Petroleum Science 21 (2024) 2329–2344

Contents lists available at ScienceDirect

Petroleum Science

journal homepage: www.keaipublishing.com/en/journals/petroleum-science

Original Paper

Deep learning CNN-APSO-LSSVM hybrid fusion model for feature optimization and gas-bearing prediction

Jiu-Qiang Yang [a], Nian-Tian Lin [a, b, *], Kai Zhang [a], Yan Cui [a], Chao Fu [c], Dong Zhang [d]

[a] College of Earth Sciences and Engineering, Shandong University of Science and Technology, Qingdao, 266590, Shandong, China
[b] Laboratory for Marine Mineral Resources, Qingdao National Laboratory for Marine Science and Technology, Qingdao, 266071, Shandong, China
[c] Transportation Institute of Inner Mongolia University, Hohhot, 010000, Inner Monggol, China
[d] Key Laboratory of Gas Hydrate, Qingdao Institute of Marine Geology, Ministry of Natural Resources, Qingdao, 266237, Shandong, China

ARTICLE INFO

Article history:
Received 16 September 2023
Received in revised form
12 December 2023
Accepted 23 February 2024
Available online 28 February 2024

Edited by Jie Hao and Meng-Jiao Zhou

Keywords:
Multicomponent seismic data
Deep learning
Adaptive particle swarm optimization
Convolutional neural network
Least squares support vector machine
Feature optimization
Gas-bearing distribution prediction

ABSTRACT

Conventional machine learning (CML) methods have been successfully applied for gas reservoir prediction. Their prediction accuracy largely depends on the quality of the sample data; therefore, feature optimization of the input samples is particularly important. Commonly used feature optimization methods increase the interpretability of gas reservoirs; however, their steps are cumbersome, and the selected features cannot sufficiently guide CML models to mine the intrinsic features of sample data efficiently. In contrast to CML methods, deep learning (DL) methods can directly extract the important features of targets from raw data. Therefore, this study proposes a feature optimization and gas-bearing prediction method based on a hybrid fusion model that combines a convolutional neural network (CNN) and an adaptive particle swarm optimization-least squares support vector machine (APSO-LSSVM). This model adopts an end-to-end algorithm structure to directly extract features from sensitive multicomponent seismic attributes, considerably simplifying the feature optimization. A CNN was used for feature optimization to highlight sensitive gas reservoir information. APSO-LSSVM was used to fully learn the relationship between the features extracted by the CNN to obtain the prediction results. The constructed hybrid fusion model improves gas-bearing prediction accuracy through two processes of feature optimization and intelligent prediction, giving full play to the advantages of DL and CML methods. The prediction results obtained are better than those of a single CNN model or APSO-LSSVM model. In the feature optimization process of multicomponent seismic attribute data, CNN has demonstrated better gas reservoir feature extraction capabilities than commonly used attribute optimization methods. In the prediction process, the APSO-LSSVM model can learn the gas reservoir characteristics better than the LSSVM model and has a higher prediction accuracy. The constructed CNN-APSO-LSSVM model had lower errors and a better fit on the test dataset than the other individual models. This method proves the effectiveness of DL technology for the feature extraction of gas reservoirs and provides a feasible way to combine DL and CML technologies to predict gas reservoirs.

1. Introduction

Conventional machine learning (CML) methods, including support vector machines (SVM) (Sen et al., 2021; Moosavi et al., 2022), random forests (RF) (Wang et al., 2020; Otchere et al., 2022) and artificial neural networks (ANN) (Brantson et al., 2019; Kalam et al., 2022), have been extensively applied in reservoir prediction to mine the internal relationships between data. These methods have shown good applicability in predicting porosity, permeability, and lithofacies (Tran et al., 2020; Shao et al., 2022; Dong et al., 2023; Zou et al., 2023), etc. For a successful prediction using these CML methods, the features of the sample data should first be optimized to remove irrelevant and redundant information from the target data and highlight the sample features that contribute significantly to it. Seismic attribute analysis technology, as an important reservoir characterization method, can highlight or enhance seismic features related to reservoirs (Hossain, 2020; Dewett et al., 2021;

* Corresponding author. College of Earth Sciences and Engineering, Shandong University of Science and Technology, Qingdao, 266590, Shandong, China.
E-mail address: linnt@sdust.edu.cn (N.-T. Lin).

https://doi.org/10.1016/j.petsci.2024.02.012

Zhumabekov et al., 2021). However, several types of seismic attributes exist and blindly inputting numerous seismic attributes into CML models can lead to information redundancy and reduce prediction accuracy. Therefore, feature optimization is required to highlight information that contributes significantly to the prediction target. Commonly used seismic attribute feature optimization methods include principal component analysis (PCA) and independent component analysis (ICA) (Lubo-Robles and Marfurt, 2019; Babiki et al., 2022; Yao et al., 2022), etc, which have been extensively applied to optimize longitudinal wave seismic attributes. Multicomponent seismic attributes (MSAs) contain more abundant reservoir information compared with longitudinal wave attributes; therefore, they have broader application prospects for reservoir prediction. However, the large amount of information in MSA data makes feature optimization more difficult. Yuan et al. (2011) optimized longitudinal and converted shear wave attributes through PCA and ICA, respectively, and then combined them using an SVM for reservoir prediction, achieving a prediction result superior to that obtained using single longitudinal wave. Yang et al. (2021) optimized MSAs sensitive to gas reservoir through cluster analysis and then constructed composite attributes through composite operations (CO) to combine longitudinal and converted shear wave attributes to complete feature optimization. Finally, the constructed composite attributes were applied to a deep neural network for reservoir prediction and good results were obtained.

The aforementioned feature optimization and reservoir prediction methods have limitations. First, the construction of composite seismic attributes using the CO requires the selection of appropriate operational methods according to the characteristics of the target area, which requires further experimental work. In areas with obvious differences in reservoir characteristics, new composite seismic attributes must be constructed for feature optimization. Moreover, their application is limited when the features of the target area are unclear. Although unsupervised learning methods, such as PCA, can still be applied to such cases, the reservoir information and geological significance provided by them are unclear, and the relationship between optimized seismic attributes and target reservoir must be re-established. Second, the steps of the above feature optimization methods are cumbersome and time-consuming, and effective information related to reservoir prediction may be lost or redundant information may be highlighted. Third, although different CML methods have demonstrated good reservoir prediction capabilities, their limitations are evident. For example, ANN are prone to overfitting during training (Srivastava et al., 2014; Nguyen et al., 2021), and the kernel function selection and parameter optimization of SVM significantly affect the prediction results (Syarif et al., 2016). Therefore, formulating more effective strategies to alleviate these problems and obtain promising predictive results is necessary.

To alleviate the problems of parameter optimization and overfitting during the training process of CML models for reservoir prediction, intelligent prediction models can be constructed by combining optimization algorithms with CML models to improve their predictive performance (Nabipour et al., 2020; Gheytanzadeh et al., 2021). In the training process of CML models, the powerful search and parameter optimization abilities of intelligent optimization algorithms could effectively alleviate overfitting problem and determine the optimal parameters for the CML models. Intelligent prediction models have been widely applied for reservoir prediction. Commonly used model construction methods combine CML methods, such as ANN, SVM, fuzzy logic systems, and Gaussian process regression (Bahadori et al., 2016; Zarei and Baghban, 2017; Kardani et al., 2018), and evolutionary algorithms, such as particle swarm optimization (PSO), genetic algorithm, and imperialist competitive algorithms for reservoir prediction (Baghban, 2016; Baghban and Adelizadeh, 2018; Bemani et al., 2020a, 2020b). To construct an intelligent prediction model, we first determine the appropriate CML method and then select the appropriate optimization algorithm to optimize its parameters to obtain an intelligent prediction model that can meet the target prediction requirements. In seismic-data-based gas reservoir predictions, limited label data can be obtained owing to the constraints of actual drilling. Therefore, least-squares SVM (LSSVM), which has strong applicability to small samples and optimization problems, was used in this study for gas-bearing prediction. The kernel parameters of the LSSVM affect its predictive performance, and if not properly selected, they can lead to unsatisfactory prediction results (Xie et al., 2019; Seyyedattar et al., 2022). This study adopted PSO algorithm with a global optimization ability to optimize the parameters of LSSVM. However, standard PSO algorithm may encounter local optimization when optimizing the model parameters (Taherkhani and Safabakhsh, 2016; Han et al., 2018). To alleviate this problem, this study improves the parameter optimization ability of PSO algorithm by simultaneously adaptively adjusting inertia weight and velocity coefficients. This paper proposes an intelligent prediction model using a combination of adaptive PSO (APSO) and LSSVM for gas-bearing prediction. The APSO algorithm alleviates the overfitting problem of PSO algorithm and adjusts relevant parameters of the LSSVM to improve its predictive performance. However, the constructed APSO-LSSVM method can only alleviate the parameter optimization problem and not its dependence on sample quality. Therefore, the features of the MSA data must be optimized before using APSO-LSSVM for MSA-based gas-bearing prediction.

As a feature optimization method, feature extraction transforms the original data into a set of features representing different attributes, which is conducive for gas reservoir prediction (Guyon et al., 2006). Deep learning (DL) methods can capture deep representations from input data (LeCun et al., 2015; Karimpouli et al., 2020; Cao et al., 2021; Wang et al., 2022). In contrast to CML methods, DL can directly extract important features of a target from original data without requiring careful feature optimization, thus simplifying the process. Convolutional neural network (CNN) has demonstrated robust performance in pattern recognition and image processing (Krizhevsky et al., 2012; Sun et al., 2020, 2021; Lou et al., 2022; Ma et al., 2023). Furthermore, CNN has demonstrated its potential for feature extraction by automatically extracting numerous robust and invariant features from raw data (Fang et al., 2020). Therefore, we applied it to the prediction of gas-bearing distribution using MSAs because it simplifies the tedious feature optimization steps and extracts important features from input variables to provide a better description for gas-bearing prediction.

This study proposes a hybrid fusion DL model that combines a CNN and APSO-LSSVM for feature optimization and gas-bearing prediction. First, the original MSAs that were sensitive to the gas reservoirs were selected through cluster analysis. Cluster analysis does not change the original attribute characteristics; that is, it is only used for feature selection. The selected MSAs were then entered into the CNN-APSO-LSSVM model for training through blind-well cross-validation, and the model parameters were adjusted to obtain the trained CNN-APSO-LSSVM model. Finally, the improved effect of the developed model on a single model was evaluated through comparing it with a single CNN and APSO-LSSVM models. To evaluate the feature extraction capability of CNN, the gas-bearing prediction result of CNN-APSO-LSSVM was compared with that obtained by inputting sensitive seismic attributes obtained from PCA and CO into the APSO-LSSVM model. In contrast to these two feature optimization methods, the CNN-APSO-LSSVM model adopts an end-to-end algorithm structure that does not need tedious feature optimization processes after cluster analysis. However, the data can be directly input into a

hybrid fusion model for prediction, which has good universality and operability. To analyze the improvement effect of APSO on LSSVM model, the model prediction performances with and without optimization using the APSO algorithm were compared. The results indicate that the proposed hybrid fusion DL model achieved better feature optimization effects and prediction results than common used feature optimization methods. It had the lowest error and highest fit among all the compared hybrid models, demonstrating excellent prediction performance of the gas-bearing distribution.

2. Methods

The key steps of feature optimization and gas-bearing prediction using the proposed hybrid fusion DL model are as follows (Fig. 1). First, the MSAs sensitive to gas reservoirs were selected using cluster analysis to reduce their dimensionality, and training and testing datasets were constructed. Subsequently, PCA, CO feature optimization, and CNN feature extraction were performed on the selected original MSAs that were sensitive to gas reservoirs. The obtained features were input into the APSO-LSSVM model for training, and three hybrid models were constructed to predict the gas-bearing probability distribution (GPD). Finally, the prediction performances of these hybrid models were evaluated using the testing dataset, and the predicted gas-bearing distribution results were comprehensively evaluated using actual drilling information.

2.1. CNN

Compared to CML methods such as SVM and RF, CNN has features such as local connections and parameter sharing owing to the existence of convolutional layers, which can effectively extract data features and alleviate the shortcomings of CML methods to some extent (Zhang et al., 2022c). When a CNN is used for seismic-data-based gas reservoir prediction, the most common method is to use curved shapes as features and input them into neural networks as images for classification (Yuan et al., 2018). However, this method has a high labor cost and low efficiency. In contrast, a one-dimensional (1D) CNN can perform convolution operations directly on 1D samples of the target reservoir and automatically extract target features from original data (Wang and Cao, 2021). Therefore, this method has higher efficiency and lower computational cost.

A 1D convolutional layer is a basic layer type in a CNN used to process data with a sequence structure (Kiranyaz et al., 2021; Xue et al., 2022). It can extract local features from sequential data and encode them into higher-level representations. A 1D convolutional layer typically contains several learnable convolutional kernels, each of which performs a 1D convolution on an input sequence to obtain a new set of feature sequences. The output is

$$y^l = f\left(w^l * x^{l-1} + b^l\right) \tag{1}$$

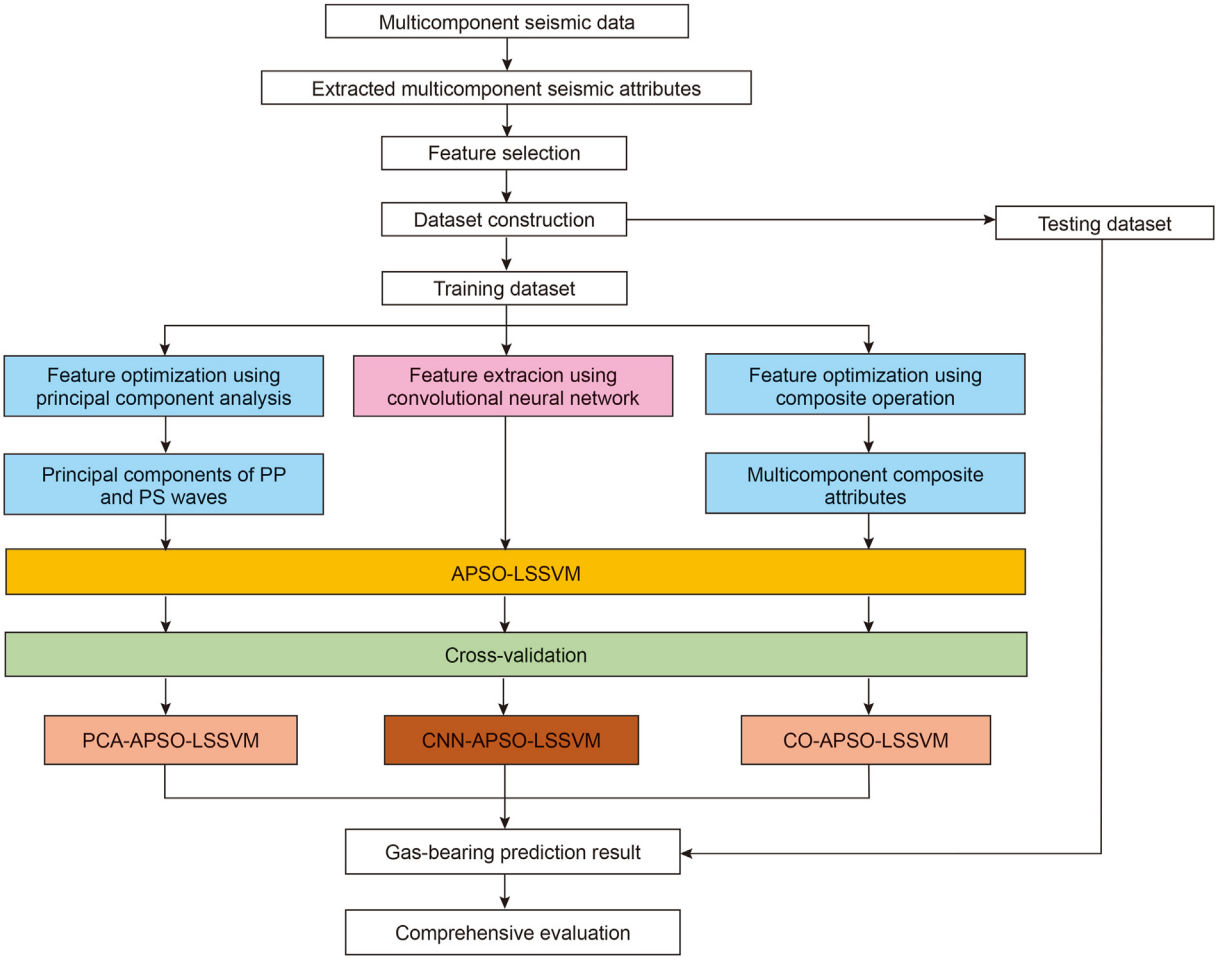

Fig. 1. Workflow for feature optimization and gas-bearing prediction using the proposed hybrid fusion model.

where x^{l-1} and y^l are the outputs of the convolutional layers $l-1$ and l, respectively; w^l is the convolutional kernel of the l^{th} convolutional layer; b^l is the bias; and * denotes the convolution operation.

$f(\cdot)$ is a non-linear activation function. This study selected the ReLU, and its expression is as follows:

$$f_{\text{ReLU}}(x) = \max(0, x) \tag{2}$$

The pooling layer is an important part of CNN and is used to reduce the space size of feature maps, thereby reducing model parameters and alleviating overfitting. The pooling operation usually involves statistical analysis of each small domain in the input feature map to obtain a summarized feature value. The mathematical expression for the pooling layer of a 1D CNN is

$$y_i = \text{pooling}(x_{i:i+l-1}) \tag{3}$$

where x and y are the input and output 1D signal sequences, respectively; l is the pooling window size; and i is the starting position of the pooling window. In this study, a max-pooling operation was adopted.

The fully connected (FC) layer receives the results of alternating outputs from the previous convolutional and pooling layers, integrates and extracts the features, rearranges them into a 1D feature vector, and transmits them to the output layer. The operation process is

$$y^r = f\left(w^r \cdot y^{r-1} + b^r\right) \tag{4}$$

where y^{r-1} and y^r are the outputs of layers $r-1$ and r, respectively; $f(\cdot)$ is the activation function; w^r is the weight matrix; b^r is the bias; and $\cdot$ is the matrix dot product.

2.2. APSO-LSSVM

After sample feature extraction by the CNN, LSSVM was used for GPD prediction. The kernel function selection in the LSSVM significantly affects its predictive performance. As the LSSVM model using radial basis function (RBF) had the best prediction effect among the different kernel functions for MSA-based GPD prediction (Yang et al., 2023b), it was selected for this study. The kernel parameters also affect the prediction accuracy. APSO was used to optimize the kernel parameters and achieve excellent prediction results.

In the classical PSO algorithm, the inertia weight and velocity coefficients are often fixed; however, with progressive iterations, they often become too large or too small, leading to local optimization and reducing the optimization ability (Chaitanya et al., 2021). To alleviate these problems, we adopted adaptive inertia weight and velocity coefficients that automatically adjusted during the iteration (Eqs. (5) and (6), respectively).

$$\begin{cases} w(t) = w_{\min} + \dfrac{(F - f_{\min}) \cdot (w_{\max} - w_{\min})}{\left(f_{\text{avg}} - f_{\min}\right)}, F \le f_{\text{avg}} \\ w(t) = w_{\min} + (w_{\max} - w_{\min}) \cdot (T - t) \Big/ T, F > f_{\text{avg}} \end{cases} \tag{5}$$

$$\begin{cases} c_1 = C_1 + \dfrac{F - f_{\text{avg}}}{f_{\text{avg}} - f_{\min}} \\ c_2 = C_2 - \dfrac{F - f_{\text{avg}}}{f_{\text{avg}} - f_{\min}} \end{cases} \tag{6}$$

where $w(t)$ is the adaptive inertia weight; c_1 and c_2 are the velocity coefficients; $f_{\min}$ and f_{avg} are the minimum and mean fitness, respectively; F is the current fitness; $w_{\max}$ and $w_{\min}$ are the maximum and minimum of the set inertial weight, respectively; t and T are the current and maximum iterations, respectively; C_1 and C_2 represent the initial minimum and maximum velocity coefficients, respectively.

During training of the APSO-LSSVM model, its kernel function parameters were set as the particle position vectors. When the fitness value satisfied the accuracy requirement, training was completed. The optimal result was set as the parameters for LSSVM (Fig. 2).

2.3. Construction and performance evaluation of the proposed hybrid fusion model

The proposed hybrid fusion model uses CNN model to automatically extract important features from input variables, which are then used in the APSO-LSSVM to obtain the GPD prediction results (Fig. 3). First, the CNN model was trained using a training dataset. The trained CNN was used to extract sensitive features of the gas reservoir using sensitive MSAs selected by cluster analysis. The extracted features and label data were then used to construct a new dataset for the APSO-LSSVM model training, wherein APSO was used to optimize model parameters. Finally, the trained APSO-LSSVM model was used to predict the GP. The trained CNN-APSO-LSSVM model was applied to the selected study area to predict GPD.

Detailed parameter information for the proposed CNN-APSO-LSSVM model is presented in Tables 1 and 2. In the DL model, model hyperparameters control the structure of the network and are set before training begins. These hyperparameters are different from the model parameters (weights) obtained during the training process and are usually set manually and adjusted through repeated experimentation. The predictive performance of the models trained with different hyperparameters varied significantly. Therefore, the selection of appropriate hyperparameters is crucial. In this study, the hyperparameters of the CNN model were obtained through previous experience and comparative experiments. The detailed parameters and training algorithms for the CNN are listed in Table 1. The CNN model consisted of input layer, feature layer (four alternately arranged convolutional and pooling layers), FC layer, and output layer. The Adam algorithm was used as the optimization algorithm.

The values of the kernel function parameters in LSSVM are crucial for model prediction performance. This study selected the RBF kernel function, which must determine the penalty factor C and kernel parameter σ. We used APSO algorithm for optimization. Table 2 lists the detailed information on the proposed APSO-LSSVM model. In contrast to the standard PSO algorithm, APSO adaptively adjusts the values of inertia weight and velocity coefficients during training; therefore, only their initial ranges need to be set. In the APSO algorithm, the swarm size was 200, maximum inertia weight $w_{\max}$ was 0.8, minimum inertia weight $w_{\min}$ was 0, initial minimum value of the velocity coefficient C_1 was 1.5, and initial maximum value of the velocity coefficient C_2 was 2.5. In the LSSVM model, the kernel function parameters C and σ were optimized using the APSO algorithm and the final APSO-LSSVM model was obtained. The specific optimization process is shown in Fig. 2.

In CML methods, cross-validation is typically adopted to train models and eliminate potential overfitting problems. However, in contrast to CML models, CNN need to extract local feature information from the input data and using conventional cross-validation will lead to data discontinuity and destroy effective feature information. Therefore, to ensure the continuity of the sample data and maximize the full use of the training samples, we adopted a blind-

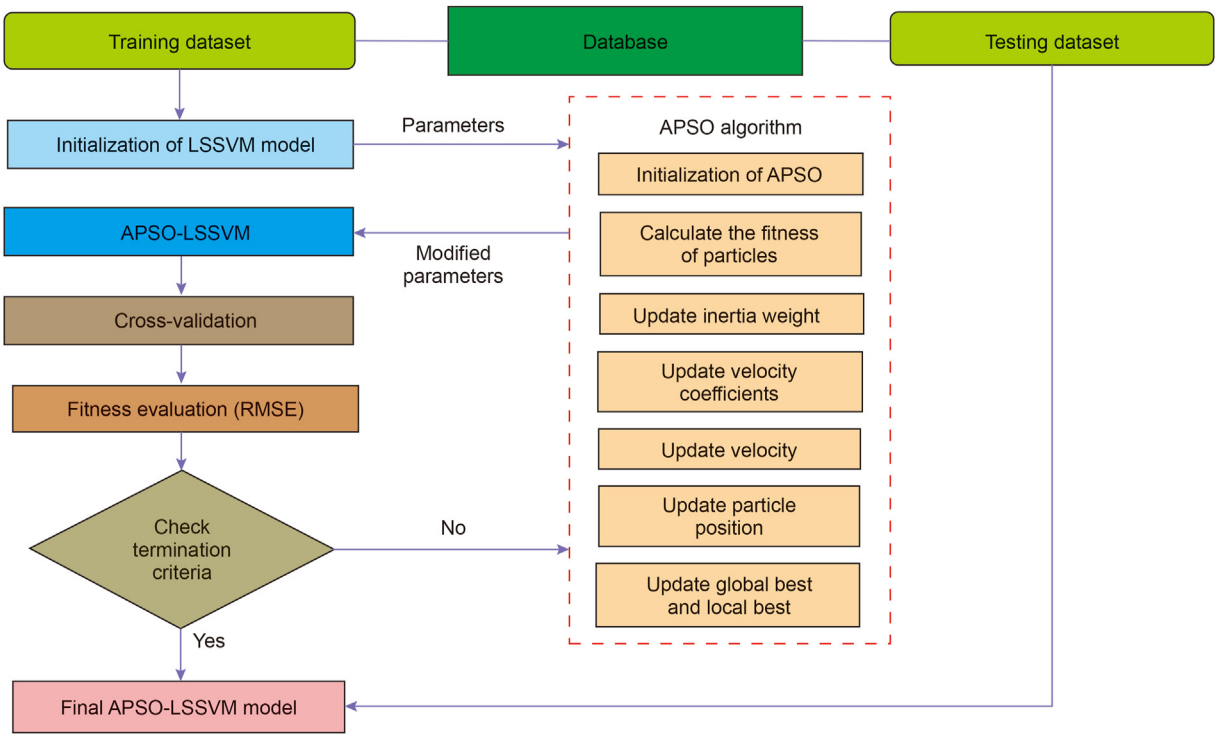

Fig. 2. Flowchart of APSO-optimized LSSVM.

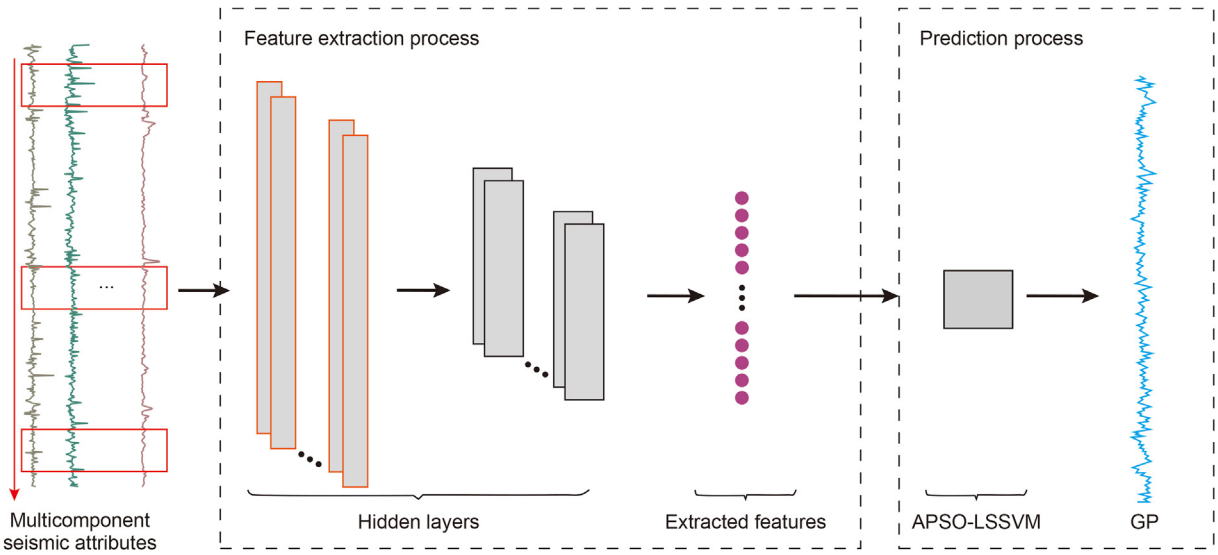

Fig. 3. Structure of the hybrid CNN-APSO-LSSVM model.

Table 1
Model hyperparameters setting of CNN model.

Model	Model hyperparameters setting	
	Structure	Other hyperparameters
CNN	1 Convolutional layer (64 filters, 4 kernel_size, ReLU activation, padding 'same'); MaxPooling (1 pooling size); 1 Convolutional layer (64 filters, 4 kernel_size, ReLU activation, padding 'same'); MaxPooling (1 pooling size); 1 Convolutional layer (128 filters, 4 kernel_size, ReLU activation, padding 'same'); MaxPooling (2 pooling size); 1 Convolutional layer (128 filters, 4 kernel_size, ReLU activation, padding 'same'); MaxPooling (2 pooling size); 1 Dense layer (1 neuron).	Optimizer is the Adam optimization algorithm; the learning rate is 0.001, the time steps are 5, the batch size is 10, the dropout is 0.2, the maximum epochs are 500, and the loss function is root mean squared error.

Table 2
Detail information of the proposed APSO-LSSVM model.

LSSVM		APSO	
Type	Value/comment	Type	Value/comment
Cross-validation	Five-fold	Swarm size	200
Kernel function	RBF	w_{max}	0.8
Kernel function parameter	C,σ	w_{min}	0
Optimization method	APSO	C_1	1.5
Iteration	500	C_2	2.5

well cross-validation method for training. Specifically, for the training dataset, sample data provided by one well each time were set aside as a validation subset, and sample data provided by the remaining wells were used as training subset. This process was repeated until all the wells were verified. Using this technique, multiple random transformations are performed on the training and validation subsets to obtain the optimal hyperparameters for constructing the CNN model. Subsequently, we used the features extracted by the CNN model and label data to form a new training dataset for APSO-LSSVM model training. The five-fold cross-validation was adopted to divide training and validation subsets and ultimately obtain the optimal parameters to construct the APSO-LSSVM model. After the two training processes are completed, the final CNN-APSO-LSSVM model was obtained.

To evaluate the prediction performance of the hybrid fusion model, we first compared its results with those of the individual APSO-LSSVM and CNN models to analyze the degree of improvement in the proposed model. Sensitive MSAs selected by cluster analysis were then processed using PCA and CO feature optimization. The obtained seismic attributes were input into the APSO-LSSVM for GPD prediction, and their results were compared with those of the hybrid model to analyze the feature extraction capability of CNN. Finally, the extracted features were input into the LSSVM model without the APSO algorithm for parameter optimization during training to analyze the impact of APSO algorithm on the model performance.

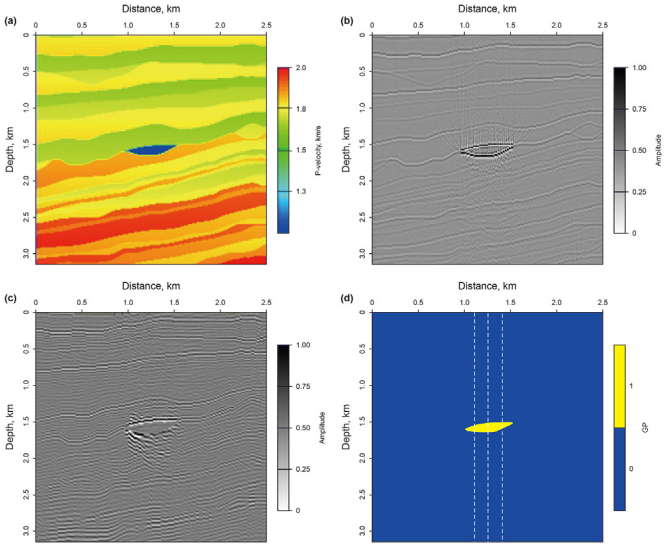

Fig. 4. Forward modeling of the wave equation of **(a)** part of the Marmousi2 model, **(b)** longitudinal and **(c)** converted shear wave migration profiles, **(d)** actual gas reservoir distribution.

3. Synthetic data examples

3.1. Data preparation and feature selection

This prediction scheme was validated using synthetic data generated by the Marmousi2 model (Martin et al., 2006). Fig. 4(a) shows the longitudinal wave velocity of a part of the Marmousi2 model that includes a gas reservoir (blue area). Staggered-grid finite-difference forward modeling was performed using the dual-phase first-order velocity stress equation (He et al., 2022) to obtain the migration profiles of the longitudinal (Fig. 4(b)) and converted shear waves (Fig. 4(c)). The yellow and blue parts in the gas profile of the Marmousi2 model (Fig. 4(d)) represent tight sandstone gas reservoirs and gas-free mudstone, respectively. Three seismic traces were selected as pseudo-wells (white dotted lines), on which the gas- and non-gas-reservoir characteristics were marked to create gas reservoir labels.

Sensitive seismic attributes were selected based on the designed scheme (Fig. 1). Features of the MSAs were selected using cluster analysis to reduce their dimensions. Note that cluster analysis does not transform the seismic attribute features. After dimension reduction, it selects the original MSAs that are sensitive to gas reservoirs and facilitates subsequent feature optimization or extraction. Six sensitive attributes, instantaneous amplitude, instantaneous frequency, instantaneous phase, RMS amplitude, average peak amplitude, and arc length (Fig. 5), were selected using cluster analysis (Zhang et al., 2019; Yang et al., 2021). To analyze the sensitivity of the selected attributes, we extracted the characteristics of different seismic attributes and label data for the three pseudo-well seismic traces (Fig. 6). Except for the instantaneous phase attribute, all other attributes were observed to contribute to gas reservoir prediction. Considering the diversity of sample data and to retain more gas reservoir characteristic information, the five contributing seismic attributes in the synthetic data were selected as input samples for the hybrid fusion DL model.

3.2. Feature optimization and gas-bearing prediction

After sample data generation, three pseudo-wells were used for CNN model training by blind-well cross-validation; that is, during each iteration, one well was set aside for validation until all three wells were validated. After network training, the predicted sample features were obtained (Fig. 7). The CNN model integrates the advantages of the original attributes, and the extracted features effectively highlight sensitive gas reservoir information while suppressing non-gas reservoir information, making the prediciton features more consistent with the label and proving the effectiveness of CNN for feature extraction. The features extracted by the CNN model and label data were then used to form a new training dataset to train the APSO-LSSVM model. During training, the APSO algorithm continuously optimized the relevant parameters of the LSSVM until the accuracy requirements were satisfied, and the entire CNN-APSO-LSSVM model training was completed.

The obtained CNN-APSO-LSSVM model accurately predicted the distribution of gas reservoirs (Fig. 8(e)) with a small error (Fig. 8(f)).

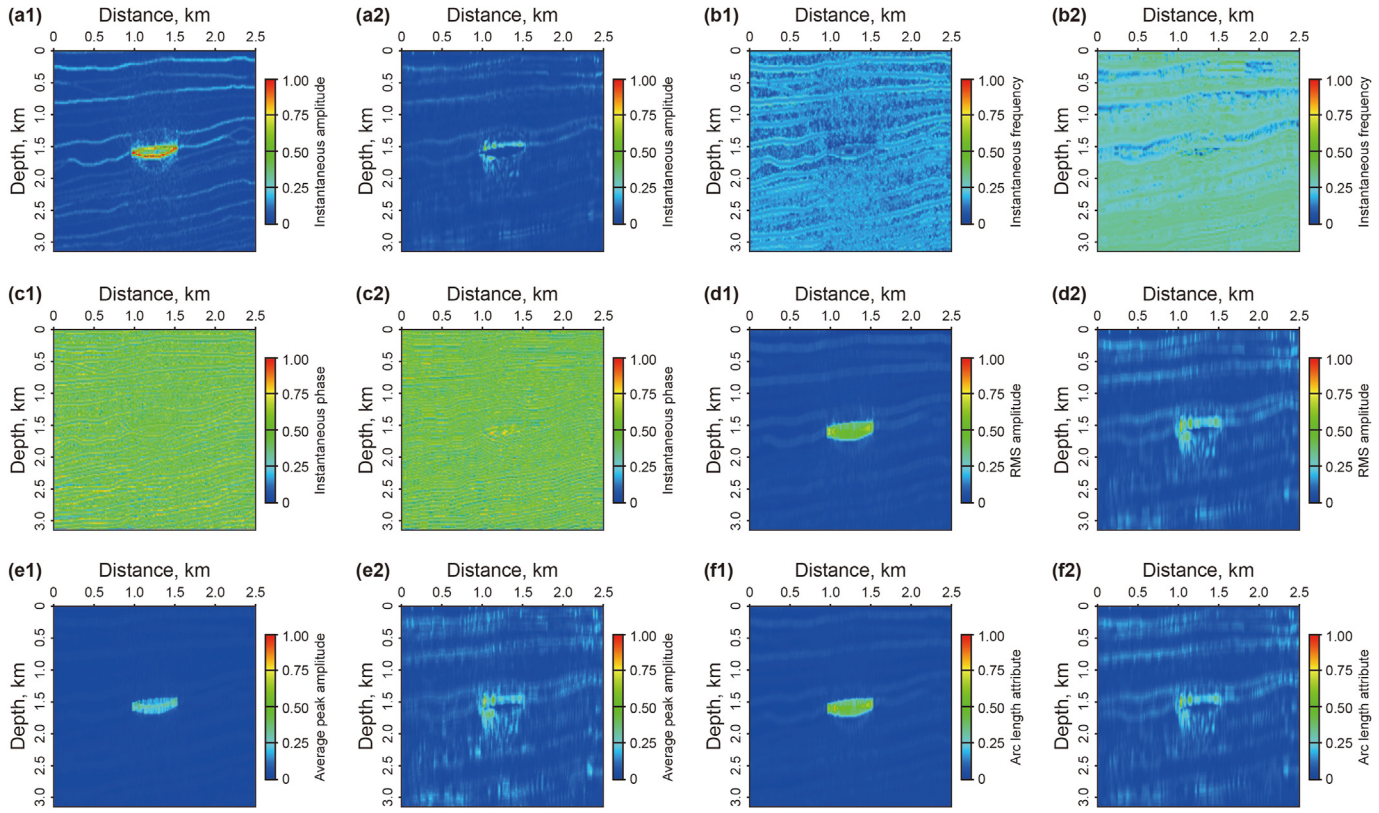

Fig. 5. Selected original MSAs sensitive to gas reservoirs in the synthetic data: **(a1)** and **(a2)** instantaneous amplitude, **(b1)** and **(b2)** instantaneous frequency, **(c1)** and **(c2)** instantaneous phase, **(d1)** and **(d2)** RMS amplitude, **(e1)** and **(e2)** average peak amplitude, and **(f1)** and **(f2)** arc length attribute. ((**a1**)–(**f1**) longitudinal waves and **(a2)**–(**f2**) converted shear waves).

To verify its advantages, we compared it with the prediction results obtained by directly inputting sensitive MSA data into the APSO-LSSVM and CNN models. The APSO-LSSVM performed poorly in gas reservoir prediction (Fig. 8(a)) and had significant errors (Fig. 8(b)) because the raw data were directly input into the model without feature optimization, and the gas reservoir features were not obvious. In contrast, the hybrid fusion model significantly improved the prediction ability of APSO-LSSVM. The CNN model made relatively accurate predictions owing to its powerful feature extraction ability (Fig. 8(c)); however, the obtained prediction results had higher errors compared to the CNN-APSO-LSSVM model, and there was some redundant information in the non-gas reservoir areas (Fig. 8(d)). The APSO-LSSVM can further optimize the features extracted by the CNN to obtain better prediction results. Overall, the constructed hybrid fusion model improved the performance of both CML and DL models, verifying its effectiveness.

3.3. Performance evaluation of the hybrid fusion model

To analyze the feature extraction capability of CNN, we compared it with the predicted results obtained by inputting sample features obtained from PCA and CO into the APSO-LSSVM model for gas-bearing prediction. The sample features optimized by PCA and CO are shown in Figs. 9 and 10. The principal components were obtained by the PCA of sensitive seismic attributes with a threshold of 90%. Finally, three principal components, each of longitudinal and converted shear wave attributes, were obtained (Fig. 9). Although the principal components obtained by the PCA synthesized the characteristics of the sensitive seismic attributes (Fig. 6) and reduced their dimensions, they still contained considerable redundant information. Three composite attributes (Fig. 10) were obtained by performing CO (Zhang et al., 2019) on sensitive seismic attributes. Compared with the principal components (Fig. 9), the three composite attributes (Fig. 10) better highlighted the sample characteristics because the latter were combined with the target reservoir characteristics to select the appropriate CO method. PCA uses the internal relationships of seismic attributes to obtain principal components, reflecting their statistical characteristics. Therefore, for a target area with known gas reservoir characteristics, the composite attributes reflected the gas reservoir characteristics better than the principal components.

We analyzed the prediction results obtained using the three feature optimization methods (Fig. 11). Although the prediction results obtained using principal components as input (PCA-APSO-LSSVM) (Fig. 11(a)) reduced redundant information and prediction error compared with the prediction results obtained using sensitive seismic attributes (Fig. 8(a)), they still contained redundant information (Fig. 11(b)), which is consistent with the sample characteristics (Fig. 9). Meanwhile, the prediction results obtained using the three composite attributes as inputs (CO-APSO-LSSVM) (Fig. 11(c)) were superior to that of PCA-APSO-LSSVM (Fig. 11(a)), with improved accuracy and significantly suppressed non-gas reservoir information (Fig. 11(d)). Furthermore, the prediction results (Fig. 11(c)) were similar to those of the CNN-APSO-LSSVM (Fig. 8(e)), with small prediction errors (Figs. 8(f) and 11(d)). This indicates that the features extracted by the CNN and those obtained from CO can better reflect the gas reservoir information. Using a CNN for feature extraction reduced the tedious steps of feature optimization present in CO and improved the feature optimization efficiency, proving its advantages in feature extraction. The feature samples obtained by the three feature optimization methods above were all input into the APSO-LSSVM model for prediction, among

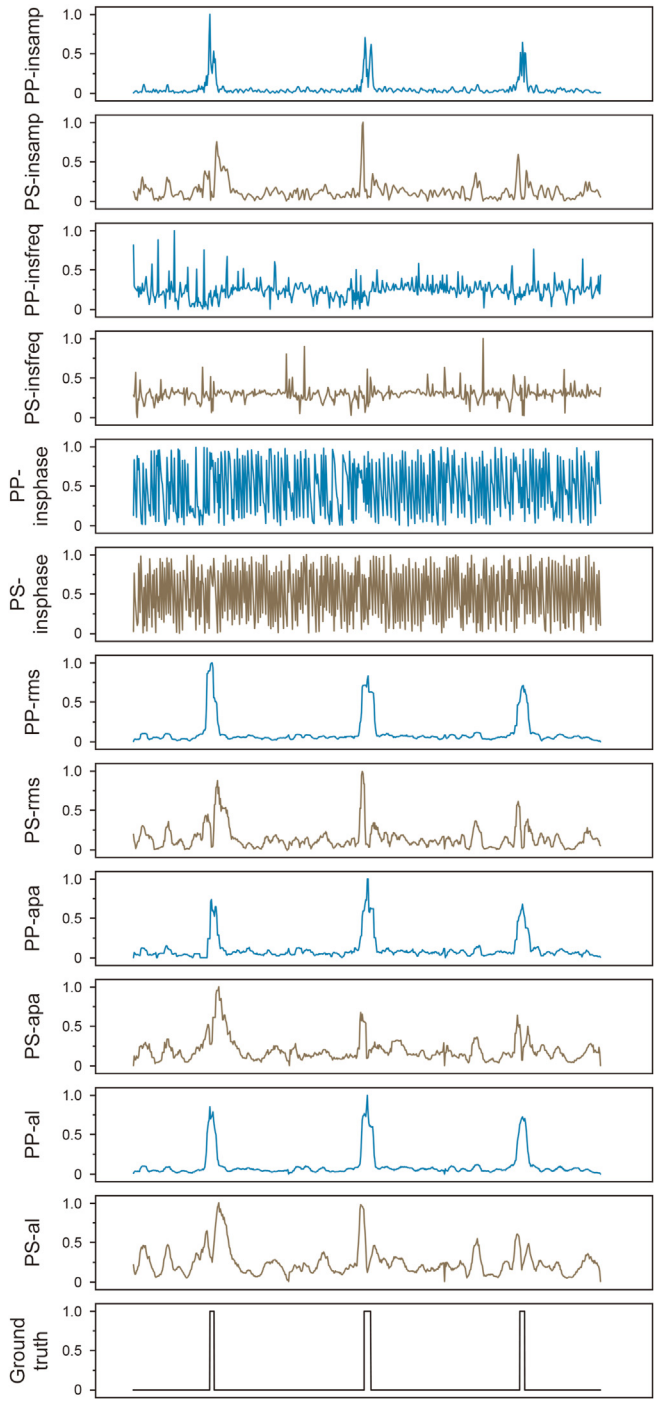

Fig. 6. Sample features of original MSAs for the three pseudo-wells in Fig. 4(d) and their label information. Note: insamp—instantaneous amplitude, insfreq—instantaneous frequency, insphase—instantaneous phase, rms—RMS amplitude, apa—average peak amplitude, al—arc length.

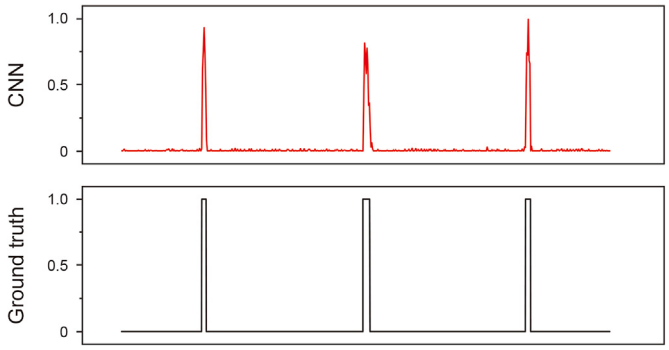

Fig. 7. Sample features for the three pseudo-wells in Fig. 4(d) predicted by CNN model and their label information.

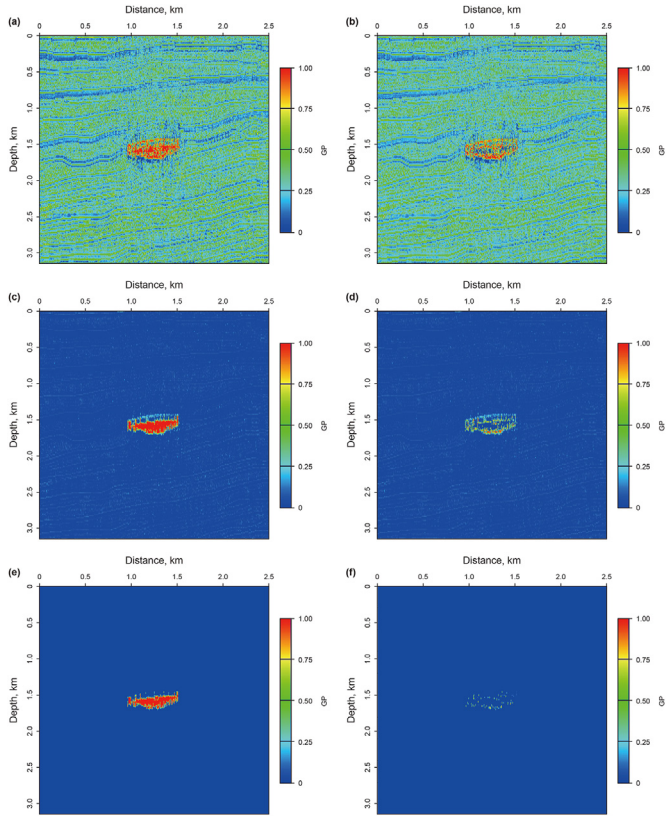

Fig. 8. Prediction results and errors of the APSO-LSSVM, CNN, and CNN-APSO-LSSVM models in the synthetic data. **(a)** Prediction result and **(b)** error of APSO-LSSVM; **(c)** prediction result and **(d)** error of CNN; **(e)** prediction result and **(f)** error of CNN-APSO-LSSVM.

which the CNN-APSO-LSSVM demonstrated the best results owing to the application of the CNN. Compared to the other two MSA optimization methods, the CNN extracted clearer sample features and improved the quality of the sample data. Moreover, the constructed CNN-APSO-LSSVM adopted an end-to-end algorithm structure that exhibited good universality and operability.

To analyze the influence of APSO algorithm on the prediction result, the sample data obtained by CO and CNN were input into the LSSVM model (CO-LSSVM and CNN-LSSVM, respectively) for gas-bearing prediction (Fig. 12). Both models yielded relatively accurate predictions because both feature optimization methods (CO and CNN) effectively extracted sensitive gas reservoir information and highlighted the gas reservoir characteristics. However, they showed larger errors compared with the results of the CO-APSO-LSSVM (Figs. 11(c) and (d)) and CNN-APSO-LSSVM (Fig. 8(e) and (f)) models with relatively unclear gas reservoir boundaries. The relationship between MSAs and gas-bearing is complicated, and high-quality input data are insufficient. An intelligent prediction model with a good performance is imperative to fully explore the relationship between them and obtain more accurate results. Optimizing the parameters of the LSSVM using the APSO algorithm

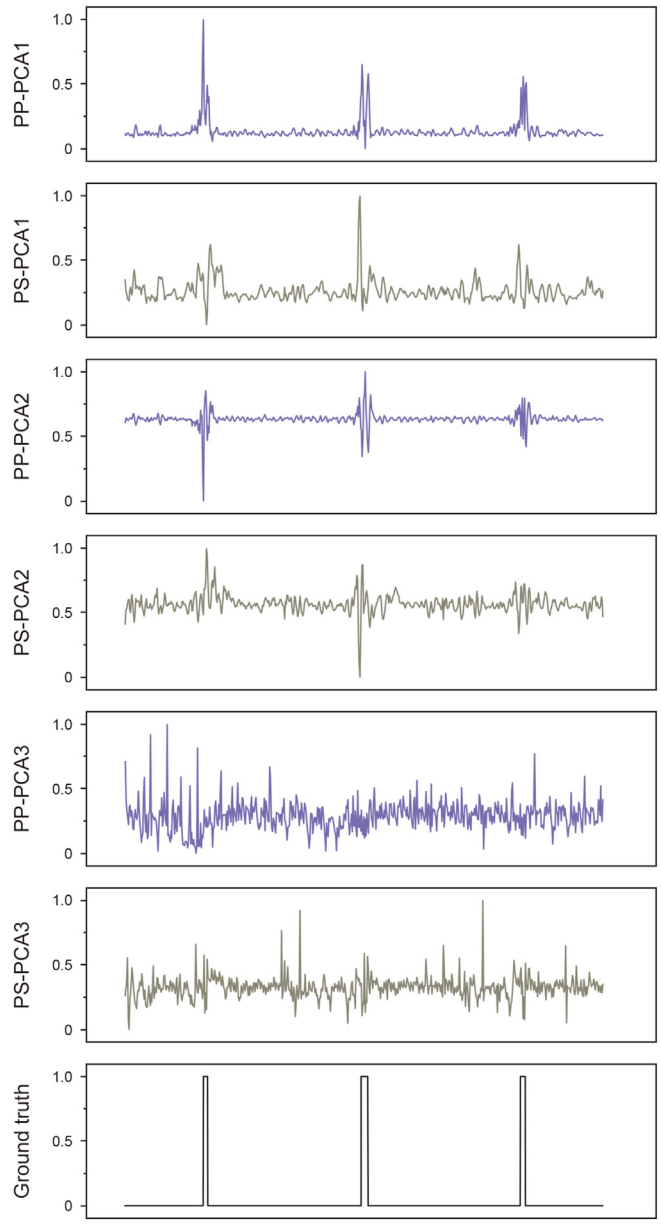

Fig. 9. Sample features for the three pseudo-wells in Fig. 4(d) optimized by PCA and their label information.

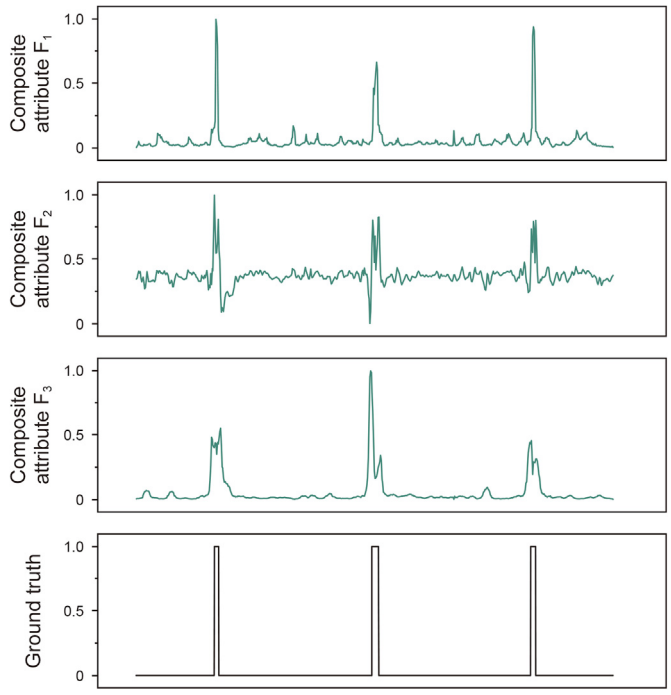

Fig. 10. Sample features for the three pseudo-wells in Fig. 4(d) optimized by CO and their label information.

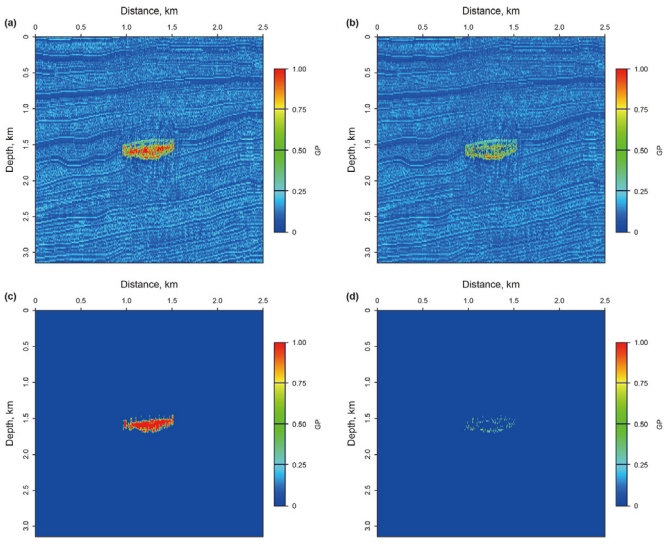

Fig. 11. Prediction results and errors of the PCA-APSO-LSSVM and CO-APSO-LSSVM models in the synthetic data. **(a)** Prediction result and **(b)** error of PCA-APSO-LSSVM; **(c)** prediction result and **(d)** error of CO-APSO-LSSVM.

can further improve its prediction and gas reservoir boundary characterization abilities, thereby improving its prediction accuracy and validating its effectiveness of the APSO algorithm. It was observed that even though the CNN-LSSVM model did not use APSO to optimize the relevant parameters during training, the obtained prediction results still had smaller errors than those of the CNN model, validating the effectiveness of integrating the two models into a hybrid fusion model. By optimizing the relevant parameters using the APSO algorithm, the hybrid fusion model performance was further improved, yielding better results.

4. Real data examples

4.1. Feature selection and dataset construction

We applied the proposed method to GPD prediction using multicomponent seismic data from the Fenggu structural area of the Western Sichuan Depression (WSD) in China. Detailed geological conditions of the WSD are available in Duan et al. (2023) and Zhang et al. (2022a). We selected $T_3X_4^6$ member of the Xujiahe Formation as the target layer for GPD prediction. Detailed multicomponent seismic data for the target layer are available in studies by Zhang et al. (2022b) and Yang et al. (2023a).

Feature selection is an important step in the use of DL methods for feature extraction and prediction. There are many types of MSAs, and the seismic attributes that contribute significantly to gas-bearing predictions were first selected using feature selection. In this study, cluster analysis was used to select the important

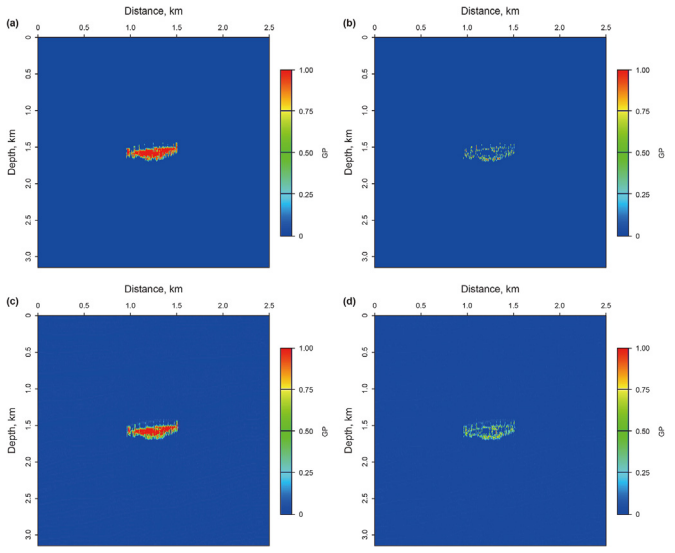

Fig. 12. Prediction results and errors of the CO-LSSVM and CNN-LSSVM models in the synthetic data. (**a**) Prediction result and (**b**) error of CO-LSSVM; (**c**) prediction result and (**d**) error of CNN-LSSVM.

attributes sensitive to gas reservoirs from the original MSAs. The six MSAs obtained through cluster analysis are shown in Fig. 13. All six MSAs provided different gas reservoir information to a certain extent. To obtain as much characteristic information of the target reservoir as possible, all six MSAs were used in the CNN-APSO-LSSVM for feature optimization and GPD prediction.

The target layer has a limited degree of exploration with only 11 wells, making it difficult to obtain more labels by relying solely on drilling data. Therefore, the following process was adopted: the GP of dry wells was set to 0, the GP of gas wells with the highest gas content was set to 1, and the GP of other gas wells was calculated proportionally through their gas content and ranged between 0 and 1. The GP around the well was determined using correlation calculations (Lin et al., 2018; Song et al., 2022). An average of approximately 20 subtrace sets were extracted around each well; each subtrace set had a length of 64 sampling points, the sliding time window had a length of 16 sampling points, and the sliding step length was one sampling point; approximately 960 sample points were extracted around each well. Using this method, additional sample data were obtained from known wells. This method can obtain the GP, rather than the gas-bearing classification. The obtained GP can contain more gas reservoir characteristic information, which aids the hybrid fusion model in GPD prediction.

4.2. Hybrid fusion model training and testing

The sample data provided by three wells (including two gas wells and one dry well) and their surrounding areas were used for testing, whereas sample data provided by the remaining eight wells and their surrounding areas were used for training. The ratio of the training to testing datasets was approximately 7:3 (Anifowose et al., 2016). For the training dataset, the training and validation subsets were divided by blind-well cross-validation, and the CNN model was trained to determine the parameters of the model. The CNN model for feature extraction was obtained after the errors satisfied the accuracy requirements. A new training dataset was built using the features extracted by CNN to train APSO-LSSVM model, wherein APSO was used to optimize the relevant parameters until the error met the accuracy requirements, ultimately obtaining the CNN-APSO-LSSVM model for predicting GPD. The detailed parameter settings of the CNN-APSO-LSSVM model are presented in Tables 1 and 2.

After training, the CNN-APSO-LSSVM model was applied to testing dataset to evaluate its effectiveness. To test the advantages of the hybrid fusion model over single models, we compared its prediction performance with those of CNN and APSO-LSSVM models (Fig. 14). Error analysis of the testing dataset showed that CNN-APSO-LSSVM improved the prediction performance of CNN and APSO-LSSVM. For the predicted GP of CNN-APSO-LSSVM

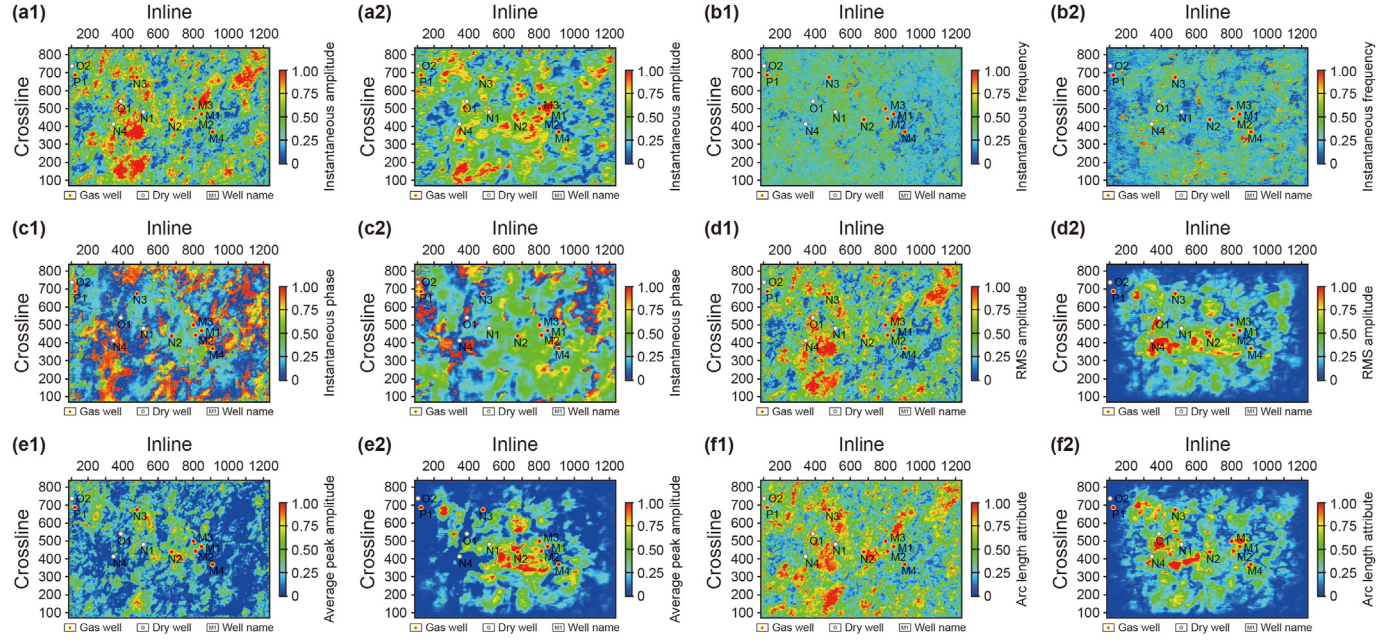

Fig. 13. Selected original MSAs sensitive to gas reservoirs in the real data: (**a1**) and (**a2**) instantaneous amplitude, (**b1**) and (**b2**) instantaneous frequency, (**c1**) and (**c2**) instantaneous phase, (**d1**) and (**d2**) RMS amplitude attribute, (**e1**) and (**e2**) average peak amplitude, and (**f1**) and (**f2**) arc length attribute. ((**a1**)–(**f1**) show longitudinal waves and (**a2**)–(**f2**) show converted shear waves).

model, 46% of the test data points had a relative error of <5%, compared to 37% and 28% for the CNN and APSO-LSSVM models, respectively. Moreover, CNN and APSO-LSSVM models had 2–4 times more data points with relative errors >20% than the CNN-APSO-LSSVM model. This proves the superior prediction accuracy of the CNN-APSO-LSSVM compared to the CNN and APSO-LSSVM models.

4.3. Prediction results and performance analysis

The GPD prediction results of the APSO-LSSVM (Fig. 15 (a)) show a poor gas reservoir boundary and inaccurate prediction of gas wells. This is because the six original MSAs were directly inputted into the APSO-LSSVM for GPD prediction without feature optimization, which seriously affected the learning ability of the CML model, making it impossible to fully utilize its advantages to explore the intrinsic characteristics of the input samples, thereby reducing its prediction ability. Feature optimization is a crucial step in the CML model for make non-linear and successful predictions for practical applications. Therefore, it is important to obtain sample data with clear characteristics by optimizing the features of the original MSAs for gas-bearing predictions. Compared with the original MSAs (Fig. 13), the CNN model fully extracted the gas reservoir characteristics and achieved better GPD results (Fig. 15 (b)). However, considerable redundant information were observed in the predicted results. Nevertheless, the results were considerably superior to those of the APSO-LSSVM model, because compared with the latter, the CNN model could better learn the distribution characteristics and internal relationships of the samples owing to the existence of convolutional layers. The prediction results of CNN-APSO-LSSVM (Fig. 15(c)) further highlight the gas-bearing information compared to those of the CNN (Fig. 15(b)). The gas reservoir boundaries were clearly delineated, and the results better reflected the variation characteristics of the gas reservoirs. Moreover, among the three models, the gas reservoir information described by the proposed CNN-APSO-LSSVM was similar to that described by the CNN, indicating its considerable influence on the proposed hybrid fusion model. Although APSO-LSSVM has the advantage of a small sample size, it requires high-quality sample data. Feature optimization of MSA using a CNN improves the quality of sample data and the prediction performance of APSO-LSSVM.

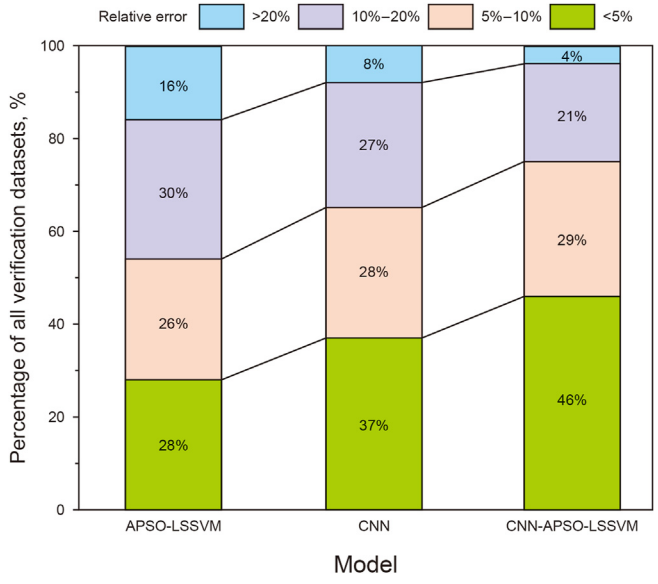

Fig. 14. Stacked histogram of GP prediction with relative error distribution of the APSO-LSSVM, CNN, and CNN-APSO-LSSVM models for the testing dataset.

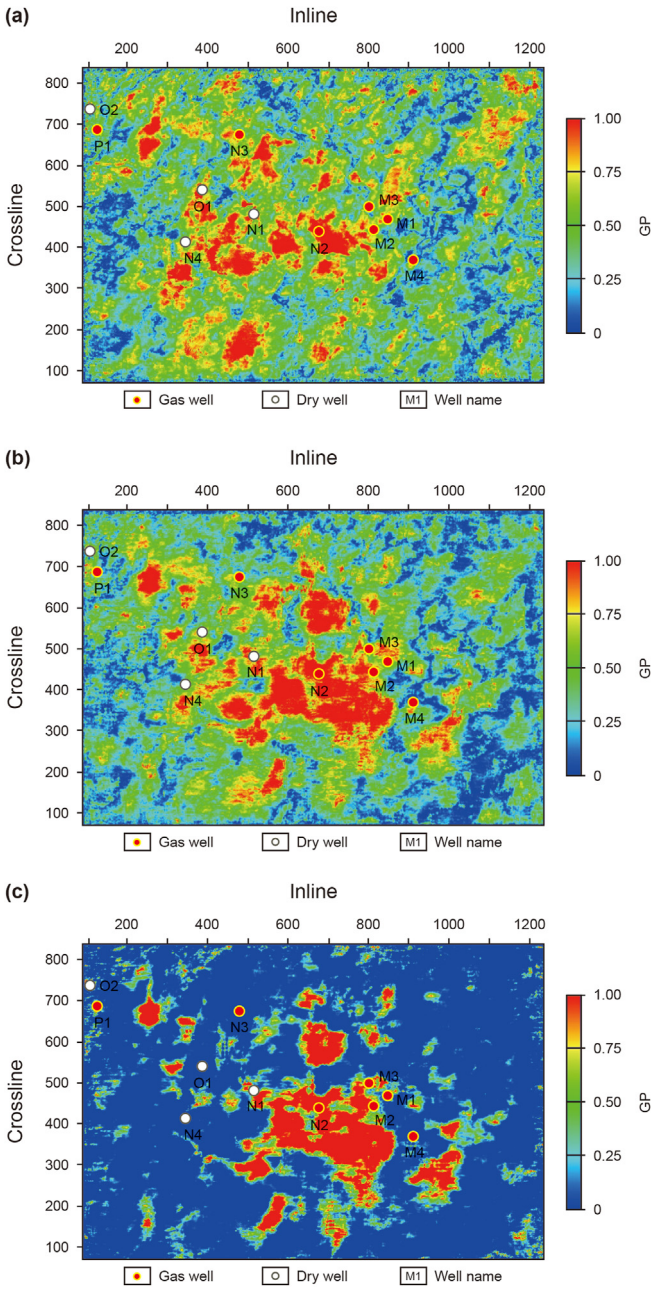

Fig. 15. GPD prediction results of the **(a)** APSO-LSSVM, **(b)** CNN, and **(c)** CNN-APSO-LSSVM models.

5. Discussion

5.1. Comparison of different feature optimization methods

The CNN-APSO-LSSVM model constructed in this study includes two processes: feature optimization and intelligent prediction. In the feature optimization process, we used a CNN for feature extraction of MSAs. To analyze the feature extraction capability of CNN, the prediction results of CNN-APSO-LSSVM were compared with those obtained using PCA-APSO-LSSVM and CO-APSO-LSSVM (Fig. 16). The prediction results obtained by the former (Fig. 16(a))

were better than those obtained by the APSO-LSSVM model using the original MSAs (Fig. 15(a)), with slightly improved consistency with actual drilling data; however, the improvement was not significant. The prediction results of CO-APSO-LSSVM (Fig. 16(b)) improved considerably compared with those of APSO-LSSVM and PCA-APSO-LSSVM, with abundant gas information. CNN-APSO-LSSVM (Fig. 15(c)) depicts the gas reservoir boundary more clearly and improves the prediction accuracy compared to PCA-APSO-LSSVM and CO-APSO-LSSVM, indicating the effectiveness of CNN for the feature optimization of sample data under real conditions compared to the two feature optimization methods, PCA and CO. For the synthetic data, the prediction results of CO-APSO-LSSVM and CNN-APSO-LSSVM were similar; however, for the real data, significant differences were observed in their prediction results. The CNN-APSO-LSSVM showed better gas reservoir boundary characterization ability than CO-APSO-LSSVM, indicating that CNN has better feature extraction ability than the CO feature optimization method, further highlighting the advantages of CNN. When applied to complex data (real data), a CNN can still use its powerful advantages to directly extract features and simplify feature optimization while obtaining more effective gas reservoir information.

5.2. Performance analysis of intelligent prediction models

By comparing the prediction results of different feature optimization methods, we verified the advantages of the CNN model for feature optimization. The APSO algorithm was adopted for the intelligent prediction process of the hybrid fusion model. We further analyzed the impact of APSO algorithm on the performance of LSSVM model (Fig. 17). The intelligent prediction model optimized by APSO has faster error reduction and model convergence than that optimized by PSO during training. This indicates that the APSO optimization algorithm improved the optimization ability of PSO, enabling faster and more accurate optimization of LSSVM.

We compared the performance of the prediction models obtained by entering the features extracted by CNN into the unoptimized LSSVM model (CNN-LSSVM), the LSSVM model optimized by PSO (CNN–PSO-LSSVM), and the LSSVM model optimized by APSO (CNN-APSO-LSSVM) (Table 3). The prediction results of LSSVM optimized using APSO (CNN-APSO-LSSVM) have a higher goodness of fit ($R^2 = 0.9739$) than those of the LSSVM optimized using PSO (PSO-LSSVM) ($R^2 = 0.9717$), indicating that the APSO-LSSVM model exhibits better predictive performance than the PSO-LSSVM on the testing dataset. Combining their performance on the training dataset (Fig. 17), although the PSO-LSSVM also reached a lower error level during training, it used more iterations, and its test performance was inferior to that of the APSO-LSSVM. This indicates that, compared with the PSO algorithm, the APSO algorithm can better perform global optimization during training to obtain the optimal model parameters, and the obtained intelligent prediction model performs better.

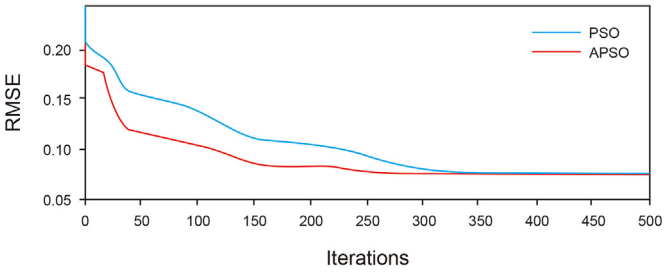

Fig. 17. Performance comparison of LSSVM optimization using PSO and APSO algorithms on training dataset.

Table 3
Performance comparison of CNN-APSO-LSSVM, PSO-LSSVM and LSSVM on the testing dataset.

Model	Performance indicators		
	MSE	RMSE	R^2
CNN-LSSVM	0.0109	0.1042	0.9693
CNN–PSO-LSSVM	0.0099	0.0995	0.9717
CNN-APSO-LSSVM	0.0095	0.0976	0.9739

In addition, compared to the CNN-LSSVM model, which was not optimized using intelligent algorithms (MSE = 0.0109, RMSE = 0.1042, $R^2 = 0.9693$), the CNN-APSO-LSSVM model had lower error (MSE = 0.0095, RMSE = 0.0976) and higher fit ($R^2 = 0.9739$), significantly improving the prediction performance. For further analysis, we input the sample features extracted from the CO and CNN optimizations into the LSSVM model (without APSO algorithm optimization) for GPD prediction (Fig. 18). The prediction results showed unclear gas reservoir boundaries; in particular, the GP obtained by the CO-LSSVM (Fig. 18(a)) was lower than that obtained by the CO-APSO-LSSVM (Fig. 16(b)). Parameter optimization of the LSSVM using the APSO algorithm improved its prediction ability, and the obtained prediction results further highlighted the gas reservoir characteristics. The features of the MSAs obtained by the CO differed from those extracted by the CNN model. When using the two different feature datasets for GPD prediction, the prediction results obtained using the APSO-optimized intelligent models (CO-APSO-LSSVM and CNN-APSO-LSSVM) were better than those obtained without optimization (CO-LSSVM and CNN-LSSVM), verifying the effectiveness of the APSO-LSSVM.

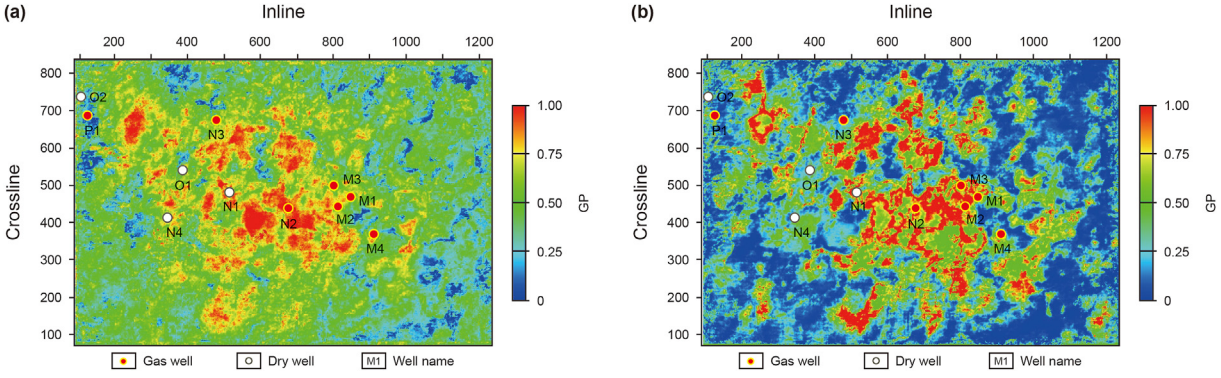

Fig. 16. GPD prediction results of the **(a)** PCA-APSO-LSSVM and **(b)** CO-APSO-LSSVM models.

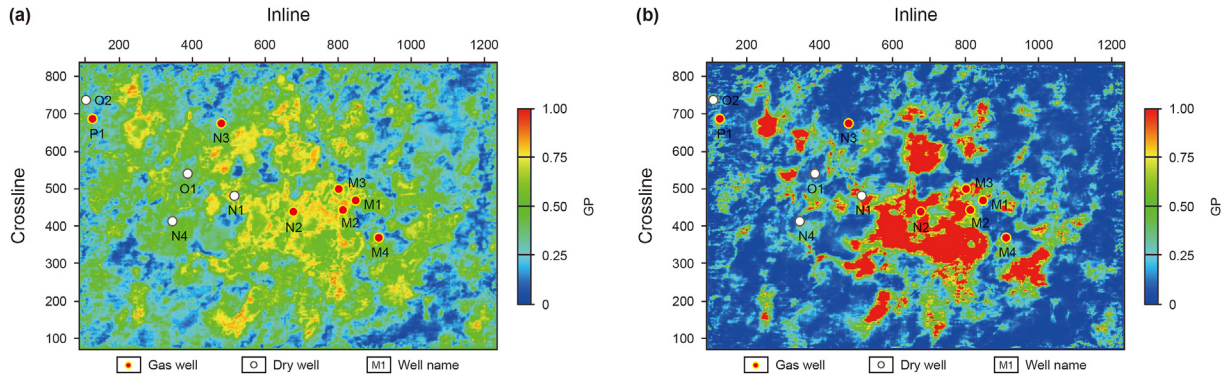

Fig. 18. GPD prediction results of the **(a)** CO-LSSVM and **(b)** CNN-LSSVM models.

5.3. Comprehensive evaluation of hybrid fusion model performance

We further analyzed the coincidence between the prediction results of hybrid models and actual drilling (Table 4), which showed poor consistency for the APSO-LSSVM model, with an overall coincidence of 6/11. The coincidence of PCA-APSO-LSSVM improved slightly, mainly in the dry well, indicating that redundant information was reduced after PCA optimization. The coincidence between the prediction results of CO-APSO-LSSVM and CNN-APSO-LSSVM was similar, indicating the effectiveness of CO and CNN in feature optimization and extraction, respectively. However, the direct feature extraction ability of CNN simplifies the complex process used by CO for feature optimization, proving that the former is more effective. Note that both models predict gas well N3 as a dry well because it has the lowest gas content of all gas wells; therefore, the GP obtained through correlation calculation is also the lowest, resulting in inaccurate predictions in both scenarios. Furthermore, the coincidence of CO-LSSVM was lower than that of CO-APSO-LSSVM, verifying the effectiveness of the APSO algorithm. Although the coincidence of the prediction results of CNN-LSSVM and CNN-APSO-LSSVM was similar, the distribution prediction results for the entire region showed that the CNN-APSO-LSSVM could depict a clearer gas reservoir boundary (Fig. 15(c)). Thus, the effectiveness of the DL method for feature extraction and the CML method for gas-bearing distribution prediction was demonstrated.

To analyze the prediction performances of the different hybrid models more intuitively, we compared their errors on the testing dataset (Table 5). The CNN-APSO-LSSVM model improved the prediction ability of both APSO-LSSVM and CNN with an increased R^2 of 10.51% and 3.20%, respectively. The fitting degree of the APSO-LSSVM was significantly improved, as CML models could successfully predict reservoir distribution, mainly using high-quality sample data. Compared to the PCA-APSO-LSSVM (MSE = 0.0337, RMSE = 0.1836, R^2 = 0.9283) and CO-APSO-LSSVM (MSE = 0.0097, RMSE = 0.0983, R^2 = 0.9731) models, the error of CNN-APSO-LSSVM model was smaller (MSE = 0.0095, RMSE = 0.0976), and the fitting degree was higher (R^2 = 0.9739). Although all three models used APSO-LSSVM for prediction, their prediction performances were different, and CNN-APSO-LSSVM had better prediction performance, demonstrating the advantage of CNN for feature extraction. The MSE and RMSE of the CO-APSO-LSSVM model were lower than those of the PCA-APSO-LSSVM model, indicating better feature optimization by CO than by PCA. However, CO is highly dependent on the reservoir characteristics and is not as applicable as PCA optimization. Moreover, compared to CNN-LSSVM and CO-LSSVM, APSO effectively improved the prediction ability of the LSSVM. The CO-APSO-LSSVM showed reduced MSE and RMSE by 14.16% and 7.44%, respectively, compared with the CO-LSSVM. The CNN-APSO-LSSVM showed reduced MSE and RMSE by 12.84% and 6.33%, respectively, compared with the CNN-LSSVM. The prediction model optimized by APSO has a lower error, further validating the efficiency of the CNN-APSO-LSSVM model.

Furthermore, the APSO-LSSVM and CNN models both used the six original attributes for GPD prediction, and CNN had a better predictive performance. Even with the PCA optimization of the

Table 5
Performance of different models for the testing dataset.

Model	Performance indicators		
	MSE	RMSE	R^2
APSO-LSSVM	0.0417	0.2042	0.8813
PCA-APSO-LSSVM	0.0337	0.1836	0.9283
CO-LSSVM	0.0113	0.1062	0.9678
CO-APSO-LSSVM	0.0097	0.0983	0.9731
CNN	0.0129	0.1136	0.9437
CNN-LSSVM	0.0109	0.1042	0.9693
CNN-APSO-LSSVM	0.0095	0.0976	0.9739

Table 4
Coincidence between prediction results of different hybrid models and actual drilling data (the inconsistent prediction results of each model are marked in bold).

Model	Well name											Coincidence		
	M1	M2	M3	M4	N1	N2	N3	N4	O1	O2	P1	Total	Gas	Dry
Actual	Gas	Gas	Gas	Gas	Dry	Gas	Gas	Dry	Dry	Dry	Gas			
APSO-LSSVM	Gas	Gas	**Dry**	**Dry**	Dry	Gas	Gas	**Gas**	**Gas**	Dry	**Dry**	6/11	4/7	2/4
PCA-APSO-LSSVM	Gas	Gas	Gas	**Dry**	**Gas**	Gas	**Dry**	Dry	Dry	Dry	**Dry**	7/11	4/7	3/4
CO-LSSVM	Gas	Gas	Gas	Gas	Dry	Gas	**Dry**	Dry	Dry	Dry	**Dry**	9/11	5/7	4/4
CO-APSO-LSSVM	Gas	Gas	Gas	Gas	Dry	Gas	**Dry**	Dry	Dry	Dry	Gas	10/11	6/7	4/4
CNN	Gas	Gas	Gas	Gas	**Gas**	Gas	Gas	Dry	**Gas**	Dry	Gas	9/11	6/7	3/4
CNN-LSSVM	Gas	Gas	Gas	Gas	Dry	Gas	**Dry**	Dry	Dry	Dry	Gas	10/11	6/7	4/4
CNN-APSO-LSSVM	Gas	Gas	Gas	Gas	Dry	Gas	**Dry**	Dry	Dry	Dry	Gas	10/11	6/7	4/4

input MSAs, the prediction results still exhibited large errors, indicating better feature extraction and prediction abilities of the CNN for the original data with unclear sample features, which can further improve prediction accuracy. These analyses indicate that the proposed CNN-APSO-LSSVM model has the highest fitting degree and lowest error, thereby proving its efficiency and feasibility.

5.4. Strengths and limitations of the proposed model

GDP prediction using multicomponent seismic data is challenging because of the complex and unclear relationships between them. The use of CML models for gas-bearing prediction requires optimization of the features of sample data to remove irrelevant and redundant information, highlighting the sample features that contribute significantly to the target data. In the feature optimization process of MSAs, commonly used methods require manual intervention. During this process, insufficient subjective awareness and experience may lead to the loss of effective information related to reservoir prediction or highlight redundant information unrelated to reservoir prediction, affecting the effect of feature optimization, and thus the prediction effect of CML models.

In this study, a CNN model was used to perform feature extraction instead of the aforementioned feature optimization process, and the CNN model was combined with the APSO-LSSVM model to build an end-to-end hybrid fusion model. The model directly inputs the features extracted by CNN into the CML model for training and prediction, thereby reducing manual intervention, simplifying the steps of feature optimization, and achieving excellent feature optimization results. Moreover, the parameter optimization of CML models during training is also an important issue. This study used the APSO algorithm to adaptively determine the optimal parameters of the LSSVM based on its powerful optimization ability. During the optimization process, the APSO algorithm adaptively adjusts the inertia weight and velocity coefficients, reduces the influence of improper parameter settings, and improves its optimization ability. In general, the CNN-APSO-LSSVM model constructed in this study improved the prediction accuracy of GPD from two aspects: feature optimization and intelligent prediction. The constructed model has adaptability and strong learning ability, and can automatically learn the features and rules of data during the training process, thereby simplifying the problem of manually setting features and constraints.

However, the proposed method has some limitations. When using a CNN model for feature extraction, there may be a risk of suppressing some key small-scale features while reducing redundant information. This is due to their inherent characteristics. A CNN extracts features using a local receptive field and weight sharing, which helps reduce redundant information. However, in some cases, this mechanism may neglect key small-scale features. To alleviate this problem, we carefully selected hyperparameters through blind-well cross-validation based on the characteristics of MSAs, and determined the network structure of CNN model through extensive experimental work. By adjusting network structure and model parameters, the risk of suppressing key small-scale features can be reduced to a certain extent. The results of this study also indicate that the CNN effectively reduces redundant information, and the CNN-APSO-LSSVM achieves an excellent prediction effect. However, this phenomenon cannot be disregarded. In future studies, we will consider introducing an attention mechanism into the CNN model to improve its ability to capture key small-scale features. In addition, the proposed APSO algorithm can effectively improve the prediction performance of LSSVM. In future studies, we plan to search for the key hyperparameters of the CNN to obtain a topology that can better match MSAs and further improve the feature extraction capability of CNN.

In gas-bearing predictions using seismic data, the quality of sample data is crucial for the construction and training of DL model. We hope to obtain abundant sample data to train DL model fully and improve its predictive ability. However, available sample data are often limited by the degree of exploration of the target area. In future studies, we will consider enhancing the diversity of the sample data and improving the generalizability of the developed model by generating synthetic data.

6. Conclusions

This study utilizes the advantages of CML and DL methods in the prediction of GPD using multicomponent seismic data, and develops a method that combines them for GPD prediction. This method improves prediction accuracy from the two aspects of feature optimization and intelligent prediction. In the feature optimization process, a CNN can fully utilize its powerful feature extraction capability to effectively extract gas reservoir characteristics contained in MSAs. During the prediction process, the APSO algorithm is used to optimize the LSSVM model, improving its gas-bearing prediction ability. By applying this method, the following conclusions were drawn:

(1) The GPD prediction performance of the constructed CNN-APSO-LSSVM model was superior to that of CNN and APSO-LSSVM models. The CNN and APSO-LSSVM demonstrated 2–4 times more test data points with relative errors >20% than that of the CNN-APSO-LSSVM. This method effectively reduces the error of a single model and improves the prediction performance of CML and DL models. In particular, the ability of CNN to extract deep-level attribute features improves the quality of sample data, which significantly improves the prediction ability of APSO-LSSVM.
(2) The feature optimization of MSAs using the CNN model can extract gas reservoir features better than those using the PCA and CO feature optimization methods. The extracted gas reservoir information was inputted into the APSO-LSSVM for GPD prediction, which achieved better prediction results than the PCA-APSO-LSSVM and CO-APSO-LSSVM models.
(3) The APSO algorithm improved the GPD prediction performance of LSSVM. The CNN-APSO-LSSVM reduced the MSE and RMSE by 12.84% and 6.33%, respectively, compared with the CNN-LSSVM. The CO-APSO-LSSVM reduced the MSE and RMSE by 14.16% and 7.44%, respectively, compared to the CO-LSSVM. Thus, the APSO algorithm improves the prediction performance of the LSSVM for different feature optimization sample data.

In general, this study demonstrates the feasibility of using DL technology for feature extraction and combining it with the CML method for gas-bearing prediction. This method reduces the interference of manual feature optimization and simplifies the steps of feature optimization. The constructed deep hybrid fusion model has excellent universality and operability. The effective implementation of this method can provide a reference for predicting gas-bearing distributions in other regions or datasets.

CRediT authorship contribution statement

Jiu-Qiang Yang: Conceptualization, Methodology, Resources, Validation, Writing – original draft. **Nian-Tian Lin:** Conceptualization, Data curation, Funding acquisition, Methodology, Project administration. **Kai Zhang:** Formal analysis, Funding acquisition, Methodology. **Yan Cui:** Formal analysis, Validation, Writing – review & editing. **Chao Fu:** Funding acquisition, Writing – review &

editing. **Dong Zhang:** Writing – review & editing.

Declaration of competing interest

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

Acknowledgements

We are grateful to the editors and reviewers for their thorough reading and many constructive comments on this paper. The research was funded by the Natural Science Foundation of Shandong Province (ZR2021MD061; ZR2023QD025), China Postdoctoral Science Foundation (2022M721972), National Natural Science Foundation of China (41174098), Young Talents Foundation of Inner Mongolia University (10000–23112101/055), and Qingdao Postdoctoral Science Foundation (QDBSH20230102094).

References

Anifowose, F., Khoukhi, A., Abdulraheem, A., 2016. Investigating the effect of training–testing data stratification on soft computing techniques: an experimental study. J. Exp. Theor. Artif. Intell. 29 (3), 517–535. https://doi.org/10.1080/0952813X.2016.1198936.

Babiki, I., Elsaadany, M., Sajid, M., et al., 2022. Evaluation of principal component analysis for reducing seismic attributes dimensions: Implication for supervised seismic facies classification of a fluvial reservoir from the Malay Basin, offshore Malaysia. J. Pet. Sci. Eng. 217, 110911. https://doi.org/10.1016/j.petrol.2022.110911.

Baghban, A., 2016. Application of the ANFIS strategy to estimate vaporization enthalpies of petroleum fractions and pure hydrocarbons. Petrol. Sci. Technol. 34 (15), 1359–1366. https://doi.org/10.1080/10916466.2016.1202975.

Baghban, A., Adelizadeh, M., 2018. On the determination of cetane number of hydrocarbons and oxygenates using Adaptive Neuro Fuzzy Inference System optimized with evolutionary algorithms. Fuel 230, 344–354. https://doi.org/10.1016/j.fuel.2018.05.032.

Bahadori, A., Baghban, A., Bahadori, M., et al., 2016. Computational intelligent strategies to predict energy conservation benefits in excess air controlled gas-fired systems. Appl. Therm. Eng. 102, 432–446. https://doi.org/10.1016/j.applthermaleng.2016.04.005.

Bemani, A., Baghban, A., Mohammadi, A.H., 2020a. An insight into the modeling of sulfur content of sour gases in supercritical region. J. Pet. Sci. Eng. 184, 106459. https://doi.org/10.1016/j.petrol.2019.106459.

Bemani, A., Baghban, A., Mosavi, A., et al., 2020b. Estimating CO_2-Brine diffusivity using hybrid models of ANFIS and evolutionary algorithms. Eng. Appl. Comput. Fluid Mech. 14 (1), 818–834. https://doi.org/10.1080/19942060.2020.1774422.

Brantson, E.T., Ju, B.S., Ziggah, Y.Y., et al., 2019. Forecasting of horizontal gas well production decline in unconventional reservoirs using productivity, soft computing and swarm intelligence models. Nat. Resour. Res. 28 (3), 717–756. https://doi.org/10.1007/s11053-018-9415-2.

Cao, W., Guo, X.B., Tian, F., et al., 2021. Seismic velocity inversion based on CNN-LSTM fusion deep neural network. Appl. Geophys. 18 (4), 499–514. https://doi.org/10.1007/s11770-021-0913-3.

Chaitanya, K., Somayajulu, D.V.L.N., Krishna, P.R., 2021. Memory-based approaches for eliminating premature convergence in particle swarm optimization. Appl. Intell. 51, 4575–4608. https://doi.org/10.1007/s10489-020-02045-z.

Dewett, D.T., Pigott, J.D., Marfurt, K.J., 2021. A review of seismic attribute taxonomies, discussion of their historical use, and presentation of a seismic attribute communication framework using data analysis concepts. Interpretation 9 (3), B39–B64. https://doi.org/10.1190/INT-2020-0222.1.

Dong, S.Q., Sun, Y.M., Xu, T., et al., 2023. How to improve machine learning models for lithofacies identification by practical and novel ensemble strategy and principles. Petrol. Sci. 20 (2), 733–752. https://doi.org/10.1016/j.petsci.2022.09.006.

Duan, Z.X., Liu, Y.F., Lou, Z.H., et al., 2023. Tight gas accumulation caused by overpressure: insights from three-dimensional seismic data in the western Sichuan Basin, southwest China. Geoenergy Sci. Eng. 223, 211589. https://doi.org/10.1016/j.geoen.2023.211589.

Fang, Z.C., Wang, Y., Peng, L., et al., 2020. Integration of convolutional neural network and conventional machine learning classifiers for landslide susceptibility mapping. Comput. Geosci. 139, 104470. https://doi.org/10.1016/j.cageo.2020.104470.

Gheytanzadeh, M., Baghban, A., Habibzadeh, S., et al., 2021. Insights into the estimation of capacitance for carbon-based supercapacitors. RSC Adv. 11 (10), 5479–5486. https://doi.org/10.1039/d0ra09837j.

Guyon, I., Gunn, S., Nikravesh, M., et al., 2006. Feature Extraction: Foundations and Applications. Springer Berlin, Heidelberg.

Han, H.G., Lu, W., Hou, Y., 2018. An adaptive-PSO-based self-organizing RBF neural network. IEEE Transact. Neural Networks Learn. Syst. 29 (1), 104–117. https://doi.org/10.1109/TNNLS.2016.2616413.

He, B.S., Yao, X.R., Shao, X.Q., 2022. Source-free P-SV converted-wave reverse-time migration using first-order velocity-dilatation-rotation equations. Front. Earth Sci. 10, 749462. https://doi.org/10.3389/feart.2022.749462.

Hossain, S., 2020. Application of seismic attribute analysis in fluvial seismic geomorphology. J. Pet. Explor. Prod. Technol. 10 (1), 1009–1019. https://doi.org/10.1007/s13202-019-00809-z.

Kalam, S., Yousuf, U., Abu-Khamsin, S.A., et al., 2022. An ANN model to predict oil recovery from a 5-spot waterflood of a heterogeneous reservoir. J. Pet. Sci. Eng. 210, 110012. https://doi.org/10.1016/j.petrol.2021.110012.

Kardani, M.N., Baghban, A., Sasanipour, J., et al., 2018. Group contribution methods for estimating CO2 absorption capacities of imidazolium and ammonium-based polyionic liquids. J. Clean. Prod. 203, 601–618. https://doi.org/10.1016/j.jclepro.2018.08.127.

Karimpouli, S., Tahmasebi, P., Saenger, E.H., 2020. Coal cleat/fracture segmentation using convolutional neural networks. Nat. Resour. Res. 29 (3), 1675–1685. https://doi.org/10.1007/s11053-019-09536-y.

Kiranyaz, S., Avci, O., Abdeljaber, O., et al., 2021. 1D convolutional neural networks and applications: a survey. Mech. Syst. Signal Process. 151, 107398. https://doi.org/10.1016/j.ymssp.2020.107398.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. Adv. Neural Inf. Process. Syst. 25 (2), 1097–1105.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539.

Lin, N.T., Zhang, D., Zhang, K., et al., 2018. Predicting distribution of hydrocarbon reservoirs with seismic data based on learning of the small-sample convolution neural network. Chin. J. Geophys. 61 (10), 4110–4125. https://doi.org/10.6038/cjg2018J0775 (in Chinese).

Lou, Y.H., Li, S.Z., Li, S.J., et al., 2022. Seismic volumetric dip estimation via multi-channel deep learning model. IEEE Trans. Geosci. Rem. Sens. 60, 4511014. https://doi.org/10.1109/TGRS.2022.3190911.

Lubo-Robles, D., Marfurt, K.J., 2019. Independent component analysis for reservoir geomorphology and unsupervised seismic facies classification in the Taranaki Basin, New Zealand. Interpretation 7 (3), SE19–SE42. https://doi.org/10.1190/INT-2018-0109.1.

Ma, X., Yao, G., Zhang, F., et al., 2023. 3-D Seismic fault detection using recurrent convolutional neural networks with compound loss. IEEE Trans. Geosci. Rem. Sens. 61, 5909815. https://doi.org/10.1109/TGRS.2023.3275951.

Martin, G.S., Wiley, R., Marfurt, K.J., 2006. Marmousi2: an elastic upgrade for Marmousi. Lead. Edge 25 (2), 113–224. https://doi.org/10.1190/1.2172306.

Moosavi, N., Bagheri, M., Nabi-Bidhendi, M., et al., 2022. Fuzzy support vector regression for permeability estimation of petroleum reservoir using well logs. Acta Geophys. 70, 161–172. https://doi.org/10.1007/s11600-021-00700-8.

Nabipour, N., Mosavi, A., Baghban, A., et al., 2020. Extreme learning machine-based model for solubility estimation of hydrocarbon gases in electrolyte solutions. Processes 8 (1), 92. https://doi.org/10.3390/pr8010092.

Nguyen, H., Bui, H.B., Bui, X.N., 2021. Rapid determination of gross calorific value of coal using artificial neural network and particle swarm optimization. Nat. Resour. Res. 30 (1), 621–638. https://doi.org/10.1007/s11053-020-09727-y.

Otchere, D.A., Ganat, T.O.A., Ojero, J.O., et al., 2022. Application of gradient boosting regression model for the evaluation of feature selection techniques in improving reservoir characterisation predictions. J. Pet. Sci. Eng. 208 (part E), 109244. https://doi.org/10.1016/j.petrol.2021.109244.

Sen, S., Abioui, M., Ganguli, S.S., et al., 2021. Petrophysical heterogeneity of the early Cretaceous Alamein dolomite reservoir from North Razzak oil field, Egypt integrating well logs, core measurements, and machine learning approach. Fuel 306, 121698. https://doi.org/10.1016/j.fuel.2021.121698.

Seyyedattar, M., Zendehboudi, S., Butt, S., 2022. Relative permeability modeling using extra trees, ANFIS, and hybrid LSSVM-CSA Methods. Nat. Resour. Res. 31 (1), 571–600. https://doi.org/10.1007/s11053-021-09950-1.

Shao, R.B., Xiao, L.Z., Liao, G.Z., et al., 2022. Multitask learning based reservoir parameters prediction with geophysical logs. Chin. J. Geophys. 65 (5), 1883–1895. https://doi.org/10.6038/cjg2022P0177 (in Chinese).

Song, Z.H., Yuan, S.Y., Li, Z.M., et al., 2022. KNN-based gas-bearing prediction using local waveform similarity gas-indication attribute—an application to a tight sandstone reservoir. Interpretation 10 (1), SA25–SA33. https://doi.org/10.1190/INT-2021-0045.1.

Srivastava, N., Hinton, G., Krizhevsky, A., et al., 2014. Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15, 1929–1958.

Sun, J., Innanen, K.A., Huang, C., 2021. Physics-guided deep learning for seismic inversion with hybrid training and uncertainty analysis. Geophysics 86, R303–R317. https://doi.org/10.1190/geo2020-0312.1.

Sun, J., Niu, Z., Innanen, K.A., et al., 2020. A theory-guided deep-learning formulation and optimization of seismic waveform inversion. Geophysics 85, R87–R99. https://doi.org/10.1190/geo2019-0138.1.

Syarif, I., Prugel-Bennett, A., Wills, G., 2016. SVM parameter optimization using grid search and genetic algorithm to improve classification performance. Telkomnika 14 (4), 1502–1509. https://doi.org/10.12928/telkomnika.v14i4.3956.

Taherkhani, M., Safabakhsh, R., 2016. A novel stability-based adaptive inertiaweight for particle swarm optimization. Appl. Soft Comput. 38, 281–295. https://doi.org/10.1016/j.asoc.2015.10.004.

Tran, H., Kasha, A., Sakhaee-Pour, A., et al., 2020. Predicting carbonate formation permeability using machine learning. J. Pet. Sci. Eng. 195, 107581. https://

doi.org/10.1016/j.petrol.2020.107581.
Wang, J., Cao, J.X., 2021. Data-driven S-wave velocity prediction method via a deep-learning-based deep convolutional gated recurrent unit fusion network. Geophysics 86 (6). https://doi.org/10.1190/geo2020-0886.1, 1ND–Z3.
Wang, P., Chen, X.H., Wang, B.F., et al., 2020. An improved method for lithology identification based on a hidden Markov model and random forests. Geophysics 85 (6), IM27–IM36. https://doi.org/10.1190/GEO2020-0108.1.
Wang, Y.Q., Wang, Q., Lu, W.K., et al., 2022. Seismic impedance inversion based on cycle-consistent generative adversarial network. Petrol. Sci. 19 (1), 147–161. https://doi.org/10.1016/j.petsci.2021.09.038.
Xie, W., Wang, Y.C., Liu, X.Q., et al., 2019. Nonlinear joint PP-PS AVO inversion based on improved Bayesian inference and LSSVM. Appl. Geophys. 16 (1), 64–76. https://doi.org/10.1007/s11770-019-0750-9.
Xue, Y.R., Guo, M.J., Feng, L.Y., et al., 2022. High resolution Radon transform inversion based on one dimensional convolutional neural network. Chin. J. Geophys. 65 (9), 3610–3622. https://doi.org/10.6038/cjg2022P0350 (in Chinese).
Yang, J.Q., Lin, N.T., Zhang, K., et al., 2023a. A data-driven workflow based on multisource transfer machine learning for gas-bearing probability distribution prediction: a case study. Geophysics 88, B163–B177. https://doi.org/10.1190/geo2022-0726.1.
Yang, J.Q., Lin, N.T., Zhang, K., et al., 2023b. An improved small-sample method based on APSO-LSSVM for gas-bearing probability distribution prediction from multicomponent seismic data. Geosci. Rem. Sens. Lett. IEEE 20, 7501705. https://doi.org/10.1109/LGRS.2023.3254531.
Yang, J.Q., Lin, N.T., Zhang, K., et al., 2021. Reservoir characterization using multi-component seismic data in a novel hybrid model based on clustering and deep neural network. Nat. Resour. Res. 30, 3429–3454. https://doi.org/10.1007/s11053-021-09863-z.
Yao, Z.Q., Yang, F., Jianatayi, D., et al., 2022. Application of multi-attribute matching technology based on geological models for sedimentary facies: a case study of the 3rd member in the Lower Jurassic Badaowan Formation, Hongshanzui area, Junggar Basin, China. Petrol. Sci. 19 (1), 116–127. https://doi.org/10.1016/j.petsci.2021.10.008.
Yuan, S.Y., Liu, J.W., Wang, S.X., et al., 2018. Seismic waveform classification and first-break picking using convolution neural networks. Geosci. Rem. Sens. Lett. IEEE 15 (2), 272–276. https://doi.org/10.1109/LGRS.2017.2785834.
Yuan, Y., Liu, Y., Zhang, J.Y., et al., 2011. Reservoir prediction using multi-wave seismic attributes. Earthq. Sci. 24, 373–389. https://doi.org/10.1007/s11589-011-0800-8.
Zarei, F., Baghban, A., 2017. Phase behavior modelling of asphaltene precipitation utilizing MLP-ANN approach. Petrol. Sci. Technol. 35 (20), 2009–2015. https://doi.org/10.1080/10916466.2017.1377233.
Zhang, G.Z., Yang, R., Zhou, Y., et al., 2022a. Seismic fracture characterization in tight sand reservoirs: a case study of the Xujiahe Formation, Sichuan Basin, China. J. Appl. Phys. 203, 104690. https://doi.org/10.1016/j.jappgeo.2022.104690.
Zhang, K., Lin, N.T., Fu, C., et al., 2019. Reservoir characterization method with multi-component seismic data by unsupervised learning and colour feature blending. Explor. Geophys. 50 (3), 269–280. https://doi.org/10.1080/08123985.2019.1603078.
Zhang, K., Lin, N.T., Yang, J.Q., et al., 2022b. Predicting gas bearing distribution using DNN based on multi-component seismic data: a reservoir quality evaluation using structural and fracture evaluation factors. Petrol. Sci. 19 (4), 1566–1581. https://doi.org/10.1016/j.petsci.2022.02.008.
Zhang, Y., Yao, L.Z., Zhang, L., et al., 2022c. Fault diagnosis of natural gas pipeline leakage based on 1D-CNN and self-attention mechanism. In: IEEE 6th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC). https://doi.org/10.1109/IAEAC54830.2022.9930063.
Zhumabekov, A., Liu, Z., Portnov, V., et al., 2021. Integrating the geology, seismic attributes, and production of reservoirs to adjust interwell areas: a case from the Mangyshlak Basin of West Kazakhstan. Appl. Geophys. 18 (3), 420–430. https://doi.org/10.1007/s11770-021-0907-1.
Zou, C.F., Zhao, L.X., Hong, F., et al., 2023. A comparison of machine learning methods to predict porosity in carbonate reservoirs from seismic-derived elastic properties. Geophysics 88 (2), B101–B120. https://doi.org/10.1190/GEO2021-0342.1.