DOI:10.1190/INT-2021-0213.1

Multiple seismic attributes fusion approach with support vector regression and forward simulation for sand body prediction and sedimentary facies interpretation — A case of the X gas field in Xihu Sag

Yuming Liu[1], Pengfei Xie[1], Dongping Duan[2], Yuanrong Yao[3], Peipei Liu[1], Wenjun Li[2], Bo Chen[2], Wei Wang[2], and Jiagen Hou[1]

Abstract

Marine exploration and production play a vital role in petroleum industries. It is difficult to acquire sufficient well data in marine settings, so seismic data become the most important interpretation data in research. In general, the seismic data in marine exploration have low quality because of the deep depth from the surface. To partly address this, the support vector regression (SVR) algorithm is proposed to fuse multiple seismic attributes for sand thickness prediction. First, we use forward modeling to establish virtual wells for improving the training data set. Second, we select the optimal attributes by correlation analysis. Third, we apply the SVR algorithm to learn the relationship between seismic attributes and sand thickness. Fourth, we use the SVR model to predict the sand thickness between wells by calculating a fused attribute. The results indicate that the fusion attribute with SVR has a higher correlation coefficient with sand thickness than the original attributes by statistical method. The approach can be widely used for improving the seismic interpretation quality of the research area with few wells.

Introduction

With the development of petroleum exploration and development, marine hydrocarbon shows great potential. Marine seismic data play an important role in marine hydrocarbon characterization (Du, 2018; Brankovic et al., 2021; Zhang and Zhang, 2021), but the quality of the data is affected by the depth, weather conditions at sea, geologic features, and seismic acquisition (Bredesen et al., 2020; Chan et al., 2021; Qi et al., 2021). The processing methods of seismic data can improve the quality of seismic data for characterizing the subsurface geologic distribution. The sand bodies distribution controls the distribution of oil and gas, so an appropriate processing method is important for seismic interpretation (Ruig and Hubbard, 2006).

In general, researchers combine well logs and seismic data to predict the subsurface distribution. The dense well spacing and high-quality seismic data are useful for prediction, but some fields are drilled with sparse well spacing and seismic data in deep depth (Gassaway and Richgels, 1984; Anifowose et al., 2016). To solve the problem, different kinds of methods have been proposed (Wankui et al., 2006; Li et al., 2017; Ghaderpour, 2019). The statistical methods based on the amplitude and frequency of seismic data are a conventional way to predict the reservoir distribution. The key point of statistical methods is the mapping relationship between seismic attributes and sand thickness. With the mapping, the sand thickness can be calculated based on seismic attributes. Statistical methods rely on several wells, whereby more wells enables more reliable mapping relation with low uncertainty. The marine oilfields often have no sufficient wells for statistical analysis. Seismic inversion is proposed to predict an acoustic impedance (AI) attribute according to well logs and seismic data (Armitage and Stright, 2010; Bakke et al., 2013; Bhattacharya et al., 2016). The attribute directly shows the spatial sand distribution, so geologists can calculate the sand thickness. Forward simulation also is an approach to predicting the sand bodies distribution (Ostrander, 1984; Ahmad and Rowell, 2012; Chen, 2015; Yue et al., 2019; Luanxiao et al.,

[1]China University of Petroleum (Beijing), State Key Laboratory of Petroleum Resources and Prospecting, Beijing, China and China University of Petroleum (Beijing), College of Geosciences, Beijing, China. E-mail: liuym@cup.edu.cn (corresponding author); 2019310047@student.cup.edu.cn (corresponding author); 2020210080@student.cup.edu.cn; houjg63@cup.edu.cn.

[2]Shanghai Branch of CNOOC (China) Co. Ltd., Shanghai, China. E-mail: duandp@cnooc.com.cn; liwj41@cnooc.com.cn; chenbo8@cnooc.com.cn; weiwang@cnooc.com.cn.

[3]China University of Petroleum (Beijing), College of Geosciences, Beijing, China and Tianjin Branch of CNOOC (China) Co. Ltd., Tianjin, China. E-mail: yaoyr5@cnooc.com.cn.

Manuscript received by the Editor 22 November 2021; revised manuscript received 19 March 2022; published ahead of production 5 June 2022; published online 22 August 2022. This paper appears in *Interpretation*, Vol. 11, No. 1 (February 2023); p. SA33–SA45, 13 FIGS., 6 TABLES.
http://dx.doi.org/10.1190/INT-2021-0213.1.

2021). The workflow of forward simulation assumes a sand bodies model and then the seismic response can be calculated by the sand bodies model and seismic convolution theory. Finally, the calculated seismic response is revised until the differences between the seismic response and the real seismic data reach convergence. However, the sand bodies have extensive heterogeneity and the sand bodies features are not easily characterized by low-resolution seismic data. In particular, the inversion method and forward simulation often produce a high uncertainty interpretation. Based on the statistical interpretation methods, researchers use a machine-learning algorithm to examine the relation between sand thickness and seismic attributes. A machine-learning algorithm, support vector machine (SVM) that can get the nonlinear relationship between data by kernel function, is applied to fit with well data. Initially proposed by Cortes and Vapnik (1995), SVM has been widely applied to oil and gas characterization of reservoirs, lithofacies recognition, and reservoir parameter prediction (Dong et al., 2014; Wang et al., 2014). The SVM can build an optimal classification model, but the deviation of the model will be large in continuous variables.

In this paper, we apply the SVR algorithm with forward simulation to the X gas field in Xihu Sag on the continental shelf of the East China Sea. Based on the well log and seismic data, the nonlinear mapping relation between the optimized seismic attributes and sand thickness is established by the SVR model. The SVR model can predict the spatial distribution of sand bodies at the target stratum. The predicted sand thickness distribution provides a beneficial basis for improving the gas field development plan.

This paper is organized as follows. The "Geologic setting" section presents the geologic setting of the work area, including the introduction of data and the location of the study area. In the "Principles" section, the principle of the method with SVR and the fusion method is explained. This is followed by the "Optimal attribute selection" section with an optimal attribute selection using correlation analysis. In the "Attribute fusion" section, the details of attribute fusion are shown, including an extension of the training set by forward modeling and attribute fusion result analysis. The "Sand prediction and sedimentary interpretation" section predicts the distribution of sand thickness and sedimentary facies based on well logs. In the "Discussion" section, we compare the proposed method with the general interpretation method and application based on SVM. Finally, the "Conclusion" section summarizes the most important findings.

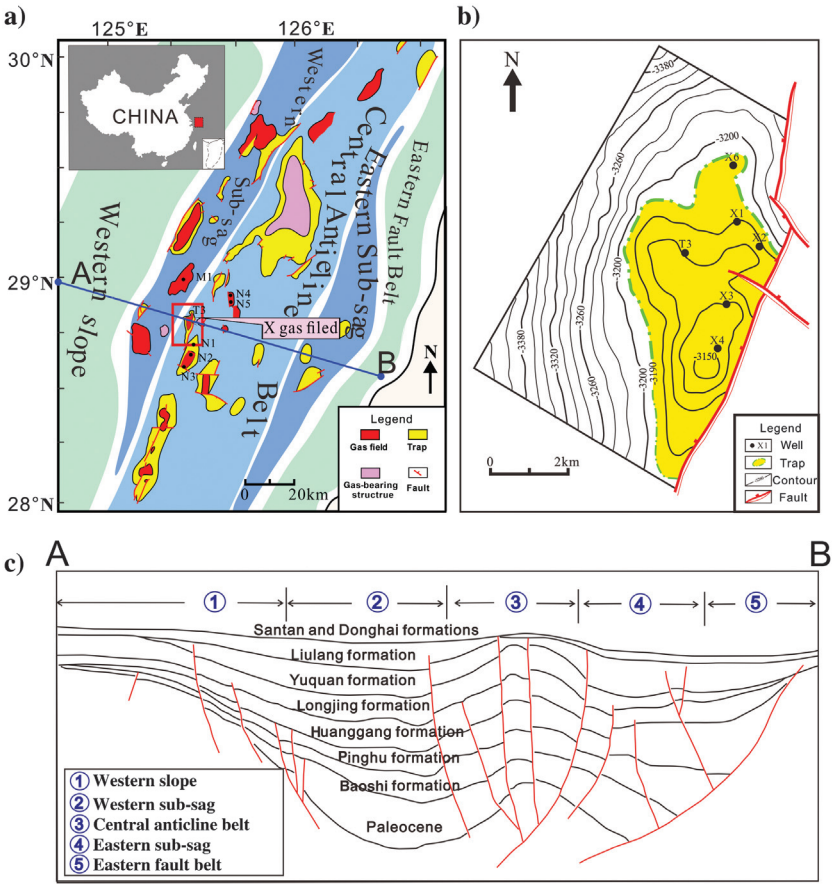


Figure 1. (a) Location and tectonic divisions of Xihu Sag. The X gas field is located in the red rectangle. (b) Structural features of the study area and position of wells. (c) Structural cross section of the Xihu Sag (section A-B is shown in [a]).

Geologic setting

The X gas field is close to Shanghai and Ningbo City in Zhejiang Province. From the structural position, the field is located south of the central anticline belt in Xihu Sag. Xihu Sag belongs to a basin on the continental shelf of the East China Sea (Figure 1). The study area is an anticline with a north-northeast distribution and the east and west subdepressions provide hydrocarbon in multiple directions. The field has favorable oil and gas support conditions. There are good structural trap conditions in the field. The tectonic high point is located at the south of the work area and the east is steeper than the west. The major axis and minor axis of the target stratum are 7.8 and 3.0 km, respectively. The area of the trap is 4.2–14.7 km^2 and the closure amplitude ranges from 31 to 37 m.

The strata of the X gas field are divided into the Paleogene Eocene Pinghu Formation and Oligocene Huagang Formation, where the latter is the main oil-bearing bed. There are a total of 12 sand groups in the Huagang Formation. The main gas-bearing beds in the X gas field are the H5 and H6 sand groups of the

DOI:10.1190/INT-2021-0213.1

Huagang Formation. The H5 and H6 sand groups are divided into four subzones H5a, H5b, H6a, and H6b. The subzone thickness is relatively stable in the H5 and H6 sand groups, whereas the formation thickness changed little. The main gas-producing beds in the X gas field are H5a, H5b, and H6a, and the burial depth of these beds exceeds 3000 m.

In this study, the H5 sand group is selected for example. There is an anticline in the middle of the target stratum, and a total of 11 normal faults with a small fault distance are developed on the east side of the study area. The thickness of the H5 sand group is 40–50 m. The sand groups presented an aggradational stable distribution.

The seismic data of the X gas field are poststack seismic data. The phase of poststack seismic data is approximately zero. The vertical sample interval is 2 ms. The spacing of inline and crossline is 25 m. The frequency spectrum of the seismic data (Figure 2) shows that the effective frequency bandwidth of the seismic data in the H5 is wide, where the minimum frequency is 10 Hz and the maximum frequency is 55 Hz (Table 1). The dominant frequency is 28 Hz and the tuning thickness is 28 m. Thus, the seismic data can only reach the recognition requirement of subzones.

Principles

SVR

Support vector regression (SVR) is a very important branch of SVM application. Compared with linear regression, SVR combines with SVM and regression to solve nonlinear regression problems (Cortes and Vapnik, 1995). Linear regression is suitable for big data, which needs a large number of statistical samples. Due to the few wells in the marine field, linear regression is not available.

With a dichotomy problem taken for example, a given data sample is classified under a given decision boundary. When the given decision boundary is $wx_i + b = 0$, the boundaries at two sides (two dotted lines) are $wx_i + b = 1$ and $wx_i + b = -1$, respectively, where all properties except b are identical. The parameters w and b will be optimized by the model training, where the parameter w means a slope of the boundary function or vector and b is the intercept. The following information can be obtained from Figure 3a: the distance between two boundaries is $2/\|w\|$ and the parameter $\|w\|$ is the length of the vector w. The distances from the optimal classification plane to two boundaries are $1/\|w\|$, and the SVM method aims to obtain the minimum length of w, which is expressed as

$$\begin{cases} \min \frac{1}{2}\|w\|^2 \\ \text{s.t. } y_i(wx_i + b) \geq 1, \quad \forall\, i \end{cases}. \tag{1}$$

These formulas mean that the length of w is minimum when all sample points (x, y) are beyond two boundaries:

$$\begin{cases} \min \frac{1}{2}\|w\|^2 \\ \text{s.t. } |y_i - (wx_i + b)| \leq e, \quad \forall\, i \end{cases}. \tag{2}$$

In the formula of SVR, the boundary width is controlled using the parameter e. Formula 2 aims to

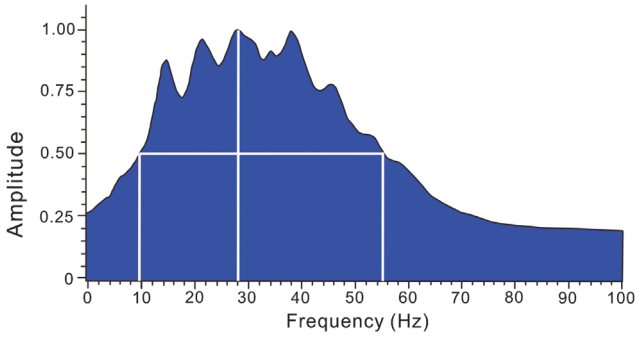


Figure 2. The spectrum of seismic data of H5 in the X gas field.

Table 1. The parameters of the seismic data spectrum of H5 in the X gas field.

Dominant frequency	Effective bandwidth
28 Hz	10–55 Hz

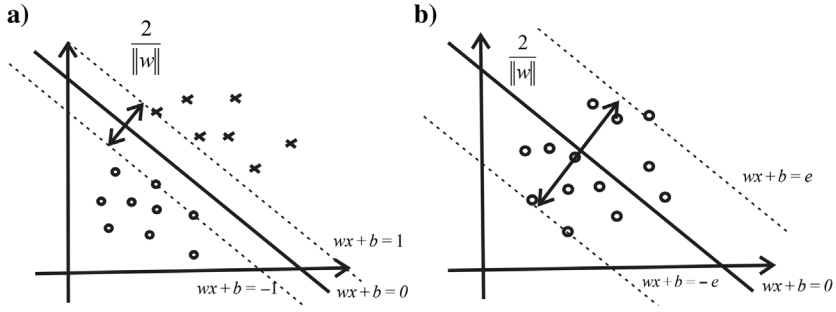


Figure 3. Diagrams of SVM and SVR. (a) Diagram of SVM. For a dichotomy problem of lithology, circles are sand and forks are mud. The x- and y-axes represent two attributes of sand or mud. The dashed lines signify the boundary between two kinds of data and the solid line represents the best interface (when data have multiple dimensions, the interface will become a hyperplane). (b) Diagram of SVR. The circle signifies a kind of data. The x- and y-axes represent two attributes of data (e.g., seismic attribute and sand thickness). The dotted lines signify the boundary of the data and the solid line represents the best regression interface. In these diagrams, the parameter $2/w$ is the distance between two boundaries.

reach the minimum length of w when all sample points are distributed within a fixed bandwidth. Different from SVM classification, the sample points of SVR are finally classified into one class in the SVR (Figure 3b).

For the optimal classification planes by SVR, two or multiple classes of sample points are not separated the most as SVM does, but instead, the total deviation of all sample points from the optimal classification plane is the minimum. The minimum total deviation means that all data points distribute on both sides of the plane (three dimensions or more) or line (two dimensions), which means that the SVR algorithm makes samples with low-biased error (the bias comes from the observed value and prediction value; the details are shown in Kutner et al., 1996). In addition, SVR generates a low-biased error for nonlinear model samples and multicollinear variables (multiple dimensional data, e.g., rocks with porosity, spontaneous potential, gamma-ray (GR) attributes, or more). In this manuscript, the parameter details of SVR are as follows.

The type of kernel function is a radial basis function, the degree of the kernel function is three, the kernel coefficient is the reciprocal number of feature numbers, and the constant term is set as 0.0. The tolerance and penalty coefficient that we selected is the default value.

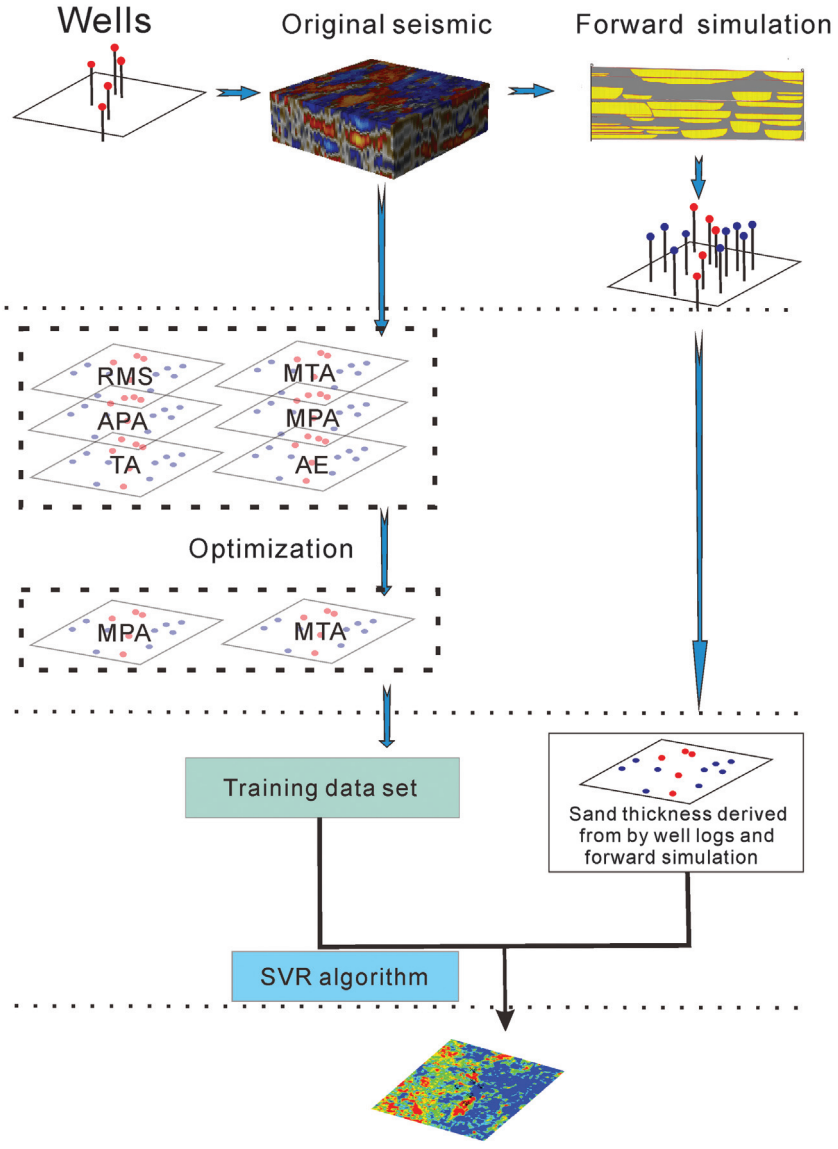


Figure 4. Workflow of multiple seismic attributes fusion based on SVR and forward modeling for sand prediction. In brief, the method proposed has three parts: (1) multiple attributes are extracted from original seismic data in wells position. Well logs and seismic data can get the hypothetical lithologic model by forward simulation. (2) The better attributes are optimized by correlation analysis (rms: root-mean-square amplitude; APA: average peak amplitude; TA: total amplitude; MTA: maximum trough amplitude; MPA: maximum peak amplitude; and AE: average energy). (3) The better seismic attributes can be used as a training data set and the sand thickness of wells can be the label. The SVR algorithm will learn the mapping relation between the label data set and the training data set. (4) The mapping relation of seismic attribute surface to sand thickness is used to predict the sand thickness between wells.

Multiple seismic attributes intelligence fusion

Due to the few wells in the study area, a forward simulation is applied to establish more wells for the SVR algorithm. By the SVR algorithm, the relation between multiple seismic attributes and sand thickness can be extracted and applied to predict the sand thickness.

The sand thickness is mainly predicted in four steps (Figure 4) through the multiple seismic attributes fusion approach based on the geologic model. First, we establish the forward model based on the original seismic data and well logs. The virtual wells are inserted in the forward model and the corresponding well logs are generated. Second, after the optimal seismic attribute assessment, the maximum trough amplitude with the highest correlation in the target stratum is selected. The maximum trough amplitude attributes in the upper and lower zones are selected as surrounding rock attributes, and the three attributes are input as the training data set and test data set. Third, we extract the sand thickness values in actual wells and virtual wells at the target stratum in the study area. The sand thickness is selected as labels of the training set and test data set in machine learning. In this process, the nonlinear mapping between the surrounding rock and the target stratum will be learned by the SVR model. Fourth, this nonlinear mapping relation is applied to fuse all attributes for sand thickness prediction.

DOI:10.1190/INT-2021-0213.1

The fused attribute is a single map that can quantitatively characterize the sand thickness.

Optimal attribute selection

Principles of optimal selection

Seismic data are the response of the reservoir physical property, which includes a significant amount of information about the lithofacies and fluid. The changes in the seismic waveform characterize the changes in the physical properties (e.g., impedance). In general, geologists extract the horizontal attributes and analyze the correlation between the horizontal attributes and sand logging interpretation. A high correlation makes the attribute available to predict the sand bodies. In the attributes analytical process, multiple seismic attributes are extracted and applied. We select the reservoir-related attributes from all seismic attributes.

There are three principles in the optimal selection of seismic attributes: (1) select the attributes that have no significant correlation, (2) select the attributes correlated well with the target parameters, and (3) remove attributes with noise information (seismic attributes with low correlation coefficient) and ensure the minimum loss of useful information.

Optimal selection of seismic attributes

Due to the low frequency of seismic data, we cannot clearly recognize lithofacies in the section. We can predict the sand bodies distribution from statistical features of seismic attributes. For example, the root-mean-square value of amplitude attributes is usually used to represent the sand thickness. Seismic attributes have sensitivities to reservoir properties. Thus, the optimal selection of seismic attributes is of great significance for guiding reservoir prediction.

In the formation H5a, 13 seismic surface attributes are calculated by the software Petrel. The correlations between seismic attribute values at six well points in each subzone of the X gas field and the corresponding sand thickness values are analyzed, and the Pearson correlation coefficients are shown in Table 2. By comparing the correlations between seismic attributes and sand thickness, the seismic attributes with higher correlation coefficients are optimally selected to predict the distribution of sand bodies. Four seismic attributes selected from 13 attributes, maximum trough amplitude value, average trough amplitude value, total absolute amplitude, and root-mean-square amplitude in H5a showed high correlations (correlation coefficients >0.6).

Attribute fusion

Forward seismic simulation and establishment of virtual wells

The few wells make the training data set of the SVR method unconvincing, so the forward model is established through the seismic forward simulation method. In the forward model, we set some virtual wells to extract the properties as well logs, which increases the number of training data sets.

In a forward simulation, we can calculate the seismic data using a geologic lithofacies model, elastic parameters, and wavelet. Based on the statistical information of well logs, the density and velocity of sand and mud can be acquired. The AI parameters (the product of density and velocity) and the seismic wavelet are the base of forward modeling. The final forward model is a new seismic convolution model, specifically in the following three steps.

Selection of AI and wavelet parameters

To decrease the difference between the original seismic data and the forward model, the calculated AI parameters of sand and mud are from the well logs

Table 2. The Pearson correlation coefficients between the seismic attributes and sand thickness of H5 in the X gas field.

Type of seismic attributes	Rms	MTA	MPA	TAA	AA	APA	ATA	AIP	ARS	AE	ST
Root-mean-square amplitude (rms)	1	—	—	—	—	—	—	—	—	—	—
Maximum trough amplitude (MTA)	0.621	1	—	—	—	—	—	—	—	—	—
Maximum peak amplitude (MPA)	0.813	0.862	1	—	—	—	—	—	—	—	—
Total absolute amplitude (TAA)	0.866	0.672	0.655	1	—	—	—	—	—	—	—
Average amplitude (AA)	0.806	0.672	0.347	0.750	1	—	—	—	—	—	—
Average peak amplitude (APA)	0.465	0.722	0.901	0.826	0.324	1	—	—	—	—	—
Average trough amplitude (ATA)	0.996	0.886	0.687	0.866	0.663	0.920	1	—	—	—	—
Average instantaneous phase (AIP)	0.226	0.436	0.322	0.312	0.521	0.336	0.376	1	—	—	—
Average reflection strength (ARS)	0.684	0.322	0.421	0.402	0.543	0.412	0.393	0.812	1	—	—
Average energy (AE)	0.765	0.261	0.275	0.277	0.711	0.283	0.287	0.751	0.942	1	—
Sand thickness (ST)	0.620	0.810	0.416	0.652	0.084	0.537	0.620	0.125	0.237	0.065	1

The selected attributes are shown in bold.

DOI:10.1190/INT-2021-0213.1

in the H5 sand group and are regarded as the corresponding impedance parameters in the forward simulation. The AI is the convolutional result of density and velocity. Here, the mud density ranges from 2.4 to 2.65 g/cm^3 and the velocity ranges from 2300 to 2350 m/s. The sand density ranges from 2.3 to 2.4 g/cm^3 and the velocity ranges from 3050 to 3100 m/s. The wavelet also is an important parameter in forward modeling. We calculate the average wavelet in the seismic data of formation H5 as a modeling parameter and the wavelet is shown in Figure 5 and the parameters are shown in Table 3.

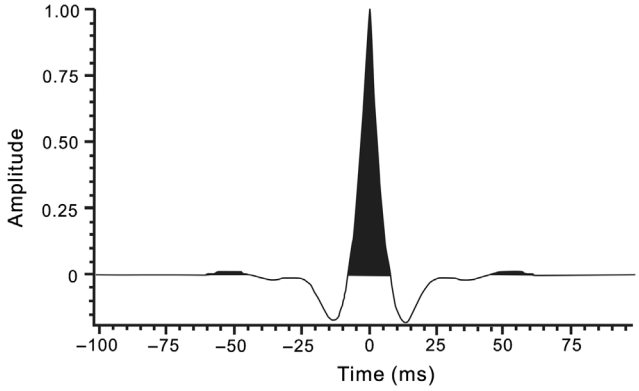


Figure 5. Statistical Ricker wavelet of the H5 sand group in the X gas field.

Table 3. The parameters of the statistical Ricker wavelet of H5 in the X gas field.

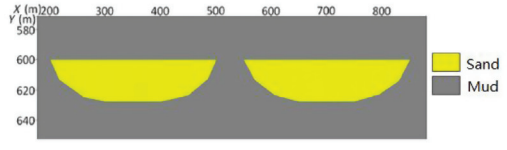
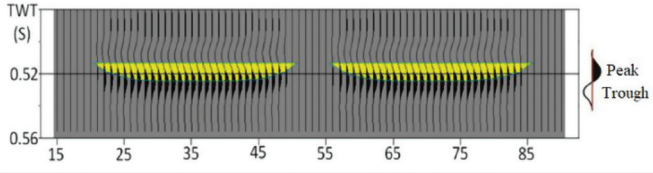
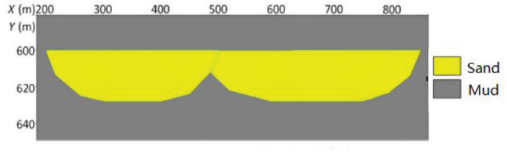
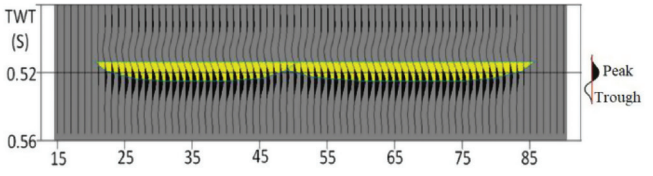
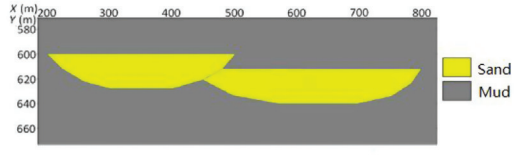
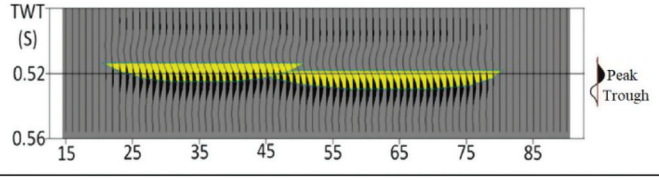
Length	Sample rate	Central frequency	Phase
128 ms	2 ms	25 Hz	0

Forward response analysis of a conceptual model

The sand bodies pattern is the basement of the lithofacies model and the seismic data are the response of the lithofacies model. We calculate the conceptual model and its seismic response for the foundation of forward modeling. In a different sedimentary background, the conceptual model has a different pattern. The core analysis (Figure 6) shows that the H5 formation is a shallow-water sedimentary system close to land. Most of the sediments in the study area are fine sandstone with a positive rhythm and gray mudstone. The gray mudstone represents a sedimentary layer above water and positive rhythm is an indication of a channel or bar. The thick sand body that buries from 3226 to 3255 m with a box GR log is a typical bar. More details of the sedimentary facies are not shown here; the key point of our paper is the seismic processing and interpretation method.

According to the analysis of sedimentary facies, three conceptual modes with a special pattern and the corresponding seismic response (Table 4) are

Table 4. Forward seismic response of different channel models.

Lithology model	Synthetic seismic response
	
	
	

designed. The seismic response is calculated as follows. (1) Forward seismic response of an isolated river channel in a single geologic era: this mode developed in a channel belt. The seismic response shows that the sand bodies' thickness gradually reduces as the seismic amplitude weakens. When the seismic amplitude reaches the minimum value, the sandstone turns into mud. (2) The seismic response of the intersection channels in a single geologic era: the seismic response shows that the seismic amplitude gradually weakens as the sand bodies' thickness reduces. At the intersection where the two channels meet, the sand thickness is thin and the amplitude is weak. (3) Seismic response of the intersection channels in the different geologic era: this mode is a common phenomenon in the channel sedimentary process. Two channels have different elevations because of the different generation time and the waveform in the intersection channel has a dislocation.

Seismic response characteristic analysis

The initial forward model is designed based on the logging interpretation of well T3 and well X2. By combining the well spacing, conceptual mode, and original seismic response, the initial forward response is established. In addition, the AI parameters of sand bodies between wells are the average impedance in formation H5, ranging from 7000 to 10,000 g/cm^3·m/s. The wavelet used in forward simulation is the statistical wavelet of H5, with the dominant frequency of 28 Hz.

If the resolution is sufficient, the single sand body can be recognized from the original seismic data. The seismic data in the marine setting make it difficult to recognize sand bodies clearly. In forward modeling, we obtain the finer result in parts as follows:

1) Establish a lithologic model (Figure 7a).
2) Calculate the corresponding seismic response (Figure 8a) based on the convolution mode.

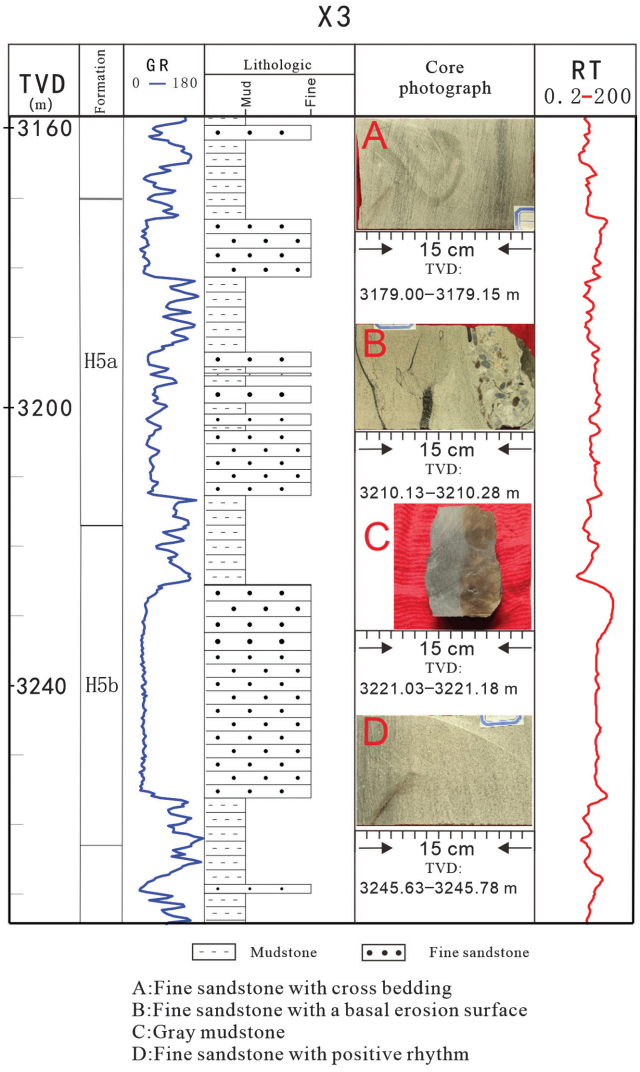


Figure 6. Lithologic features of well X3 in formation Huagang 5, showing the log (GR and RT) and core description; the well location is shown in Figure 1b.

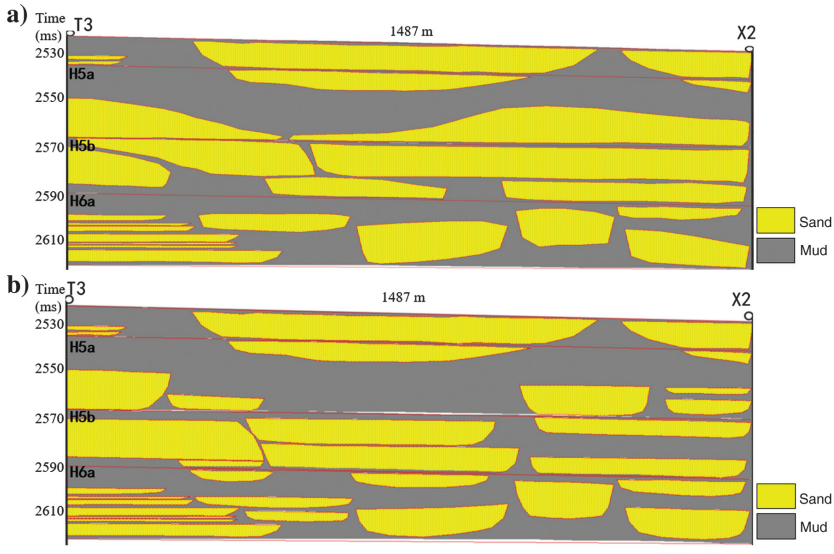


Figure 7. Lithologic model of section T3-X2. (a) Initial hypothesis model derived by the seismic response. (b) Final lithologic model for forward modeling of section T3-X2. The yellow is the sand facies and gray represents the mud facies.

DOI:10.1190/INT-2021-0213.1

3) Compare the seismic response in the second step with the actual seismic data (Figure 8b).
4) If the error is not able to reach the requirement, the lithologic model is revised manually.
5) Run steps 2–4 to decrease the error between the forward response and initial seismic data.
6) When the error reaches the requirement, stop the forward simulation.

The final lithologic model and forward seismic response are shown in Figures 7b and 8c. In visual, the final seismic response (Figure 8c) is still different from the original seismic data (Figure 8b), especially the well T3 of H5. There are two reasons: (1) the original seismic data are affected by the weather in marine and acquisition technology, and different noises cause significant difficulties in seismic data. (2) The lithologic

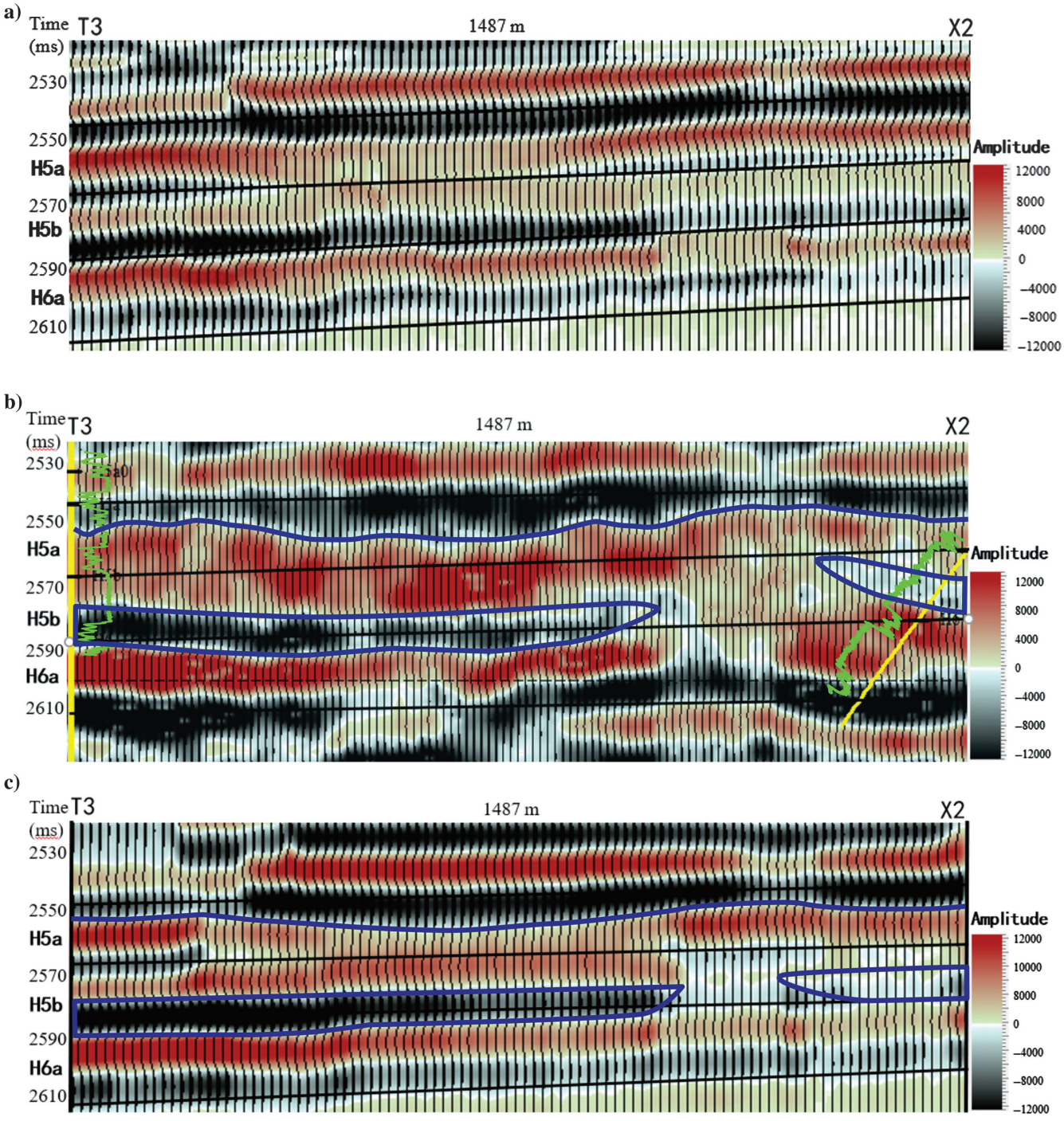


Figure 8. (a) Initial forward seismic response: synthetic seismic response of Figure 7a and the forward seismic response calculated by Figure 7a has a big bias with original poststack seismic data (b). (b) Original poststack seismic data of the study area: yellow lines are well T3 and X2 and the blue lines draw the outline of seismic wiggle. (c) Final forward seismic response: the synthetic seismic response of Figure 7b and the final forward seismic response fit original seismic data well. The blue lines draw the outline of seismic wiggle.

DOI:10.1190/INT-2021-0213.1

model replies to the well-logging interpretation and has the same resolution as wells. Even though the forward response is different from the initial seismic response in the position of the wells, the forward model with the low error that meets the conditional wells can characterize the sand bodies' distribution accurately. Based on the final lithology model generated by forward simulation, we set 16 wells for extension of the training data set in SVR.

The virtual wells (16 in total, as shown in Table 5) with a spacing of 90 m are inserted in the forward model, and the sand thickness values at the virtual well points in the H5a and H5b are extracted as labels of training data, as shown in Table 5.

Attributes fusion and result verification

Through the optimal attribute selection, the result shows that the maximum trough amplitude and sand thickness have the highest correlation coefficient, so the maximum trough amplitude is used for predicting the sand thickness. The 16 virtual wells are set as the training data set and six actual wells are set as the test data set. The test data set is used to certify the prediction result.

After the operation using the SVR method, the multiple seismic attributes fusion results (Figure 9) of formations H5a and H5b are obtained. The sand thickness of the test data set in formations H5a and H5b is shown in Table 6; the difference between the fusion attribute and sand thickness is small. It can be discovered through the correlation analysis (Figure 10) that the multiple seismic attributes fusion results have a high correlation with the sand thickness and the square of correlation coefficients is 0.9276 and 0.8597. Compared with the linear regression (Pearson correlation in Table 2), the correlation coefficient is much higher (the highest value of the Pearson correlation coefficient is 0.81, whereas the square of the correlation coefficient is higher than 0.8597). Figure 10a and 10b shows a low P-value ($P < 0.01$). It proves that the multiple seismic attributes fusion method is more suitable to predict the distribution of sand bodies in H5a and H5b.

Sand prediction and sedimentary interpretation

The fusion attribute has a high correlation with sand thickness in wells, thus it can be used to predict the sand thickness distribution. Based on the sand interpretation data of well logs, the distribution of sand thickness in H5a and H5b can be characterized (Figure 11). From Figure 11a, the sand bodies show a continuous and wide distribution and the sand thickness is horizontally nonuniform. The six wells all drill the sand bodies, the sand thickness ranges from 2.83 to 43.46 m, and the average thickness is 20.85 m. The thick sand bodies concentrate in the southwest of the study area, whereas the sand bodies in the northeast are thinner than those in the southwest. In the middle, thick sand bodies are locally developed with a short

Table 5. Sand thickness at well point.

Well type	Well name	Thickness of H5a (m)	Thickness of H5b (m)
Virtual well	XN1	27.16	26.75
	XN2	24.64	26.94
	XN3	9.09	31.25
	XN4	8.68	32.00
	XN5	8.78	24.64
	XN6	11.59	33.83
	XN7	13.20	33.82
	XN8	12.29	34.42
	XN9	11.33	34.18
	XN10	8.28	33.70
	XN11	2.81	14.60
	XN12	20.07	21.67
	XN13	20.42	22.01
	XN14	19.77	22.58
	XN15	11.15	22.94
	XN16	13.70	22.94
Actual well	T3	27.40	26.20
	X1	7.31	15.48
	X2	2.83	22.83
	X3	24.10	30.53
	X4	43.33	27.37
	X6	8.30	20.31

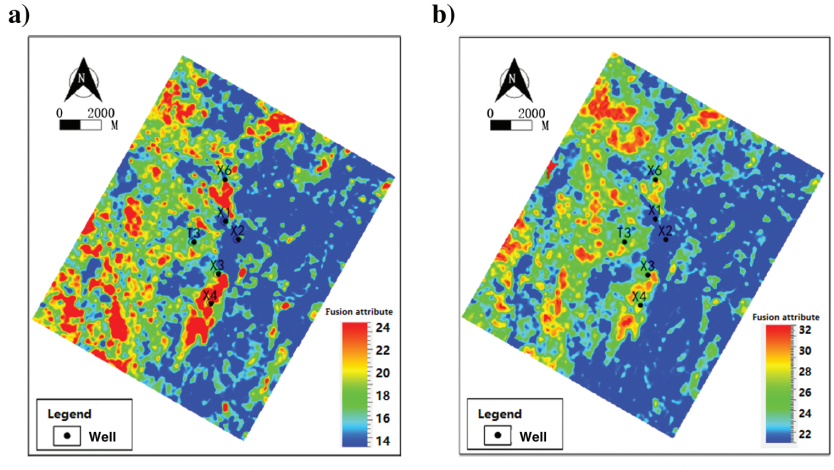


Figure 9. Results of multiple seismic attributes fusion in the X gas field. (a) Prediction result of H5a sand thickness by SVR. (b) Prediction result of H5b sand thickness by SVR. The label of the data set is the sand thickness and the results of SVR are attributed to be related to sand thickness.

band distribution. From Figure 11b, the sand bodies show a continuous and wide distribution and the sand thickness is relatively horizontally uniform. The six wells all drill the sand bodies, the sand thickness ranges from 15.48 to 30.53 m, and the average thickness is 24.48 m. The thick sand bodies are distributed in the middle, east, and north of the study area.

Table 6. Comparison table of the sand thickness and fusion attribute of H5a and H5b in the X gas field.

Well	Sand thickness of H5a (m)	Fusion attribute of H5a	Sand thickness of H5b (m)	Fusion attribute of H5b
T3	27.4	21.1	26.2	25.01
X1	7.31	13.71	15.48	18.1
X2	2.83	7.3	22.83	22.52
X3	24.1	22.12	30.53	30.61
X4	43.33	33.21	27.37	25.33
X6	8.3	16.51	20.31	24.02

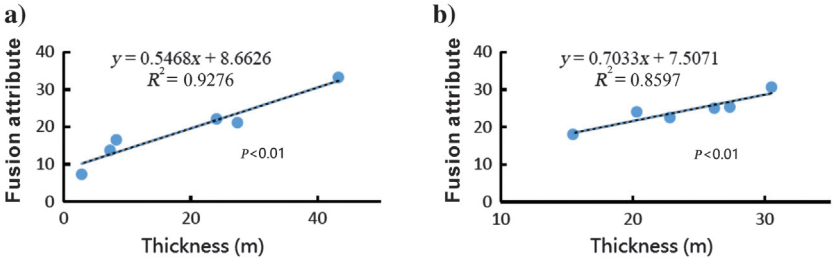


Figure 10. Correlation analysis of the sand thickness of test wells and seismic multiattributes fusion results in the X gas field. Sand thickness and fusion attribute of six actual wells are shown in Table 6. (a) Correlation of fusion attribute with thickness in H5a. (b) Correlation of fusion attribute with thickness in H5b.

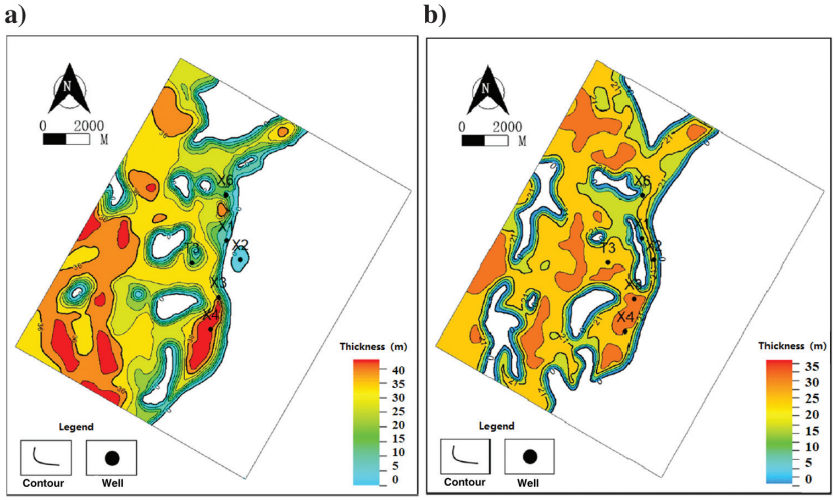


Figure 11. Sand thickness maps in the X gas field. The blank areas represent thickness 0. (a) Predicted sand thickness map of H5a with fusion attribute by SVR (Figure 9) and correlation coefficient in Figure 10. (b) Predicted sand thickness map of H5b with fusion attribute by SVR (Figure 9) and correlation coefficient in Figure 10. A high correlation coefficient makes the fusion attribute become sand thickness directly.

Based on the fusion seismic attribute and well-logging interpretation, the composite channels boundary can be predicted under the guidance of sedimentary knowledge. Then, we can recognize the bars and overbanks base on channels. The sedimentary interpretation shows that the three channels of H5a meet from the northeast, north, and northwest directions. Eighteen bars are identified in the channels, which are mainly located at the center of each channel, with a length of 500–2000 m and a width of 200–1000 m. The dams show potato-like or narrow banded distribution (Figure 12a–12c). The channels of H5b have a similar shape as H5a. Sixteen bars are identified in the river channels and are mainly distributed at the center of each river channel, with a length of 300–1500 m and a width of 200–1000 m. Most of them show narrow banded distribution (Figure 13a–13c).

Discussion

The method that we proposed helps the low-frequency marine seismic interpretation. When we cannot characterize the reservoir feature by the original post-stack data, we can fuse useful information, such as statistical attributes and sand thickness by well-log interpretation. The machine-learning algorithm is a good way to delve into the complicated mapping relation between training data and labels. And we can predict the labels of test data by the mapping relation.

Compared with the general seismic interpretation, the proposed method combines with the seismic attributes and machine-learning algorithm. General seismic interpretation selects the optimal attribute that has the highest correlation to sand thickness in wells. And then, the optimal attribute can be used to characterize the sand distribution. However, all statistical attributes in marine seismic data have no good correlation to sand thickness by linear regression; it is not available for seismic interpretation in low-quality seismic data.

SVM has been applied to seismic interpretation for many years and has a good effect. However, the algorithm is limited by the quantity of the data set, especially in marine settings. The drilling cost in marine exploration is too expensive, so several wells are not sufficient to use SVM for interpretation. The method that we proposed adds some virtual wells by seismic forward modeling. And SVR is more appropriate for a few training data

DOI:10.1190/INT-2021-0213.1

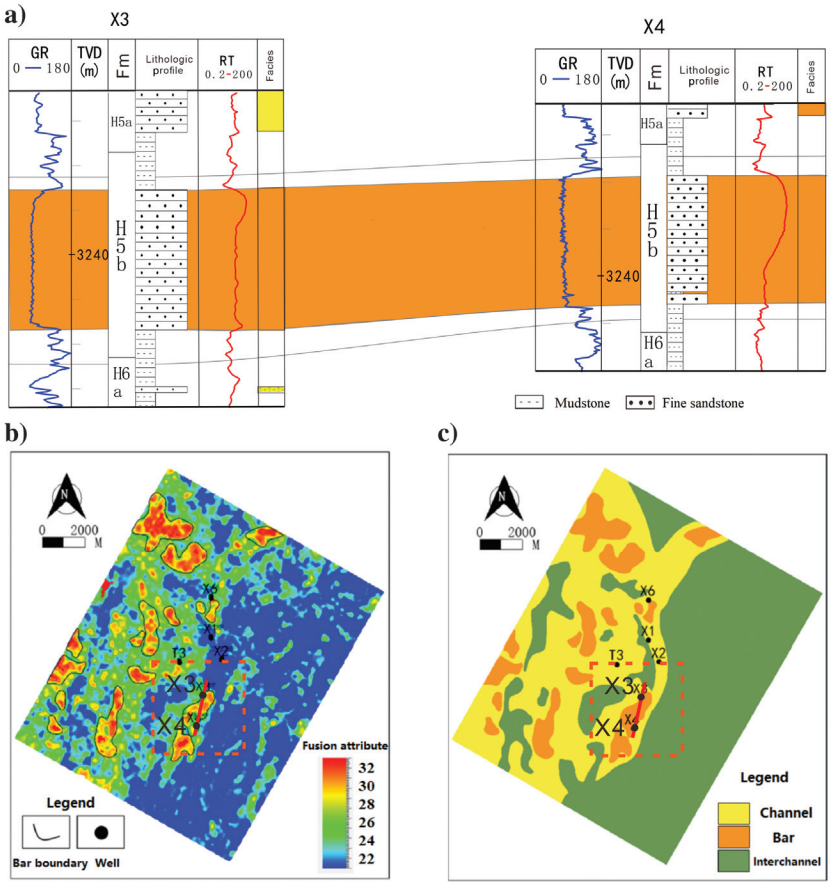


Figure 12. Sedimentary facies interpretation of H5a based on the fused attribute in the X gas field. (a) The interpretation section based on the well logging of H5a, (b) the feature of the sand bar of H5a based on the fusion attribute, and (c) sedimentary interpretation of H5a.

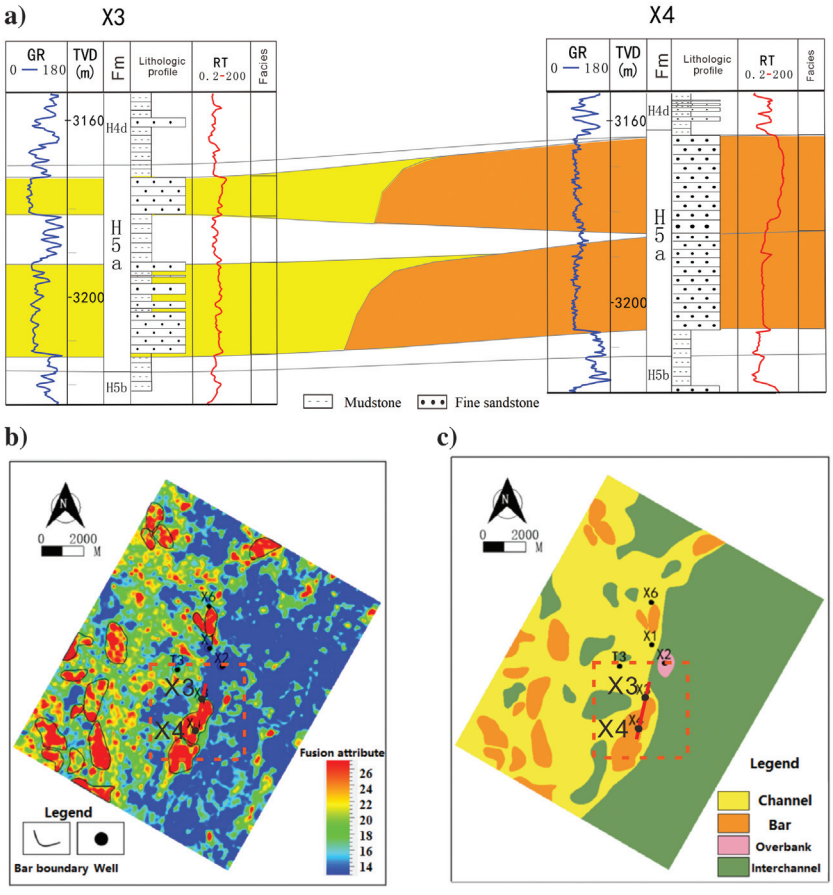


Figure 13. Sedimentary facies interpretation of H5b based on the fused attribute in the X gas field. (a) The interpretation section based on the well logging of H5b, (b) the feature of the sand bar of H5b based on the fusion attribute, and (c) sedimentary interpretation of H5b.

DOI:10.1190/INT-2021-0213.1

sets than SVM. To solve the low-frequency seismic data and few wells condition of marine settings, the proposed method is better than general seismic interpretation and other methods based on SVM.

Conclusion

The multiple seismic attributes fusion method with SVR provides an important way to help the interpretation of marine oilfields with few wells and low-quality seismic data. In this process, forward seismic simulation creates a significant amount of effective wells to improve the training data set of SVR. The SVR algorithm extracts the nonlinear mapping relationship between sand thickness and seismic attributes, which is then applied to predict the distribution of sand thickness. The results show that sand thickness prediction has a high correlation with the real sand thickness. The method is reliable and can be applied to sand thickness prediction and sedimentary interpretation. Compared with the general seismic attributes interpretation, the method acquires a new fusion attribute that has a higher correlation with sand thickness. The SVR algorithm provides a good way to extract useful information from seismic attributes, whereas linear regression uses a single optimal attribute for interpretation. Thus, the fused attribute provides more information than linear regression. In particular, the attributes of shoulder beds are added to the SVR learning process. Based on the fusion attribute and wells interpretation, a convincing sedimentary facies interpretation can be acquired by researchers. It is helpful to analyze the reservoir architecture and direct the development plan. But in the forward simulation, it requires much time to revise the lithologic model. In the future, we have to improve the speed of establishing the prior lithologic model. And the processing method that can improve the quality of seismic data may be an available way to improve the geologic interpretation of the method.

Acknowledgments

This work was supported by National Natural Science Foundation of China (nos 42172154 and 42072146) and Science Foundation of China University of Petroleum, Beijing (no. 2462020YXZZ022).

Data and materials availability

Data associated with this research are confidential and cannot be released.

References

Ahmad, M., and P. Rowell, 2012, Application of spectral decomposition and seismic attributes to understand the structure and distribution of sand reservoirs within Tertiary rift basins of the Gulf of Thailand: The Leading Edge, **31**, 630–634, doi: 10.1190/tle31060630.1.

Anifowose, F., S. Adeniye, A. Abdulraheem, and A. Al-Shuhail, 2016, Integrating seismic and log data for improved petroleum reservoir properties estimation using non-linear feature-selection based hybrid computational intelligence models: Journal Petroleum Science Engineering, **145**, 230–237, doi: 10.1016/j.petrol.2016.05.019.

Armitage, D. A., and L. Stright, 2010, Modeling and interpreting the seismic-reflection expression of sandstone in an ancient mass-transport deposit dominated deep-water slope environment: Marine and Petroleum Geology, **27**, 1–12, doi: 10.1016/j.marpetgeo.2009.08.013.

Bakke, K., I. A. Kane, O. J. Martinsen, S. A. Petersen, T. A. Johansen, S. Hustoft, F. H. Jacobsen, and A. Groth, 2013, Seismic modeling in the analysis of deep-water sandstone termination styles: AAPG Bulletin, **97**, 1395–1419, doi: 10.1306/03041312069.

Bhattacharya, S., T. R. Carr, and M. Pal, 2016, Comparison of supervised and unsupervised approaches for mudstone lithofacies classification: Case studies from the Bakken and Mahantango-Marcellus Shale, USA: Journal of Natural Gas Science and Engineering, **33**, 1119–1133, doi: 10.1016/j.jngse.2016.04.055.

Brankovic, M., E. Gildin, R. L. Gibson, and M. E. Everett, 2021, A machine learning-based seismic data compression and interpretation using a novel shifted-matrix decomposition algorithm: Applied Sciences, **11**, 4874, doi: 10.3390/app11114874.

Bredesen, K., E. Dalgaard, A. Mathiesen, R. Rasmussen, and N. Balling, 2020, Seismic characterization of geothermal sedimentary reservoirs: A field example from the Copenhagen area, Denmark: Interpretation, **8**, no. 2, T275–T291, doi: 10.1190/INT-2019-0184.1.

Chan, Y., D. Duan, L. Zhang, and F. Ding, 2021, Quantitative characterization of the depositional system in Gas field A, Pinghu slope belt, Xihu Sag and its bearing on periodicity of sea level changes: Marine Geology & Quaternary Geology, **41**, 12–21, doi: 10.16562/j.cnki.0256-1492.2020081302.

Chen, G., 2015, Fault identification by logging-seismic combination in Lamadian oilfield: Special Oil & Gas Reservoirs, **22**, 131–134.

Cortes, C., and V. Vapnik, 1995, Support vector networks: Machine Learning, **20**, 273–297.

Dong, Y., X. Zhu, B. Xian, K. Cheng, and P. Wang, 2014, Mapping sediment-dispersal characteristics using seismic geomorphology: Late Paleogene to Neogene, Qinan Sag, Huanghua Depression, China: Marine and Petroleum Geology, **54**, 180–197, doi: 10.1016/j.marpetgeo.2014.03.008.

Du, X., 2018, Progress of seismic exploration technology in offshore China: Geophysical Prospecting for Petroleum, **57**, 321–331.

Gassaway, G. S., and H. J. Richgels, 1984, SAMPLE: Seismic amplitude measurement for primary lithology estimation: AAPG Bulletin, **2**, 479.

Ghaderpour, E., 2019, Multichannel antileakage least-squares spectral analysis for seismic data regularization beyond aliasing: Acta Geophysica, **67**, 1349–1363, doi: 10.1007/s11600-019-00320-3.

Kutner, M. H., C. J. Nachtsheim, W. Wasserman, and J. Neter, 1996, Applied linear statistical models: American Statistical Association.

Li, W., D. Yue, G. Hu, T. Fan, and X. Fang, 2017, Frequency-segmented seismic attribute optimization and sandbody distribution prediction: An example in North Block, Qinghuangdao 32-6 Oilfield: Oil Geophysical Prospecting, **52**, 121–130, doi: 10.13810/j.cnki.issn.1000-7210.2017.01.017.

Luanxiao, Z., Z. Caifeng, C. Yuanyuan, S. Wenlong, W. Yirong, C. Huaizhen, and G. Jianhua, 2021, Fluid and lithofacies prediction based on integration of well-log data and seismic inversion: A machine-learning approach: Geophysics, **86**, no. 4, M151–M165, doi: 10.1190/geo2020-0521.1.

Ostrander, W. J., 1984, Plane-wave reflection coefficients for gas sands at non-normal angles of incidence: Exploration Geophysics, **15**, 193, doi: 10.1071/EG984193a.

Qi, P., G. Guo, Y. Ren, M. Cui, and X. Wang, 2021, Geological characterization of the Eocene Pinghu movement in the Xihu Sag and its hydrocarbon geological significance: Geoscience, **35**, 1098–1105, doi: 10.19657/j.geoscience.1000-8527.2021.045.

Ruig, M., and S. M. Hubbard, 2006, Seismic facies and reservoir characteristics of a deep-marine channel belt in the Molasse foreland basin, Puchkirchen Formation, Austria: AAPG Bulletin, **90**, 735–752, doi: 10.1306/10210505018.

Wang, G., T. R. Carr, Y. Ju, and C. Li, 2014, Identifying organic-rich Marcellus Shale lithofacies by support vector machine classifier in the Appalachian basin: Computers & Geosciences, **64**, 52–60, doi: 10.1016/j.cageo.2013.12.002.

Wankui, G., P. Yanming, K. Fanzhong, G. Yu, and B. Marion, 2006, Large well-spacing crosswell seismic imaging for deep gas reservoir mapping and description: 76th Annual International Meeting, SEG, Expanded Abstracts, 364–368, doi: 10.1190/1.2370277.

Yue, D., W. Li, W. Wang, G. Hu, H. Qiao, J. Hu, M. Zhang, and W. Wang, 2019, Fused spectral-decomposition seismic attributes and forward seismic modelling to predict sand bodies in meandering fluvial reservoirs: Marine and Petroleum Geology, **99**, 27–44, doi: 10.1016/j.marpetgeo.2018.09.031.

Zhang, P. F., and S. H. Zhang, 2021, Neural network seismic prediction of sand and mudstone lithology of Pinghu Formation in Xihu Sag: Geophysical and Geochemical Exploration, **45**, 1014–1020, doi: 10.11720/wtyht.2021.1297.

Biographies and photographs of the authors are not available.

DOI:10.1190/INT-2021-0213.1