Introduction

The tight sandstone reservoir has the characteristics of low porosity, low permeability and strong heterogeneity. The relationship between the reservoir and its seismic response is complex and nonlinear. A common method of gas-bearing prediction is to extract gas-sensitive seismic attributes from seismic data (Goloshubin, et al., 2006). However, there is no one-to-one relationship between the gas-bearing property and the seismic attributes. Seismic attenuation response which is usually considered to be gas-bearing characteristic of seismic response, can be caused by thin layer interference (Yuan, et al., 2017). Another common gas-bearing prediction method is to obtain the elastic parameters which can also be further converted to fluid indicator by pre-stack inversion. However, the accuracy of the gas-bearing prediction methods driven by a certain model is affected by the accuracy of the model itself. Linear simplification of wave equation can cause long-offset seismic information which is gas-sensitive loss (Goodway, et al., 1997). Artificial neural network (ANN) methods have great potential for solving the gas-bearing prediction problem because its ability to uncover the complex nonlinear relationship by establishing a nonlinear network model (LeCun, et al., 2015). ANN methods are imperfect in interpretability. In addition, the ANN-based gas-bearing prediction rely on big data, which means a large number of samples and labels are needed to train the network.

An improved method according to the original kNN (Cover and Hart, 2003) that can output gas-bearing probability is proposed in this study. It can predict gas-bearing distribution based on sample database made from borehole-side seismic traces and the corresponding interpreted gas-bearing curves obtained from the measured well logs. The method has a simple principle and strong interpretability, as well as no need for large samples to train the network. The method outputs the gas-bearing probability instead of gas-bearing property classification, which is more adaptive to gas-bearing prediction. It can also avoid the gas-bearing information loss during the prediction process.

Method

First, use a sliding time window to divide pre-stack seismic data from borehole-side traces into samples. The samples’ categories are determined by the corresponding gas-bearing curves. All the samples form a sample database. When the method is applied to prediction, divide the data to be predicted into patches which are in the same size as the samples. Measure the similarity between the patches to be predicted and each sample from the sample database with a given similarity evaluation criteria. Here, we employ the L1 distance as similarity evaluation criteria, which can be described as follows:

$$R^{(i)} = \text{SUM} \left( d^{(i)}_{\text{sample}} - d_{\text{predict}} \right),$$

where $R^{(i)}$ represents the similarity between the patch to be predicted and the $i^{th}$ sample, $\text{SUM}(\bullet)$ represents the summation operation, $d^{(i)}_{\text{sample}}$ represents the $i^{th}$ sample from the database, and $d_{\text{predict}}$ represents the patch to be predicted. Then output the gas-bearing probability approximated by the occurrence frequency of the gas-bearing samples in the first $k$ samples which are the most similar to the patch to be predicted according to the similarity evaluation criteria. The workflow of the method is shown in Figure 1. The idea that outputting the probability results instead of classification results has the potential of extension to multiclassification problem.
Examples

First, a numerical model whose structure is obtained from the MarmousiII model filled with petrophysical parameters extracted from the real work area is made to test the method. The gas-bearing profile of the numerical model is shown in Figure 2a. The yellow parts represent tight sandstone gas reservoir, and the blue parts represent gas-free mudstone. The synthetic per-stack gathers of the model is generated based on the Kuster-Toksoz (KT) model of petrophysics and Aki-Richard approximate formula, as shown in Figure 2b. The dominant frequency of the used seismic wavelet is 35 Hz. There are 24 offsets ranging from 500 m to 4100 m. The 82nd, 114th, 174th and 294th trace is randomly selected as pseudo-wells which are marked by the white dashed lines in Figure 2a. When the method is tested on the numerical model, the length of the sliding time window is 16 sampling points, while the width which is determined by the number of offsets is 24 sampling points, and the sliding step is 1 sampling point. When $k$ equals 5, the result of kNN which output gas-bearing property classification is shown in Figure 3a. Accurate reservoir location can only be characterized near the wells, which leads to the significant loss of the gas-bearing information. Therefore, it cannot obtain the complete reservoir morphology. It is because that the kNN method output the classification result depending on the category of the most samples in the first $k$ samples which are the most similar to the data to be predicted. It could lead to valid information loss. The subsurface space without gas usually larger than the reservoir volume, which could exacerbate the problem. The method solves the problem by changing the output of the kNN method into gas-bearing probability according to the characteristics of the gas-bearing prediction. The result of the method is shown in Figure 3b. The shape and location of the reservoir are characterized accurately, which will help further study on the reservoir. Besides, the gas-bearing probability is approximated by the occurrence frequency of the gas-bearing samples in the first $k$ samples, which are the most similar to data to be predicted which makes the analysis on the predicted result more conveniently. The red circles mark out the bigger probability of the gas-bearing appears in the center of the reservoir or near the wells, which proves the rationality of the approximation. The method is tested on noisy seismic data with 10% random noise. The result is shown in Figure 3c. The gas-bearing prediction probability on the profile is reduced slightly. However, the shape and location of the reservoir are still clear, which demonstrates that the method is robust in the presence of noise.
Figure 2 Synthetic data test. (a) Gas-bearing profile, (b) synthetic pre-stack seismic gathers with 337 CDPs and 24 offsets from 500 m to 4100 m.

Figure 3 Comparisons among gas-bearing results. (a) Gas-bearing property classification result, (b) gas-bearing probability result, (c) gas-bearing probability result from the test on noisy seismic data with 10% random noise.

A field data from northern China including multiple logging data is chosen to test the application potential of the method. The field data is a well-through section with 735 CDPs. Its pre-stack seismic gathers with 735 CDPs is shown in Figure 4a. The sampling interval is 1 ms, and there are 16 offsets in each gather ranging from 500 m to 4100 m. The field data also contains two horizons and five wells. The target reservoir of the work area is tight sandstone reservoir. The well-through post-stack seismic profile is shown in Figure 4b. Red fold lines represent the gas-bearing curves interpreted from five wells. Blue curves represent horizons L1 and L2. The favorable area of the reservoirs development is dispersed between horizons L1 and L2. Post-stack seismic profile shows that there is only one wave trough between two horizons. When \( k \) equals 5, the predicted result is shown in Figure 4c. The kNN-based gas-bearing prediction distribution has a good match with the gas-bearing curves obtained from well-log interpretation. What is more, it has a certain continuity near the wells’ location. The prediction result also indicates that there is more gas dispersed around horizon L2, which is a known favorable area of reservoir development. There is few gas above the horizon L1, since cover layer distribute around horizon L1. The kNN-based predicted result is basically consistent with the existing geological law of the work area, which demonstrates that the proposed method has a certain rationality.
Conclusions

In this paper, the kNN-based method is applied to predict gas-bearing distribution in tight sandstone reservoir. The kNN method has a simple principle and strong interpretability. No network needs to be trained and therefore, it does not rely on big data. According to the characteristics of gas-bearing prediction problem, kNN method is improved by outputting gas-bearing probability instead of gas-bearing property classification results. It can reserve the valid information. Outputting probability result can also make the further reservoir analysis convenient. Testing on synthetic data demonstrates that the method has good ability to characterize the location and shape of tight sandstone reservoir. At the same time, it is robust in the presence of noise. For field data application, the result of the method matches to interpreted gas-bearing curves obtained from the measured well logs at the corresponding well locations and the predicted gas-bearing property has a certain continuity around the wells, which is in good agreement with the geological law of the work area.

Acknowledgements

This research is supported by the National Natural Science Foundation of China (41974140), the Major Scientific Research Program of Petrochina Science and Technology Management Department “Comprehensive Seismic Prediction Technology and Software Development of Natural Gas” (2019B-0607), and the Scientific Research & Technology Development Project of China National Petroleum Corporation (2017D-3504).

References