

Article

A Carbonate Reservoir Prediction Method Based on Deep Learning and Multiparameter Joint Inversion

Xingda Tian ¹, Handong Huang ^{1,*}, Suo Cheng ², Chao Wang ³, Pengfei Li ² and Yaju Hao ⁴

¹ State Key Laboratory of Petroleum Resources and Prospecting, College of Geophysics, China University of Petroleum-Beijing, Beijing 102249, China; 2018310201@student.cup.edu.cn

² PetroChina Tarim Oil Field Company, Korla 841000, China; chengs-tlm@petrochina.com.cn (S.C.); lipf-tlm@petrochina.com.cn (P.L.)

³ China National Petroleum Corporation, Exploration and Development Institution of Tarim Oilfield, Korla 841000, China; wangchao-tlm@petrochina.com.cn

⁴ School of Geophysics and Measurement-Control Technology, East China University of Technology, Nanchang 330013, China; haoyj@ecut.edu.cn

* Correspondence: hhd@cup.edu.cn

Abstract: Deep-water carbonate reservoirs are currently the focus of global oil and gas production activities. The characterization of strongly heterogeneous carbonate reservoirs, especially the prediction of fluids in deep-water presalt carbonate reservoirs, exposes difficulties in reservoir inversion due to their complex structures and weak seismic signals. Therefore, a multiparameter joint inversion method is proposed to comprehensively utilize the information of different seismic angle gathers and the simultaneous inversion of multiple seismic parameters. Compared with the commonly used simultaneous constrained sparse-pulse inversion method, the multiparameter joint inversion method can characterize thinner layers that are consistent with data and can obtain higher-resolution presalt reservoir results. Based on the results of multiparameter joint inversion, in this paper, we further integrate the long short-term memory network algorithm to predict the porosity of presalt reef reservoirs. Compared with a fully connected neural network based on the backpropagation algorithm, the porosity results are in better agreement with the new log porosity curves, with the average porosity of the four wells increasing from 89.48% to 97.76%. The results show that the method, which is based on deep learning and multiparameter joint inversion, can more accurately identify porosity and has good application prospects in the prediction of carbonate reservoirs with complex structures.

Keywords: carbonate reservoirs; LSTM neural network; prestack inversion; parameter prediction



Citation: Tian, X.; Huang, H.; Cheng, S.; Wang, C.; Li, P.; Hao, Y. A Carbonate Reservoir Prediction Method Based on Deep Learning and Multiparameter Joint Inversion. *Energies* **2022**, *15*, 2506. <https://doi.org/10.3390/en15072506>

Academic Editor: Carlos Miguel Costa

Received: 7 March 2022

Accepted: 28 March 2022

Published: 29 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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

The presalt sequence has favourable hydrocarbon source conditions, reservoir conditions and hydrocarbon accumulation conditions, all of which have great potential for petroleum prospecting. However, oil and gas exploration in deep-water areas is high-cost and high-risk, which limits exploration processes. In addition to the general characteristics of carbonate rocks [1,2], presalt targets are affected by the thickness of the salt rocks, and their geophysical sensitivity properties and changing characteristics put forward higher requirements for the seismic prediction and description accuracy of high-quality reservoirs.

Deep-water presalt carbonate reservoirs have high exploration risks and high investment costs [3]. This type of carbonate reservoir has strong heterogeneity and severe lateral variation, which makes reservoir prediction and reservoir description a major technical challenge at present [4–7]. Therefore, based on presalt geology, well logging, and prestack and poststack seismic data, research on reservoir heterogeneity and reservoir-fluid-containing geophysical characteristics was conducted to optimize the sensitive petrophysical parameters and explore the distribution of oil and gas in presalt geological tar-

gets [8,9]. Such work is particularly important for the optimization of exploration targets and investment decisions.

Prestack seismic inversion retains the characteristics that the seismic reflection amplitude varies with the offset or incident angle and can provide more sensitive and effective data volume results. Under the constraint of log data, prestack seismic inversion research on seismic P-wave and S-wave velocity, density and other elastic parameters is carried out to obtain high-precision elastic parameters that can reflect the lateral changes in reservoirs. Crucially, the spatial distribution of complex oil and gas reservoirs must be studied to enable researchers to create detailed descriptions of complex oil and gas reservoirs [10,11]. Therefore, under the constraint that data must be logged, the inversion of prestack seismic elastic parameters is carried out by using abundant prestack seismic data to obtain high-precision P-wave and S-wave velocities that can reflect the lateral changes in reservoirs. On this basis, the P-wave and S-wave velocity ratio section, Poisson's ratio section, gas-zone indicator section and fluid-factor section can be obtained, all of which can provide more information for geological interpretation, in addition to realizing quantitative predictions of oil and gas reservoirs, narrowing the solution options, improving the drilling success rate and reducing exploration risk.

The prestack inversion used in industrial production generally employs partial stacking data to establish an initial model, and then through the production of synthetic seismic records, the obtained synthetic seismic records are continuously adjusted, and the least squares are calculated with respect to the original seismic data; if the error requirement is satisfied, then the inversion result is desirable [12,13]. When solving for porosity, the result is generally based on data-driven or rock physical modelling methods. The former approach uses polynomial fitting, local regression and other methods to solve the relationship between porosity and property parameters [14], whereas the latter requires the use of various rock physical models to obtain the porosity [15–17]. However, many assumptions are commonly used in rock physical modelling, most of which are not completely suitable for real underground situations. This uncertainty is especially strong when the reservoir is located in a complex structure. To overcome this uncertainty and better restore the distribution of porosity and other parameters, deep learning technology is integrated based on high-precision inversion. Because the seismic signal is a time-domain signal that has been collected and processed, the signal's amplitude changes with time. In fact, potential relationships and connections can exist between seismic signals. Unfortunately, a convolutional neural network is not good at dealing with such problems and is more suitable for image recognition, which is not applicable to the time-series situation. LSTM is a kind of neural network that is suitable for solving serialization data problems and thus is better suited for processing seismic data. Therefore, in this paper, we adopt an LSTM network to analyse and restore the real distribution of underground porosity as much as possible and provide a basis for subsequent exploration, development and well deployment.

Deep learning was proposed by Geoffrey Hinton in 2006 [18]. Its core is a multi-layer perceptron that contains multiple hidden layers. The main feature of deep learning is to allow the neural network to mine the characteristics of data, abstracting from low-dimensional features to higher-dimensional machine-space dimensions. In general, convolutional neural networks are suitable for processing images and other problems with a spatial dimension [19]. However, because seismic signals are typically analysed in the time dimension, a convolutional neural network is not suitable for dealing with such problems [20]. Here, we utilize long short-term memory (LSTM) based on a recurrent neural network (RNN) in deep learning to address such problems [21]. J. Hopfield proposed the initial Hopfield network of RNN [22]; the core element of this approach is similar to the increase in experience consciousness during the process of growth. The network has a certain experience memory function, which is realized by the energy function of output feedback to input. M. Jordan et al. constructed a recurrent network and neural network together in 1986 [23], and J. Elman et al. proposed a simple recurrent network

(SRN) in 1990 [24], which introduces a reverse input in the hidden layer to optimize the entire neural network.

Hochreiter and Schmidhuber proposed LSTM in 1997 [25]. The biggest difference between LSTM and RNN is that RNN uses the form of continuous multiplication of partial derivatives during backpropagation, whereas LSTM uses the form of continuous addition, which can better avoid the problem of a vanishing gradient [26]. The LSTM is mainly characterized by the fact that it is provided with an input gate, a forget gate and an output gate, where the input gate determines how much input of the current time step is kept for a memory unit, the forget gate determines how many states of the memory unit of a previous time step are held for the memory unit of the current time step, and the output gate determines how many memory cell states of that current time step are output. The three gates are controlled by different signals. The gate-control signal is generally transformed by the sigmoid activation function. The output value of the gate control ranges from zero to one. Zero is similar to the closing of the gate, one simulates the opening of the gate, and the transition from zero to one indicates that a portion of the data is allowed to pass. In recent years, LSTM and its hybrid versions have been used in many industries [27–31], such as in the field of medicine for the prediction of malaria abundance [32], acoustic modelling in the field of language recognition [33], trend classification in the financial field [34], power prediction in the field of mechanical engineering [35] and exploratory research on the performance and hyperparameters of deep learning [36–38]. However, few applied studies exist with a focus on joint prediction based on LSTM and multiparameter inversion of reservoirs. Based on the above situation, in this paper, we conduct corresponding research on complex deep-water presalt carbonate reef reservoirs. The main contributions are as follows:

- a. In this paper, based on the results of drilling and scanning electron microscopy analyses, different types of reservoir parameters are classified, and the results of porosity, lithology and geological outcrops are combined to divide the reservoirs into three types, thereby laying a foundation for determining the sensitive parameters of high-quality reservoirs.
- b. After reservoir classification, rock physical analysis is carried out on the main target reservoirs for the three different types of reservoirs, and it is determined that the Lamé coefficient and the v_p/v_s values are the sensitive attribute parameters, whereas the corresponding parameter constraint values are provided for the subsequent multiparameter reservoir prediction.
- c. The multiparameter joint inversion algorithm can obtain finer results than the common simultaneous constrained sparse-pulse inversion, which is more suitable for describing the horizon information of presalt reservoirs, especially for thin layers, and can reflect the real structure of complex reservoirs.
- d. The sensitive attribute parameter data volume obtained based by the multiparameter joint inversion algorithm can be used as a volume constraint for porosity inversion. Combining the porosity information of the well log as the trace constraint beside the well, the deep learning algorithm is used to construct a mapping relationship that is suitable for complex carbonate rock work areas. The whole process avoids complicated processes, such as building a rock physical model. The final results, verified by new wells, show that the method proposed in this paper can obtain results with a high degree of agreement. The new method has been applied in actual oil fields, and new exploration wells have detected oil, allowing the approach to be further improved and applied in the future.

The structure of this article is as follows. Section 2 explains the principle of multiparameter joint inversion and LSTM neural networks. In Section 3, for the presalt carbonate reef reservoir, using the proposed method to invert and obtain a volume of highly reliable porosity data provides a reference for later exploration and development. Section 4 discusses the advantages and disadvantages of the proposed method, the hyperparam-

eter selection of deep learning and directions for future improvement. Section 5 gives the conclusions.

2. Theory and Methods

2.1. Multiparameter Joint Inversion Based on Multiangle Gathers Information

The traditional prestack inversion method usually uses logged data interpolation and extrapolation to establish the initial model of the inversion. However, this method can cause large deviations in the inversion results when only a few wells are present or complex structures exist. Prestack multiparameter joint inversion is used to conduct research to improve the resolution of prestack inversion and to effectively control multiple parameters and the multiparameter joint inversion method at the same time, which can ultimately improve the efficiency and accuracy of the inversion. The seismic model response can be expressed as:

$$S(V_p, V_s, \rho) = \mathbf{W} * R \quad (1)$$

where V_p is the P-wave velocity, V_s is the S-wave velocity, ρ is the density and D is the actual seismic angle gather records. R is the reflection coefficient calculated using the Zoeppritz formula, and $\mathbf{W}$ is the seismic wavelet vector.

The sequence of reflection coefficients for the P-wave is R_{pp} , where R_{pp} is a multivariate function of the incident angle θ , and the model parameters, $m_1, m_2, \dots, m_k$. The actual prestack angle gather, S , in the time domain can be expressed according to Equation (2).

$$S(\theta, m_1, m_2, \dots, m_k) = \mathbf{w} * R_{pp}(\theta, m_1, m_2, \dots, m_k) \quad (2)$$

The actual prestack angle gather vector expression is $\mathbf{s} = (s_1, s_2, \dots, s_{m+L-2})^T$, where $\mathbf{w} = (w_1, w_2, \dots, w_m)^T$ represents a seismic wavelet vector of length m . L represents the number of sampling points for seismic records. After Fourier transformation and Taylor expansion of S_r at S , we obtain Equation (3) and transform this equation as a vector product, as given in Equation (4).

$$S_r(\theta, m_{1r}, \dots, m_{kr}) = S(\theta, m_1, \dots, m_k) + \mathbf{w} \frac{\partial R_{pp}(\theta, m_1, \dots, m_k)}{\partial m_1} \Delta m_1 + \dots + \mathbf{w} \frac{\partial R_{pp}(\theta, m_1, \dots, m_k)}{\partial m_k} \Delta m_k \quad (3)$$

$$S_r(\theta, m_{1r}, \dots, m_{kr}) - S(\theta, m_1, \dots, m_k) = \left[\mathbf{w} \frac{\partial R_{pp}(\theta, m_1, \dots, m_k)}{\partial m_1}, \dots, \mathbf{w} \frac{\partial R_{pp}(\theta, m_1, \dots, m_k)}{\partial m_k} \right] \begin{bmatrix} \Delta m_1 \\ \vdots \\ \Delta m_k \end{bmatrix} \quad (4)$$

Equation (4) is the difference between the synthetic prestack-angle gather record and the actual prestack-angle gather at a single incident angle, which makes the model parameters V_p , V_s and ρ highly multisolved. Therefore, to obtain the three main model parameters, V_p , V_s and ρ , a system of equations composed of equations under different incident angles should be established to reduce the multisolution problem caused by prestack inversion. Taking the angle gathers with the three angles θ_1 , θ_2 and θ_3 as an example, $\Delta s = s_r(\theta) - s(\theta)$, and thus Equation (5), can be obtained, which can also be written as Equation (6).

$$\begin{bmatrix} \Delta s_1 \\ \Delta s_2 \\ \Delta s_3 \end{bmatrix} = \begin{bmatrix} \mathbf{w} \frac{\partial R_{pp}(\theta_1)}{\partial m_1} & \mathbf{w} \frac{\partial R_{pp}(\theta_1)}{\partial m_2} & \mathbf{w} \frac{\partial R_{pp}(\theta_1)}{\partial m_3} \\ \mathbf{w} \frac{\partial R_{pp}(\theta_2)}{\partial m_1} & \mathbf{w} \frac{\partial R_{pp}(\theta_2)}{\partial m_2} & \mathbf{w} \frac{\partial R_{pp}(\theta_2)}{\partial m_3} \\ \mathbf{w} \frac{\partial R_{pp}(\theta_3)}{\partial m_1} & \mathbf{w} \frac{\partial R_{pp}(\theta_3)}{\partial m_2} & \mathbf{w} \frac{\partial R_{pp}(\theta_3)}{\partial m_3} \end{bmatrix} \begin{bmatrix} \Delta m_1 \\ \Delta m_2 \\ \Delta m_3 \end{bmatrix} \quad (5)$$

$$\Delta \mathbf{s} = \mathbf{L} \Delta \mathbf{m} \quad (6)$$

At the beginning of the inversion, an initial model, $\mathbf{m}_0$, must be given to participate in the subsequent iterative calculation. By continuously revising the model parameters, the

approximate real underground medium or rock model parameters can finally be obtained through iterative calculation. The specific iterative calculation process can be expressed as:

$$\Delta \mathbf{s} + \mathbf{L} \mathbf{m}_0 = \mathbf{L}(\Delta \mathbf{m} + \mathbf{m}_0) \quad (7)$$

According to Equation (7), the observed seismic data are composed of the prestack seismic data based on the initial model, the first-order partial derivative matrix of the model parameters and the model parameters that are continuously revised during the inversion process. According to the various forward analytical formulas used in the actual prestack inversion, the model parameters, $\mathbf{m}$, will also vary, and then the first-order partial derivative matrix, $\mathbf{L}$, is further obtained according to the selection of the model parameters.

Regularization parameters for prestack inversion can be derived using Bayesian theory. Statistically, the likelihood function can be used to describe the degree of similarity between two variables. The likelihood function is used in prestack inversion to describe the similarity between actual prestack data and synthetic records. The difference between the two is the noise, $\mathbf{n}$, as in Equation (8); assuming that the noise is a Gaussian distribution with mean 0, the distribution of the posterior probability can be obtained as Equation (9).

$$\mathbf{n} = \mathbf{s} - \mathbf{L} \mathbf{m} \quad (8)$$

$$P(\mathbf{m} | \mathbf{s}) \propto \exp \left[-\frac{1}{2} (\sigma_n^2)^{-1} (\mathbf{L} \mathbf{m} - \mathbf{s})^T (\mathbf{L} \mathbf{m} - \mathbf{s}) + (\sigma_m^2)^{-1} \mathbf{m}^T \mathbf{m} \right] \quad (9)$$

where σ_n^2 is the mean square error of the noise, and σ_m^2 is the covariance of the model parameter variables [39]. The expression of the regularization parameter, λ , under the framework of Bayesian theory can be obtained as:

$$\lambda = \sigma_n^2 (\sigma_m^2)^{-1} \quad (10)$$

Based on the convolution model and the Aki–Richards approximate formula, the objective function of prestack inversion can be established by the approximation formula:

$$f(V_p, V_s, \rho) = \|S(V_p, V_s, \rho) - D\| = \sum_{i=0}^{n-1} \left(S(V_p, V_s, \rho)_i^\Delta - D_i \right)^2 \quad (11)$$

Taylor expansion of $S(V_p, V_s, \rho)$ is carried out, and the higher-order terms containing more than the second order are omitted. Equation (12) can be obtained and substituted into Equation (11), and then the derivatives of ΔV_p , ΔV_s and $\Delta \rho$ can be taken. As a result, we obtain Equation (13).

$$S(V_p, V_s, \rho)_i^\Delta = S(V_p, V_s, \rho)_i + \sum_{k=0}^{n-1} \Delta V_{p_k} \frac{\partial S_i}{\partial V_{p_k}} + \sum_{k=0}^{n-1} \Delta V_{s_k} \frac{\partial S_i}{\partial V_{s_k}} + \sum_{k=0}^{n-1} \Delta \rho_k \frac{\partial S_i}{\partial \rho_k} \quad (12)$$

$$-dS_i + \sum_{k=0}^{n-1} \Delta V_{p_k} \frac{\partial S_i}{\partial V_{p_k}} + \sum_{k=0}^{n-1} \Delta V_{s_k} \frac{\partial S_i}{\partial V_{s_k}} + \sum_{k=0}^{n-1} \Delta \rho_k \frac{\partial S_i}{\partial \rho_k} = 0 \quad (13)$$

From the above equations, it can be seen that the perturbations ΔV_p , ΔV_s and $\Delta \rho$ of the three elastic parameters are solved by Equation (13). In this way, the underdetermined equation must have a strong multiplicity. To reduce the difficulty of inversion, a joint inversion method of multiangle seismic data is adopted to invert these three parameters. The final objective function is shown in Equation (14). According to the above equations, $\Delta V_{p_k} \sum_{k=0}^{n-1} \frac{\partial S_i}{\partial V_{p_k}}$, $\Delta V_{s_k} \sum_{k=0}^{n-1} \frac{\partial S_i}{\partial V_{s_k}}$ and $\Delta \rho_k \sum_{k=0}^{n-1} \frac{\partial S_i}{\partial \rho_k}$ can be calculated, and the perturbations of the three elastic parameters, ΔV_p , ΔV_s and $\Delta \rho$, can also be obtained. To obtain the final result, Bayesian theory is utilized in the prestack inversion. The solution is achieved by solving

the adaptive damping factor, λ , and changing the original matrix appropriately. Combining Equations (5) and (14), the objective can be obtained.

$$\begin{bmatrix} \Delta v_p \sum_{k=0}^{n-1} \frac{\partial S_i}{\partial V_{p_k}} \\ \Delta v_s \sum_{k=0}^{n-1} \frac{\partial S_i}{\partial V_{s_k}} \\ \Delta \rho \sum_{k=0}^{n-1} \frac{\partial S_i}{\partial \rho_k} \end{bmatrix} = \begin{bmatrix} \frac{\sec^2(\theta_1)}{A} & \frac{\sin^2(\theta_1)}{B} & \frac{\cos^2(\theta_1)}{C} \\ \frac{\sec^2(\theta_2)}{A} & \frac{\sin^2(\theta_2)}{B} & \frac{\cos^2(\theta_2)}{C} \\ \frac{\sec^2(\theta_3)}{A} & \frac{\sin^2(\theta_3)}{B} & \frac{\cos^2(\theta_3)}{C} \end{bmatrix}^{-1} \begin{bmatrix} dS_1 \\ dS_2 \\ dS_3 \end{bmatrix} \quad (14)$$

$$A = \sec^2(\theta_1) + \sec^2(\theta_2) + \sec^2(\theta_3) \quad (15)$$

$$B = \sin^2(\theta_1) + \sin^2(\theta_2) + \sin^2(\theta_3) \quad (16)$$

$$C = \cos^2(\theta_1) + \cos^2(\theta_2) + \cos^2(\theta_3) \quad (17)$$

2.2. Porosity Prediction Method Based on LSTM and Multiparameter Joint Inversion

A recurrent neural network is a neural network with memory. Figure 1 is a structure diagram of a simple recurrent network [40]. There are five neurons in the input layer, three nodes in the hidden layer and three neurons in the output layer. The structure of an RNN is similar to that of a fully connected neural network [41,42]. The main difference is the memory element in the middle. The existence of the memory element makes the output, y of n , not only determined by the previous x but also related to the state of n in the previous time step. The result of the memory is circulated into the current time step for common operation, and the weight in each connection is determined by the neurons, ω , connected before and after.

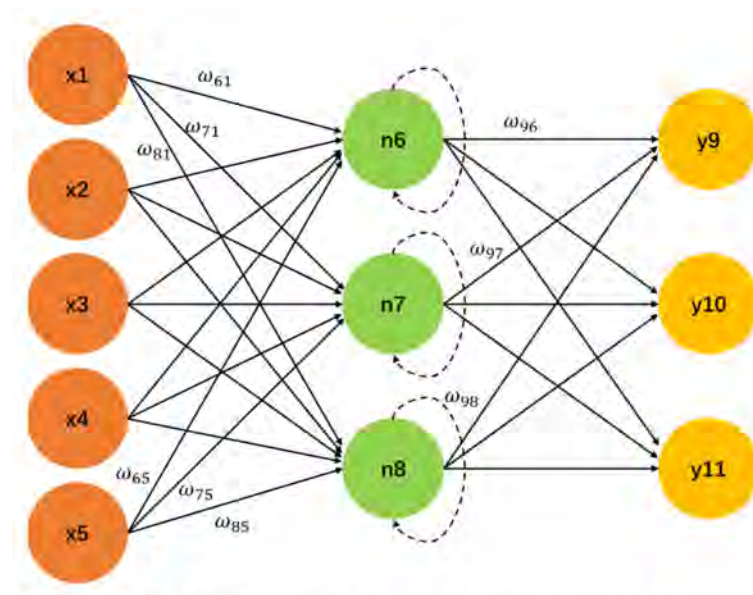


Figure 1. Structure diagram of the simple recurrent network.

Figure 2 shows the network structure of LSTM. The structure is composed of multiple recurrent units stacked together. The framework of the main part is based on RNN, but the significant difference is that there are some gated technologies in the internal units, that is, forget gates, input gates and output gates. The gating mechanism mainly controls the path of information propagation and improves the ability of the hidden layer to learn deep information. Similar to RNN, the result, h_{t-1} , of the previous time step will be combined with the input, x_t , of the current time step. After the operation, h_t will be obtained, and then h_t will be stored and calculated together with x_{t+1} of the next time step. The method will

continue running in this way until the end. ε in the figure represents the sigmoid function, which is used in addition to the tanh function of the hyperbolic tangent in the unit.

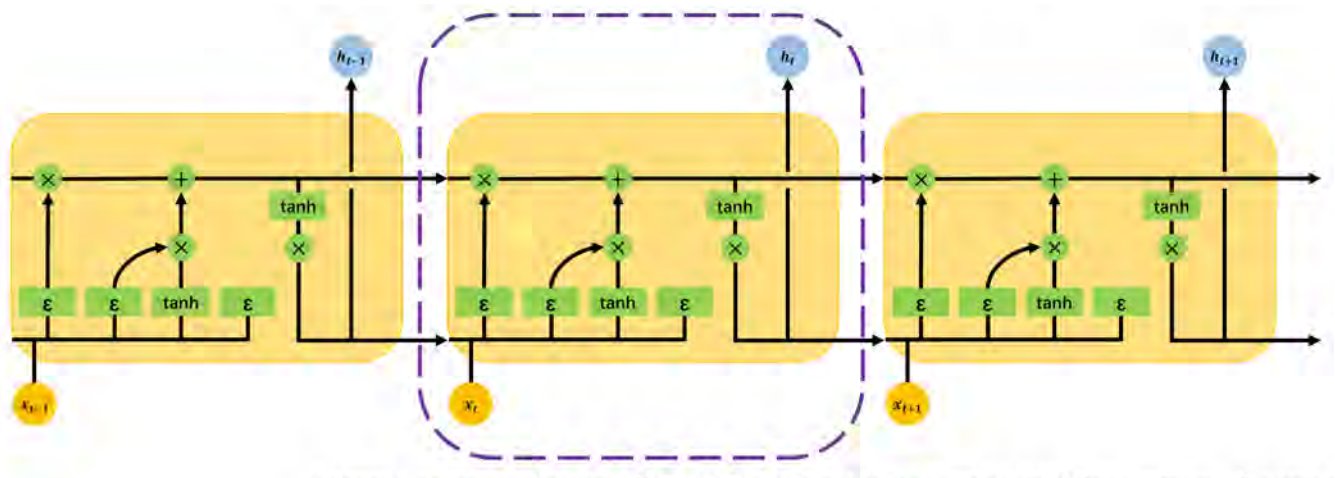


Figure 2. Structure diagram of LSTM. The part inside the purple dotted line represents the core part of the current network operation. The left side outside the purple dotted line is the operation of the previous time period, and the right side is the operation of the next time period.

The forget gate function is

$$f_t = \varepsilon(W_f[h_{t-1}, x_t] + b_f) \quad (18)$$

where W_f is the weight matrix of the forgetting gate, and b_f is the bias term of the forgetting gate. ε represents the activation function, and ε generally takes $\varepsilon(x) = (1 + e^{-x})^{-1}$ as the activation function.

The input gate is used to calculate the data, i_t , saved to the memory unit at the current time step:

$$i_t = \varepsilon(W_i[h_{t-1}, x_t] + b_i) \quad (19)$$

where W_i represents the weight matrix of the input gate, h_{t-1} is the output of the network at the previous time step, x_t is the input of the current time step and b_i is the bias term of the input gate.

The new information, $\tilde{c}_t$, saved to the memory unit generated by the transformation of the currently input information is:

$$\tilde{c}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (20)$$

The activation function, $\tanh$, in the above formula can limit the result to the $[-1, 1]$ interval, and W_c and b_c are the weight matrix and the bias term, respectively.

The calculation formula of the memory value, c_t , of the current time step is as follows:

$$c_t = f_t \circ c_{t-1} + i_t \circ \tilde{c}_t \quad (21)$$

where $\circ$ represents the Hadamard product. In this equation, the item before the addition sign represents the information filtered through the forget gate, and the latter item represents the new information added.

The information formula output by the output gate is as follows, where W_o represents the weight matrix of the output gate, and b_o represents the bias term of the output gate.

$$o_t = \varepsilon(W_o[h_{t-1}, x_t] + b_o) \quad (22)$$

The output result, h_t , of the current time step is:

$$h_t = o_t \circ \tanh(c_t) \quad (23)$$

According to the mechanical relationship of elastic parameters in rock physics, the Lamé coefficient conversion formula is:

$$\lambda_L = \frac{E\nu}{(1+\nu)(1-2\nu)} \quad (24)$$

where E is Young's modulus, and ν is Poisson's ratio. Young's modulus and Poisson's ratio are:

$$E = \rho V_s^2 \frac{4-3\gamma^2}{1-\gamma^2} \quad (25)$$

$$\nu = \frac{2-\gamma^2}{2(1-\gamma^2)} \quad (26)$$

where $\gamma = V_p/V_s$; then, we can obtain λ_L according to Equations (24)–(26),

$$\lambda_L = \rho(V_p^2 - 2V_s^2) \quad (27)$$

Mapping between reservoir elastic parameters and reservoir attributes can be established using rock physics based on the LSTM network. The data with high resolution and multiple parameters obtained from inversion are fixed with multiangle prestack-angle gathers, and the prior log information is used to establish constraints. The long short-term memory neural network is used to iteratively correct the model until a reservoir property parameter volume with a high coincidence rate is obtained.

3. Results

The H Basin has a structural pattern of uplift and depression, as shown in Figure 3. From the shore to the sea, the basin can be divided into three tectonic units: a coastal uplift zone, a sunken zone and a platform. It is a typical passive continental margin salt-bearing basin, which is mainly divided into two oil-gas systems: the presalt system and the postsalt system. The G group of the target interval belongs to the Lower Cretaceous depression sequence. The depth of the target layer is between 5000 and 7500 m, and there is a thick salt rock layer above the target layer. Based on the above geological background, it is confirmed that the working area belongs to a deep-water presalt carbonate reservoir. The main frequency of the seismic data is approximately 14 Hz, and the frequency band is approximately 5–25 Hz. The main frequency belongs to the seismic signal of the low dominant frequency and low frequency bands. The prestack data volumes are stacked into seismic data volumes with three partial-angle gathers in the range of 3–12°, 12–21° and 21–30°. The lithology is complex and diverse. According to the scanning electron microscopy (SEM) analysis of the main reservoirs, as shown in Figure 4, the lithology of the G1 layer mainly shows high porosity and medium permeability, with distributed stromatolites and granular rocks. The G2 layer mainly shows medium-high porosity and low-medium permeability and is mainly distributed in spherical microbial limestone. The G3 layer is mainly characterized by medium porosity, low permeability and extra-low permeability and is mainly distributed in striated limestone. The focus of this study is on the G1 and G2 layers.

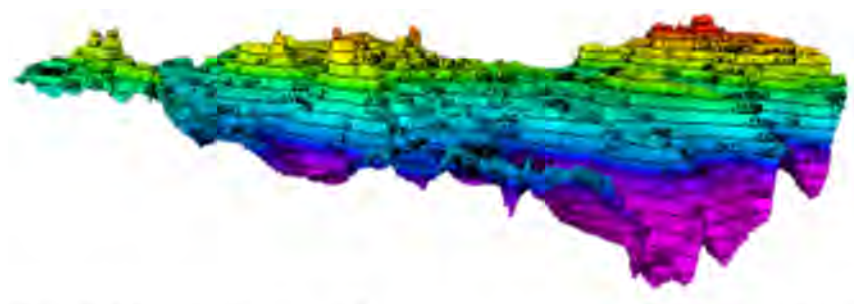


Figure 3. Schematic diagram of the structure of a layer in the H Basin with a structural pattern of uplift and depression.



Figure 4. SEM results of the main reservoirs. The rows, from top to bottom, represent the G1 layer, G2 layer and G3 layer, respectively.

Reef-shoal facies and intershoal depressions are developed in the presalt carbonate rocks in the study area, the reservoir heterogeneity is strong and the lateral variation is large. According to the observation of cores and outcrops, analysis of the geological conditions and prior information, such as sediments, physical parameters and lithology, the reservoirs are divided into three types: type I reservoirs are reefs, type II reservoirs are beaches and type III reservoirs are platform facies, as shown in Table 1. The reefs mainly develop stromatolites with good reservoir properties and porosities greater than 14%, the beach facies mainly develop grainstone and pellet limestone with porosities of 10% to 14% and medium reservoir properties, and the platform facies mainly develop laminated bindstone with relatively poor reservoir properties and porosities less than 10%.

Table 1. Reservoir classification table based on the analysis of geological conditions, such as cores and outcrops, as well as prior information, such as sediments, physical parameters and lithology.

Type	Porosity	Lithology	Categories
Type I	$14\% < \text{Por}$	stromatolite	reef
Type II	$10\% < \text{Por} \leq 14\%$	grainstone, pellet limestone	beach
Type III	$\text{Por} \leq 10\%$	laminated bindstone	platform

By optimizing the physical parameters of the rock, finally, the G1 and G2 layers can be sufficiently distinguished into type I, type II and type III reservoirs through the intersection of the Lamé coefficient and porosity and the intersection of V_p/V_s and porosity, which is less sensitive than the Lamé coefficient, as shown in Figure 5. The rock-physical statistics show that the intersection of porosity and the Lamé coefficient can better distinguish the three types of reservoirs. Type I reservoirs show a low Lamé coefficient and low V_p/V_s , type II reservoirs show a medium Lamé coefficient and medium V_p/V_s , and type III reservoirs show a high Lamé coefficient and high V_p/V_s . To obtain the porosity data volume, the relationship between porosity and Lamé coefficient must be established. The parameters obtained from the petrophysical analysis can provide relevant parameter constraint thresholds for subsequent multiparameter interactions and reservoir predictions.

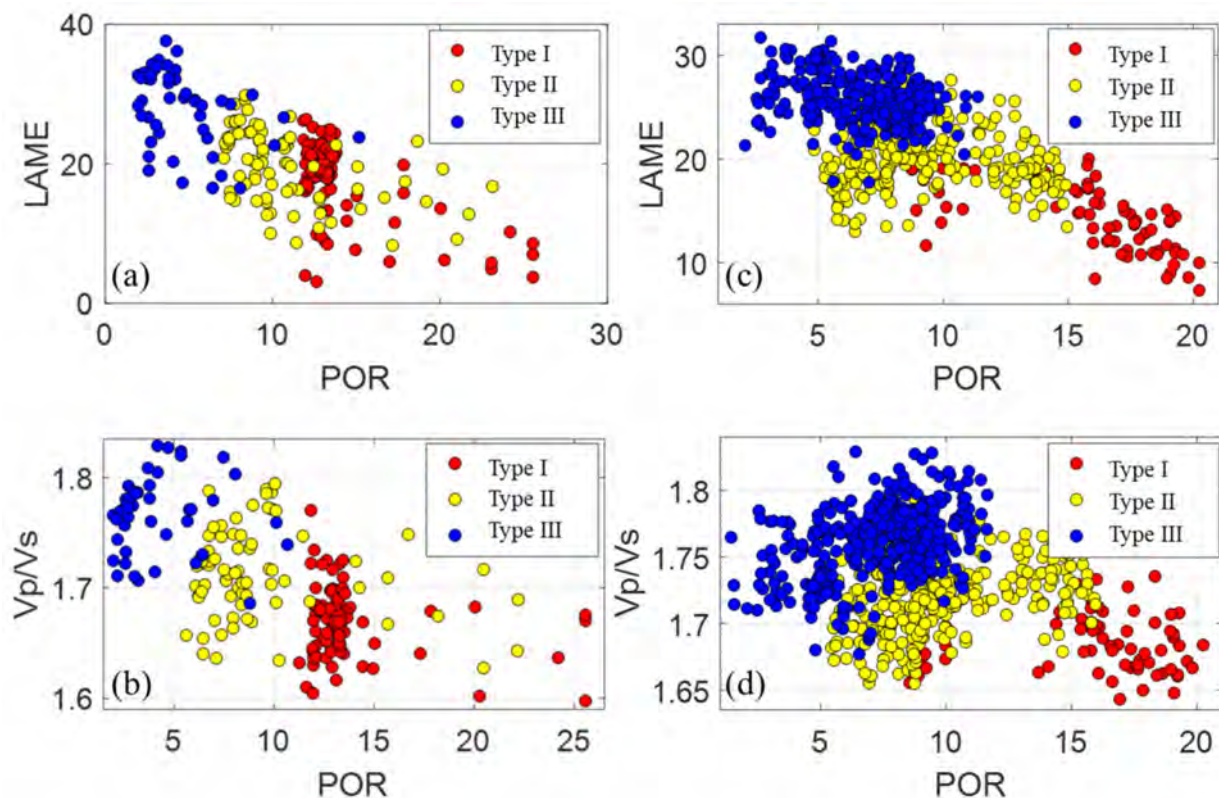


Figure 5. (a) The intersection of the Lamé coefficient and porosity of the G1 layer, (b) the intersection of V_p/V_s and porosity of the G1 layer, (c) the intersection of the Lamé coefficient and porosity of the G2 layer and (d) the intersection of V_p/V_s and porosity of the G2 layer.

Combining the P-wave, S-wave and density log data of the wells in the study area allows for the construction of a low-frequency model. The initial low-frequency model of the prestack three-parameter inversion established by the seismic data is shown in Figure 6. The initial model established by this method varies in different target intervals. The corresponding low-frequency initial models are different, with highly apparent lateral variation features.

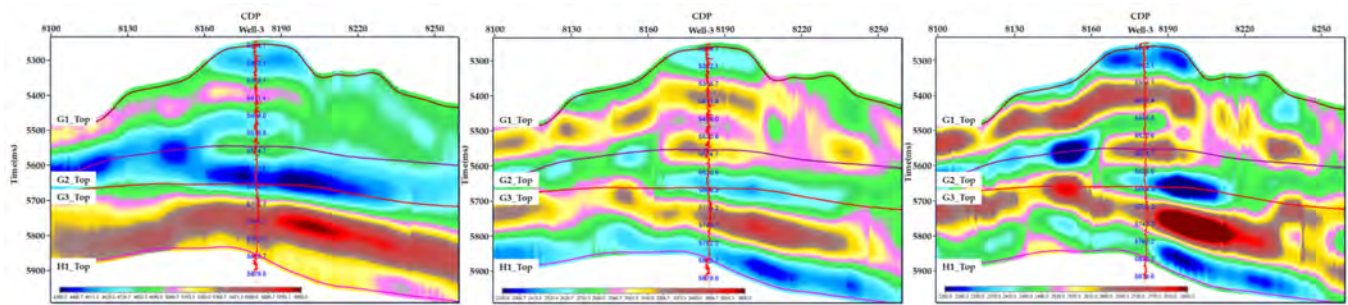


Figure 6. Low-frequency models for prestack multiparameter joint inversion. From left to right are the initial inversion models based on P-wave velocity, S-wave velocity and density, respectively.

The Lamé coefficient inversion results are obtained by using the prestack multiparameter inversion method in the study area. Figure 7 is the profile of the Lamé coefficient results obtained from the actual data inversion in the study area. Well 5 is a new well without shear wave information that does not participate in the prestack inversion work; therefore, Well 5 can be used here for testing of the inversion effect. From the variation of the gamma GR curve and the real resistivity RT curve of the Well 5 section, it can be seen that the reservoir corresponds well to the log curve and the inversion result profile. The Lamé coefficient inversion profile is in good agreement with the actual reservoir, which can better characterize the changes in oil-bearing reservoirs.

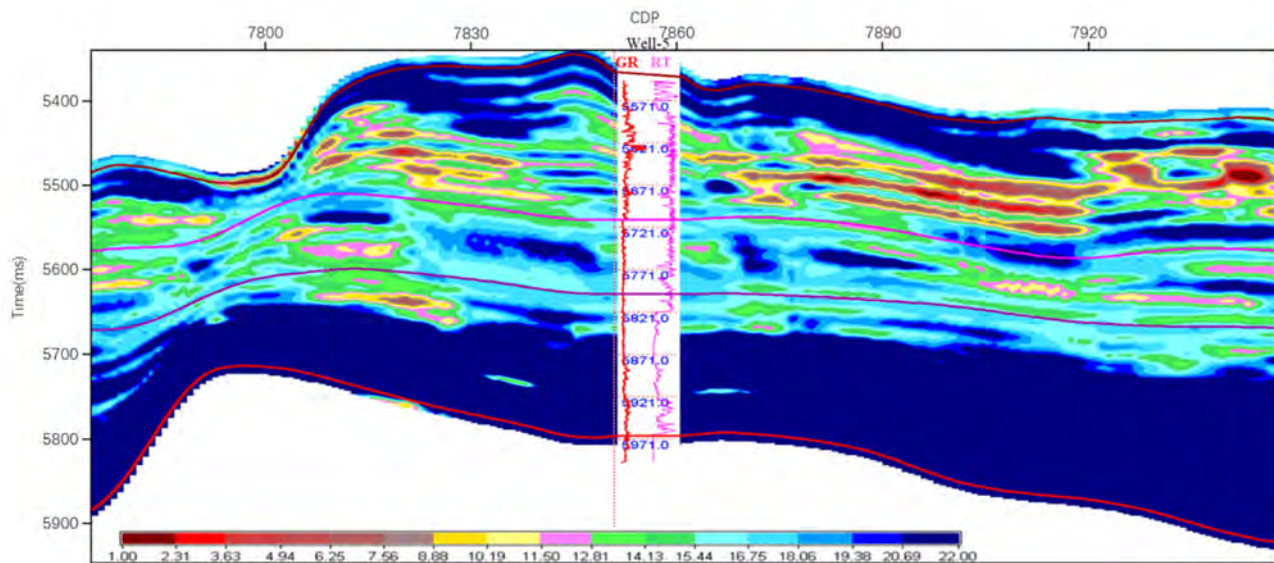


Figure 7. Profile of the Lamé coefficient inversion result. Note that Well 5 is the well that was not involved in the inversion.

The common inversion method for software used in industry is the simultaneous sparse-pulse inversion method. The two inversion results are compared in Figure 8. The upper figure is the Lamé coefficient inversion profile obtained by constrained sparse-pulse inversion, and the lower figure is the Lamé coefficient inversion profile obtained by multiparameter joint inversion. From the Figure 8, the Lamé coefficients can better characterize the change in oil-bearing reservoirs. From left to right, the reservoir varies from thick to thin before thickening again. Although the constrained coefficient pulse inversion can reflect the changing trend of reservoir thickness, it is obvious that the inversion results in the figure below can obtain thinner reservoirs and high-resolution results, which proves that the new inversion method can obtain more accurate inversion results.

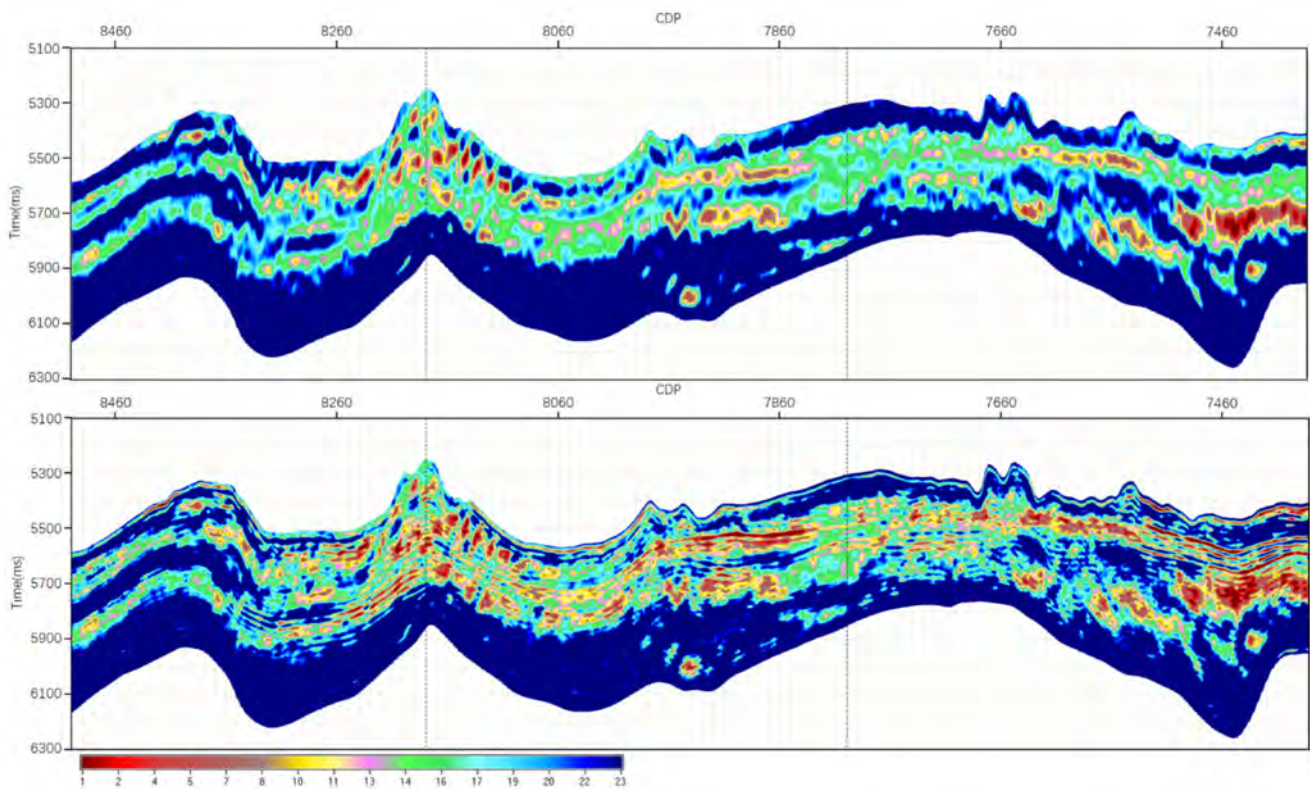


Figure 8. Comparison results of the constrained sparse spike inversion and the multiparameter joint inversion of the Lamé coefficient profile. The upper figure is the constrained sparse spike inversion result, and the lower figure is the multiparameter joint inversion result.

The Lamé coefficient inversion data volume and seismic and log information obtained by the method in this paper are used as joint prior constraints. Because the magnitudes of the data are different, the input data must be normalized, and the output result is the porosity data volume. The model convergence analysis in this paper adopts the mean square error method the adaptive moment estimation (Adam) algorithm for optimization. The initial learning rate is 0.005. To prevent the gradients from exploding, the gradient threshold is set to 1. During learning, an optimization operator is used to join the neural network so that after a certain training time, the learning rate will be multiplied by a coefficient between 0 and 1; it is expected that more accurate results will be obtained in the later stage of learning. After repeated testing and optimization, the results of the LSTM model test are shown in Figure 9. The verified actual data and the model-calculated data exhibit changes with a consistent trend. The training situation of the model is shown in Figure 10. It can be seen that the model is relatively good after approximately 40 iterations, and the loss is almost stable by approximately 110 iterations. Table 2 shows the adopted architecture, hyperparameters and errors of the applied LSTM. At this time, the RMSE value is 0.4603, the MSE value is 0.2119 and the R^2 value is 0.9830. The final inversion LSTM network structure consists of one input layer, three hidden layers and an output layer. The number of hidden units is 28, and gradient descent is used for convergence.

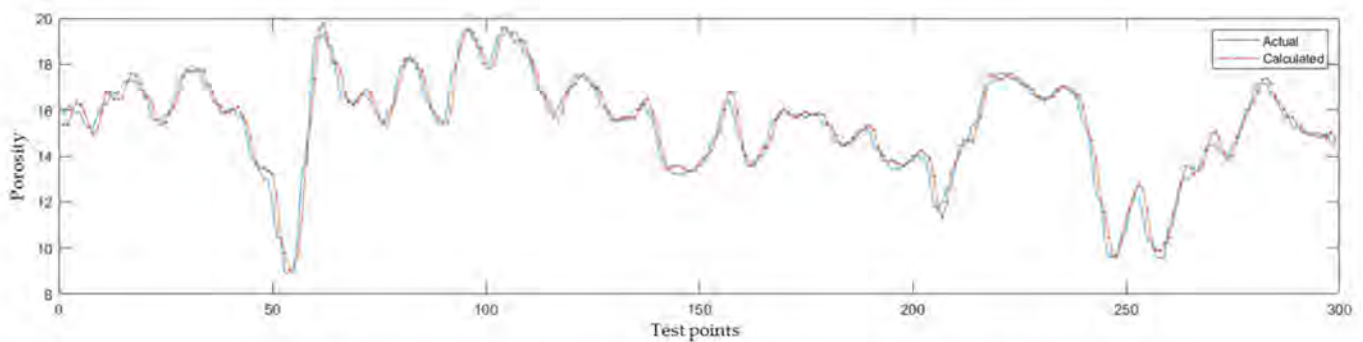


Figure 9. Comparison of the model-generated porosity and validated porosity.

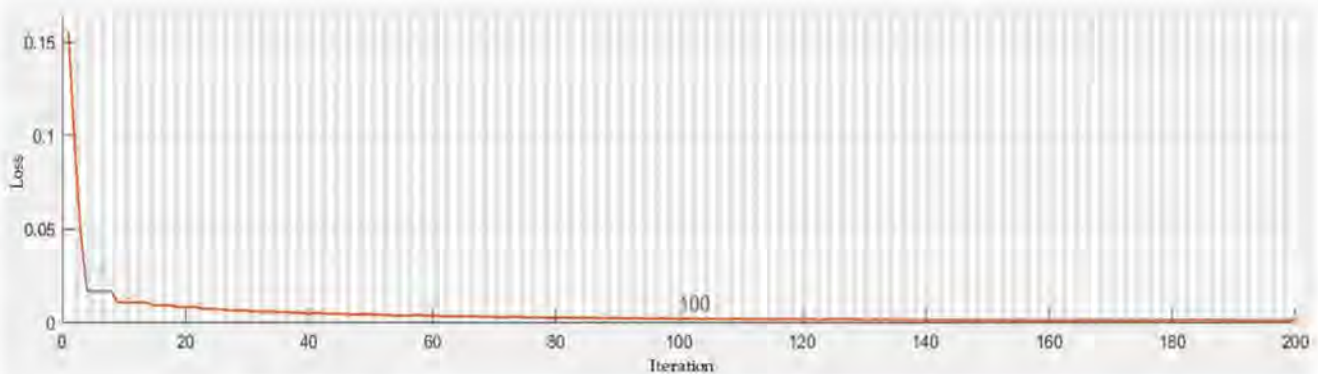


Figure 10. Convergence of the model as the number of iterations is increased.

Table 2. Architecture, hyperparameters and errors of the applied LSTM.

Architecture	Number	Hyperparameters	Strategy	Errors	Value
Input layer	1	Initial Learn Rate	0.001	RMSE	0.4603
Hidden layer	3	Drop Factor	0.2	R ²	0.9830
LSTM unit	28	State Function	tanh	MSE	0.2119
Output layer	1	Gate Function	sigmoid	-	-

To obtain high-precision porosity inversion results, in this article, we use a deep neural network (DNN) based on the backpropagation (BP) algorithm and LSTM network to test the data volume. The BP neural network is a multilayer feed-forward neural network trained by the error backpropagation algorithm. The topological structure of the BP neural network model can be divided into three layers: the input layer, hidden layer and output layer. The model learns according to given input samples and output samples and according to certain training criteria, such as the minimum mean square error. The error between the actual output value and the expected output value of the network is calculated, and error backpropagation is continuously carried out to adjust the weight of each layer of the network to minimize the error and complete the purpose of learning. DNNs based on the BP algorithm are widely used in predictive deep learning frameworks. This model has a very strong nonlinear fitting ability, and it can fit almost any function, so we chose it for comparison in this manuscript.

Based on the data for the four new wells, different methods are used for comparison. Figure 11 is a comparison of the porosity of the four new wells not included in the deep learning activity and with the porosity predicted by the DNN. The blue curve is the porosity calculated by the inversion based on the DNN with the BP algorithm; the light red curve is the actual porosity data, where each segment of the red dotted line represents the

information of each well of the four new verification wells; and the purple curve represents the porosity error. Figure 12 is a comparison of the porosities of the four new wells not involved in deep learning with the porosities predicted by the LSTM network, in which the blue curve is the porosity calculated by inversion based on the LSTM algorithm, the light red curve is the actual porosity data and the purple curve represents the porosity error. Because the third new well with porosity information is a well outside the work area, that is, the well with the abscissa ranging from 423 to 515 in Figures 11 and 12, it is mainly used with the three other new wells to reflect the porosity of the two methods. The inversion effect of the porosity is not calculated in the final calculation of the statistical rate in the work area. It can be seen from the comparison results of Figures 11 and 12 that the porosity errors of the four wells are reduced. Although DNN can obtain good results and the average porosity coincidence rate obtained by calculation is approximately 89.48%, the porosity inversion results based on the LSTM algorithm have an average porosity coincidence rate of 97.76%, which is better than that of the deep neural network and is more suitable for deep learning inversion in this work area.

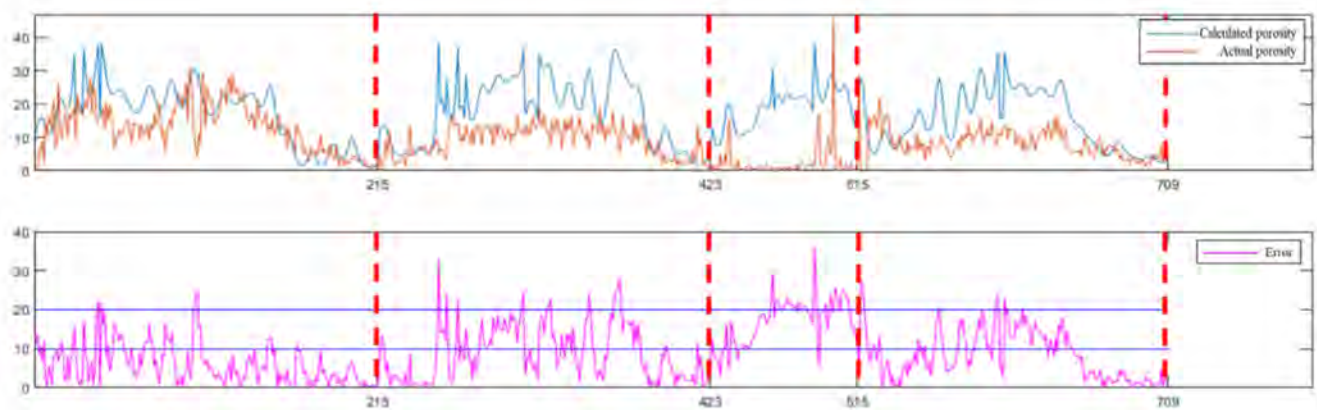


Figure 11. Comparison of the actual porosity of the four new wells with the porosity calculated by the deep neural networks based on the BP algorithm.

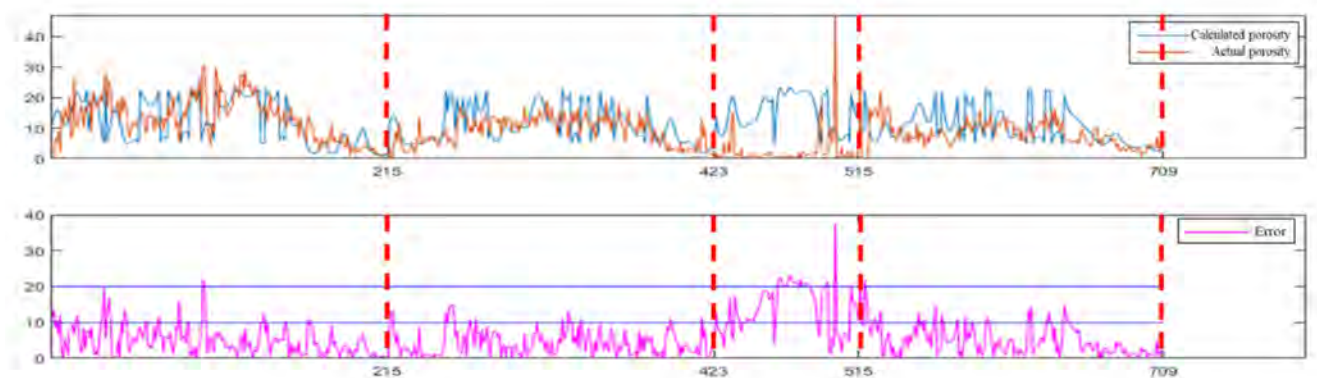


Figure 12. Comparison of actual porosity from the four new wells and porosity calculated by LSTM inversion.

The obtained Lamé coefficient inversion data volume, the original logging porosity, the seismic trace gather and other information are used as prior information for deep learning inversion. Based on the Lamé coefficient result profile of the multiparameter joint inversion in the upper image of Figure 13 and the porosity result of the multiparameter joint inversion with LSTM in the lower image, Well 6 has logging porosity data but lacks a constraint inversion process, which is used to verify the accuracy of the porosity inversion data volume. By extracting the well-tie profile, it can be seen that the porosity data volume based on deep learning LSTM are horizontally distributed in layers, and the change in the

inversion porosity data volume in the vertical direction at Well-6 is consistent with the changing trend of the porosity logging curve. The porosity inversion profile and the highly sensitive Lamé inversion profile have the same change trend, which can better depict the underground geological conditions. From the results of deep learning, the method based on multiparameter joint inversion and LSTM can obtain reasonable prediction results from reservoir parameters.

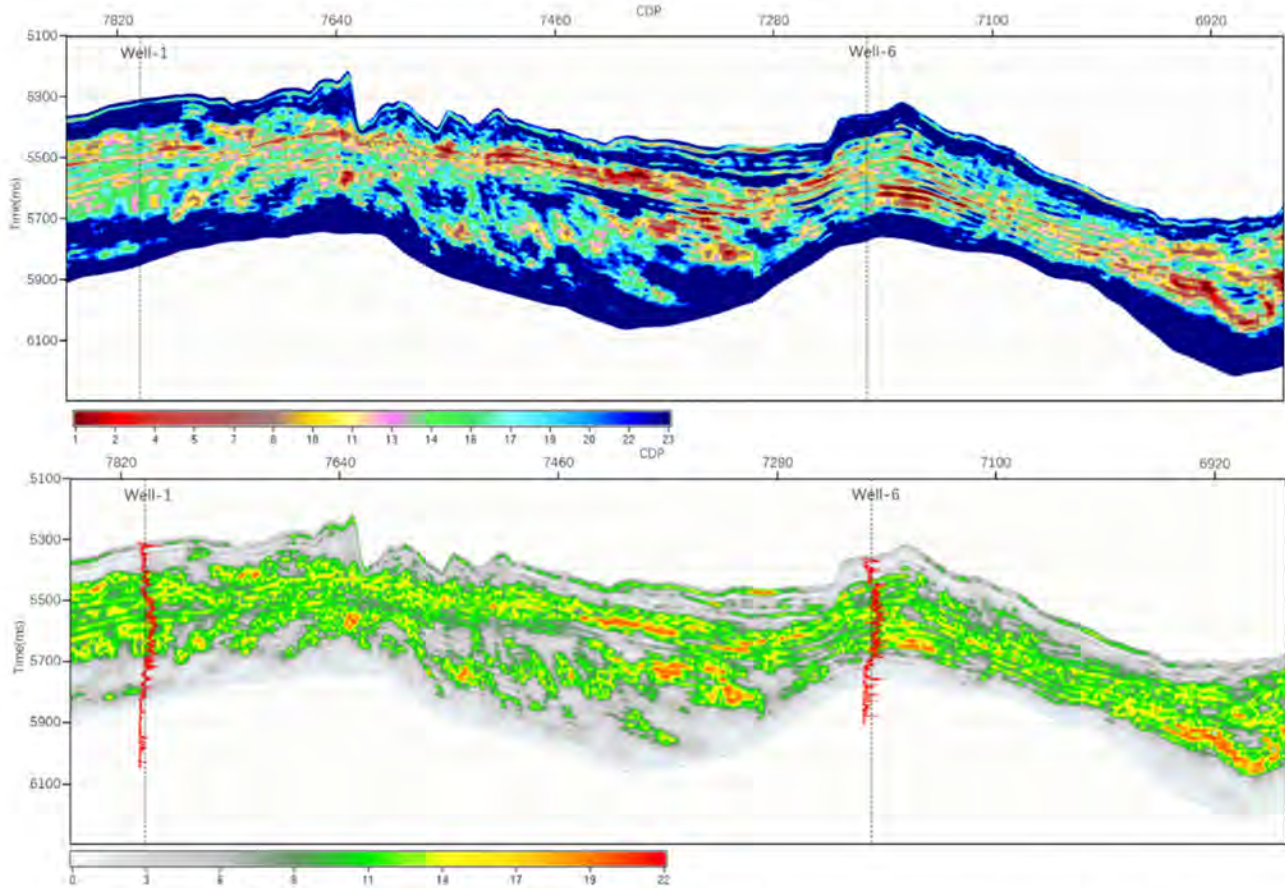


Figure 13. Comparing the Lamé coefficient profile calculated by inversion with the porosity profile calculated by inversion, Well 6 is a new well with porosity information that can be used to verify the inversion effect.

4. Discussion

When the parameters are solved in reservoir prediction, especially in the case of logged information constraints, the constrained sparse-pulse inversion method is often used, which can better reflect the lateral changes in the reservoir, despite the fact that the resolution of this method is low. When the underground reservoir structure is not so complex or the reservoir is thick, this inversion method can be used for reservoir prediction. However, the multiparameter joint inversion method based on deep learning, as used in this article, can be better applied in the case of carbonate reservoirs under salt, which have a low main frequency and complex underground reservoir structure.

The fitting problem between different parameters is basically low-dimensional. The traditional method is to use Gaussian fitting or polynomial fitting to solve the problem directly. The information that can be used is relatively singular. The use of LSTM in deep learning can obtain better results with greater accuracy. The multiparameter joint inversion method based on deep learning, as proposed in this paper, can improve the accuracy of reservoir prediction and increase the resolution and stability of the inversion results, which improves the consistency of the multiparameter inversion results. The results show

that multiparameter joint inversion can reflect the thin-layer changes of the underground structure, which is helpful for a new round of exploration and development of the oilfield in the future and for setting new well positions.

In the network construction process of LSTM based on multiparameter inversion, the selection process of hyperparameters also has an important impact on the final accuracy of porosity prediction. The following discusses four typical examples of hyperparameter tuning processes. Figure 14a,b shows the results when the learning rate is 0.1 and 0.001, respectively, under the same conditions of other hyperparameters. It can be seen that the calculated porosity and the actual porosity have similar trends, but the error in (a) is much larger than that in (b), and the coincidence rate in (a) is approximately 58.63%. Both (b) and (c) in Figure 14 are the results of changing the number of LSTM hidden units, whereas the learning rate and other parameters remain unchanged. Among them, the overall coincidence rate of (b) is 91.28%, and the coincidence rate of (c) is approximately 87.40%. It can be seen that changing the number of LSTM hidden units changes the final coincidence rate result, although the impact is not as large as the learning rate. The difference between (d) and (b) is as follows: after setting a better learning rate and number of LSTM units, by setting the dropout factor function, the learning accuracy increases with the number of iterations in the later stage of training. At this time, the coincidence rate of (d) reaches 93.15%. From the discussion of the above examples, it can be concluded that when the network is optimized for hyperparameters, the learning rate has a greater impact on the final result, followed by the number of LSTM hidden units, and the model accuracy can be increased by setting the drop factor, although the impact is smaller than that of the previous two parameters.

Because hyperparameter optimization involves many intrinsic parameters of neural networks, it is an optimization problem of parameter combination, so it cannot be optimized by methods such as gradient descent; thus, there will be great uncertainty in the optimization process. Currently, most of the hyperparameters in the application are manually configured, which is also closely related to the experience of the researchers. Most of the selection process is based on experience and compares multiple results to select the results of the hyperparameters corresponding to the best results. In addition, the optimization of hyperparameters often depends on the quality of the results used to determine the quality of the parameters. However, when faced with a large amount of data, each time, the process of solving for the results takes considerable time, which makes the optimization process of hyperparameters very time-consuming. The above process of hyperparameter selection and optimization needs to be further studied in the future.

The particle swarm optimization method can search several random solutions and then iterate to determine the optimal solution. Particle swarm optimization can be introduced into the hyperparameter optimization process of deep learning to obtain the most suitable hyperparameters for the model. Alternatively, a Bayesian optimization selection method for hyperparameters can also be considered, and the next optimal hyperparameter combination with maximum probability can be predicted according to the existing hyperparameter results. The above two optimization algorithms or similar optimization algorithms obtain results with smaller errors in the construction of a deep learning framework.

Because the LSTM is composed of input gates, output gates and forget gates, the computer used for calculations must meet stronger requirements for calculation efficiency. If the amount of data to be processed is particularly large, some other networks, such as gated recurrent units (GRUs), can also be considered to reduce the calculation load. In this case, fewer parameters can be used in the GRU network, and the proportion of forgetting versus remembering information can be selected by using one gate-control signal. Due to the structural characteristics of LSTM itself, it is difficult to parallelize, although new technologies may serve to optimize it in the future. Therefore, we can also consider using other networks for analysis in the future, which may offer better results.

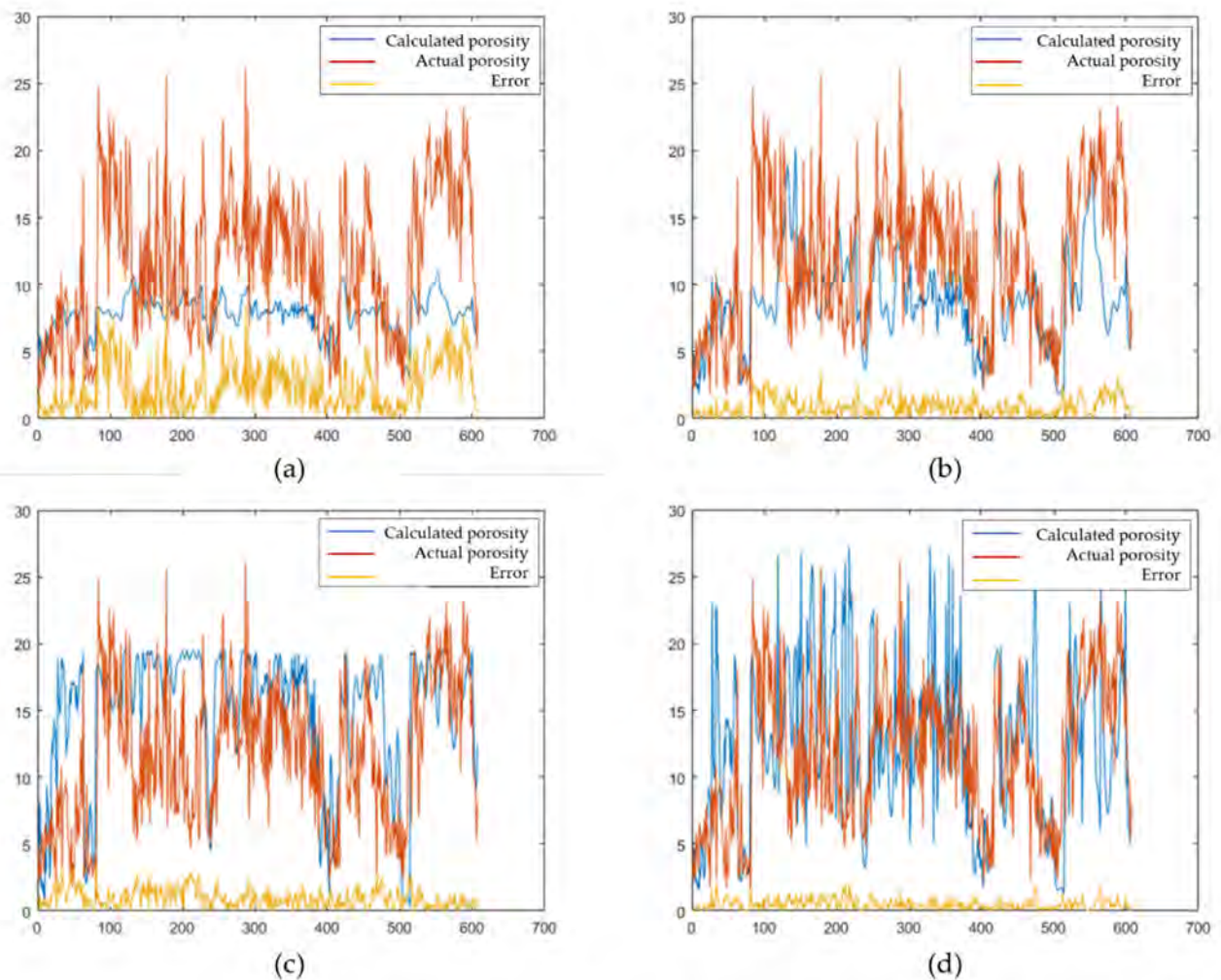


Figure 14. Four typical examples encountered in the hyperparameter selection process. There are different learning rates between (a,b), different numbers of hidden units between (b,c), and different dropout factor functions between (b,d).

In the future, we will be able to refer to the new method proposed in this article to solve problems associated with reservoir prediction, going beyond analysis of known carbonate reservoirs, e.g., traditional sandstone reservoirs. If we encounter an oil field during the development period or a work area with significant well information, then due to the increase in prior information and data, the method in this article should be able to obtain better results with greater efficiency than the conventional prediction method.

5. Conclusions

1. Carbonate reservoirs are highly heterogeneous and difficult to predict. Joint deep learning and multiparameter joint inversion technology can simultaneously use multiple pieces of prior information for inversion and can restore real underground geological structure information.
2. In the quantitative prediction of carbonate reservoirs, precision and stability are improved by using prestack multiparameter inversion, which lays a foundation for the calculation of high-quality prestack parameters and reduces the uncertainty of prediction results. The inversion results can be used as the original input, and the porosity values of the Lamé coefficient inversion data volume and log data can be used as common constraints to obtain better results of reservoir physical parameters.
3. Recurrent neural networks can deal with complex reservoir prediction work and can be used for feature mapping in high-dimensional spaces. LSTM technology can obtain

the inversion results of porosity in high-dimensional spaces through the role of gate control, e.g., a forgetting gate, where the coincidence rate of porosity and logged porosity results reached approximately 97.76% in this study.

4. Because the method proposed in this study is in the early period of exploration and development and the oil reservoir of interest is buried under deep water, the associated drilling costs are high. Few well data are available for the existing work area, with only eight existing wells in the area. In addition, of the existing well data, fewer than eight wells possess shear wave information, meaning that there is little data available for prestack multiparameter inversion, which will affect the inversion results. With continued progress in the areas of exploration and development, more well data will become available, and more log data can be used to constrain the multiparameter joint inversion method presented in this paper because the increase in what will become prior information will yield better inversion results before subsequent future activities.
5. Due to the variety of deep learning methods available today, different network frameworks have their own advantages and disadvantages. The selection of the number of layers in the network, the change in input data and the selection of hyperparameters will affect the final prediction result. In future research, all these attributes can be improved based on the method proposed in this paper. Importantly, although relevant well data are lacking, by combining the results of deep learning and multiparameter joint inversion, predictions for the main target interval in the early stage of exploration and development can deliver thin-layer prediction results that can serve as a reference. The prediction results for porosity are also close to the results obtained from log curves, which reflects the advantages of the proposed method and thus provides a new approach for the prediction of complex carbonate reservoirs in the future.
6. Because artificial intelligence technology, computer hardware and computing power are continuously being advanced, new artificial intelligence technologies will be available in the future, which will better help seismologists and geologists to carry out reservoir prediction work, thereby reducing the costs of exploration and interpretation work and delivering increasingly accurate reservoir prediction results.

Author Contributions: Conceptualization, X.T. and H.H.; methodology, X.T. and H.H.; software, X.T. and H.H.; validation, X.T., H.H., S.C., C.W., P.L. and Y.H.; formal analysis, X.T., H.H. and Y.H.; investigation, X.T., S.C. and C.W.; resources, X.T. and H.H.; data curation, X.T., H.H. and S.C.; writing—original draft preparation, X.T.; writing—review, X.T., H.H., P.L. and Y.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was jointly supported by the Natural Science Foundation of China with grant number 41974124 and 42004114, and the National Science Foundation of Jiangxi Province with grant number 20202BAB211010.

Data Availability Statement: Data sharing not applicable. The data used to support the findings of this study have not been made available because the author did not obtain permission to share the data.

Acknowledgments: The authors would like to thank the National Natural Science Foundation of China for their financial support during this research. We would also like to sincerely thank the reviewers for their helpful and constructive comments that clearly contributed to improve this article.

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

References

1. Tavakoli, V. *Carbonate Reservoir Heterogeneity: Overcoming the Challenges*; Springer Nature: Berlin/Heidelberg, Germany, 2019; ISBN 3-030-34773-7.
2. Kerans, C. Karst-Controlled Reservoir Heterogeneity in Ellenburger Group Carbonates of West Texas. *AAPG Bull.* **1988**, *72*, 1160–1183.
3. Eberli, G.P.; Baechle, G.T.; Anselmetti, F.S.; Incze, M.L. Factors Controlling Elastic Properties in Carbonate Sediments and Rocks. *Lead. Edge* **2003**, *22*, 654–660. [[CrossRef](#)]

4. Roehl, P.O.; Choquette, P.W. *Carbonate Petroleum Reservoirs*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012; ISBN 1-4612-5040-4.
5. Lucia, F.J. Rock-Fabric/Petrophysical Classification of Carbonate Pore Space for Reservoir Characterization. *AAPG Bull.* **1995**, *79*, 1275–1300.
6. Lucia, F.J.; Kerans, C.; Jennings, J.W. Carbonate Reservoir Characterization. *J. Pet. Technol.* **2003**, *55*, 70–72. [[CrossRef](#)]
7. Garland, J.; Neilson, J.; Laubach, S.E.; Whidden, K.J. Advances in Carbonate Exploration and Reservoir Analysis. *Geol. Soc. Lond. Spec. Publ.* **2012**, *370*, 1–15. [[CrossRef](#)]
8. Wright, V.P.; Barnett, A.J. An Abiotic Model for the Development of Textures in Some South Atlantic Early Cretaceous Lacustrine Carbonates. *Geol. Soc. Lond. Spec. Publ.* **2015**, *418*, 209–219. [[CrossRef](#)]
9. Riding, R. Microbial Carbonates: The Geological Record of Calcified Bacterial–Algal Mats and Biofilms. *Sedimentology* **2000**, *47*, 179–214. [[CrossRef](#)]
10. Ahr, W.M. *Geology of Carbonate Reservoirs: The Identification, Description and Characterization of Hydrocarbon Reservoirs in Carbonate Rocks*; John Wiley & Sons: Hoboken, NJ, USA, 2011; ISBN 1-118-21038-7.
11. Moore, C.H.; Wade, W.J. *Carbonate Reservoirs: Porosity and Diagenesis in a Sequence Stratigraphic Framework*; Newnes: London, UK, 2013; ISBN 0-444-53832-1.
12. Adler, F.; Baina, R.; Soudani, M.A.; Cardon, P.; Richard, J.-B. Nonlinear 3D Tomographic Least-Squares Inversion of Residual Moveout in Kirchhoff Prestack-Depth-Migration Common-Image Gathers. *Geophysics* **2008**, *73*, VE13–VE23. [[CrossRef](#)]
13. Köhl, H.; Sacchi, M.D. Least-Squares Wave-Equation Migration for AVP/AVA Inversion. *Geophysics* **2003**, *68*, 262–273. [[CrossRef](#)]
14. Kováčik, J. Correlation between Poisson’s Ratio and Porosity in Porous Materials. *J. Mater. Sci.* **2006**, *41*, 1247–1249. [[CrossRef](#)]
15. Hill, R. The Elastic Behaviour of a Crystalline Aggregate. *Proc. Phys. Soc. Sect. A* **1952**, *65*, 349. [[CrossRef](#)]
16. Kuster, G.T.; Toksöz, M.N. Velocity and Attenuation of Seismic Waves in Two-Phase Media: Part I. Theoretical Formulations. *Geophysics* **1974**, *39*, 587–606. [[CrossRef](#)]
17. Smith, T.M.; Sondergeld, C.H.; Rai, C.S. Gassmann Fluid Substitutions: A Tutorial. *Geophysics* **2003**, *68*, 430–440. [[CrossRef](#)]
18. Hinton, G.E.; Salakhutdinov, R.R. Reducing the Dimensionality of Data with Neural Networks. *Science* **2006**, *313*, 504–507. [[CrossRef](#)] [[PubMed](#)]
19. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet Classification with Deep Convolutional Neural Networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*. [[CrossRef](#)]
20. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-Cnn: Towards Real-Time Object Detection with Region Proposal Networks. *Adv. Neural Inf. Process. Syst.* **2015**, *28*. [[CrossRef](#)]
21. Hopfield, J.J. Neurons with Graded Response Have Collective Computational Properties like Those of Two-State Neurons. *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 3088–3092. [[CrossRef](#)]
22. Hochreiter, S.; Schmidhuber, J. LSTM Can Solve Hard Long Time Lag Problems. *Adv. Neural Inf. Process. Syst.* **1996**, *9*, 473–479.
23. Jordan, M.I. An Introduction to Linear Algebra in Parallel Distributed Processing. *Parallel Distrib. Process.* **1986**, *1*, 365–422.
24. Elman, J.L. Finding Structure in Time. *Cogn. Sci.* **1990**, *14*, 179–211. [[CrossRef](#)]
25. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]
26. Hochreiter, S. The Vanishing Gradient Problem during Learning Recurrent Neural Nets and Problem Solutions. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* **1998**, *6*, 107–116. [[CrossRef](#)]
27. Koller, O.; Camgoz, N.C.; Ney, H.; Bowden, R. Weakly Supervised Learning with Multi-Stream CNN-LSTM-HMMs to Discover Sequential Parallelism in Sign Language Videos. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *42*, 2306–2320. [[CrossRef](#)] [[PubMed](#)]
28. Choi, H.K. Stock Price Correlation Coefficient Prediction with ARIMA-LSTM Hybrid Model. *arXiv* **2018**, arXiv:1808.01560.
29. Bappy, J.H.; Simons, C.; Nataraj, L.; Manjunath, B.S.; Roy-Chowdhury, A.K. Hybrid Lstm and Encoder–Decoder Architecture for Detection of Image Forgeries. *IEEE Trans. Image Process.* **2019**, *28*, 3286–3300. [[CrossRef](#)]
30. Tang, J.; Shu, X.; Yan, R.; Zhang, L. Coherence Constrained Graph LSTM for Group Activity Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**. [[CrossRef](#)]
31. Barzegar, R.; Aalami, M.T.; Adamowski, J. Short-Term Water Quality Variable Prediction Using a Hybrid CNN–LSTM Deep Learning Model. *Stoch. Environ. Res. Risk Assess.* **2020**, *34*, 415–433. [[CrossRef](#)]
32. Santosh, T.; Ramesh, D.; Reddy, D. LSTM Based Prediction of Malaria Abundances Using Big Data. *Comput. Biol. Med.* **2020**, *124*, 103859. [[CrossRef](#)]
33. Graves, A.; Jaitly, N.; Mohamed, A.R. Hybrid Speech Recognition with Deep Bidirectional LSTM. In Proceedings of the 2013 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU), Olomouc, Czech Republic, 8–12 December 2013.
34. Kwon, D.-H.; Kim, J.-B.; Heo, J.-S.; Kim, C.-M.; Han, Y.-H. Time Series Classification of Cryptocurrency Price Trend Based on a Recurrent LSTM Neural Network. *J. Inf. Process. Syst.* **2019**, *15*, 694–706.
35. Neshat, M.; Nezhad, M.M.; Abbasnejad, E.; Mirjalili, S.; Gropi, D.; Heydari, A.; Tjernberg, L.B.; Garcia, D.A.; Alexander, B.; Shi, Q. Wind Turbine Power Output Prediction Using a New Hybrid Neuro-Evolutionary Method. *Energy* **2021**, *229*, 120617. [[CrossRef](#)]
36. Rijdsdijk, J.; Wu, L.; Perin, G.; Picek, S. Reinforcement Learning for Hyperparameter Tuning in Deep Learning-Based Side-Channel Analysis. *IACR Trans. Cryptogr. Hardw. Embed. Syst.* **2021**, *2021*, 677–707. [[CrossRef](#)]
37. Koutsoukas, A.; Monaghan, K.J.; Li, X.; Huan, J. Deep-Learning: Investigating Deep Neural Networks Hyper-Parameters and Comparison of Performance to Shallow Methods for Modeling Bioactivity Data. *J. Cheminform.* **2017**, *9*, 42. [[CrossRef](#)] [[PubMed](#)]

-
38. Pareek, V.; Chaudhury, S. Deep Learning-Based Gas Identification and Quantification with Auto-Tuning of Hyper-Parameters. *Soft Comput.* **2021**, *25*, 14155–14170. [[CrossRef](#)]
 39. Tian, X.; Huang, H.; Gao, J.; Luo, Y.; Zeng, J.; Cui, G.; Zhu, T. Pre-Stack Seismic Data-Driven Pre-Salt Carbonate Reef Reservoirs Characterization Methods and Application. *Minerals* **2021**, *11*, 973. [[CrossRef](#)]
 40. Thomsen, L. Weak Elastic Anisotropy. *Geophysics* **1986**, *51*, 1954–1966. [[CrossRef](#)]
 41. Staudemeyer, R.C.; Morris, E.R. Understanding LSTM—A Tutorial into Long Short-Term Memory Recurrent Neural Networks. *arXiv* **2019**, arXiv:1909.09586.
 42. Jalayer, M.; Orsenigo, C.; Vercellis, C. Fault Detection and Diagnosis for Rotating Machinery: A Model Based on Convolutional LSTM, Fast Fourier and Continuous Wavelet Transforms. *Comput. Ind.* **2021**, *125*, 103378. [[CrossRef](#)]