Journal of
Marine Science and Engineering

MDPI

Article

Evaluation of Gas Hydrate Saturation Based on Joint Acoustic–Electrical Properties and Neural Network Ensemble

Donghui Xing [1,2], Hongfeng Lu [1,2,*], Lanchang Xing [3], Chenlu Xu [1,2], Jinwen Du [1,2], Xinmin Ge [4] and Qiang Chen [5,6,*]

1 Guangzhou Marine Geological Survey, China Geological Survey, Ministry of Natural Resources, Guangzhou 511458, China; xingdh126@126.com (D.X.)
2 National Engineering Research Center of Gas Hydrate Exploration and Development, Guangzhou 511458, China
3 College of Control Science and Engineering, China University of Petroleum (East China), Qingdao 266580, China
4 School of Geosciences, China University of Petroleum (East China), Qingdao 266580, China
5 The Key Laboratory of Gas Hydrate, Qingdao Institute of Marine Geology, Ministry of Natural Resources, Qingdao 266237, China
6 Laboratory for Marine Mineral Resources, Qingdao Marine Science and Technology Center, Qingdao 266237, China
* Correspondence: gmgslhf@126.com (H.L.); chenqiang_hds@126.com (Q.C.)

Abstract: Natural gas hydrates have great strategic potential as an energy source and have become a global energy research hotspot because of their large reserves and clean and pollution-free characteristics. Hydrate saturation affecting the electrical and acoustic properties of sediments significantly is one of the important parameters for the quantitative evaluation of natural gas hydrate reservoirs. The accurate calculation of hydrate saturation has guiding significance for hydrate exploration and development. In this paper, experiments regarding methane hydrate formation and dissociation in clay-bearing sediments were carried out based on the Ultrasound Combined with Electrical Impedance (UCEI) system, and the measurements of the joint electrical and acoustic parameters were collected. A machine learning (ML)-based model for evaluating hydrate saturation was established based on electrical–acoustic properties and a neural network ensemble. It was demonstrated that the average relative error of hydrate saturation calculated by the ML-based model is 0.48%, the average absolute error is 0.0005, and the root mean square error is 0.76%. The three errors of the ensemble network are lower than those of the Archie formula and Lee weight equation. The ML-based modeling method presented in this paper provides insights into developing new models for estimating the hydrate saturation of reservoirs.

Keywords: gas hydrate; hydrate-bearing sediments; joint acoustic–electrical properties; hydrate saturation; neural network ensemble

check for updates

Citation: Xing, D.; Lu, H.; Xing, L.; Xu, C.; Du, J.; Ge, X.; Chen, Q. Evaluation of Gas Hydrate Saturation Based on Joint Acoustic–Electrical Properties and Neural Network Ensemble. *J. Mar. Sci. Eng.* **2024**, *12*, 2163. https://doi.org/10.3390/jmse12122163

Academic Editor: Atilla Incecik

Received: 17 October 2024
Revised: 22 November 2024
Accepted: 25 November 2024
Published: 27 November 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Natural gas hydrates (NGHs), also known as "combustible ice", are ice-like solid compounds with a cage structure formed under special conditions such as low temperature and high pressure [1–3]. NGHs are considered to be an effective clean energy source in the future [4], which are usually distributed in continental permafrost and marine sediments and also widely recorded in the geological record [5,6]. Their huge energy and economic value has made countries all over the world devote themselves to the research, exploration, and development of natural gas hydrates.

Hydrate saturation is one of the main parameters for the evaluation of hydrate-bearing reservoirs. Due to the higher acoustic velocity and resistivity of solid hydrates relative to the fluid in sediments, hydrate-bearing sediments tend to show relatively high acoustic velocity and resistivity [7]. There are many logging methods used in hydrate formation,

such as resistivity, acoustic, nuclear magnetic, dielectric, density, and neutron methods. However, conventional hydrate saturation evaluation methods based on logging data still use evaluation methods for oil and gas fields, mainly based on the resistivity method and sound velocity method. The resistivity method mainly includes the Archie formula [8], modified Archie formula with mud, Indonesia equation, and dual-water model [9]. The velocity of sound method mainly includes the time-averaged equation [10–12], wood equation [13], Lee weight equation [14], equivalent medium theory (EMT) [15], and improved Biot–Gassmann formula (BGTL) [16]. However, these models based on core testing have certain limitations in field applications [17]. For example, Archie's formula in the resistivity method is applicable to pure pore-filled sandstone reservoirs [18,19], and the empirical coefficient in the formula can be affected by mineral composition, porosity, pore water composition, and concentration, so it needs to be calibrated according to the actual reservoir situation. There are unknown mineralogical and bulk modulus parameters in the sonic method, and it is necessary to assume that the shape of the gas hydrate in sediment pores (such as particle cementation, particle wrapping, particle support, and pore filling) is known, and this cannot be directly measured in the logging field or obtained from sample examination in the lab [20].

Data-driven models, such as artificial neural networks and deep learning models, have been widely used to solve such engineering problems in recent years [21–23]. Neural networks have great applicability when using logging data to calculate hydrate saturation. Their high self-learning and self-adaptive ability makes them more advantageous when dealing with the complex nonlinear mapping problem between hydrate saturation and electroacoustic parameters. These advantages attract some researchers to use neural networks for saturation prediction [24–29]. Actually, the data used in the training network are logging data. Affected by the resolution of logging tools and logging depth, the amount of physical information entered into the network model is limited, and the data themselves have large errors. Moreover, the target value of hydrate saturation in the network training is obtained by the NMR method, which has low accuracy when used in the logging field.

A test system called UCEI (Ultrasound Combined with Electrical Impedance) for conducting methane hydrate experiments in unconsolidated sediments using joint electrical–acoustic parameters was developed in Xing et al. [30]. In this paper, experiments regarding methane hydrate formation and dissociation were carried out based on the UCEI system, and joint electrical and acoustic parameters were collected. A machine learning (ML)-based model for evaluating hydrate saturation was established based on electrical–acoustic properties and a neural network ensemble, which provides insights into developing new models of the joint interpretation of the electrical and sonic logging data for estimating the hydrate saturation of reservoirs.

2. Materials and Methods

2.1. Materials

The experiments were conducted on the UCEI system (shown in Figure 1), and the details of the system and experimental procedure were presented in Xing et al. [30], not repeated here for brevity. The experimental materials include natural marine sands, montmorillonites, methane gas, NaCl, and deionized water. Natural marine sands with a grain diameter of 180–250 μm and montmorillonite with a median particle size of 2.6 μm were used to simulate the unconsolidated marine sediments and clay, respectively. A 3.5% NaCl solution and methane gas were used to synthesize methane hydrates. The marine sands were washed with deionized water several times before putting them into a drying oven. The oven worked at 150 °C for at least 12 h to dry the washed sands. The dried sands were then mixed with a certain amount of montmorillonite, e.g., 10% of the sands by volume, to simulate a clay-bearing sediment. The sediment was wetted by NaCl solution and filled layer by layer into a sample vessel with 120 mm internal diameter and 250 mm internal height. The initial pressure was 6 MPa, and the temperature was gradually reduced from 10.4 °C to 1 °C.

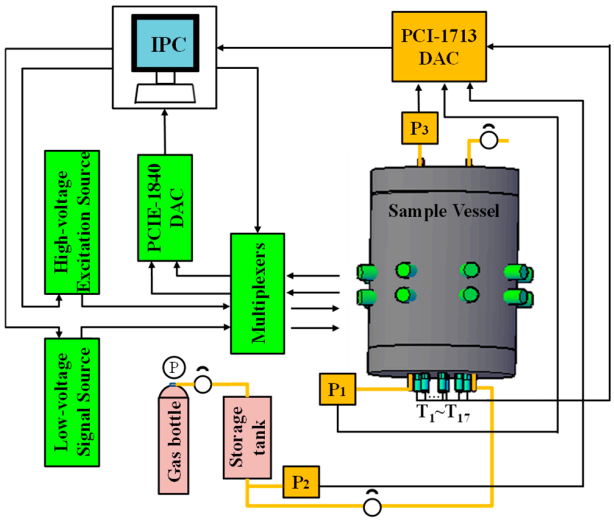


Figure 1. A schematic diagram of the structure of the UCEI experimental system.

2.2. Test of Joint Acoustic–Electrical Properties

The distribution of hydrates in the reactor presents the characteristics of uneven spatial distribution. To obtain the spatial distribution information of hydrates, both acoustic and electrical parameters were measured with UEC (Ultrasonic–Electrical Compound) sensors, and the saturation of hydrates was calculated according to the principle of gas consumption [30].

As shown in Figure 2, the UEC sensors consist of both ultrasonic (denoted by U) and electrical (denoted by E) components, and each of the two sensors with connecting lines in the figure together form an active electrode pair or ultrasonic transceiver unit, which was used to measure the electrical signals of the individual electrode pairs and ultrasonic waves from the transceiver unit. During the formation of gas hydrates, electrical and acoustic data were measured every 15 min. In this experiment, the electrical data consist of 5984 impedance mode values and 5984 absolute phase angle values at 22 test frequency points in four groups of different measurement orientations, and the acoustic data consist of 272 sound velocity data points in four groups of different measurement orientations.

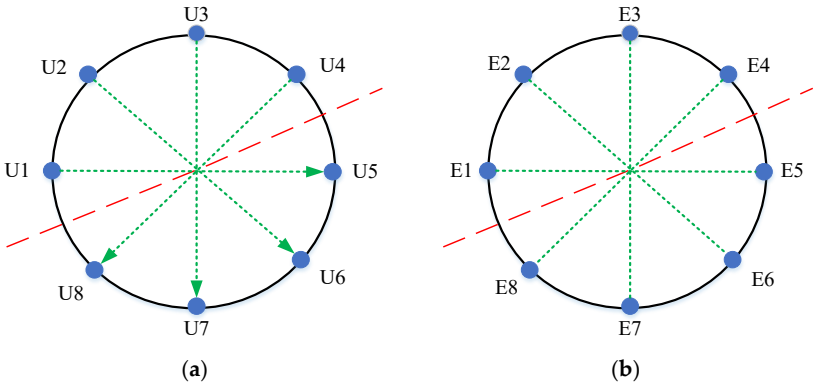


Figure 2. An illustration of the working mode for UEC sensors. (**a**) The ultrasonic part. (**b**) The electrical part.

3. Modeling Methods

3.1. Theory of Modeling with NN Ensemble

A BP (Back-Propagation) neural network includes three or more layers. Each layer of the network is connected from the input layer to the output layer in turn, and the connection is weighted to form the topological structure of the network, as shown in Figure 3.

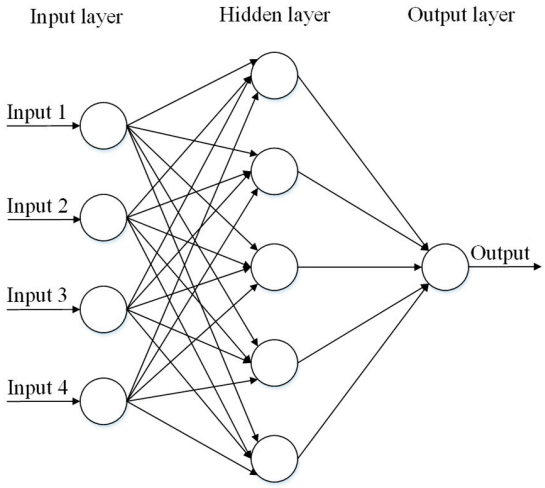


Figure 3. BP network.

BP neural networks are common algorithms for training network weights since they have good performance in nonlinear mapping and are self-learning and self-adaptive. However, traditional BP networks use the gradient descent method for training network weights, which will inevitably lead to a local minimum and slow convergence speed, which affects the improvement in network generalization ability to a certain extent. In view of the problems of a single neural network, Hansen et al. put forward the idea of a neural network ensemble [31]. Schapire created the boosting algorithm to realize neural network integration [32]. In practical use, it is found that the algorithm has disadvantages: it requires prior knowledge of the weak learner; that is, it needs to know the lower limit for the algorithm to learn correctly. In order to solve this problem, Freund et al. created an adaptive boosting (AdaBoost) algorithm based on a boosting algorithm [33]. The algorithm has an efficiency close to the efficiency of the boosting algorithm, and it is suitable for dealing with practical problems and avoids the overfitting problem to a great extent. In the first mock exam, the Adaboost algorithm and neural network and classification and regression tree have better classification or regression effects than the single model [34–37]. Based on this, this study uses the AdaBoost algorithm to integrate a BP neural network and uses an integrated BP neural network model to predict hydrate saturation.

Let the training sample set be $D = \{(x_1, y_1), (x_2, y_2), \cdots, (x_m, y_m)\}$, the number of weak learners be T, the weak learner algorithm be G(x), and the final strong learner be H(x).

The weight distribution of training samples is initialized:

$$D_1 = (w_{11}, w_{12}, \cdots w_{1m}); w_{1i} = \frac{1}{m}; i = 1, 2, 3, \cdots m. \tag{1}$$

For iteration rounds, k = 1, 2, ……, T:

(a) Using the sample set with weight D_k to train the data, the weak learner $G_k(x)$ is obtained.

(b) The maximum error in the training set is obtained:

$$E_k = \max|y_i - G_k(x_i)| \quad i = 1, 2, 3, \cdots m. \tag{2}$$

(c) For each sample, its relative error is calculated:

$$e_{ki} = \frac{|y_i - G_k(x_i)|}{E_k}. \tag{3}$$

(d) The regression error rate of the weak learner on the training set is calculated:

$$e_k = \sum_{i=1}^{m} w_{ki} e_{ki}. \tag{4}$$

(e) The weight coefficient of weak learner a_k is obtained:

$$a_k = \frac{e_k}{1 - e_k}. \tag{5}$$

(f) The weight distribution of the updated sample set is $D_{k+1} = (w_{k+1,1}, w_{k+1,2}, \cdots w_{k+1,m})$:

$$w_{k+1,i} = \frac{w_{ki}}{Z_k} a_k^{1-e_{ki}}. \tag{6}$$

where Z_k is the normalization factor, $Z_k = \sum_{i=1}^{m} w_{ki} a_k^{1-e_{ki}}$.

(g) The final strong learner is built:

$$\mathrm{H}(x) = \sum_{t=1}^{T} (\ln \frac{1}{a_t}) G_k(x). \tag{7}$$

The main goal of the Adaboost algorithm is to improve the sample weight with a large training error in the training set and the output weight of the strong learner with a strong prediction effect and reduce the sample weight with a small training error in the training set and the output weight of the weak learner with a weak prediction effect. There are two parameters to be considered in the Adaboost algorithm. One is the number of weak learners T, which is the number of iterations. The other is the preset prediction error threshold E. When determining these two parameters, the minimum root mean square error of the test results is taken as the constraint condition.

In this paper, the modeling processes include the construction of a sample set and the determination of structural parameters. The electrical parameters were composed of the impedance modulus and phase angle at 22 test frequency points in four different measurement directions. The acoustic parameters were sound velocity in four different measurement directions. These experimental data were used as the input of the BPNN-Ada model. The overall temperature and pressure in the reactor can be calculated by the gas equation of state, which can be used as the output of the BPNN-Ada model. Based on the above parameters, an electrical model and an acoustic model were established by using the integrated BP neural network to calculate hydrate saturation. On this basis, the weights of the output results of the model were allocated according to the error of the model. The combined strategy of the weighted average was used to fuse the electrical model and the acoustic model to establish the joint electrical–acoustic saturation calculation model, compared with the conventional saturation calculation model. Figure 4 presents a flow chart of the hydrate saturation calculation model based on acoustic–electrical characteristic parameters.

3.2. Construction of Sample Sets

The quality of the sample sets is very important in the construction of the network model. The construction of the sample set includes data preprocessing, data feature extraction, partitioning the sample set, and data normalization.

(1) Data preprocessing. The first step is to preprocess the experimental data due to the presence of noises and outliers in the raw data, mainly represented by abnormal data. For the extreme large and small values, the mean interpolation correction method was used. The wavelet denoising and moving average filtering approach proposed in Xing et al. was used to denoise the data [30].

(2) Feature selection. Computational complexity increases exponentially with the increase in data dimension for the input layer of the network. Therefore, it is necessary to select features from the input data so as to compress the data dimension while retaining the intrinsic features of the raw data. The PCA (principal component analysis) method used in this work is a linear dimensionality reduction method [38], which uses a linear projection method to map the high-dimensional data to a low-dimensional space to

maximize the variance in the projected data. The resulting principal components are linear combinations of the original variables, but they are uncorrelated to each other. A whole sample set was constructed from the selected features (the input of the model) and the corresponding hydrate saturation (the output of the model).

(3) Partitioning sample sets. To obtain an efficient neural network model, it is necessary to divide the whole dataset into three parts, i.e., the training set, verification set, and test set. In this work, half of the samples were used for training the network, a quarter were used for validation, and the remaining samples were used for testing.

(4) Data normalization. Data normalization may eliminate the influence of dimension and accelerate the convergence speed of the network. In this paper, the raw data x are mapped to the value $\overline{x}$ of the interval [0, 1] through the max-min normalization method. The calculation formula is as follows:

$$\overline{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{8}$$

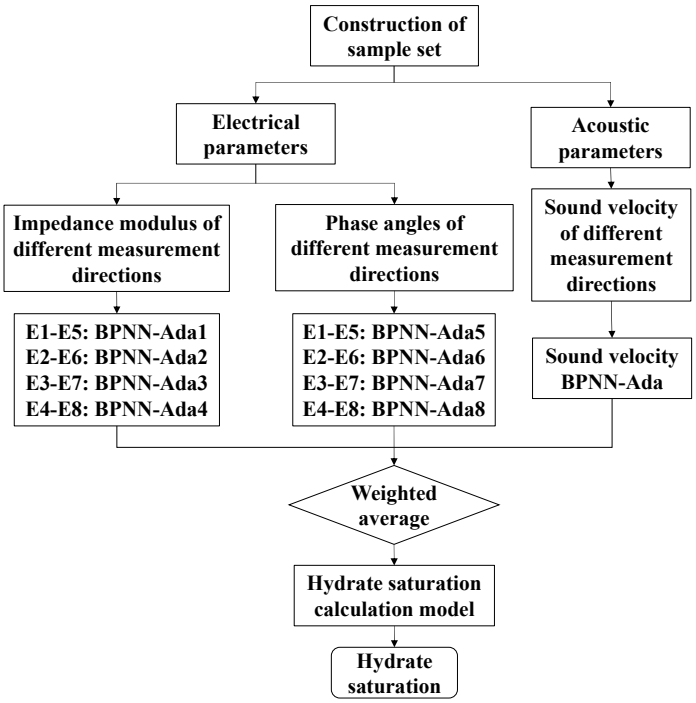


Figure 4. Flow chart of hydrate saturation calculation model based on acoustic–electrical characteristic parameters.

3.3. Determination of Model Parameters

A three-layer BP neural network was selected to establish neural network models for hydrate saturation calculation [39]. Three models were built with the input as the broadband impedance modulus value, broadband phase angle, and sound velocity.

The activation function of the hidden layer of the BP neural network employs a tangent sigmoid function, whose input value can be any value, and the output value ranges from −1 to 1. The activation function of the network output layer adopts a linear function, and its input value and output value can be arbitrary. The learning method of the neural network uses the momentum gradient descent method [40], whereby a portion of the preceding weight adjustment is superimposed upon the weight adjustment calculated according to the error, thus forming the weight adjustment amount for the current time.

$$\Delta w(n) = -\eta \Delta E(n) + \alpha \Delta w(n-1) \tag{9}$$

where α is the momentum coefficient, usually $0 < \alpha < 1$; η is the learning rate, which is in the range of 0.001~10. The momentum factor introduced by the method is, in fact, equivalent to the damping term, which serves to reduce the oscillation trend in the learning process, enhance convergence, diminish the sensitivity of the network to the local details of the error surface, and effectively prevent the network from falling into a local minimum.

The number of nodes in the hidden layer of the BP neural network significantly affects the accuracy of network prediction. If the number of nodes is insufficient, the network will not learn effectively, so it needs to increase the amount of network training, which will also affect the performance of the network. Too many nodes will increase the training time of the network, which will lead to the problem of overfitting, which will affect the generalization ability of the network. In order to optimize the number of nodes in the hidden layer, a reference formula was employed to determine the approximate distribution range of the optimal number of nodes. Subsequently, the trial-and-error method was utilized to ascertain the optimal number of nodes in the hidden layer, with the objective of minimizing the specified error. The following formula can be used to identify the optimal number of nodes in the hidden layer [41]. In this formula, L represents the number of nodes in the hidden layer, n denotes the number of nodes in the input layer, m is the number of nodes in the output layer, and α is a constant value between 1 and 10.

$$L = \sqrt{n + m} + \alpha \tag{10}$$

In order to analyze and evaluate the prediction effect of the model more comprehensively, the following three error indexes are used for evaluation:

The mean relative error (MRE):

$$\mathrm{RE_i} = \frac{\mathrm{V_{Pi}} - \mathrm{V_{Ai}}}{\mathrm{V_{Ai}}} \times 100\% \tag{11}$$

$$\mathrm{MRE} = \frac{1}{\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} |\mathrm{RE_i}| \times 100\% \tag{12}$$

The mean average error (MAE):

$$\mathrm{MAE} = \frac{1}{\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} |\mathrm{V_{Pi}} - \mathrm{V_{Ai}}| \tag{13}$$

The root mean square error (RMSE):

$$\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} (\mathrm{RE_i})^2} \times 100\% \tag{14}$$

4. Results and Discussion

4.1. Hydrate Formation and Dissociation Process

Figure 5 shows the time traces of the pressure and temperature during the hydrate formation and dissociation process. The initial pressure and temperature are 6 MPa and 10.4 °C, respectively. The experimental process can be observed from Figure 5: during the 0~6.3 h stage, with the gradual decrease in temperature in the reactor, methane hydrate is formed slowly, and the hydrate saturation rises to 4.64%. During the period of 6.3~30 h, methane hydrate is formed in a large quantity, and the hydrate saturation increases rapidly to 17.60%. During the period of 30~43 h, the hydrate formation process basically ends, and the hydrate saturation finally reaches 18.28%.

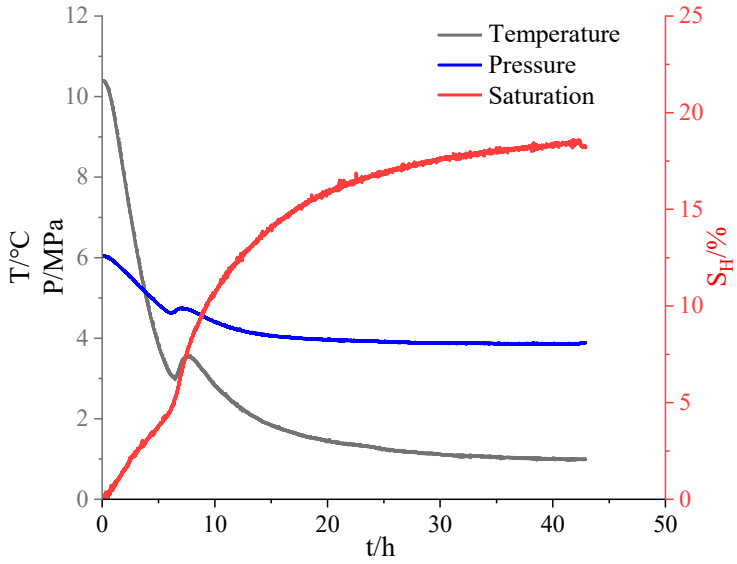


Figure 5. Time traces of pressure, temperature, and hydrate saturation.

There are two possible reasons for the low saturation of the hydrate generated in the experiment: (1) the presence of clay minerals filling part of the pore throat of the sediment makes it difficult for the hydrate to fill all the pore spaces, and (2) with the continuous formation of methane hydrate, a dense hydrate layer is formed in the pores of the sediment [42], thus limiting the contact between methane gas and pore water, resulting in low hydrate saturation.

4.2. Analysis of Acoustic Characteristics

The ultrasonic signals of each receiving probe were collected during the experiment. Figure 6 shows the variation curve of the P-wave velocity of sediments with different hydrate saturations. The difference in velocity measured by different sensors reveals the uneven distribution of gas hydrates in clay-bearing sediments.

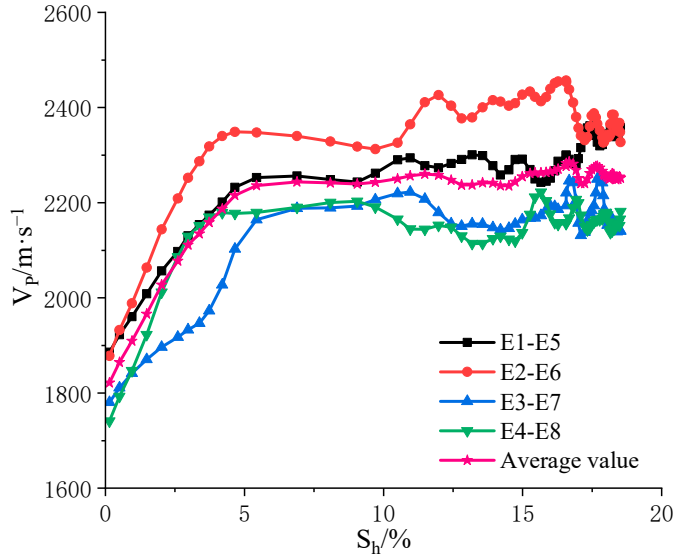


Figure 6. Variation curve of P-wave velocity with hydrate saturation during methane hydrate formation.

The initial average P-wave velocity of hydrate-free sediments is 1821 m/s. With the formation of methane hydrate, the sound velocity begins to increase sharply. When the saturation reaches 1%, the average P-wave velocity reaches 1966 m/s. The reason is that in loose sediments, sediment particles are not cemented, and the skeleton is easily affected

by the formation of hydrate. A small amount of hydrate may increase the cementation degree of the sediment particles and make the sound velocity rise sharply. With the further formation of methane hydrate, the sound velocity of hydrate-bearing sediments continues to increase. When the saturation reaches 3.8%, the average P-wave velocity reaches 2158 m/s. After that, the hydrate saturation continues to increase, and the rise in sound velocity become smaller. It is speculated that the hydrate distribution mode changed into contact or suspension mode, which results in a smaller effect of the hydrate on the elastic modulus. When the hydrate saturation is stable at 18.5%, the average P-wave velocity is 2253 m/s, which is 432 m/s higher than the initial P-wave velocity without the hydrate.

4.3. Analysis of Electrical Characteristics

During the experiment, the electrical signals of each pair of working electrodes are collected. Based on the principle of impedance measurement, the measured signals are processed and calculated, and then the impedance modulus and phase angle are obtained. The two parameters cover the resistance and capacitance information of the measured medium and can comprehensively describe the resistance and capacitance characteristics of the measured medium.

Figure 7 shows the impedance mode value and phase angle dispersion characteristic curve of methane hydrate in different measurement directions in the measurement range of 5~40 kHz at the stable stage after the formation of methane hydrate. It can be seen from the figure that within the test frequency range, the impedance mode value gradually decreases with the increase in frequency. The absolute value of the phase angle decreases first and then increases with an increase in frequency, and the turning point frequency is about 1 kHz. The above phenomenon can be explained by "the complex dispersion characteristics of different test bands are dominated by different polarization mechanisms". In clay-bearing sediments, there are two possible conductive paths of applied electric fields: one is through pore water; the other is along the double electric layer (DEL) of a clay mineral surface, which is called surface conductivity [43–45]. When the frequency is between 5 Hz and 1 kHz, surface conductivity plays the leading role. With an increase in frequency, the polarization effect of the DEL gradually weakens, the storage capacity weakens, and the absolute value of the phase angle decreases. When the applied electric field frequency in the experiment is between 1 kHz and 40 kHz, interface polarization plays the leading role. With an increase in the test frequency, the directional movement of conductive ions in pore water gradually lags behind the variation in the applied electric field, the interface polarization effect gradually increases, and the absolute value of the phase angle increases.

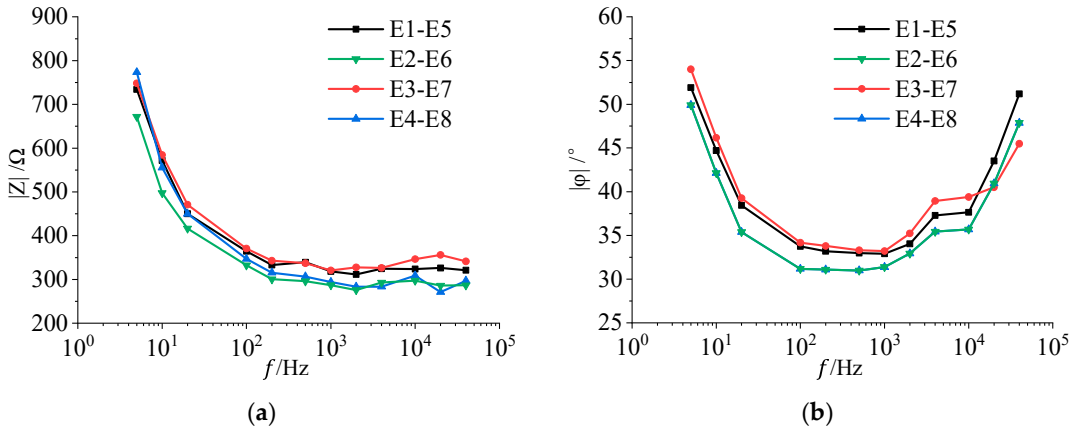


Figure 7. Dispersion curves of electrical parameters in different measurement directions at stable stage (SH = 18%) after hydrate formation. (**a**) Impedance mode value; (**b**) absolute value of phase angle.

Figure 8 shows the change curve of the impedance mode value and phase angle with hydrate saturation at the characteristic frequency point (1 kHz) of sensor pairs in different

measuring directions during hydrate formation. It can be seen from the figure that each pair of sensors tests the electrical parameters of hydrate-bearing sediments in the reactor from different measuring directions. The test data reflect the uneven distribution of hydrates in the cross-section of the reactor, and the average value of the test data of each pair of sensors reflects the average distribution of hydrates in the cross-section. The impedance mode value and phase angle show a nonlinear growth trend with the increase in hydrate saturation, which confirms the results of Schmutz et al. [46]. The change process of the impedance modulus can be divided into two stages: in the initial stage, the hydrate saturation is below 10%, and the salt drainage effect is significant, which makes the impedance modulus increase slowly. With the increase in hydrate saturation, the high resistivity of the hydrate and the plugging effect of the hydrate on sediment pores gradually occupy the dominant position and continue to strengthen, resulting in a rapid increase in the impedance modulus.

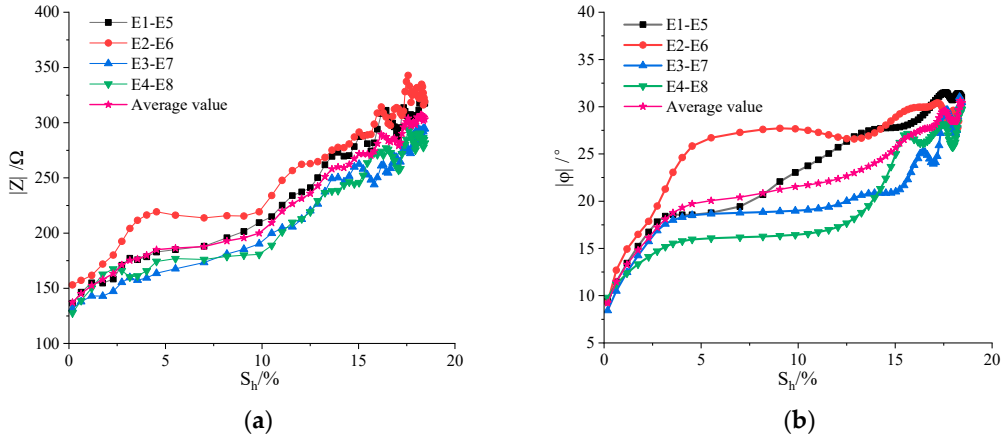


Figure 8. Variation curve of impedance modulus and phase angle (1 kHz) with hydrate saturation during hydrate formation. (**a**) Impedance mode value; (**b**) absolute value of phase angle.

4.4. Modeling and Performance Assessment

The whole dataset for this model covers a total of 5984 impedance mode values, 5984 absolute phase angle values at 22 test frequency points, and 272 sound velocity data points. The data were measured in four different measurement directions and 68 different saturations, and a quarter, that is, 17 of which, were used for validation.

Taking the impedance mode value of the E1–E5 sensor group as an example, the singular value of the original impedance modulus value is removed first, and the data are denoised using the combination of wavelet denoising and moving average filtering. Then, principal component analysis is carried out for the preprocessed impedance mode values at 22 measuring frequency points. As shown in Figure 9, the contribution rate of the first principal component is 69.84%, and the contribution rates of the second and third principal components are 6.43% and 5.30%. In this work, a cumulative contribution rate of 95% of the principal components is selected, and the number of principal components is finally determined to be seven. The data are reduced from 22 dimensions to 7 dimensions. At the same time, the correlation between impedance modulus values under different frequencies is eliminated.

Next, the PCA data are divided into the training set and test set according to the ratio of 3:1. There are 612 impedance modulus values in the training set and 204 impedance modulus values in the test set.

Finally, the sample set is normalized according to Equation (8). So far, the sample set of the network is determined, and the network starts training.

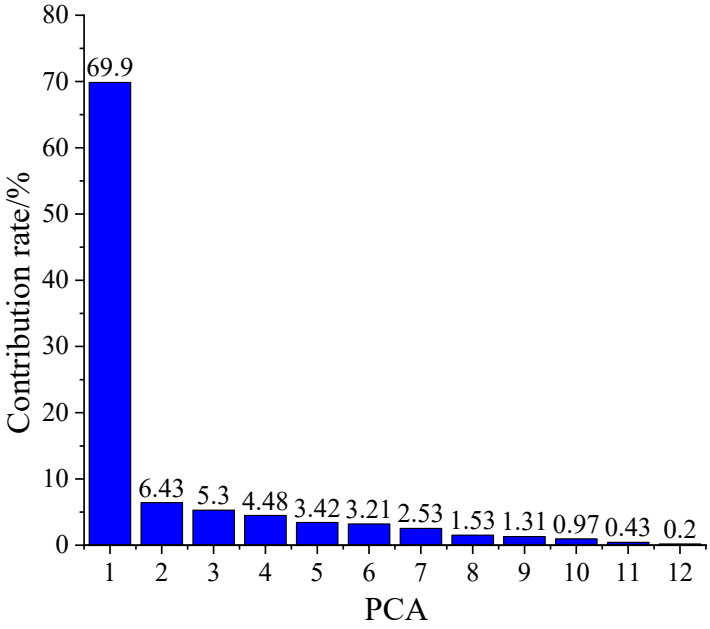


Figure 9. Principal component analysis of broadband impedance modulus of E1–E5 sensor.

According to Equation (10), the range of the hidden layer node number L is set as 4~13. The network with different hidden layer node numbers is trained and tested 500 times. The optimal hidden layer node number is obtained by analyzing and comparing three error specifications (MRE, MAE, and RMSE) of the network. The average results of 500 test errors of neural networks are shown in Table 1: the overall trend in the three error specifications first decreases and then increases, and there is generally a minimum value. When L = 6, the average relative error and root mean square error reach the optimal value. Comprehensively, the number of hidden layer nodes is set to six, and the hydrate saturation BP network model based on the impedance modulus for the E1–E5 sensor pair is preliminarily obtained.

Table 1. Test error table of E1–E5 impedance modulus network with different numbers of hidden layer nodes.

Number of Hidden Layer Nodes (L)	MRE (%)	MAE	RMSE (%)
4	11.13	0.0109	40.76
5	10.06	0.0108	36.54
6	9.74	0.0112	35.26
7	10.06	0.0118	36.23
8	11.35	0.0126	40.98
9	11.13	0.0136	39.80
10	10.21	0.0130	36.39
11	11.46	0.0143	40.80
12	11.73	0.0149	41.70
13	11.88	0.0159	41.93

The integrated BP neural network is integrated based on the optimal BP network structure mentioned above. When two important parameters (the number of weak learning machines T and the prediction error threshold E) in the AdaBoost algorithm are determined, the minimum root mean square error of the test results is taken as the constraint condition (shown in Table 2). Finally, the number of weak learners based on the E1–E5 measurement azimuth impedance modulus is 20, and the error threshold is 0.2.

Table 2. Root mean square error of networks with different parameters.

Error Threshold	Number of Weak Learners of Integrated BP Neural Network Based on AdaBoost Algorithm									
	10	20	30	40	50	60	70	80	90	100
0.05	1.87	2.43	2.32	1.70	2.18	2.41	2.26	2.03	2.38	1.83
0.10	3.50	3.34	2.03	1.87	2.17	2.36	2.08	2.14	1.99	1.98
0.20	3.48	1.68	2.53	2.66	2.36	2.07	2.24	2.06	1.95	2.09
0.30	2.39	2.20	3.70	1.86	2.68	1.95	2.45	2.29	2.23	2.50

A BPNN-Ada saturation calculation model based on the E1–E5 measurement azimuth broadband impedance modulus is established. The same method is used to obtain the other measurement azimuth impedance modulus values and network parameters of phase angle and sound velocity. All network parameters are shown in Table 3.

Table 3. Parameters of different network models.

Sample Set	BPNN-Ada	Number of Hidden Layer Nodes	Number of Weak Learning Machines	Error Threshold
Broadband impedance modulus	E1–E5 E2–E6 E3–E7 E4–E8	6 5 10 11	20	0.2
Broadband phase	E1–E5 E2–E6 E3–E7 E4–E8	7 6 10 8	30	0.1
P-wave velocity	U1–U5, U2–U6, U3–U7, U4–U8	7	40	0.1

According to the absolute error of the BPNN-Ada model prediction results obtained from the training of the impedance modulus value, phase angle, and sound speed of four measurement directions, the weights of the output results of nine groups of models are assigned. According to the new weight distribution results, the combined strategy of the weighted average of the output values of nine groups of models is carried out (refer to Equations (15)–(17)), and the final electroacoustic joint model is obtained. The model can input the broadband impedance modulus, broadband phase angle, and sound velocity of four measurement directions simultaneously, and the input information of the model is abundant.

$$\beta_i = 1 - \frac{|V_{Pi} - V_A|}{V_A} \tag{15}$$

$$\lambda_i = \frac{\beta_i}{\sum_{i=1}^{9} \beta_i} \tag{16}$$

$$S_h = \sum_{i=1}^{9} \lambda_i V_{Pi} \tag{17}$$

where V_{Pi} is the predicted saturation value of the BPNN-Ada model of different sample sets, V_A is the actual saturation value, β_i is the accuracy coefficient, and λ_i is the weighting coefficient.

The test samples are input into the electrical–acoustic joint model to obtain the predicted saturation and compare it with the actual saturation. As shown in Figure 10, the predicted values of the model are all limited to ±5%.

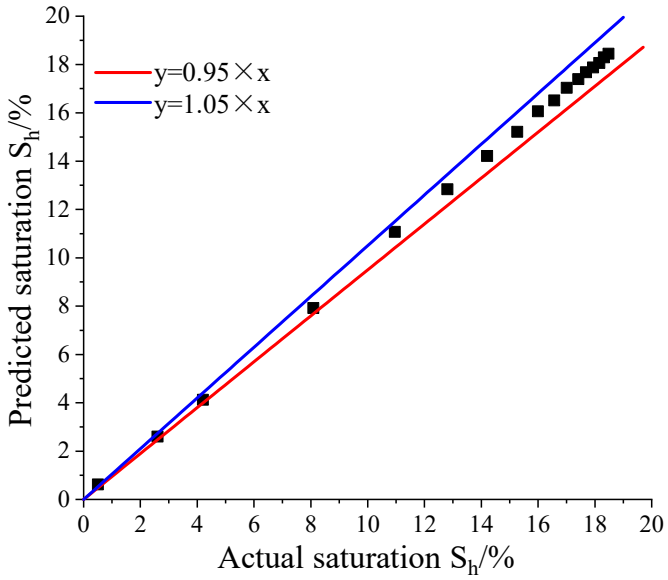


Figure 10. Comparison of the 17 points of predicted saturations and actual saturations by combined electroacoustic model.

Three error specifications are used to compare the prediction results of the acoustic–electrical joint model, Archie's equation model, and the weighted equation by Lee et al. [47]. The comparison results are shown in Table 4. It can be seen that the calculation error of the acoustic–electrical joint model based on the integrated BP neural network is far less than that of Archie's equation and the weighted equation by Lee. The reasons may come from the following aspects.

Table 4. Comparison of electrical model prediction errors.

Method	MRE (%)	MAE	RMSE (%)
Archie equation	8.41	0.0090	12.24
Weighted equation by Lee	14.47	0.0175	20.37
Joint acoustic–electrical model based on BPNN-Ada	0.48	0.0005	0.76

(1) For the sediment with uniform hydrate distribution, resistivity and sound velocity show isotropy to a large extent. Archie's equation for resistivity and the weighted equation by Lee for sound velocity mentioned above can accurately predict hydrate saturation, especially in the sediment with higher saturation. However, when the hydrate is unevenly distributed in the unconsolidated argillaceous sediment layer, the sediment layer will show anisotropy [48]. The calculation of hydrate saturation by the above isotropic method will lead to a large error. In the presence of clay, the electroacoustic parameters of hydrate-bearing sediments cannot be directly combined with the above formula for saturation calculation, so the clay correction method should be considered.

(2) Both the resistivity method and sound velocity method only consider part of the physical properties of the reservoir, which is not suitable for complex formation conditions with the occurrence of clay. In the acoustic model, the influence of the clay bonding degree and hydrate distribution mode should be considered. Meanwhile, for the electrical equation, the different conductive paths introduced by clay should maybe be introduced into the model calculation. Moreover, there are many parameters in the equation, which need to be calibrated according to the actual data. And many errors transfer in the process, which indirectly reduces the accuracy of the model.

(3) The model based on BPNN-Ada is more advantageous when dealing with the complex nonlinear mapping relationship between electroacoustic parameters and saturation. The input information of the model is more exhaustive. The acoustic–electrical parameters have different properties but play a complementary role, and the phase angle parameter which is not considered in the conventional calculation model is introduced. Selecting the weighted average fusion strategy to fuse the electroacoustic BPNN-Ada model in different measurement directions can effectively overcome the adverse impact of the anisotropy of the sediment layer caused by uneven hydrate spatial distribution, enhancing the reliability of the model and prediction accuracy.

5. Conclusions

In this study, formation and dissociation experiments of methane hydrate in simulated marine sediments were carried out. Characteristic parameters such as the impedance modulus, phase angle, and sound velocity were obtained, and the influence of hydrate saturation on the electrical–acoustic characteristic parameters was analyzed. Based on the integrated BP neural network (BPNN-Ada) and the electroacoustic characteristic parameters of methane hydrate sediments in different measurement directions, the electroacoustic joint calculation model of hydrate saturation was established and compared with the Archie formula and Lee weight equation. The results show that the three error standards of the combined model are lower than the calculation errors of the Archie formula and Lee weight equation. Considering different but complementary parameters, the model is more suitable for evaluating hydrate saturation in clay-bearing sediments. The comprehensive application of experimental data in different measurement directions, through the use of the appropriate fusion strategy, may allow us to overcome the adverse effects of the uneven distribution and anisotropy of hydrates on the accuracy of saturation calculation.

In this paper, we preliminarily verify the feasibility of experimental and computational methods. However, only a single sand–clay ratio was used in this experiment to simulate clay-bearing sediments, and different sand–clay ratios or grain sizes should be considered in subsequent studies. In addition, there are areas where the modeling can be optimized, such as the introduction of k-fold cross-validation in the dataset setup and the use of randomized and grid searches to determine hyperparameters.

Author Contributions: Conceptualization, D.X. and L.X.; methodology, D.X.; software, L.X.; validation, H.L., C.X. and J.D.; formal analysis, X.G.; investigation, Q.C.; data curation, Q.C.; writing—original draft preparation, D.X.; writing—review and editing, L.X. and Q.C.; supervision, H.L.; project administration, H.L.; funding acquisition, C.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Key-Area Research and Development Program of Guangdong Province (2023B1111050014), the National Natural Science Foundation of China (No. 42476232), the Shandong Provincial Natural Science Foundation (ZR2024ME090), the Key Program of Marine Economy Development (Six Marine Industries) Special Foundation of Department of Natural Resources of Guangdong Province (GDNRC [2024]48), and the Youth Research Team Project of the National Engineering Research Center of Gas Hydrate Exploration and Development (NERC2024003).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets generated and analyzed during the current study are available from the corresponding author by request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Sloan, E.D. *Clathrate Hydrate of Natural Gases*; Marcel Dekker, Inc.: New York, NY, USA, 1998.
2. Ruppel, C.; Boswell, R.; Jones, E. Scientific results from Gulf of Mexico Gas Hydrates Joint Industry Project Leg1drilling: Introduction and overview. *Mar. Pet. Geol.* **2008**, *25*, 819–829. [CrossRef]

3. Makogon, Y.F. Natural gas hydrates—A promising source of energy. *J. Nat. Gas Sci. Eng.* **2010**, *2*, 49–59. [CrossRef]
4. Lee, M.W.; Collett, T.S. In-situ gas hydrate hydrate saturation estimated from various well logs at the Mount Elbert Gas Hydrate Stratigraphic Test Well, Alaska North Slope. *Mar. Pet. Geol.* **2011**, *28*, 439–449. [CrossRef]
5. Kvenvolden, K.A. A primer on gas hydrates. *US Geol. Surv. Prof. Pap.* **1993**, *1570*, 555–561.
6. Argentino, C.; Conti, S.; Fioroni, C.; Fontana, D. Evidences for Paleo-Gas Hydrate Occurrence: What We Can Infer for the Miocene of the Northern Apennines (Italy). *Geosciences* **2019**, *9*, 134. [CrossRef]
7. Collett, T.S.; Ehlig-Economides, C.A. Detection and Evaluation of the In-Situ Natural Gas Hydrates in the North Slope Region, Alaska. In Proceedings of the SPE California Regional Meeting, Ventura, CA, USA, 23–25 March 1983; p. 11673.
8. Archie, G.E. The electrical resistivity log as an aid in determining some reservoir characteristics. *Trans. AIME* **1942**, *146*, 54–62. [CrossRef]
9. Clavier, C.; Coates, G.; Dumanoir, J. Theoretical and Experimental Bases for the Dual-Water Model for Interpretation of Shaly Sands. *SPE J.* **1984**, *24*, 153–168. [CrossRef]
10. Wyllie, M.R.J.; Gregory, A.R.; Gardner, G.H.F. An experimental investigation of factors affecting elastic wave velocities in porous media. *Geophysics* **1958**, *23*, 459–493. [CrossRef]
11. Pearson, C.F. Natural gas hydrate deposits: A review of in situ properties. *J. Phys. Chem.* **1983**, *87*, 4180–4185. [CrossRef]
12. Pearson, C.; Murphy, J.; Halleck, P.; Hermes, R.; Mathews, M. Sonic and resistivity measurements on Berea sandstone containing tetrahydrofuran hydrates: A possible analogue to natural gas hydrate deposits. In Proceedings of the Fourth International Conference on Permatrost, Fairbanks, AK, USA, 17–22 July 1983; Volume 87, pp. 937–978.
13. Wood, A.B.; Lindsay, R.B. A Textbook of Sound. *Phys. Today* **1956**, *9*, 37. [CrossRef]
14. Lee, M.W.; Hutchinson, D.R.; Collett, T.S.; Dillon, W.P. Seismic velocities for hydrate-bearing sediments using weighted equation. *J. Geophys. Res. Solid Earth* **1996**, *101*, 20347–20358. [CrossRef]
15. Helgerud, M.B.; Dvorkin, J.; Nur, A.; Sakai, A.; Collett, T. Elastic-wave velocity in marine sediments with gas hydrates: Effective medium modeling. *Geophys. Res. Lett.* **1999**, *26*, 2021–2024. [CrossRef]
16. Lee, M.W. Modified Biot-Gassmann Theory for Calculating Elastic Velocities for Unconsolidated and Consolidated Sediments. *Mar. Geophys. Res.* **2002**, *23*, 403–412. [CrossRef]
17. Shankar, U.; Riedel, M. Assessment of gas hydrate saturation in marine sediments from resistivity and compressional-wave velocity log measurements in the Mahanadi Basin, India. *Mar. Petrol. Geol.* **2014**, *58*, 265–277. [CrossRef]
18. Riedel, M.; Collett, T.S.; Kim, H.S.; Bahk, J.J.; Kim, J.H.; Ryu, B.J.; Kim, G.Y. Large-scale depositional characteristics of the Ulleung Basin and its impact on electrical resistivity and Archie-parameters for gas hydrate saturation estimates. *Mar. Petrol. Geol.* **2013**, *47*, 222–235. [CrossRef]
19. Lee, M.W.; Collett, T.S. Scale-dependent gas hydrate saturation estimates in sand reservoirs in the Ulleung Basin, East Sea of Korea. *Mar. Petrol. Geol.* **2013**, *47*, 195–203. [CrossRef]
20. Dugarov, G.A.; Duchkov, A.A.; Duchkov, A.D. Drobchik ANLaboratory validation of effective acoustic velocity models for samples bearing hydrates of different type. *J. Nat. Gas Sci. Eng.* **2019**, *63*, 8–46. [CrossRef]
21. Masroor, M.; Emami Niri, M.; Rajabi-Ghozloo, A.H.; Sharifinasab, M.H.; Sajjadi, M. Application of machine and deep learning techniques to estimate NMR-derived permeability from conventional well logs and artificial 2D feature maps. *J. Pet. Explor. Prod. Technol.* **2022**, *12*, 2937–2953. [CrossRef]
22. Dou, P.; Huang, C.; Han, W.; Hou, J.; Zhang, Y. Time Series Remote Sensing Image Classification Using Feature Relationship Learning. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 1–13. [CrossRef]
23. Tang, R.; Li, F.; Shen, F.; Gan, L.; Shi, Y. Fast Forecasting of WaterFilled Bodies Position Using Transient Electromagnetic Method Based on Deep Learning. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 1–13. [CrossRef]
24. Jeong, T.; Byun, J.; Choi, H.; Yoo, D. Estimation of gas hydrate saturation in the Ulleung basin using seismic attributes and a neural network. *J. Appl. Geophys.* **2014**, *106*, 37–49. [CrossRef]
25. Bhowmick, D.; Shankar, U.; Maiti, S. Revisiting Supervised Learning in the context of Predicting Gas Hydrate Saturation. In *Extended Abstracts, Proceedings of the 78th EAGE Conference & Exhibition, Vienna, Austria, 30 May–2 June 2016*; European Association of Geoscientists & Engineers: Utrecht, The Netherlands, 2016.
26. Gholanlo, H.H.; Amirpour, M.; Ahmadi, S. Estimation of Water Saturation by using Radial Based Function Artificial Neural Network in Carbonate Reservoir: A Case Study in Sarvak Formation. *Petroleum* **2016**, *2*, 166–170. [CrossRef]
27. Kim, S.; Lee, K.; Lee, M.; Ahn, T.; Lee, J.; Suk, H.; Ning, F. Saturation Modeling of Gas Hydrate Using Machine Learning with X-Ray CT Images. *Energies* **2020**, *13*, 5032. [CrossRef]
28. Singh, H.; Seol, Y.; Myshakin, E.M. Automated Well-Log Processing and Lithology Classification by Identifying Optimal Features Through Unsupervised and Supervised Machine-Learning Algorithms. *SPE J.* **2020**, *25*, 2778–2800. [CrossRef]
29. Singh, H.; Seol, Y.; Myshakin, E.M. Prediction of gas hydrate saturation using machine learning and optimal set of well-logs. *Comput. Geosci.* **2021**, *25*, 267–283. [CrossRef]
30. Xing, L.; Zhu, T.; Niu, J.; Liu, C.; Wang, B. Development and validation of an acoustic-electrical joint testing system for hydrate-bearing porous media. *Adv. Mech. Eng.* **2020**, *12*, 168781402090898. [CrossRef]
31. Hansen, L.K.; Salamon, P. Neural network ensembles. *IEEE Trans. Pattern Anal. Mach. Intell.* **1990**, *12*, 993–1001. [CrossRef]
32. Schapire, R.E. The Strength of Weak Learnability. *Mach. Learn.* **1990**, *5*, 197–227. [CrossRef]

33. Freund, Y.; Schapire, R.E. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *J. Comput. Syst. Sci.* **1997**, *55*, 119–139. [CrossRef]
34. Saghafi, H.; Arabloo, M. Modeling of CO_2 solubility in MEA, DEA, TEA, and MDEA aqueous solutions using AdaBoost-Decision Tree and Artificial Neural Network. *Int. J. Greenh. Gas Control.* **2017**, *58*, 256–265. [CrossRef]
35. Yarveicy, H.; Ghiasi, M.M.; Mohammadi, A.H. Performance evaluation of the machine learning approaches in modeling of CO_2 equilibrium absorption in Piperazine aqueous solution. *J. Mol. Liq.* **2017**, *255*, 375–383. [CrossRef]
36. Ji, X.; Yang, B.; Tang, Q. Acoustic Seabed Classification Based on Multibeam Echosounder Backscatter Data Using the PSO-BP-AdaBoost Algorithm: A Case Study From Jiaozhou Bay, China. *IEEE J. Ocean. Eng.* **2020**, *46*, 509–519. [CrossRef]
37. Nguyen, P.T.; Ha, D.H.; Jaafari, A.; Nguyen, H.D.; Van Phong, T.; Al-Ansari, N.; Prakash, I.; Le, H.V.; Pham, B.T. Groundwater Potential Mapping Combining Artificial Neural Network and Real AdaBoost Ensemble Technique: The DakNong Province Case-study, Vietnam. *Int. J. Environ. Res. Public Health* **2020**, *17*, 2473. [CrossRef]
38. Kemsley, E.K. Discriminant analysis of high-dimensional data: A comparison of principal components analysis and partial least squares data reduction methods. *Chemom. Intell. Lab. Syst.* **1996**, *33*, 47–61. [CrossRef]
39. Hecht-Nielsen, R. Neurocomputing: Picking the human brain. *Spectr. IEEE* **1988**, *25*, 36–41. [CrossRef]
40. Rumelhart, D.E.; Mcclelland, J.L.; PDP Research Group. Parallel Distributed Processing. *Encycl. Database Syst.* **1986**, *1*, 45–76.
41. Wang, R.B.; Xu, H.Y.; Li, B.; Feng, Y. Research on Method of Determining Hidden Layer Nodes in BP Neural Network. *Comput. Technol. Dev.* **2018**, *28*, 5, (In Chinese with English Abstract). [CrossRef]
42. Kumar, A.; Sakpal, T.; Roy, S.; Kumar, R. Methane hydrate formation in a test sediment of sand and clay at various levels of water saturation. *Can. J. Chem.* **2015**, *93*, 1742–1773. [CrossRef]
43. Revil, A. Effective conductivity and permittivity of unsaturated porous materials in the frequency range 1 mHz-1GHz. *Water Resour. Res.* **2013**, *49*, 306–327. [CrossRef]
44. Revil, A. On charge accumulation in heterogeneous porous rocks under the influence of an external electric field. *Soc. Explor. Geophys.* **2013**, *78*, D271–D291. [CrossRef]
45. Revil, A. Spectral induced polarization of shaly sands: Influence of the electrical double layer. *Water Resour. Res.* **2012**, *48*, W02517.1–W02517.23. [CrossRef]
46. Schmutz, M.; Revil, A.; Vaudelet, P.; Batzle, M.; Viñao, P.F.; Werkema, D.D. Influence of oil saturation upon spectral induced polarization of oil-bearing sands. *Geophys. J. Int.* **2010**, *183*, 211–224. [CrossRef]
47. Lee, M.W. Biot–Gassmann theory for velocities of gas hydrate-bearing sediments. *Geophysics* **2002**, *67*, 1711–1719. [CrossRef]
48. Cook, A.E.; Anderson, B.I.; Rasmus, J.; Sun, K.; Li, Q.; Collett, T.S.; Goldberg, D.S. Electrical anisotropy of gas hydrate-bearing sand reservoirs in the Gulf of Mexico. *Mar. Pet. Geol.* **2012**, *34*, 72–84. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.