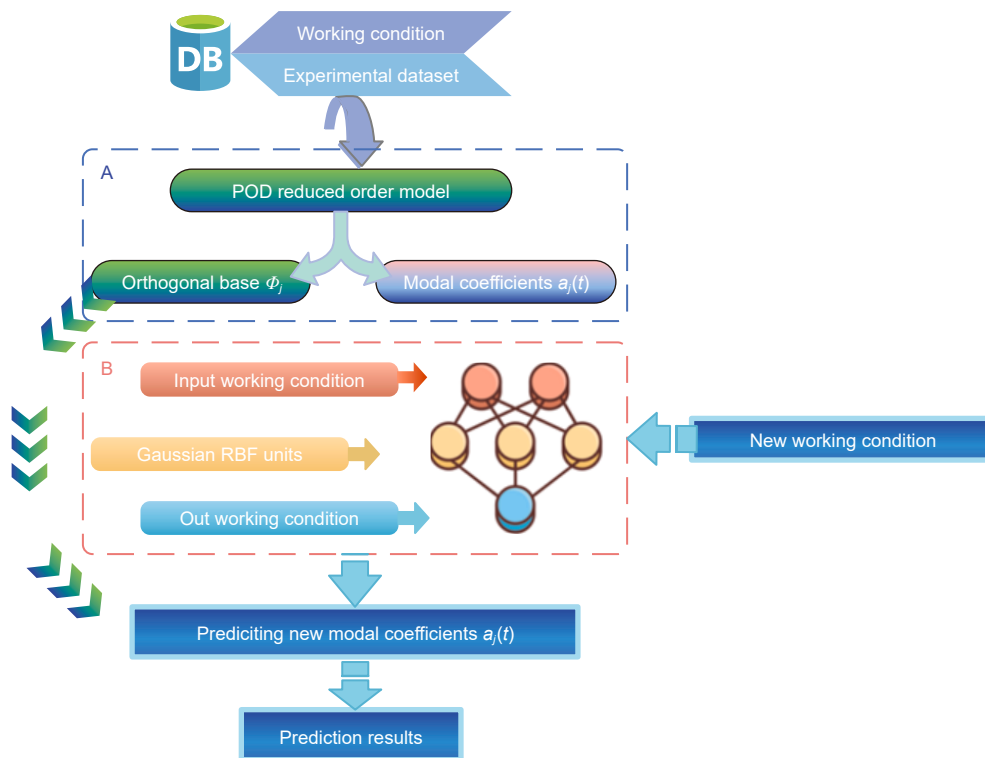


Fig. 3. Diagram of POD-RBF method.

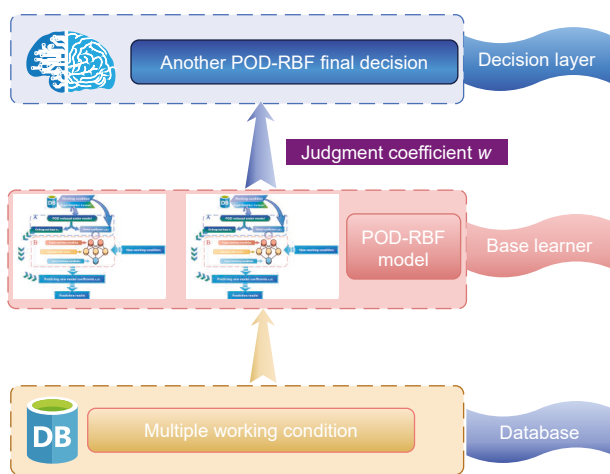


Fig. 4. Stacked ensemble learning model based on POD-RBF base learners.

Section 3.2, construct n base learners. This paper proposed a weighting method for measuring the ability of a single base learner, which assigns different judgment coefficients w_i to each base learner; Finally, an independent RBF network was used to construct the final “decision layer”, which utilizes the weighted output predicted values of the base learner and the corresponding output true values under the working conditions for learning, in order to integrate and improve the learning ability of a single base learner.

There are differences in the accuracy of different base learners, so their impact on the final accuracy of ensemble learning models is also different. This article introduces judgment coefficients in stacked ensemble learning models to improve model accuracy. The introduction to the judgment coefficient is placed in Section 4.3.3.

4. Result and discussion

4.1. Thermodynamic characteristics of hydrate in oil-containing liquid

Understanding the thermodynamic behavior of hydrates under conditions closer to those in oil and gas fields is crucial for mitigating hydrate blockages in deep-sea pipelines (Sloan, 2003). In the early stages of oil and gas production, the oil content of extracted fluids is usually high, but it decreases as production continues. The thermodynamic stability of hydrates varies with different oil-to-water ratios (Babadagli, 2007). Therefore, it is essential to investigate the shift in hydrate phase equilibrium curves at different oil contents.

With the continued development of an oil and gas field in the South China Sea, the oil content in the produced fluid has shown a declining trend (Huang et al., 2017). Forecasts indicate that the oil content will decrease to 60 vol% by 2031, 50 vol% by 2033, and further to 40 vol% by 2040 (Li et al., 2020a). To optimize the injection strategy of thermodynamic hydrate inhibitors, this study systematically investigated the thermodynamic behavior of natural gas hydrates under oil content conditions of 60 vol%, 50 vol%, and 40 vol%, aiming to clarify the influence of oil content variation on hydrate phase equilibrium. To better simulate real-field conditions, no emulsifiers were added in this study; instead, oil–water emulsions were formed solely under the influence of shear forces. In such systems, a relatively stable emulsion can be achieved through the reversible diffusion of oil molecules between droplets, as described by Kelvin’s and Raoult’s laws (Zhang et al., 2018; Sakai, 2020). Typically, the phase occupying the larger volume distributes continuously and forms the continuous phase, while the phase with a smaller volume exists as dispersed droplets, forming the dispersed phase (Zhang et al., 2018). The conventional

isochoric methods (Lee et al., 2014) were used to measure the three-phase equilibrium of hydrates formed from a gas mixture (methane, ethane, propane, n-butane, isobutane, isopentane, carbon dioxide, and nitrogen), as shown in Fig. 5. The results indicate that, in the oil-water system, the phase equilibrium curve of natural gas hydrate with the C1–C5 component shifts to lower temperatures and higher pressures compared to pure water, suggesting that the presence of oil reduces the hydrate stability region. The extent of this shift varies with oil content, with the 50 vol% oil content showing the most significant shift, indicating a stronger inhibitory effect. This is likely because nonpolar natural gas molecules dissolve more easily in the nonpolar oil phase than in the polar water phase (Azam et al., 2019). The presence of the oil-water interface creates resistance to gas diffusion into the water phase. This resistance primarily arises from the van der Waals forces between oil and gas molecules, which are attractive in nature, making it difficult for gas molecules to enter the water phase and form a cage-like structure with water molecules. As a result, gas molecules must overcome an additional energy barrier to enter the water phase, significantly reducing the effective concentration of dissolved gas. At the same time, when a gas mixture partially dissolves in oil, the component of the gas changes. Heavier hydrocarbons are more soluble in the oil phase, leaving behind more light components to form hydrates. These effects collectively shrink the thermodynamic stability region (Meng et al., 2025). At an oil content of 50 vol%, the system approaches the critical point of phase inversion, potentially forming a dynamic mixed structure. In this state, interfacial tension fluctuates frequently, causing constant reconstruction of the oil-water interface and resulting in high instability. This instability complicates the diffusion path for gas molecules crossing the oil-water interface, significantly increasing the diffusion resistance. Gas molecules need to overcome higher energy barriers to enter the water phase and form clathrate structures with water molecules, thereby reducing their concentration in the water phase (Zhang et al., 2018). As a result, the thermodynamic stability of the system in this oil content range is further reduced. In practical engineering, a small amount of thermodynamic inhibitor can effectively prevent hydrate formation in oil-containing systems, thereby reducing both economic costs and environmental impact.

To further analyze the effect of oil content on the phase equilibrium of natural gas hydrates under different pressures, the decrease in hydrate equilibrium temperature relative to pure water was calculated for various oil contents, denoted as ΔT (Liu et al., 2022) (the calculation method is shown in Fig. 2). The results presented in Fig. 5(b) show that at low to medium pressures (below 4 MPa), the impact of oil content on the hydrate equilibrium temperature is minimal, and the inhibition effect was limited. However, at higher pressures (above 4 MPa), the effect became more pronounced, with a significant enhancement in inhibition. This phenomenon could be further understood through Henry's law, which states that gas solubility in liquids increases linearly with pressure. At higher pressures, non-polar gases such as methane become more soluble in non-polar solvents like oil due to their larger Henry's constants (Liu et al., 2024). Once dissolved, these gas molecules could be retained in the oil phase through relatively strong van der Waals interactions, which resist their diffusion into the aqueous phase (Meng et al., 2025). This reduces the amount of gas available in the water phase for hydrate formation, thereby, shifting the hydrate equilibrium toward lower temperatures and higher pressures. However, beyond a certain pressure threshold, as the concentration of gas in the oil phase continues to rise, the gas reaches saturation. At this point, the system approaches a "maximum stable state," where ΔT reaches its maximum value. Further increases in pressure could be causing some of the dissolved gas to escape from the oil phase, leading to a slight increase in gas concentration in the water phase. This causes the phase equilibrium point to shift towards higher temperatures, and ΔT decreases (Liu et al., 2024). Therefore, in actual oil and gas field operations, the impact of oil on hydrate thermodynamics under different pressures should be carefully considered. Through the optimization of the type and dosage of thermodynamic inhibitors, the inhibition of hydrate formation can be enhanced in a cost-effective and efficient way.

4.2. The effect of gas/oil components on thermodynamic characteristics

In addition to oil content, the component of natural gas is a key factor in determining hydrate equilibrium conditions. While many

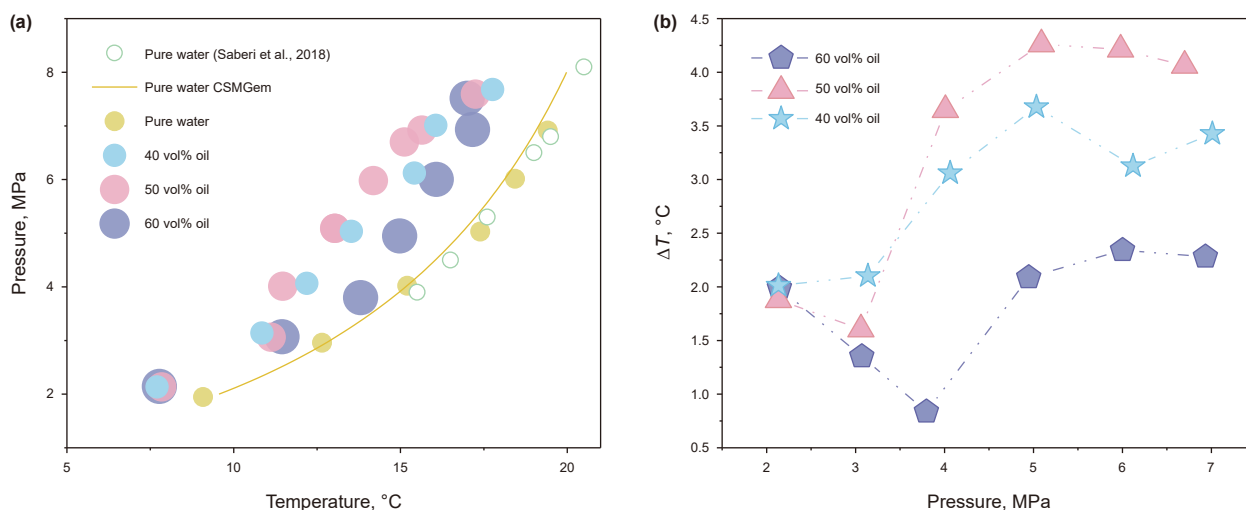


Fig. 5. Thermodynamic characteristics of natural gas hydrate with the C1–C5 component in oil-water systems with varying oil contents. (a) Phase equilibrium diagram of natural gas hydrates, in which the green line represents the natural gas hydrates in pure water measured by Saberi et al. (2018). (b) Influence of oil content on the phase equilibrium temperature of natural gas hydrates. The C1–C5 component includes methane, ethane, propane, n-butane, isobutane, isopentane, carbon dioxide, and nitrogen.

studies have discussed the thermodynamic characteristics of different gas components forming hydrates in pure water systems, the impact of different gas components on hydrate phase equilibrium in the presence of oil remains underexplored systematically (Menezes et al., 2020; Gong et al., 2022). As discussed in Section 4.1, at an oil content of 50 vol%, the influence of the oil phase on hydrate equilibrium temperature was most significant. Therefore, this section further investigated the thermodynamic characteristics of natural gas hydrates at 50 vol% oil content with the C1–C5 component, which includes methane, ethane, propane, n-butane, isobutane, isopentane, CO₂, and nitrogen, as well as the C1–C3 component, which includes methane, ethane, propane, CO₂, and nitrogen. This analysis aimed to reveal the specific impact of gas component on phase equilibrium temperature in the presence of oil.

As shown in Fig. 6(a), the phase equilibrium curve of natural gas hydrate with the C1–C3 component in a pure water system shifted significantly to lower temperatures and higher pressures compared to the C1–C5 component. This was due to the increasing ability of hydrocarbon gases to form hydrates with water as molecular weight increased, following the order of methane, ethane, propane, and isobutane. Additionally, n-butane was able to enter the hydrate lattice and form type II hydrates in the presence of methane (Belosludov et al., 2020). Consequently, the natural gas hydrate with the C1–C5 component, which included isobutane and n-butane, achieved phase equilibrium at higher temperatures and lower pressures. This finding aligned with previous experimental results (Smith et al., 2017). However, at 50 vol% oil content, the shift of the phase equilibrium curve of natural gas hydrate with the C1–C5 component to lower temperatures and higher pressures became significantly greater, while that with the C1–C3 component remained minimal. This indicated a more pronounced inhibitory effect of oil on the formation of hydrate with the C1–C5 component. This effect could be attributed to the larger hydrocarbon molecules (n-butane and isobutane) in the C1–C5 component having stronger nonpolar interactions with the oil phase, resulting in higher solubility in the oil and increased difficulty in combining with water molecules to form hydrates (Azam et al., 2019). Thus, the inhibitory effect of the oil phase on these larger hydrocarbons significantly altered the conditions for hydrate phase equilibrium, leading to a more pronounced shift of the equilibrium curve to lower temperatures and higher pressures.

In order to further analyze the impact of different gas components on the phase equilibrium of natural gas hydrates at 50 vol% oil content under varying pressures, the decreases of hydrate phase equilibrium temperatures of different gas components at 50 vol% oil content compared with that of the pure water system were calculated and denoted as ΔT (Liu et al., 2022) (The calculation method was shown in Fig. 2; the percentages represented the relative difference in ΔT between the two gas components at the same pressure.) Additionally, the relative differences in the inhibition effect on phase equilibrium temperature of hydrate between that with the C1–C5 component and that with the C1–C3 component at the same pressure were assessed. From Fig. 6(b), at 50 vol% oil content and lower pressures, the ΔT of the hydrates formed by the two gas components were small, and the relative differences were not significant. However, at higher pressures, the ΔT increased for both hydrates with two components, and the hydrate with the C1–C5 component showed a more significant rise and a greater relative difference. This may be due to the fact that the increased pressure enhances the stability of the oil-water interface, while higher pressures facilitate the dissolution of larger hydrocarbon molecules (n-butane and isobutane) in the oil phase, further obstructing gas diffusion into the water phase (Azam et al., 2019). Hence, the ΔT of the hydrate with the C1–C5 component, containing larger hydrocarbon molecules, exhibited significant changes with increasing pressure. In practical oil and gas production, substituting fewer gas components for multi-component gases to predict hydrate phase equilibrium and calculate the required dosage of thermodynamic inhibitors may lead to inhibitor overdose, thereby increasing production costs. Furthermore, the impact of pressure on hydrate phase equilibrium should be thoroughly considered when optimizing inhibitor addition strategies, combining these two factors to enhance economic efficiency and reduce environmental impact.

4.3. A prediction method based on a stacked ensemble learning model

4.3.1. Dataset introduction

This article obtained a total of 28 sets of test results for phase equilibrium temperature and pressure of the hydrate with the C1–C5 component under different initial pressures, with oil content of 60 vol%, 50 vol%, 40 vol%, and 0 vol% (non-oil phase),

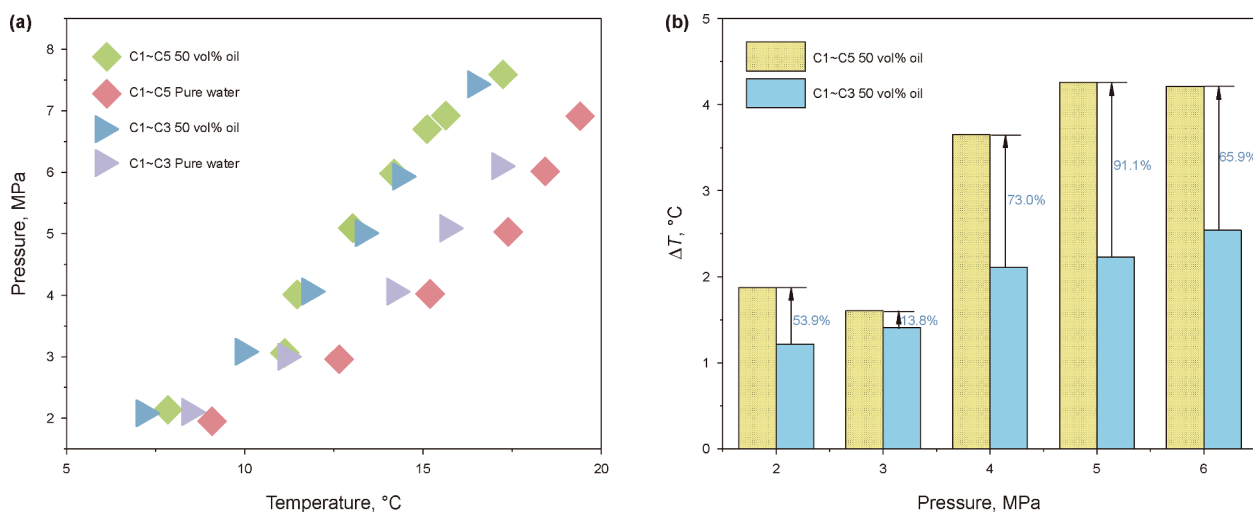


Fig. 6. Thermodynamic characteristics of hydrates formed by different natural gas components in oil-water systems. (a) Three-phase equilibrium diagram of natural gas hydrate. (b) Effect of different gas components on the phase equilibrium temperature of natural gas hydrate. The C1–C5 component includes methane, ethane, propane, n-butane, isobutane, isopentane, carbon dioxide, and nitrogen. The C1–C3 component includes methane, ethane, propane, carbon dioxide, and nitrogen.

Table 3
Phase equilibrium temperature and phase equilibrium pressure under different oil content and initial pressure conditions.

Oil content	Initial pressure, MPa	Phase equilibrium temperature, °C	Phase equilibrium pressure, MPa
0.6	7.694	17.003	7.514
0.6	7.050	17.162	6.934
0.6	6.086	16.075	6.002
0.6	5.014	14.977	4.947
0.6	3.820	14.808	3.797
0.6	3.114	11.449	3.070
0.6	2.177	7.771	2.144
0.5	7.759	17.250	7.591
0.5	7.106	15.646	6.922
0.5	6.291	15.552	6.180
0.5	5.044	14.773	4.964
0.5	4.083	13.038	4.017
0.5	3.111	11.125	3.061
0.5	2.162	7.842	2.133
0.4	7.823	17.763	7.678
0.4	7.177	16.060	7.011
0.4	6.234	15.420	6.120
0.4	5.147	13.529	5.033
0.4	4.150	12.193	4.064
0.4	3.194	10.847	3.139
0.4	2.159	7.704	2.133
0	7.742	17.659	7.590
0	6.897	19.414	6.897
0	6.011	18.429	6.011
0	5.031	17.454	5.031
0	4.034	15.511	4.023
0	2.992	12.772	2.972
0	2.033	9.375	2.019

through the experiment in Section 2. In Table 3, oil content and initial pressure are used as independent variables, while phase equilibrium temperature and pressure are used as dependent variables.

Although a total of 178 experimental datasets were obtained during the full experimental campaign, only 28 representative data points were selected for model training and testing. This was a deliberate design choice aimed at simulating real-world scenarios where acquiring large-scale hydrate phase equilibrium data is challenging due to the high cost, operational complexity, and time-consuming nature of high-pressure experiments. The intention was to evaluate the model's robustness and extrapolation capacity under limited-data conditions. To ensure prediction reliability, we implemented 7-fold cross-validation, employed the

POD-RBF method for dimensionality reduction and noise suppression, and developed a stacked ensemble learning model to improve stability and generalization. Furthermore, six additional experiments under previously untested conditions validated the model's ability to generalize beyond the training range, indicating its practical potential in data-scarce environments.

4.3.2. Cross validation and construction of base learners

Random screening method was used to shuffle the dataset shown in Table 3, and a 7-fold cross validation method was used to construct 7 sets of training and validation sets. Each set contains 24 data points, with the remaining 4 sets used as the validation set. Cross validation is shown in Fig. 7.

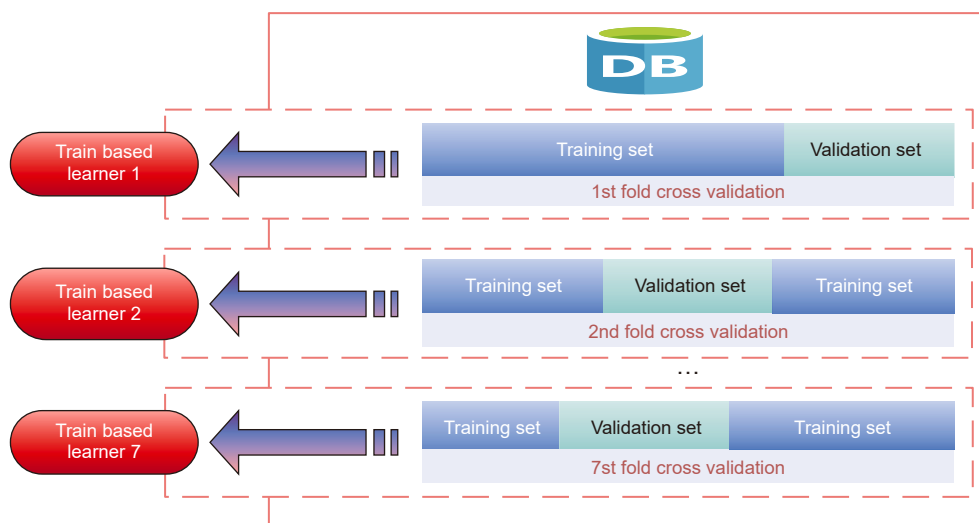


Fig. 7. Cross validation and base learning training process.

For each set of training and validation sets, the POD method shown in Section 3.1 is first used to reduce the dimensionality of the training set, extract the main features, and filter out data noise. In this paper, second-order orthogonal bases and second-order modal coefficients are extracted from each base learner dataset, and RBF neural networks are used to learn the nonlinear relationship between independent variables (oil content, initial pressure) and modal coefficients as the base learners; Substitute the independent variables in the validation set into the base learner to obtain the corresponding predicted modal coefficients, and combine them with the original orthogonal basis to obtain the predictions of the validation set data. Then, in Section 4.3.3, use the predicted and true values of the validation set data to solve the judgment coefficients of each base learner.

4.3.3. Judgment coefficients in stack ensemble learning

Assuming that the validation set matrix in the k -th fold cross validation is represented as L_k , each row in L_k corresponds to L_k^1 -phase equilibrium temperature, L_k^2 -phase equilibrium pressure, and each column corresponds to a combination of independent variables (oil content and initial pressure). The predicted validation set for each base learner is L'_k , and similarly, the predicted values for each row are L_k^1 , L_k^2 , and L_k^3 , respectively. The judgment coefficients corresponding to each dependent variable of each basis function are defined as Eq. (9):

$$\left\{ \begin{aligned} w_k^1 &= \frac{\operatorname{argmax}\left(\frac{|L_k^1 - L_k^1|}{L_k^1}\right)}{\sum_{k=1}^n \operatorname{argmax}\left(\frac{|L_k^1 - L_k^1|}{L_k^1}\right)} \\ w_k^2 &= \frac{\operatorname{argmax}\left(\frac{|L_k^2 - L_k^2|}{L_k^2}\right)}{\sum_{k=1}^n \operatorname{argmax}\left(\frac{|L_k^2 - L_k^2|}{L_k^2}\right)} \end{aligned} \right. \quad (9)$$

Each coefficient w_k^i (where k indicates the base learner and i the predicted variable) quantifies the relative reliability and predictive accuracy of a specific base learner for a particular output variable (i.e., phase equilibrium temperature and pressure). These coefficients are derived based on the performance of each base learner on the validation sets during the cross-validation process. A higher value of w_k^i implies a smaller prediction error and thus indicates that the corresponding base learner contributes more significantly to the final ensemble prediction. Conversely, lower values suggest that the learner had larger errors and, therefore, is assigned a reduced weight in the decision fusion process. For instance, in Table 4, base learner #2 exhibits perfect coefficients ($w = 1.0$) for all variables, indicating that it performed exceptionally well across all validation sets and effectively serves as a

Table 4
Corresponds to a table of judgment coefficients for three dependent variables.

Base learner serial number	w_k^1	w_k^2
1	0.294	0.460
2	1.000	1.000
3	0.273	0.125
4	0.232	0.223
5	0.230	0.207
6	0.178	0.104
7	0.1000	0.100

Table 5
Maximum relative prediction error of each variable in cross validation set.

k -th fold cross validation	Maximum relative error	
	Equilibrium temperature	Equilibrium pressure
1	0.036	0.015
2	0.022	0.034
3	0.088	0.012
4	0.093	0.004
5	0.034	0.033
6	0.091	0.088
7	0.023	0.050

reference standard. On the other hand, learners such as #6 and #7 show substantially lower coefficients, reflecting their limited predictive accuracy and minimal contribution to the final decision layer. This weighting mechanism ensures that the ensemble model prioritizes more reliable learners while suppressing the influence of less accurate ones, thereby enhancing the overall robustness, stability, and accuracy of the prediction. The introduction and analysis of the judgement coefficients thus offer both theoretical and empirical justification for the superior performance of the proposed SELM-DCC model.

4.3.4. Stacked ensemble learning model and prediction performance

After determining the base learner and judgment coefficients, a stacked ensemble learning model is constructed as shown in Fig. 4. In Fig. 4, the dot product of the output of the base learner and the judgment coefficients was used as the input for the independent decision layer RBF neural network, resulting in a more stable prediction result. Still using the same cross validation dataset as shown in Fig. 7. Each training set was input into the constructed stacked ensemble learning model to train the model, and the prediction accuracy of the stacked ensemble learning was tested using the validation dataset. This paper used the maximum relative prediction error of each dependent variable in each validation set to measure the prediction accuracy of the stacked ensemble learning model. The relative error E was defined as Eq. (10), where y was the true value and \hat{y} was the predicted value.

$$E = \frac{|y - \hat{y}|}{y} \quad (10)$$

Obtain the maximum relative prediction error of each variable in the k -th fold cross validation set as shown in Table 5. It can be seen that the maximum prediction relative error for phase

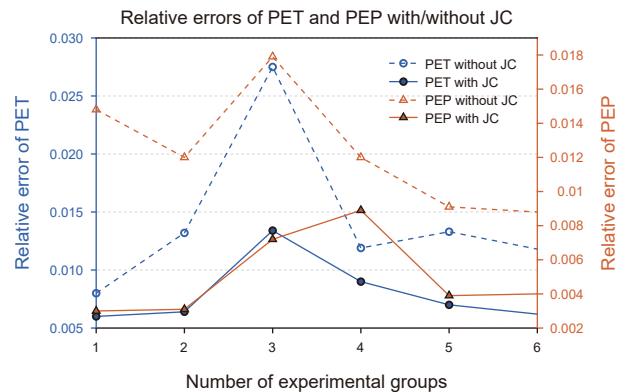


Fig. 8. The relative error size of the six groups of experiments with and without the JC (judgment coefficient).

Table 6
RBF neural network predicts the maximum relative error of phase equilibrium temperature and phase equilibrium pressure.

k-th fold cross validation	Maximum relative error	
	Equilibrium temperature	Equilibrium pressure
1	0.062	0.002
2	0.014	0.002
3	0.070	0.020
4	0.091	0.005
5	0.092	0.006
6	1.052	0.060
7	3.351	0.092

equilibrium temperature and pressure using a stack based ensemble learning model based on a reduced order model is within 10%; In order to compare and illustrate the advantages of this method, a separate RBF neural network was used to predict the maximum relative error of phase equilibrium temperature and pressure, as shown in Fig. 8. It can be seen that when using RBF neural network alone, the maximum relative error of phase equilibrium temperature prediction reached 335% at the 7th fold, and the maximum relative error at some positions was also close to 100%, indicating that the prediction results were seriously inaccurate. The comparison Tables 5 and 6 indicate that the

stacked ensemble learning algorithm has significantly improved stability and prediction accuracy.

To test the predictive ability of the ensemble learning model for new operating conditions, six additional experiments were conducted under the combined conditions of oil content of 0.35 and 0.45, initial pressures of 7.21, 6.49, and 5.31 MPa, respectively, to measure the phase equilibrium temperature and pressure. Using the stacked ensemble learning model constructed in the previous text, the phase equilibrium temperature and phase equilibrium pressure under the six experimental conditions were predicted. At the same time, to illustrate the improvement of the model's prediction ability by introducing decision coefficients, the decision coefficients of each base learner were equal to 1. The same 7-fold cross validation set was used to test the prediction performance of the stacked ensemble learning model, and the relative prediction errors under the six experimental conditions were obtained as shown in Fig. 8. From Fig. 8, it can be seen that under the condition of having a judgment coefficient, the relative error of the stacked ensemble learning model for predicting the phase equilibrium temperature is less than 1.4%, and the relative error for predicting the phase equilibrium pressure is less than 0.8%; Without the judgment coefficient, the maximum relative error of the prediction of the phase equilibrium temperature exceeds 2.8%, and the maximum relative error of the prediction of the phase equilibrium pressure reaches 1.8%. Therefore, the judgment coefficient can significantly improve the accuracy of the stacked ensemble learning model.

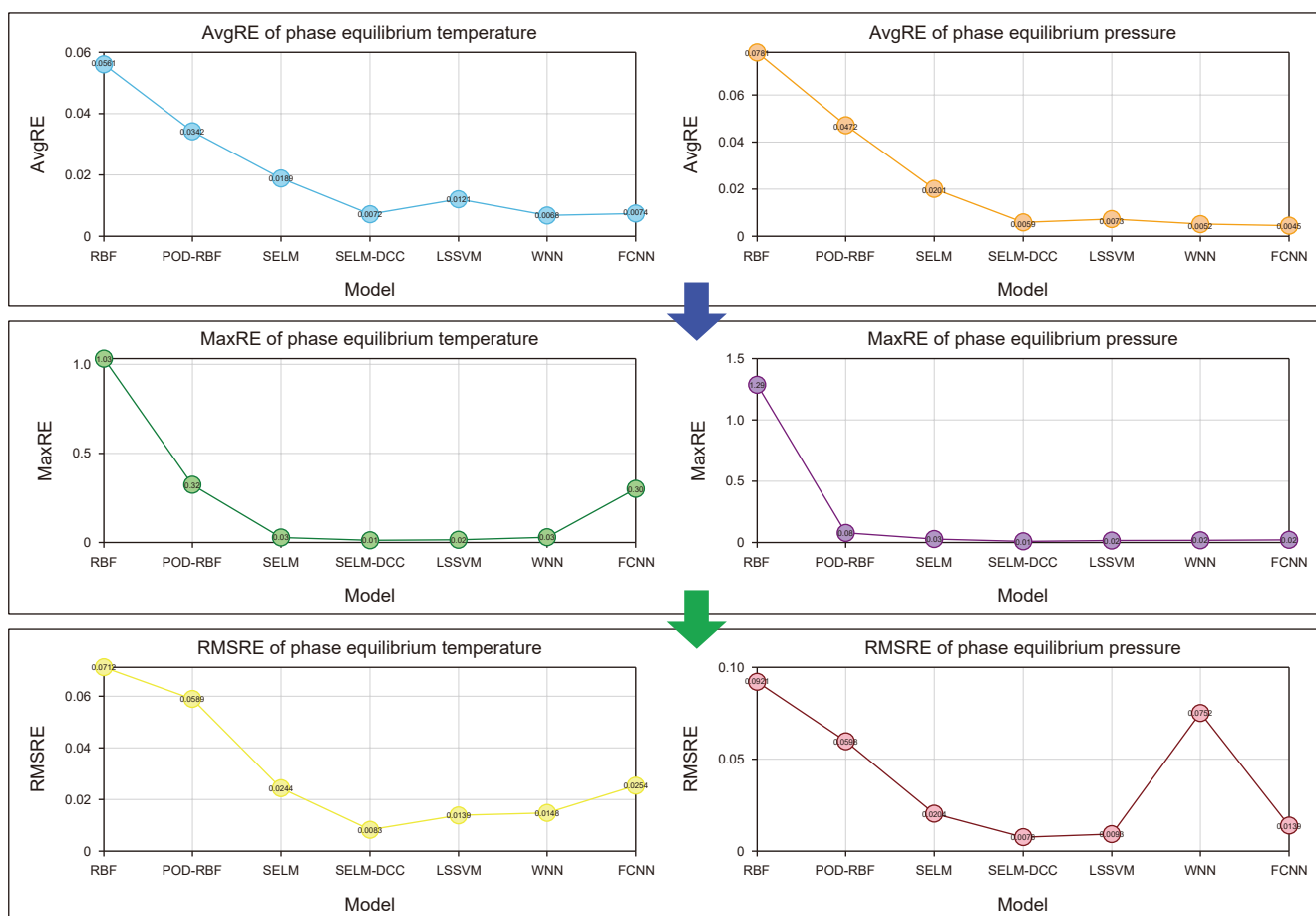


Fig. 9. Comparison of prediction accuracy of different models for phase equilibrium temperature and pressure (without denoising).

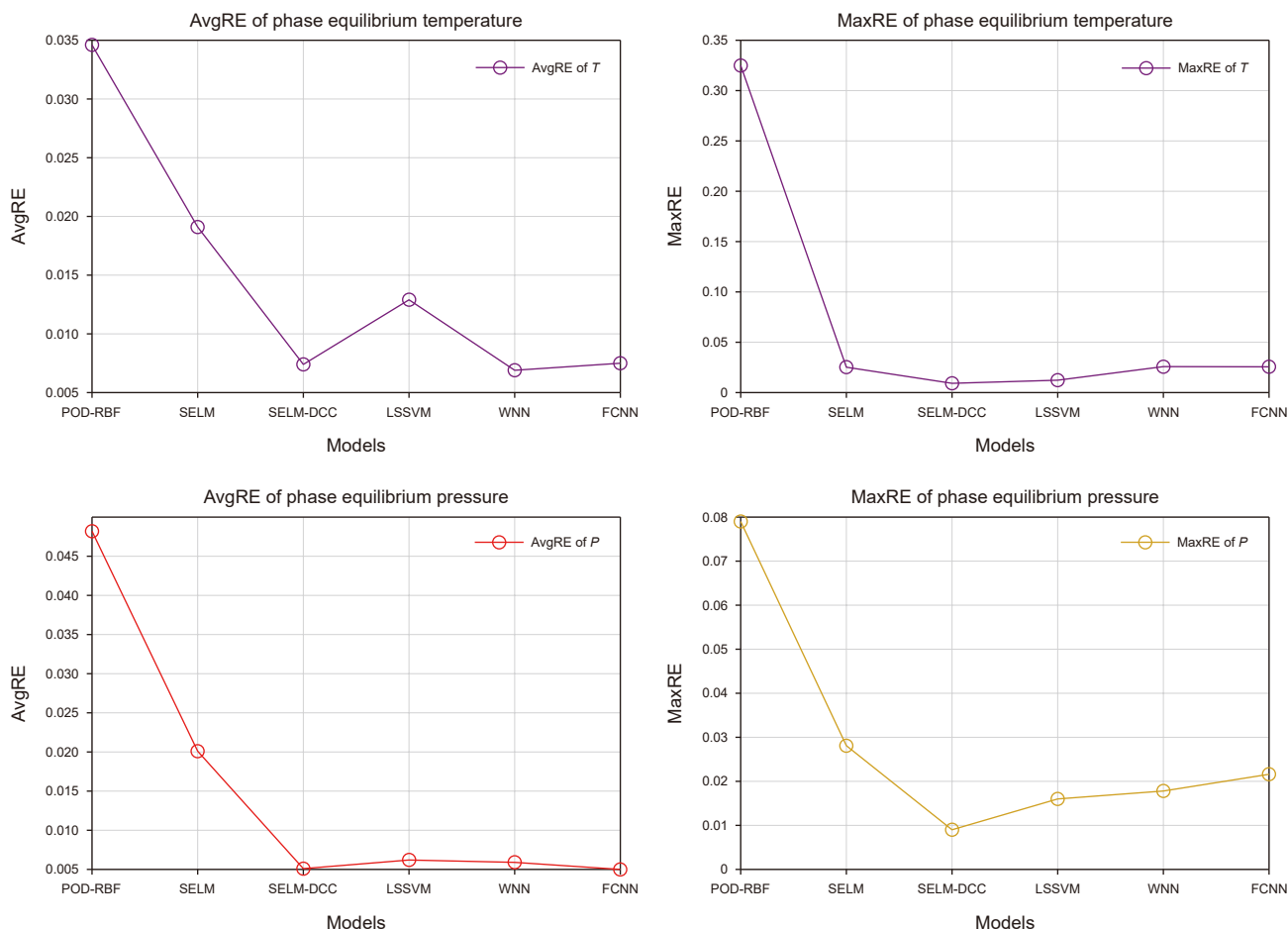


Fig. 10. Comparison of prediction accuracy for phase equilibrium temperature and pressure using six methods (after denoising).

To evaluate the performance of the proposed SELM-DCC, a comparative analysis was conducted with several widely used prediction methods. The training and testing datasets used in this comparison were the same as those described in Section 4.3.4, and the average relative error (AvgRE) and maximum relative error (MaxRE) were defined by Eqs. (11) and (12), respectively.

$$\text{AvgRE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (11)$$

$$\text{MaxRE} = \max_{1 \leq i \leq n} \left(\left| \frac{y_i - \hat{y}_i}{y_i} \right| \right) \quad (12)$$

where y_i represents the true value for the i -th data point, \hat{y}_i denotes the predicted value, and n is the total number of data points.

In addition to the AvgRE and the MaxRE, we introduce the Root Mean Square Relative Error (RMSRE) as a supplementary metric to evaluate the prediction performance of different reduced-order models. RMSRE is particularly useful because it places greater emphasis on large deviations, which makes it more sensitive to outliers or localized prediction failures. This feature helps assess the consistency and robustness of the model more thoroughly.

The Root Mean Square Relative Error (RMSRE) is introduced as an additional performance metric to complement the average and maximum relative errors. It is defined as follows:

$$\text{RMSRE} = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - \hat{y}_i}{y_i} \right)^2} \quad (13)$$

The comparison involved six prediction methods, covering a comprehensive spectrum of classical and advanced machine learning techniques commonly used for hydrate phase equilibrium prediction or similar nonlinear regression tasks, allowing for a thorough comparison in terms of accuracy, stability, and generalization. The methods included the RBF neural network (Zahedi et al., 2009), the standalone POD-RBF model (Shen et al., 2023), the Stacked Ensemble Learning Model based on the POD-RBF learner (SELM) (Abrate et al., 2023), the Stacked Ensemble Learning Model with Decision Coefficients (SELM-DCC), Least Squares Support Vector Machine (LSSVM) (Ghiasi et al., 2016), Wavelet Neural Network (WNN) (Zhu et al., 2022), and Fully Connected Neural Network (FCNN) (Xiao et al., 2024). All methods were evaluated based on their performance on the test set.

The results are presented in Fig. 9. As illustrated, the prediction performance for phase equilibrium temperature, measured by AvgRE and RMSRE, ranks as follows: RBF < POD-RBF < SELM < LSSVM < SELM-DCC. Although WNN and FCNN exhibit lower average relative errors than SELM-DCC, their maximum relative errors and RMSRE values are considerably higher, indicating greater prediction variability and reduced robustness. This suggests that while WNN and FCNN may achieve

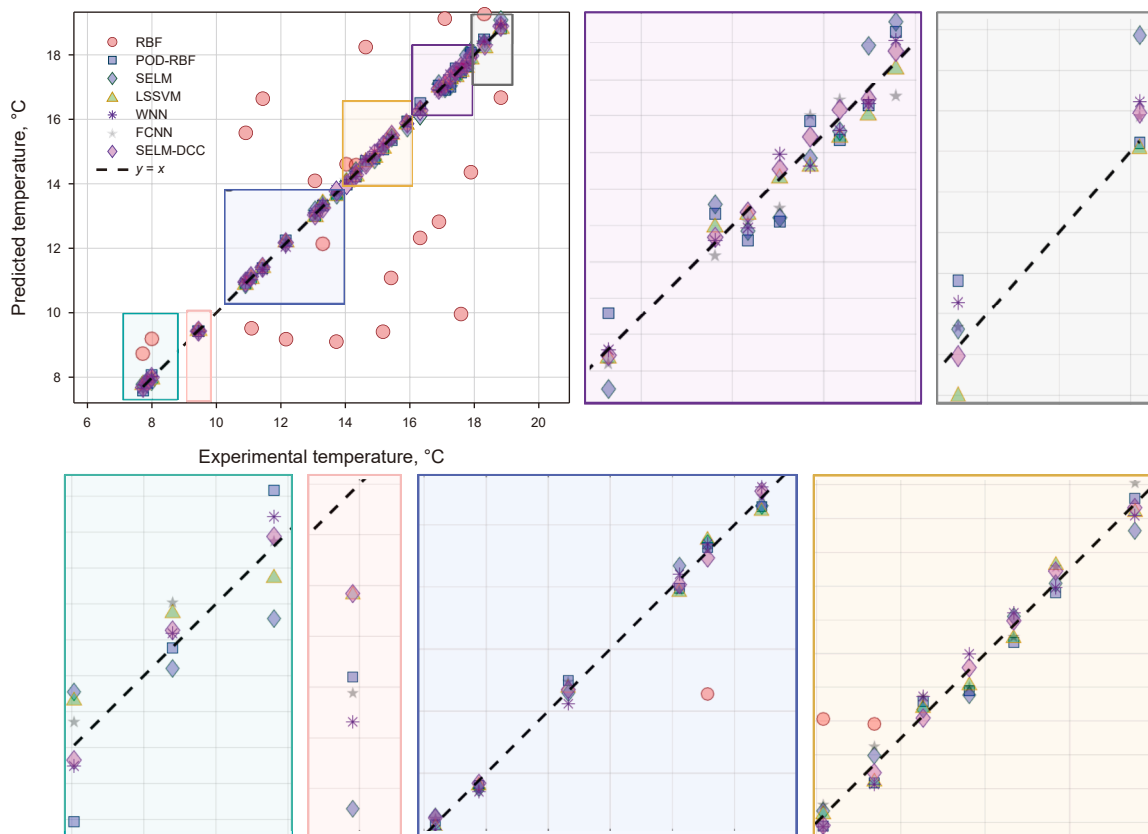


Fig. 11. Cross-comparison of phase equilibrium temperatures predicted by different models and experimental values.

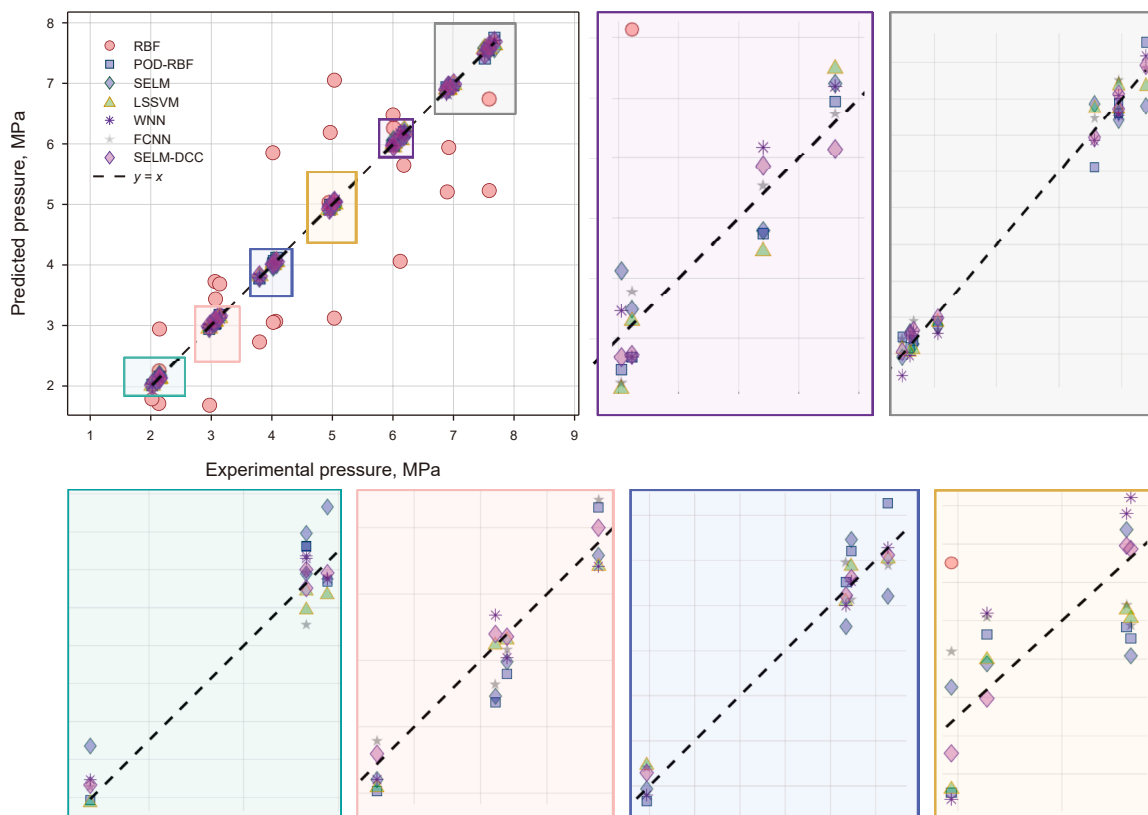


Fig. 12. Cross-comparison of phase equilibrium pressures predicted by different models and experimental values.

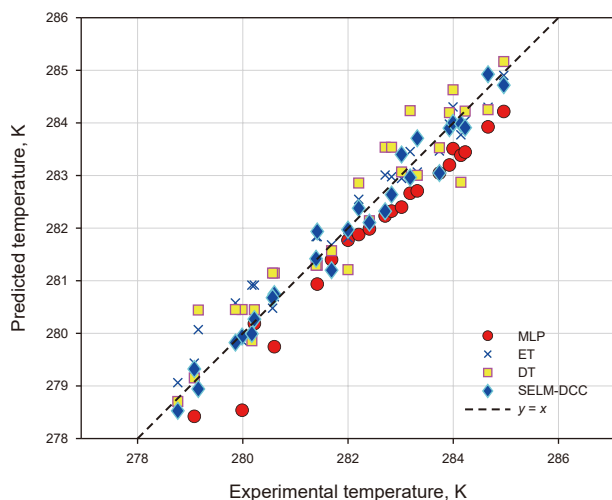


Fig. 13. Predicted equilibrium temperatures of methane hydrate in MgBr_2 brine systems using different models.

higher accuracy under certain conditions, their performance is less consistent across different scenarios.

A detailed comparison between RBF and POD-RBF models highlights the effectiveness of data denoising: the POD-RBF model, essentially a denoised version of the RBF model, significantly reduces AvgRE, MaxRE, and RMSRE values. This underscores the substantial benefits of integrating proper orthogonal decomposition (POD) as a preprocessing step.

The SELM and SELM-DCC models both incorporate POD-based denoising, whereas LSSVM, WNN, and FCNN originally do not. To investigate the impact of denoising, POD is applied to preprocess the dataset before feeding it into these three models. As shown in Fig. 10, this preprocessing step improves their AvgRE, MaxRE, and RMSRE metrics. Nevertheless, even after denoising, the RMSRE of WNN remains higher than that of SELM-DCC, reaffirming the superior stability and robustness of the SELM-DCC method.

To provide a more comprehensive evaluation of model performance, two additional comparison plots were generated to visually assess the accuracy of each model in predicting both phase equilibrium temperature and pressure. As shown in Figs. 11 and 12, Fig. 11 compares the predicted and experimental phase equilibrium temperatures, while Fig. 12 presents a similar comparison for phase equilibrium pressures. In both plots, the x-axis represents experimental values and the y-axis denotes predicted values. Ideally, data points should lie along the diagonal line $y = x$, indicating perfect prediction. The SELM-DCC model shows strong alignment with experimental results in both subfigures, reflecting its superior performance and generalization ability.

To validate the model's applicability and predictive performance, the equilibrium temperature of methane hydrate in the MgBr_2 brine system was predicted using the SELM-DCC model, with phase equilibrium pressure and salt concentration as input variables, based on the findings from Hosseini and Leonenko (2023). To assess the predictive performance of the model, the results were compared with those obtained from Extreme Randomized Trees (ET), Multilayer Perceptron (MLP), and Decision Tree (DT) models developed in the referenced study. The cross-plots of these comparisons are shown in Fig. 13. As seen in Fig. 13, the SELM-DCC model outperforms the Decision Tree and Extreme Randomized Tree models in predicting the equilibrium temperature of methane hydrate in the MgBr_2 brine system, and significantly surpasses the Multilayer Perceptron model in terms of prediction accuracy.

5. Conclusions

This study systematically investigated the hydrate phase equilibrium of natural gas under different oil contents and gas components based on the actual conditions of oil and gas fields in the South China Sea. The aim was to provide theoretical support for optimizing the injection dosage of thermodynamic inhibitors in oil and gas fields. The experimental results showed that the presence of the oil phase shifted the hydrate phase equilibrium curve toward lower temperatures and higher pressures, with the most significant suppression observed at an oil content of 50 vol%. This was attributed to the strong van der Waals forces between oil and gas, which increased the difficulty of gas molecules and water molecules forming a clathrate structure, thereby narrowing the thermodynamic stability zone. Additionally, at 50 vol% oil content, the system approached the critical point of phase inversion, with a dynamic mixed structure that reduced interface stability, leading to a stronger inhibition effect. The experiments also revealed that the oil phase had a more pronounced inhibition effect on the hydrate phase equilibrium of natural gas containing heavier hydrocarbons (such as butane and pentane). This was likely due to the stronger nonpolar interactions between the heavy hydrocarbons and the oil phase, which increased the difficulty of gas molecules and water molecules forming hydrates. Moreover, further experiments in the low-to high-pressure range (2–8 MPa) confirmed that these inhibition effects were more pronounced at higher pressures. These results indicated that using simplified natural gas components or neglecting the influence of the oil phase in hydrate phase equilibrium predictions and thermodynamic inhibitor dosage calculations could lead to inhibitor overuse, increasing production costs. Therefore, fully considering these factors when optimizing inhibitor dosage could not only improve economic efficiency but also reduce environmental impact.

Furthermore, based on the experimental data, this study proposed a stacked ensemble learning model (SELM-DCC) composed of multiple POD-RBF base learners and decision coefficients for predicting hydrate phase equilibrium in oil-containing systems. This model was compared with six other prediction methods, including RBF neural networks, independent POD-RBF models, Stacked Ensemble Learning Models (SELM) based on POD-RBF learners, Least Squares Support Vector Machines (LSSVM), Wavelet Neural Networks (WNN), and Fully Connected Neural Networks (FCNN). The results showed that the SELM-DCC model demonstrated lower maximum relative error, stronger stability, and superior denoising performance when predicting hydrate phase equilibrium temperature and pressure. Future research will focus on refining phase equilibrium testing under different conditions, expanding the model's applicability, and combining physical models with ensemble learning models to develop a “mathematical-physical” dual-driven ensemble learning model, thereby improving the model's generalization ability.

CRedit authorship contribution statement

Bing-Yue Han: Writing – original draft. **Qing-Yang Yuan:** Data curation. **Jie Wang:** Formal analysis. **Peng-Cheng Li:** Investigation. **Zheng Ling:** Conceptualization. **Lei Yang:** Project administration. **Yu Liu:** Conceptualization. **Yong-Chen Song:** Project administration. **Lun-Xiang Zhang:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (Grant Nos. 52476058 and U21B2065), the Young Elite Scientist Sponsorship Program by CAST (Grant No. YESS20220158), and the Dalian High-Level Talent Innovation Program (Grant No. 2024RY002).

References

- Abrate, N., Dulla, S., Pedroni, N., 2023. A non-intrusive reduced order model for the characterisation of the spatial power distribution in large thermal reactors. *Ann. Nucl. Energy* 184, 109674. <https://doi.org/10.1016/j.anucene.2022.109674>.
- Alavi, F., Sharifzadeh, M., 2024. Customized federated kernel regression learning for predicting natural gas hydrate equilibrium with thermodynamic inhibitors: A comprehensive study. *Chem. Eng. J.* 498, 155664. <https://doi.org/10.1016/j.cej.2024.155664>.
- Azam, M.Z., Xin, F., Song, Y., Qaraah, F.A.A., Abbas, S.Z., 2019. Rate enhancement of methane hydration in slurry of ice by phase change of water-in-oil emulsions. *Fuel* 244, 296–303. <https://doi.org/10.1016/j.fuel.2019.01.159>.
- Babadagli, T., 2007. Development of mature oil fields — A review. *J. Petrol. Sci. Eng.* 57, 221–224. <https://doi.org/10.1016/j.petrol.2006.10.006>.
- Belosludov, R.V., Zhdanov, R.K., Gets, K.V., Bozhko, Y.Y., Belosludov, V.R., Kawazoe, Y., 2020. Role of methane as a second guest component in thermodynamic stability and isomer selectivity of butane clathrate hydrates. *J. Phys. Chem. C* 124, 18474–18481. <https://doi.org/10.1021/acs.jpcc.0c05947>.
- Cha, M., Hu, Y., Sum, A.K., 2016. Methane hydrate phase equilibria for systems containing NaCl, KCl, and NH₄Cl. *Fluid Phase Equilib.* 413, 2–9. <https://doi.org/10.1016/j.fluid.2015.08.010>.
- Chee, S., Leokprasirtkul, T., Kanno, T., Osawa, O., Sudo, Y., Takekoshi, M., Yu, H., Yamamoto, K., 2014. A deepwater sandface monitoring system for offshore gas hydrate. *Offshore Technology Conference. OTC-25328-MS*.
- Chen, B., Li, K., Sun, H., Jiang, L., Yang, M., Song, Y., 2023. Promoting effect of common marine cations on hydrate dissociation and structural evolution under a static electric field. *J. Phys. Chem. B* 127, 698–709. <https://doi.org/10.1021/acs.jpcc.2c05382>.
- Dec, S.F., 2012. Surface transformation of methane–ethane sI and sII clathrate hydrates. *J. Phys. Chem. C* 116, 9660–9665. <https://doi.org/10.1021/jp301766y>.
- Du, J., Wang, X., Liu, H., Guo, P., Wang, Z., Fan, S., 2019. Experiments and prediction of phase equilibrium conditions for methane hydrate formation in the NaCl, CaCl₂, MgCl₂ electrolyte solutions. *Fluid Phase Equilib.* 479, 1–8. <https://doi.org/10.1016/j.fluid.2018.09.028>.
- Garapati, N., Anderson, B.J., 2014. Statistical thermodynamics model and empirical correlations for predicting mixed hydrate phase equilibria. *Fluid Phase Equilib.* 373, 20–28. <https://doi.org/10.1016/j.fluid.2014.03.010>.
- Geo Lim, S., Yeop Oh, C., Lee, J.-W., Cha, M., Ha Kim, S., Ra, K., Yoo, K., Yoon, J.-H., 2023. Sustainable freshwater recovery from radioactive wastewater by gas hydrate formation. *Chem. Eng. J.* 461, 141830. <https://doi.org/10.1016/j.cej.2023.141830>.
- Ghiasi, M.M., Yarveicy, H., Arabloo, M., Mohammadi, A.H., Behbahani, R.M., 2016. Modeling of stability conditions of natural gas clathrate hydrates using least squares support vector machine approach. *J. Mol. Liq.* 223, 1081–1092. <https://doi.org/10.1016/j.molliq.2016.09.009>.
- Gong, Y., Mendgaziev, R.L., Hu, W., Li, Y., Li, Z., Stoporev, A.S., Manakov, A.Y., Vinokurov, V.A., Li, T., Semenov, A.P., 2022. Urea as a green thermodynamic inhibitor of sII gas hydrates. *Chem. Eng. J.* 429, 132386. <https://doi.org/10.1016/j.cej.2021.132386>.
- Hassanpouryouzband, A., Joonaki, E., Farahani, M.V., Takeya, S., Ruppel, C., Yang, J.H., English, N.J., Schicks, J.M., Edlmann, K., Mehrabian, H., Aman, Z.M., Tohid, B., 2020. Gas hydrates in sustainable chemistry. *Chem. Soc. Rev.* 49, 5225–5309. <https://doi.org/10.1039/c8cs00989a>.
- Hosseini, M., Leonenko, Y., 2023. A reliable model to predict the methane-hydrate equilibrium: An updated database and machine learning approach. *Renew. Sustain. Energy Rev.* 173, 113103. <https://doi.org/10.1016/j.rser.2022.113103>.
- Hosseini, M., Leonenko, Y., 2024. Development of explicit models to predict methane hydrate equilibrium conditions in pure water and brine solutions: a machine learning approach. *Chem. Eng. Sci.* 286, 119603. <https://doi.org/10.1016/j.ces.2023.119603>.
- Hosseini, M., Boudreault, R., Leonenko, Y., 2025. Insights into predicting equilibrium conditions of clathrate hydrates of methane + water-soluble hydrate former. *Fuel* 384, 133945. <https://doi.org/10.1016/j.fuel.2024.133945>.
- Huang, H., Huang, B., Huang, Y., Li, X., Tian, H., 2017. Condensate origin and hydrocarbon accumulation mechanism of the deepwater giant gas field in Western South China Sea: A case study of Lingshui 17-2 gas field in Qiongdongnan Basin. *Petrol. Explor. Dev.* 44, 409–417. [https://doi.org/10.1016/S1876-3804\(17\)30047-2](https://doi.org/10.1016/S1876-3804(17)30047-2).
- Jeong, J.H., Cha, M., Jang, J., Lim, S.G., Oh, C.Y., Lee, J.-W., Park, J., Yoon, J.-H., 2022. Thermodynamic behavior and spectroscopic properties of CO and C₃H₈ mixed gas hydrates: Implications for hydrate-based gas separation. *Chem. Eng. J.* 428, 132076. <https://doi.org/10.1016/j.cej.2021.132076>.
- Khan, M.N., Warriar, P., Peters, C.J., Koh, C.A., 2018. Advancements in hydrate phase equilibria and modeling of gas hydrates systems. *Fluid Phase Equilib.* 463, 48–61. <https://doi.org/10.1016/j.fluid.2018.01.014>.
- Khan, N., Kumar, A., Johns, M.L., May, E.F., Aman, Z.M., 2023. Experimental investigation to elucidate the hydrate Anti-Agglomerating characteristics of 2-Butoxyethanol. *Chem. Eng. J.* 471, 144288. <https://doi.org/10.1016/j.cej.2023.144288>.
- Kou, X., Li, X.S., 2019. The SAFT for prediction of hydrate formation conditions of gas mixtures in the presence of methane, glycerol, ethylene glycol, and triethylene glycol. *Energy* 188, 116028. <https://doi.org/10.1016/j.energy.2019.116028>.
- Kwon, M., Lee, J.-W., Lee, H., 2014. Temperature-dependent structural transitions in methane–ethane mixed gas hydrates. *J. Phys. Chem. C* 118, 28906–28913. <https://doi.org/10.1021/jp5102219>.
- Lang, C., Chen, Z.R., Hassanpouryouzband, A., Farahani, M.V., Zhang, L.X., Zhao, J.F., Song, Y.C., 2024. Spontaneous lifting and self-cleaning of gas hydrate crystals. *ACS Nano* 18, 33671–33680. <https://doi.org/10.1021/acsnano.4c12943>.
- Lee, Y., Lee, S., Lee, J., Seo, Y., 2014. Structure identification and dissociation enthalpy measurements of the CO₂+N₂ hydrates for their application to CO₂ capture and storage. *Chem. Eng. J.* 246, 20–26. <https://doi.org/10.1016/j.cej.2014.02.045>.
- Li, H., Zhang, M., Lau, H.C., Fu, S., 2020a. China's deepwater development: sub-surface challenges and opportunities. *J. Petrol. Sci. Eng.* 195, 107761. <https://doi.org/10.1016/j.petrol.2020.107761>.
- Li, S., Wang, J., Lv, X., Ge, K., Jiang, Z., Li, Y., 2020b. Experimental measurement and thermodynamic modeling of methane hydrate phase equilibria in the presence of chloride salts. *Chem. Eng. J.* 395, 125126. <https://doi.org/10.1016/j.cej.2020.125126>.
- Li, X., Fan, S., Wang, Y., Li, G., Wang, S., Lang, X., Liu, F., 2022a. Hydrate phase equilibrium of hydrogen-natural gas blends: experimental study and thermodynamic modeling. *Fluid Phase Equilib.* 556, 113417. <https://doi.org/10.1016/j.fluid.2022.113417>.
- Li, Z., Zhang, Y., Shen, Y., Cheng, L., Liu, B., Yan, K., Chen, G., Li, T., 2022b. Molecular dynamics simulation to explore the synergistic inhibition effect of kinetic and thermodynamic hydrate inhibitors. *Energy* 238, 121697. <https://doi.org/10.1016/j.energy.2021.121697>.
- Liu, T., Tang, H., Wu, P., Wang, H., Song, Y., Li, Y., 2023. Acoustic characteristics on clayey-silty sediments of the South China Sea during methane hydrate formation and dissociation. *Energy* 282, 128978. <https://doi.org/10.1016/j.energy.2023.128978>.
- Liu, W., Hu, J., Li, X., Sun, Z., Sun, F., Chu, H., 2018. Assessment of hydrate blockage risk in long-distance natural gas transmission pipelines. *J. Nat. Gas Sci. Eng.* 60, 256–270. <https://doi.org/10.1016/j.jngse.2018.10.022>.
- Liu, X., Huang, Y., Babaei, S., Hazra, B., Ostadhassan, M., 2024. Experimental and molecular simulation of carbon dioxide solubility in hexadecane at varying pressures and temperatures. *Chem. Eng. J.* 502, 157721. <https://doi.org/10.1016/j.cej.2024.157721>.
- Liu, Z., Sun, B., Wang, Z., Zhang, J., Wang, X., 2021. Prediction and management of hydrate reformation risk in pipelines during offshore gas hydrate development by depressurization. *Fuel* 291, 120116. <https://doi.org/10.1016/j.fuel.2020.120116>.
- Liu, Z., Wang, Z., Chen, L., Chen, L., Li, X., Sun, B., 2022. Experimental and modeling investigations of hydrate phase equilibria in natural clayey-silty sediments. *Chem. Eng. J.* 449, 137557. <https://doi.org/10.1016/j.cej.2022.137557>.
- Liu, Z., Yao, H., Li, Y., Wang, L., Wang, J., Liu, Y., Song, Y., Zhang, L., 2025. Analysis of the effect of pipeline inclination on hydrate growth mechanisms. *J. Chem. Eng. Data*. <https://doi.org/10.1021/acs.jced.4c00516>.
- Men, W., Peng, Q., Gui, X., 2022. Hydrate phase equilibrium determination and thermodynamic modeling of CO₂ epoxy heterocycle water systems. *Fluid Phase Equilib.* 556, 113395. <https://doi.org/10.1016/j.fluid.2022.113395>.
- Menezes, D.É.S.d., Pessôa Filho, P.d.A., Robustillo Fuentes, M.D., 2020. Phase equilibrium for methane, ethane and carbon dioxide hydrates at pressures up to 100 MPa through high-pressure microcalorimetry: experimental data, analysis and modeling. *Fluid Phase Equilib.* 518, 112590. <https://doi.org/10.1016/j.fluid.2020.112590>.
- Meng, S., Wang, W., Wang, Z., 2025. Key factors and interaction mechanisms affecting methane hydrate nucleation in W/O systems. *Fuel* 386, 134260. <https://doi.org/10.1016/j.fuel.2024.134260>.
- Meng, Y., Han, B., Wang, J., Chu, J., Yao, H., Zhao, J., Zhang, L., Li, Q., Song, Y., 2024. Hydrate blockage in subsea Oil/Gas pipelines: characterization, detection, and engineering solutions. *Engineering*. <https://doi.org/10.1016/j.eng.2024.10.020>.
- Mesbah, M., Soroush, E., Rezakazemi, M., 2017. Development of a least squares support vector machine model for prediction of natural gas hydrate formation temperature. *Chin. J. Chem. Eng.* 25, 1238–1248. <https://doi.org/10.1016/j.cjche.2016.09.007>.
- Nait Amar, M., 2021. Prediction of hydrate formation temperature using gene expression programming. *J. Nat. Gas Sci. Eng.* 89, 103879. <https://doi.org/10.1016/j.jngse.2021.103879>.
- Park, K.H., Kim, D.H., Cha, M., 2022. Phase equilibria and spectroscopic identification of structure II hydrates with new hydrate-forming agents (cyclopropylamine and cyclopentylamine). *J. Phys. Chem. C* 126, 13585–13594. <https://doi.org/10.1021/acs.jpcc.2c03288>.
- Park, K.H., Kim, D.H., Yoon, J.-H., Cha, M., 2023. Tuning cyclopanecarboxaldehyde clathrate hydrates for enhancement of methane storage. *Chem. Eng. J.* 477, 147015. <https://doi.org/10.1016/j.cej.2023.147015>.

- Saberi, A., Alamdari, A., Shariati, A., Mohammadi, A.H., 2018. Experimental measurement and thermodynamic modeling of equilibrium condition for natural gas hydrate in MEG aqueous solution. *Fluid Phase Equilib.* 459, 110–118. <https://doi.org/10.1016/j.fluid.2017.11.034>.
- Sakai, T., 2020. Emulsifier-free emulsions. *Jpn. Soc. Colour Material* 93, 105–110. <https://doi.org/10.4011/shikizai.93.105>.
- Semenov, A.P., Mendgaziev, R.I., Stoporev, A.S., Istomin, V.A., Sergeeva, D.V., Ogienko, A.G., Vinokurov, V.A., 2021. The pursuit of a more powerful thermodynamic hydrate inhibitor than methanol. Dimethyl sulfoxide as a case study. *Chem. Eng. J.* 423, 130227. <https://doi.org/10.1016/j.cej.2021.130227>.
- Seo, Y., Lee, S., Cha, I., Lee, J.D., Lee, H., 2009. Phase equilibria and thermodynamic modeling of ethane and propane hydrates in porous silica gels. *J. Phys. Chem. B.* 113, 5487–5492. <https://doi.org/10.1021/jp810453t>.
- Shen, X., Du, C., Jiang, S., Sun, L., Chen, L., 2023. Enhancing deep neural networks for multivariate uncertainty analysis of cracked structures by POD-RBF. *Theor. Appl. Fract. Mech.* 125, 103925. <https://doi.org/10.1016/j.tafmec.2023.103925>.
- Shi, L., He, Y., Lu, J., Liang, D., 2020. Effect of dodecyl dimethyl benzyl ammonium chloride on CH₄ hydrate growth and agglomeration in oil-water systems. *Energy* 212, 118746. <https://doi.org/10.1016/j.energy.2020.118746>.
- Shi, L., He, Y., Lu, J., Hou, G., Liang, D., 2021. Anti-agglomeration evaluation and Raman spectroscopic analysis on mixed biosurfactants for preventing CH₄ hydrate blockage in n-octane + water systems. *Energy* 229, 120755. <https://doi.org/10.1016/j.energy.2021.120755>.
- Sun, S.C., Zhang, Y., Kong, Y.Y., Liu, C.L., Liu, Y.F., 2015. Preliminary study on measurement technology for hydrate phase equilibrium. *Fluid Phase Equilib.* 403, 60–69. <https://doi.org/10.1016/j.fluid.2015.06.010>.
- Sirovich, L., 1987. Turbulence and the dynamics of coherent structures. I. Coherent structures. *Q. Appl. Math.* 45, 561–571. <https://doi.org/10.1090/qam/910462>.
- Sloan, E.D., 2003. Fundamental principles and applications of natural gas hydrates. *Nature* 426, 353–359. <https://doi.org/10.1038/nature02135>.
- Sloan, E.D., 2005. A changing hydrate paradigm—from apprehension to avoidance to risk management. *Fluid Phase Equilib.* 228–229, 67–74. <https://doi.org/10.1016/j.fluid.2004.08.009>.
- Sloan, E.D., 2010. *Natural Gas Hydrates in Flow Assurance*. Gulf Professional Publishing.
- Smith, C., Pack, D., Barifcani, A., 2017. Propane, n-butane and i-butane stabilization effects on methane gas hydrates. *J. Chem. Thermodyn.* 115, 293–301. <https://doi.org/10.1016/j.jct.2017.08.013>.
- Sun, H., Wang, S., Wang, S., Song, Y., Ling, Z., Zhang, L., 2025. Rapid processing of dispersed hydrate-based purification for high-yield seawater desalination and wastewater treatment. *Desalination* 604, 118741. <https://doi.org/10.1016/j.desal.2025.118741>.
- Sun, H., Chen, J., Ji, X., Karunakaran, G., Chen, B., Ranjith, P.G., Song, Y., Yang, M., 2024. Optimizing CO₂ hydrate storage: Dynamics and stability of hydrate caps in submarine sediments. *Appl. Energy* 376, 124309. <https://doi.org/10.1016/j.apenergy.2024.124309>.
- Sun, J., Jiang, L., Chou, I.M., Nguyen, N.N., Nguyen, A.V., Chen, Y., Lin, J., Wu, C., 2023. Thermodynamic and kinetic study of methane hydrate formation in surfactant solutions: from macroscale to microscale. *Energy* 282, 128356. <https://doi.org/10.1016/j.energy.2023.128356>.
- Wang, H.J., Jin, T., 2019. Comparative Study of BP Neural Network and RBF Neural Network in Surface Reconstruction. In: 2019 4th International Conference on Mechanical, Control and Computer Engineering (ICMCCE), pp. 405–4054.
- Wang, J., Zhang, L., Hassanpouryouzband, A., Sun, X., Liu, Y., Zhao, J., Song, Y., 2025a. Thermal Marangoni natural convection enables directional transport across immiscible liquids. *Nat. Commun.* 16, 5727. <https://doi.org/10.1038/s41467-025-60930-y>.
- Wang, L., Chen, J., Ma, T., Jing, J., Lei, L., Guo, J., 2024. Experimental study of methane hydrate formation and agglomeration in waxy oil-in-water emulsions. *Energy* 288, 129945. <https://doi.org/10.1016/j.energy.2023.129945>.
- Wang, Q., Wang, Z., Li, P., Song, Y., Wang, D., 2022. Numerical modeling of coupled behavior of gas production and mechanical deformation of gas hydrate reservoir in Shenhu area, South China Sea: enlightenments for field monitoring and model verification. *Energy* 254, 124406. <https://doi.org/10.1016/j.energy.2022.124406>.
- Wang, T., Hassanpouryouzband, A., Fan, M., Mebrahtu, C., Zhang, L., Song, Y., 2025b. Organic magnetic nanoparticles catalyze CO₂ capture in hydrogen-bonded nanocages via water-driven crystallization. *Nat. Commun.* 16, 3702. <https://doi.org/10.1038/s41467-025-58734-1>.
- Wu, M., Sun, H., Liu, Q., Lv, X., Chen, B., Yang, M., Song, Y., 2025. Enhancing CO₂ sequestration safety with hydrate caps: A comparative study of CO₂ injection modes and saturation effects. *Energy* 320, 135044. <https://doi.org/10.1016/j.energy.2025.135044>.
- Xiao, C., Zhang, S., Hu, Y., Gu, X., Ma, X., Zhou, T., Jin, J., 2024. Robust optimization of geoenergy production using data-driven deep recurrent auto-encoder and fully-connected neural network proxy. *Expert Syst. Appl.* 242, 122797. <https://doi.org/10.1016/j.eswa.2023.122797>.
- Yang, M., Zhang, L., Song, W., Chen, B., Song, Y., 2024. A method of cyclic icing and melting for stable and rapid formation of hydrate: Novel strategy of hydrate-based energy storage. *J. Energy Storage* 98, 112839. <https://doi.org/10.1016/j.jest.2024.112839>.
- Zahedi, G., Karami, Z., Yaghoobi, H., 2009. Prediction of hydrate formation temperature by both statistical models and artificial neural network approaches. *Energy Convers. Manag.* 50, 2052–2059. <https://doi.org/10.1016/j.enconman.2009.04.005>.
- Zhang, L.X., Sun, M.R., Wang, T., Yang, L., Zhang, X.T., Zhao, J.F., Song, Y.C., 2022. An in-situ mri method for quantifying temperature changes during crystal hydrate growths in porous medium. *J. Therm. Sci.* 31, 1542–1550. <https://doi.org/10.1007/s11630-022-1674-x>.
- Zhang, L.X., Yang, L., Wang, J.Q., Zhao, J.F., Dong, H.S., Yang, M.J., Liu, Y., Song, Y.C., 2017. Enhanced CH₄ recovery and CO₂ storage via thermal stimulation in the CH₄/CO₂ replacement of methane hydrate. *Chem. Eng. J.* 308, 40–49. <https://doi.org/10.1016/j.cej.2016.09.047>.
- Zhang, M., Dabirian, R., Mohan, R.S., Shoham, O., 2018. Effect of shear and water cut on phase inversion and droplet size distribution in oil–water flow. *J. Energy Resour. Technol.* 141. <https://doi.org/10.1115/1.4041661>.
- Zhao, M., Chen, J.C., Xu, M.Y., 2019. Comparison of Accuracy Estimation for Weighted k-nearest Neighbor Classifiers. In: 5th International Conference on Fuzzy Systems and Data Mining (FSDM), Kitakyushu, Japan, pp. 783–791.
- Zhong, J., Wang, Z., Li, L., Guo, M., Zhang, J., Wang, F., Zhang, J., Wang, Z., 2024. Resolving hydrate inhibition mechanism: interactions between kinetic hydrate inhibitors and CH₄ bubble. *Chem. Eng. J.* 490, 151440. <https://doi.org/10.1016/j.cej.2024.151440>.
- Zhu, Y., Wang, H., Vano, K., 2022. Applying the wavelet neural network to estimate hydrogen dissolution in underground sodium chloride solutions. *Int. J. Hydrogen Energy* 47, 22720–22730. <https://doi.org/10.1016/j.ijhydene.2022.05.130>.