



Original Paper

Machine learning for carbonate formation drilling: Mud loss prediction using seismic attributes and mud loss records



Hui-Wen Pang ^{a, b}, Han-Qing Wang ^c, Yi-Tian Xiao ^c, Yan Jin ^{a, d, *}, Yun-Hu Lu ^{a, d}, Yong-Dong Fan ^{a, d}, Zhen Nie ^e

^a National Key Laboratory of Petroleum Resources and Engineering, China University of Petroleum (Beijing), Beijing, 102249, China

^b College of Science, China University of Petroleum-Beijing, Beijing, 102249, China

^c Petroleum Exploration and Production Research Institute, SINOPEC, Beijing, 102206, China

^d College of Petroleum Engineering, China University of Petroleum-Beijing, Beijing, 102249, China

^e Research Institute of Petroleum Exploration and Development, CNPC, Beijing, 100083, China

ARTICLE INFO

Article history:

Received 11 December 2022

Received in revised form

26 October 2023

Accepted 29 October 2023

Available online 13 November 2023

Edited by Jia-Jia Fei

Keywords:

Lost circulation

Risk prediction

Machine learning

Seismic attributes

Mud loss records

ABSTRACT

Due to the complexity and variability of carbonate formation leakage zones, lost circulation prediction and control is one of the major challenges of carbonate drilling. It raises well-control risks and production expenses. This research utilizes the H oilfield as an example, employs seismic features to analyze mud loss prediction, and produces a complete set of pre-drilling mud loss prediction solutions. Firstly, 16 seismic attributes are calculated based on the post-stack seismic data, and the mud loss rate per unit footage is specified. The sample set is constructed by extracting each attribute from the seismic trace surrounding 15 typical wells, with a ratio of 8:2 between the training set and the test set. With the calibration results for mud loss rate per unit footage, the nonlinear mapping relationship between seismic attributes and mud loss rate per unit size is established using the mixed density network model. Then, the influence of the number of sub-Gaussians and the uncertainty coefficient on the model's prediction is evaluated. Finally, the model is used in conjunction with downhole drilling conditions to assess the risk of mud loss in various layers and along the wellbore trajectory. The study demonstrates that the mean relative errors of the model for training data and test data are 6.9% and 7.5%, respectively, and that R^2 is 90% and 88%, respectively, for training data and test data. The accuracy and efficacy of mud loss prediction may be greatly enhanced by combining 16 seismic attributes with the mud loss rate per unit footage and applying machine learning methods. The mud loss prediction model based on the MDN model can not only predict the mud loss rate but also objectively evaluate the prediction based on the quality of the data and the model.

© 2024 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Lost circulation is a costly and time-consuming issue encountered in drilling a well and directly related to many drilling problems (Ahdaya et al., 2022; Jin et al., 2021). According to some estimates, the annual cost of lost circulation is around 2–4 billion dollars globally (Cook et al., 2011). Drilling in carbonate formations always presents more difficult lost circulation challenges. Due to the uncertainty of leakage zone (e.g. vug, fracture, and cave)

development, predicting and controlling mud loss rate is extremely hard. In all types of carbonate formations, the mud loss control of vug-fractured limestone is the most complex because the size distribution of fractures and vugs varies significantly and their combinations are unpredictable. When drilling in the vug-fractured limestone, mud loss occurs, ranging from seepage loss to total loss, which makes mud loss control the primary challenge of safe and efficient drilling (Wang et al., 2020a).

To solve the problem of lost circulation, different experts and scholars have put forward different methods. As we all know, two criteria must be met for mud loss to occur: 1. Wellbore fluid pressure in the wellbore is higher than pore pressure; 2. The formation has leakage zones and sufficient space for liquid storage.

* Corresponding author. State Key Laboratory of Petroleum Resources and Prospecting, China University of Petroleum-Beijing, Beijing, 102249, China.

E-mail address: jiny@cup.edu.cn (Y. Jin).

Considering the aforementioned two necessary conditions, two kinds of mud loss prediction and control methods are formed, respectively. The first focuses on fluid pressure control. The ceiling limit is the fracture pressure, which characterizes the initiation pressure of rock due to hydraulic pressure (Tan et al., 2020, 2021). The fracture pressure is one of the most important parameters for safe mud window design. In 1957, Hubbert and Willis proposed an empirical expression for fracture pressure calculation (Hubbert and Willis, 1957) used in the Gulf of Mexico region. Then, Matthews and Kelly (1967) proposed a similar relationship for fracture pressure. In 1969, Eaton suggested a physically based technique for calculating fracture pressure based on Poisson's ratio, known as Eaton's equation (Eaton, 1969). Daines proposed Stephen's equation (Daines, 1982), which can be directly applied to measured Poisson's ratio. Subsequently, by considering the effect of pore elasticity, some scholars proposed a new equation. The representative equations include Anderson's equation (Anderson et al., 1973), Huang's equation (Huang, 1984), and Zoback-Healy's equation (Zoback and Healy, 1984). There are other fracture pressure equations considering the effect of temperature and other physical field couplings (Hauser, 2021; Gao et al., 2021). It's important to point out that "leakage pressure" or "lost circulation pressure" can usually be considered a special case of fracture pressure (Morita et al., 1990; Chan et al., 2013). The tensile strength of a non-integrity formation (e.g., formation with fractures) is almost zero, so it is more reasonable to use leakage pressure instead of fracture pressure as the upper limit of mud density (Shi et al., 2012; Li et al., 2019).

The second is concerned with the identification of leakage zone and mud loss patterns. If there are no leakage zones, mud loss may not occur due to the higher mechanical compaction and strength of carbonate. Accurately describing the leakage zones may serve as a basis for comprehending the mud loss pattern, and it can also help the preparation of lost circulation materials. The leakage zones are characterized using a mix of core, logging, seismic, and engineering mud loss records. Compared to caves in the p-wave section derived from seismic data that are beaded and flaky with significant reflection characteristics (Zheng et al., 2019), vugs and fractures often lack substantial variation. Consequently, logging data, especially image logging data, is commonly used to diagnose vugs and fractures. Based on image logging data, Wu and Pollard (2002) proposed a method for generating a 3D borehole fracture network. Image logging and interpretation were used by Özkaya (2003) to determine a scaling factor between fracture length and aperture as well as the average fracture length. Imaging logging, when compared to seismic data, can accurately identify vugs and fractures, but it comes at a high cost. As a result, "inversion" of leakage zone characteristics using engineering records has become an acceptable approach. Dyke et al. (1995) obtained the trendline of liquid level change over time in mud pools under three types of lost circulation conditions including porosity, natural fracture, and induced fracture, through a statistical and theoretical analysis. Liétard et al. (1999) presented the relationship between mud loss volume and natural fractures aperture by analyzing the pattern of mud loss volume changing with time and integrating it with well log calibration. Later, Sawaryn et al. (2001) derived the corresponding analytical solution from Liétard et al.'s work (differential equation of mud ent incursion into the formation around the well), as well as the relationship between mud loss type and fracture permeability. Civan et al. (2002) then presented a simpler analytical solution, making the equation more general. Camacho-Velázquez et al. (2005) used the Warren-Root model to illustrate the response characteristics of the fracture-solution-pore reservoir under transient and quasi-steady flow conditions. Xia et al. (2015) carried out a method of calculating the fracture aperture

according to the mud loss rate based on the mud loss data of leakage zones. Characterizing leakage zones provides physical and geometric models for studying mud loss patterns, the foundation for the establishment of mathematical models. Lavrov et al. (2004) conducted a systematic study on the lost circulation caused by the natural fracture. Majidi et al. (2010) focused on the influence of rheological properties of drilling fluids on lost circulation in fractured formations. Gulbransen et al. (2010) established the multi-scale Mixed Finite-Element method to simulate fluid flow according to the different flow forms of vugs, caves, and fractures. Following that, several researchers took a similar approach, joining discrete fractures with the continuum matrix or discrete cave to perform mud loss simulations under various leaking zone combination conditions (Wu et al., 2011; Wang et al., 2020a, 2020b; Wei et al., 2020).

For a long time, the above two methods provide effective technical support for mud loss control. Due to the development of artificial intelligence and machine learning in recent years, some new methods for mud loss control have emerged. The data-driven method has powerful advantages in dealing with uncertainty problems and revealing useful information in drilling engineering (Noshi and Schubert, 2018). Most of the research on lost circulation has concentrated on mud loss diagnosis using mud logging data or other engineering data. Li et al. (2018) treated lost circulation as a dichotomous question, gathered geological characteristics and operational drilling parameters related to lost circulation, and conducted supervised learning using BP neural network, support vector machine, and random forest methods. Pang et al. (2022) selected 16 mud logging parameters and applied the mixed density network model to predict the mud loss rate. In addition, other scholars have also constructed the lost circulation diagnostic model by selecting different parameters and machine learning algorithms (Abbas et al., 2019; Hou et al., 2020; Sabah et al., 2021). Comparatively, the machine learning model of lost circulation diagnosis based on drilling parameters is more suitable for directing the optimization of engineering parameters for mud loss control. The pre-drilling prediction of mud loss can only use seismic data. Geng et al. (2019) took the lead in establishing the relationship between the probability of lost circulation occurrence and four seismic attributes using an ensemble algorithm and evaluating the weights of different subclasses by a majority voting algorithm. Then Ding et al. (2021) carried out a similar study based on geological research. Increasingly, engineering practice demonstrates that data analytics applications based on geological-engineering information fusion will become an effective way to solve drilling problems (Noshi and Schubert, 2018). However, there are few studies on the prediction of mud loss rate based on seismic data and the evaluation of the prediction results' uncertainty.

In this paper, taking H oilfield as an example, a complete workflow for pre-drilling mud loss rate prediction using machine learning is proposed. Firstly, 16 seismic attributes from post-stack seismic data related to leakage zones and mud loss records are extracted, and 20 typical wells covering all types of mud loss at varying depths are selected for research. Secondly, the mixed density network is used to establish the mapping relationship between borehole-side seismic attributes and unit footage mud loss rate. Then, according to the qualified relationship after the test to replace the original trace data, the 3D mud loss rate distribution and two uncertainty evaluation indices are obtained. Finally, the obstacles associated with the application of this approach are discussed.

2. Methodology

To further illustrate how seismic data can be combined with

machine learning methods to predict pre-drilling mud loss rate and assess lost circulation risk. The entire procedure consists of three parts: 1. Seismic-Well data preprocessing; 2. Model construction using machine learning; 3. 3D presentation and application.

1. **Seismic-Well data preprocessing:** ① Calculating multiple seismic attributes according to different attribute types. ② Single wells were selected to cover different mud loss type characteristics based on multivariate loss records. ③ Multiple seismic attributes of the characteristic single well are extracted, and the calibration and matching of mud loss records are carried out according to a time-depth relationship.
2. **Machine learning model construction:** ① Using the mixed density network to construct a machine learning model for mud loss rate. ② The data were divided into train sets and test sets at a ratio of 8:2, and estimate model parameters were modified to increase model precision and generalizability.
3. **3D presentation and applications:** ① The original post-stack seismic data in the depth domain is substituted with the mud loss rate distribution obtained by the machine learning for 3D visualization. ② Combined with other data, well location design and well structure optimization are performed.

The whole workflow is shown in Fig. 1.

2.1. Data analysis and processing

2.1.1. Extraction and processing of seismic attributes in a single well

Seismic attributes, the geometric, kinematic, dynamic, or statistical features of seismic waves, are derived from pre-stack or post-stack seismic data by mathematical transformations (Chopra and Marfurt, 2005). For different processing and interpretation purposes, there are many kinds of classification for seismic attributes (Brown, 1996; Chen and Sidney, 1997; Fomel, 2007). To ensure the accuracy of the calculation without sacrificing efficiency, we did not use hundreds of attributes to construct the model in a data-driven manner. Instead, we utilized a strategy that incorporated both data and experience (Geng et al., 2019; Tian et al., 2019). First, based on the recommendations of senior geophysical experts and the actual conditions of the study area, we calculated and obtained a total of 20 types of seismic attributes divided into 4 categories. On the basis of these attributes, we analyze the relationship between various attributes and mud loss using the mud loss rate as the objective function. When the absolute value of the Pearson correlation coefficient is less than 0.1, it is believed that this parameter has a negligible effect on circulation loss (Fig. 2). Therefore, 16 attributes with a strong correlation to sediment loss were selected.

For lost circulation, the existence of leakage zones is a prerequisite, so it is vital to select seismic attributes that can better describe the formation characteristics. These sixteen attributes have been able to satisfy the requirements of reservoir interpretation and are consequently appropriate for the engineering problem of lost circulation prediction. Trace AGC (Amplitude Gain Control) and structural smoothing are used to process the post-stack seismic data in order to improve the performance of the seismic data before attributes abstract. The original seismic data and different types of seismic attributes are shown in Fig. 3.

Due to the different mathematical transformation methods adopted by different attributes, the value distribution of different attribute bodies varies greatly. Table 1 shows the classification and value distribution range of different types of attributes.

Because of the large difference in the value of different attributes, it is necessary to normalize the value to prevent the "small value" swallowed by the "large value" during the later calculation.

The min-max normalization method is used to normalize the data (Ali et al., 2014):

$$x_{\text{nor}} = \frac{x - x_{\text{max}}}{x - x_{\text{min}}} \quad (1)$$

where, x_{nor} stands for normalized data, x stands for raw data, x_{min} and x_{max} stands for the maximum and minimum of the data respectively.

Normalization ensures that all seismic attributes are evenly distributed between 0 and 1. Fig. 4 depicts a box chart of the data distribution of 16 seismic attributes.

Vertical wells should be chosen as far as possible when picking typical wells to provide a unique correspondence between longitudinal data and depth. However, since the well trajectory usually does not coincide with the seismic trace, it is necessary to optimize and analyze the seismic trace data beside the well to achieve the well and seismic data matching. In this paper, the borehole is taken as the center of the circle for radial search, and the seismic trace closest to the center of the circle is selected as the seismic data to complete the matching of single well and seismic data. The schematic diagram is shown in Fig. 5.

2.1.2. Extraction and processing of mud loss records in a single well

Unlike seismic data, mud loss records rely primarily on hand-kept records during drilling. Typically, detailed mud loss-related records can be obtained from the Daily drilling report (DDR), Final well report (FWR), and Mud log records (Masterlog). Mud loss rate and mud loss volume, the key to the mud loss message, are usually used for evaluation mud loss type and preparation for lost circulation materials. However, due to the hand-kept records error, sometimes the three data sources (DDR, FWR and Masterlog) mud loss data records are not uniform. The FWR and Masterlog are results reports, whereas the DDR is a daily drilling record. To determine the mud loss calibration, the DDR mud loss record should serve as the base stone, with reference to FWR and Masterlog to verify mud loss rate and volume. Accurate determination of the mud loss depth interval and the associated mud loss rate and volume is also necessary. Fig. 6 shows the sketch of the collated mud loss engineering records.

To define the kind of mud loss, it is typical to categorize the type according to the mud loss rate. In the oil field, the monitoring of mud loss rate mainly depends on the various of outlet mud flow/mud pool volume, and then estimating the mud loss volume per hour based on monitoring results. The mud loss type is classified into seepage loss, partial loss, severe loss, and total loss by Jiang (2006) and Rabia (2002) according to the mud loss rate. Refer to Jiang and Rabia's classification criteria, considering the actual mud loss rate results, the distribution of mud loss volume of each categorized well is shown in Fig. 7. It implies the relationship between mud loss rate and volume is nonlinear, and there is no significant one-one correspondence neither the overall data nor the data for each category, due to the mud loss volume is related to the leakage zone volume. An ideal mud loss records set should not only cover different types of mud loss data at different depths, but also satisfy certain correlation between mud loss rate and mud loss volume. Thus, it is necessary to further deal with the mud loss rate.

To solve this problem, we defined a new parameter—"Mud loss rate per unit footage" to unify the mud loss rate. The mud loss rate per unit footage refers to the mud loss volume in unit drilling time through the unit thief zone, and its calculation formula is as follows:

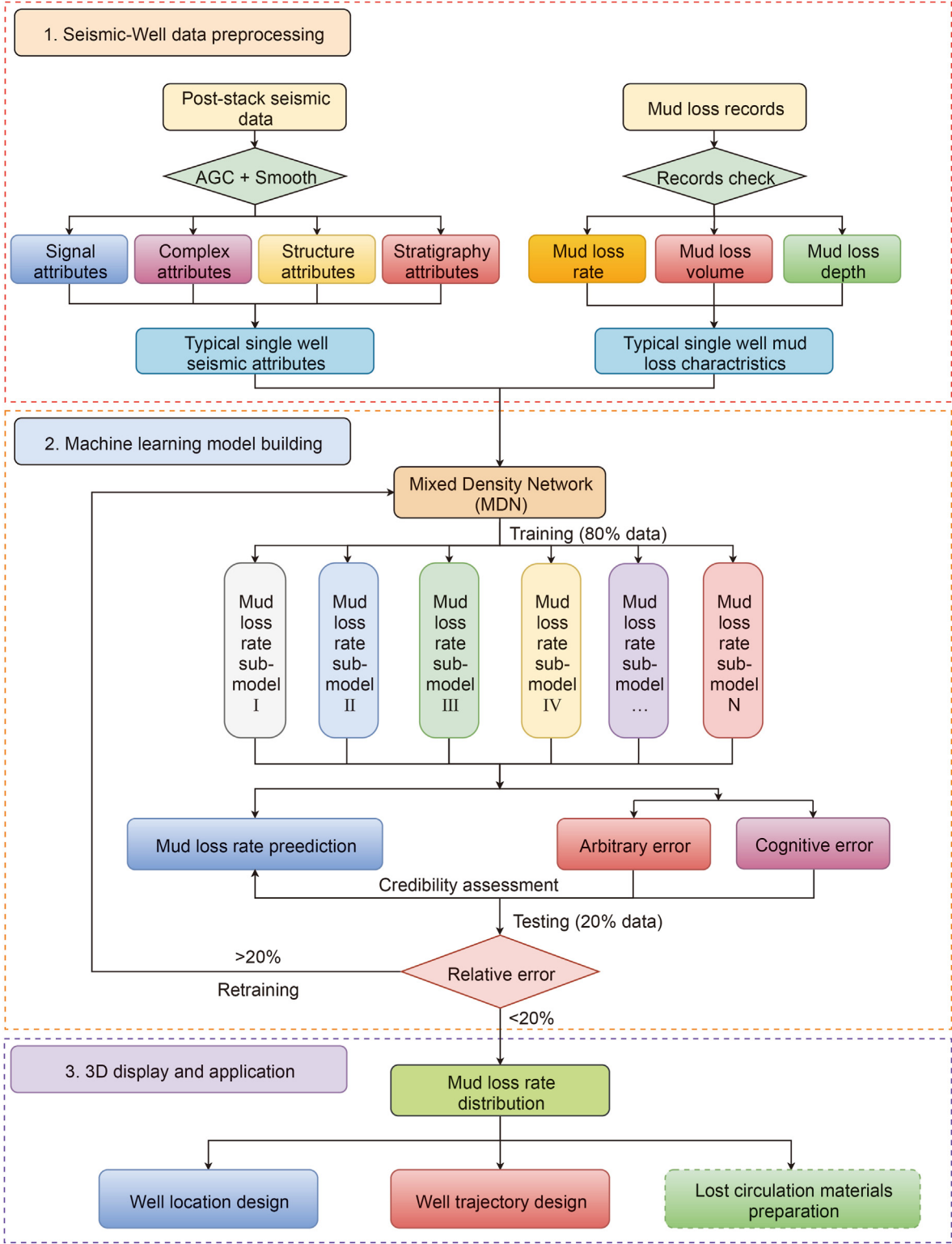


Fig. 1. The workflow of mud loss rate prediction using machine learning.

$$h = H_{\text{down}} - H_{\text{up}}$$

$$v = \frac{q}{\Delta t}$$

$$q = \frac{Q}{h}$$

where, h stands for footage during mud loss, H_{down} stands for the depth of the point when mud loss stop, H_{up} stands for the depth of the point when mud loss start, q stands for the mud loss volume per unit footage, Q stands for the total mud loss volume, of the point

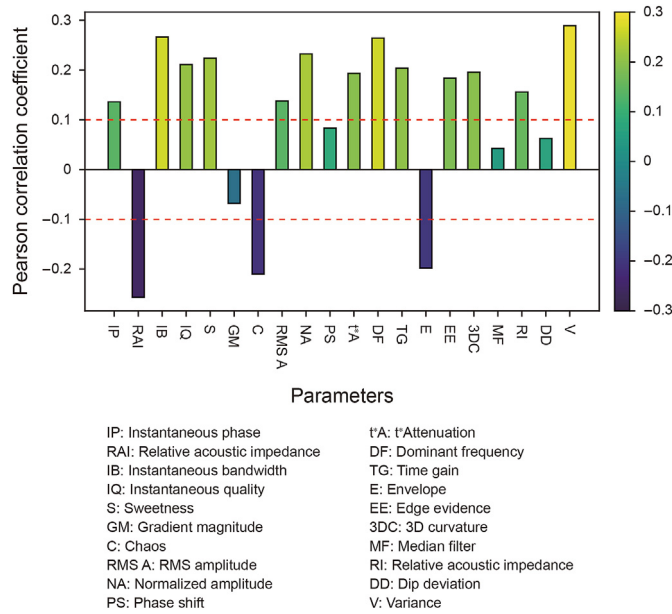


Fig. 2. Selection of seismic attributes.

when mud loss stop, v stands for the mud loss rate per unit footage, Δt stands for the time when drilling through the thief zone.

Fig. 8 illustrates the relationship between the mud loss rate per unit footage and the mud loss volume. Except for discrete data when the mud loss volume of surpasses around 500 m³, the mud loss rate per unit footage and volume for various forms of mud loss type have an approximate linear connection, that the mud loss volume rises linearly with the increase in mud loss rate per unit footage.

Comparing the two figures, using mud loss rate per unit footage to characterize mud loss characteristics has better discrimination,

and the unity of mud loss rate and mud loss volume is realized. Meanwhile, the classification of mud loss rate per unit footage and volume relationship is shown in Table 2.

2.1.3. Calibration of mud loss rate in a single well

Since seismic data are time domain data, and the single well mud loss records are depth domain data, it is necessary to unify them into the same domain to achieve calibration. By convention, seismic data interpreters transform single well data from the depth domain to the time domain using the time-depth relationship. Since the selected study area is a gentle anticlinal carbonate formation (with a dip angle of less than 5°) and no major fault/fractured zone is developed, the seismic data in the time domain are converted to the depth domain for the single well calibration. Fig. 9 shows the calibration of the mud loss type with depth point of the seismic profile of 14 single well with different characteristics in the depth domain.

2.2. Machine learning model

Predictions of mud loss rate are probabilistic in nature. Therefore, the ideal solution to this problem is to construct a model that can characterize the likelihood of the mud loss rate distribution. Neural network, the well-known machine learning model, is a nonlinear statistical modeling tool that is commonly used to simulate complicated relationships between inputs and outputs or to investigate data patterns (Schmidhuber, 2015). It has been widely used to solve many engineering problems including mud loss, thus incorporating uncertainty into the neural network of mud loss prediction may more accurately represent the probability distribution of mud loss rate.

The neural network can approximate any function according to the general approximation theorem; nevertheless, neural networks are good at fitting data with unimodal distribution but poor at fitting data with multimodal distribution. In order to improve the fitting effect of the neural network on data with multimodal

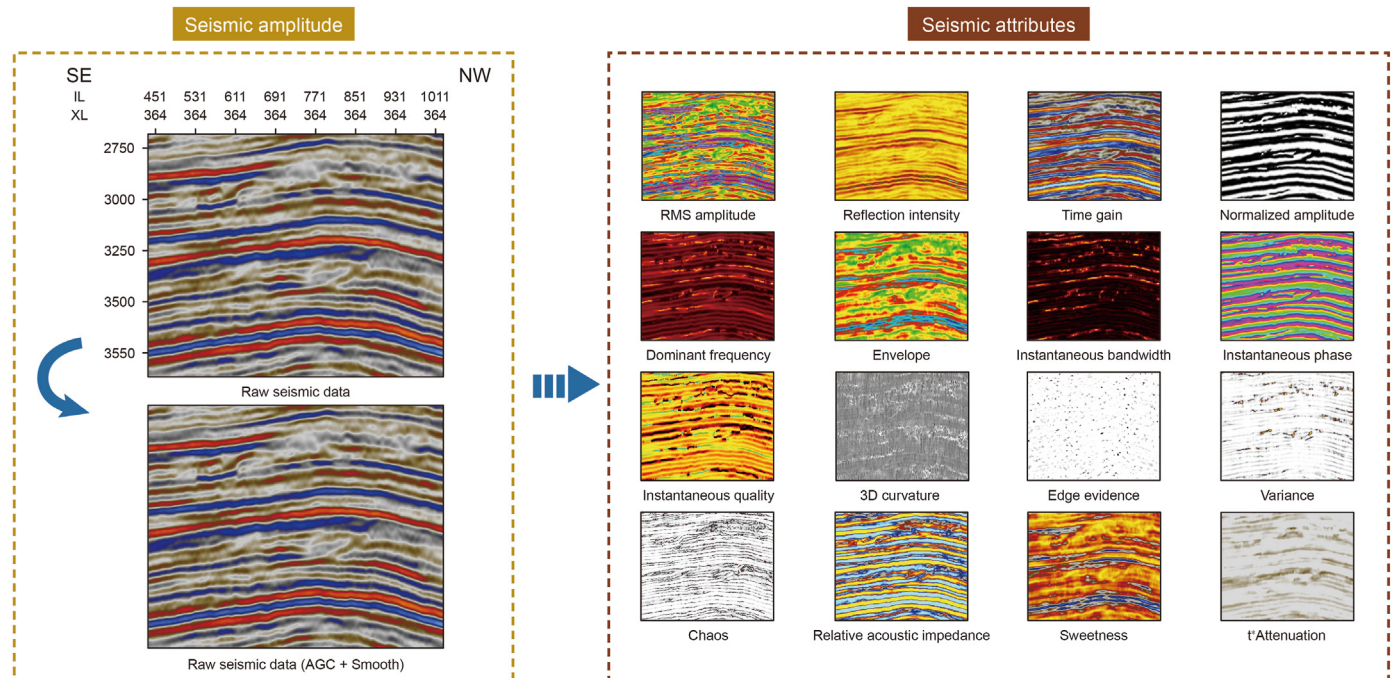


Fig. 3. Selection of seismic attributes.

Table 1
Statistics analysis of 16 attributes from seismic data.

Category	Attribute	Value
Signal processing	RMS amplitude (RMS A)	0–1.98
	Reflection intensity (RI)	0–8.52
	Time gain (TG)	–499013.91–488490.91
Complex	Normalized amplitude (NA)	–1–1
	Dominant frequency (DF)	0–0.91
	Envelope (E)	0–3.48
	Instantaneous bandwidth (IB)	0–0.27
	Instantaneous phase (IP)	–180–180
	Instantaneous quality (IQ)	0–0.29
Structure	3D curvature (3D C)	–3–3
	Edge evidence (EE)	0–50
	Variance (V)	0–1
Stratigraphy	Chaos (C)	0–1
	Relative acoustic impedance (RAI)	–24.33–24.37
	Sweetness (S)	0–3.84
	t*attenuation (t* A)	–0.34–0.27

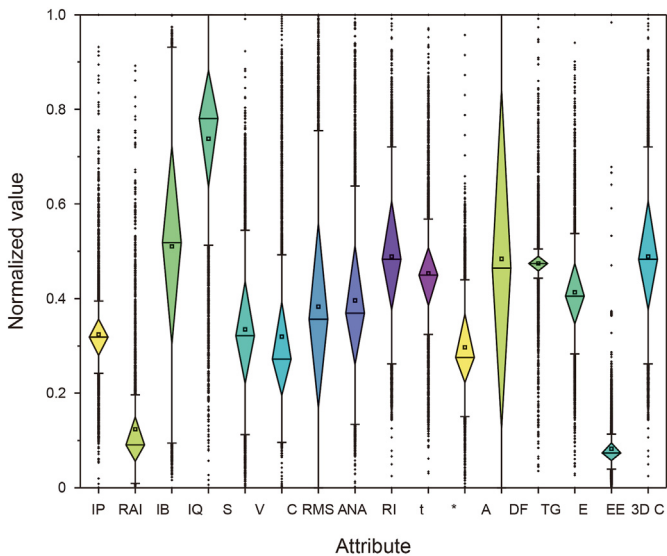


Fig. 4. Normalized distribution of different seismic attributes.

distribution, it can be assumed in the post-processing portion of the neural network that the particular solution of the objective function can be described by the single-peak distribution function, while the general solution can be obtained by stacking multiple single-peak distribution functions, and the weight can be obtained by fitting the actual data. "Mixed Density Network (MDN)" refers to the neural network that was enhanced using the aforementioned technique. The MDN consist of a feed-forward neural network whose outputs determine the parameters in a mixture density model. The mixture model then represents the conditional probability density function of the target variables, conditioned on the input vector to the neural network (Bishop, 1994). It can be used to predict the mud loss rate satisfying the characteristics of multi-peak distribution. The schematic diagram of the MDN is shown in Fig. 10. Like the traditional neural network, there are 3 main parts in the topological structure of the MDN including the input layer, hidden layer, and output layer. The difference is that MDN outputs a probability distribution instead of a constant value.

2.2.1. Mathematical model

The expression of the MDN model is as follows, where the output of a neural network is composed of parameters constructing a Gaussian mixture model (GMM) (Bishop, 1994; Choi et al., 2018):

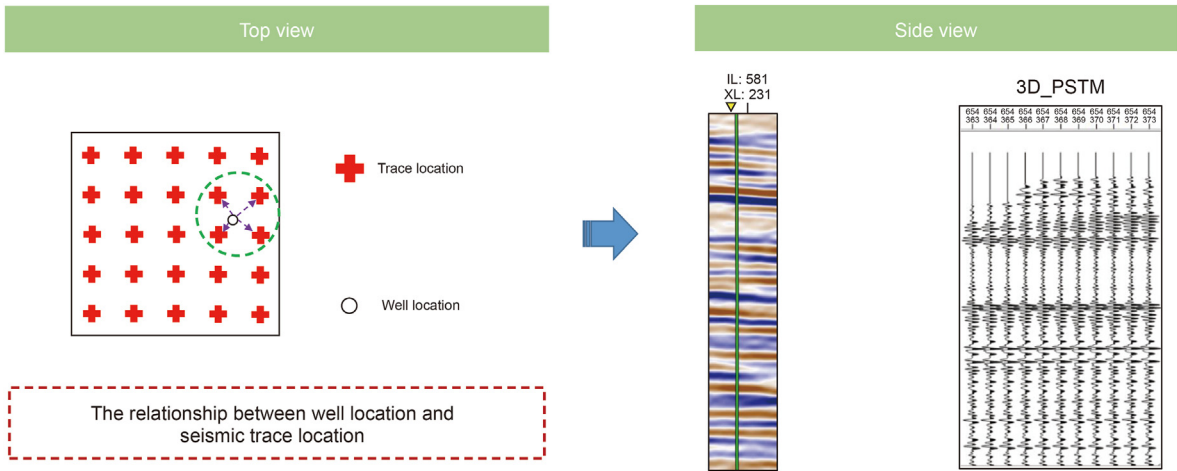


Fig. 5. The extraction of the seismic attribute from a single well.

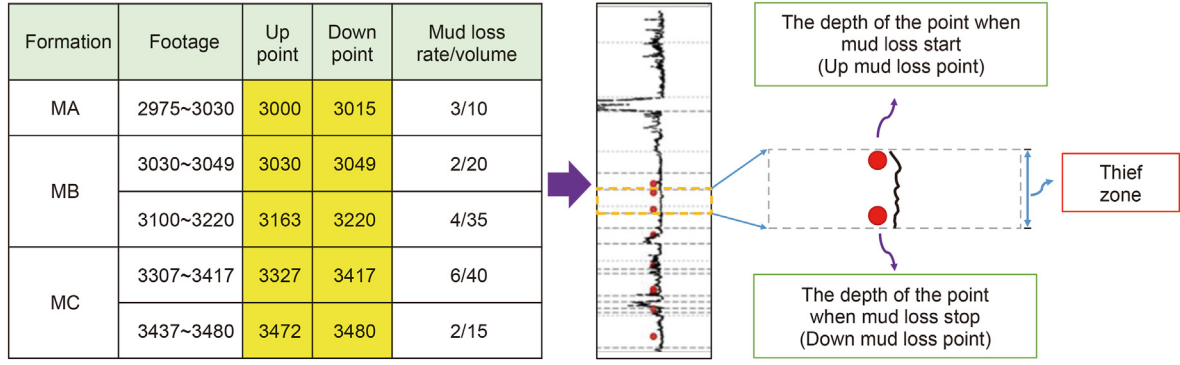


Fig. 6. The collated mud loss engineering records sketch.

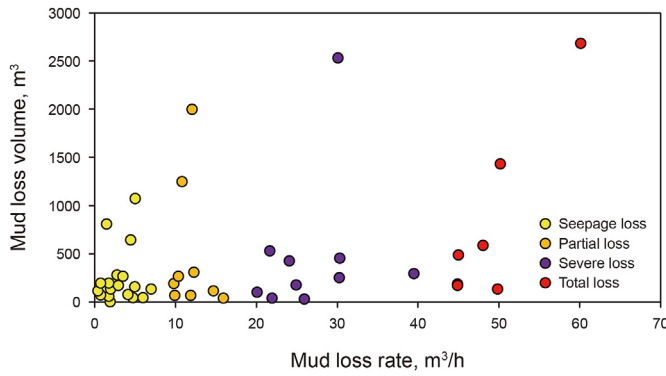


Fig. 7. The relationship between original mud loss rate and mud loss volume.

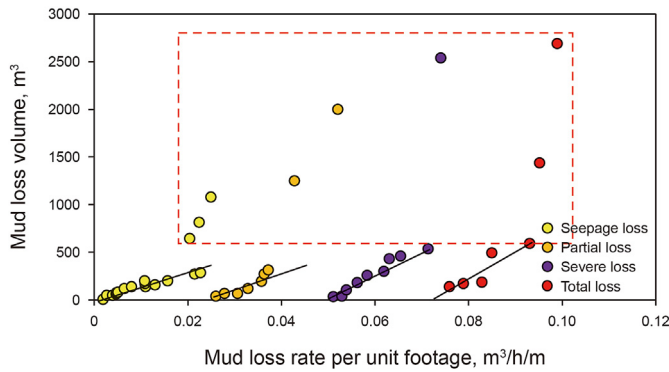


Fig. 8. The relationship between mud loss rate per unit footage and mud loss volume.

Table 2

The relationship mud loss rate per unit footage and mud loss type.

Mud loss rate per unit footage	Mud loss type
0–0.025	Seepage loss
0.026–0.05	Partial loss
0.06–0.075	Severe loss
0.076–0.1	Total loss

$$p(y|x) = \sum_{j=1}^D \omega_j N[y|\mu_j(x), \sigma_j^2(x)] \quad (5)$$

where, $p(y|x)$ is the probability distribution of y given that x take a

particular value, x and y are the seismic attributes and mud loss rate per unit footage defined in this paper, respectively. D is the number of sub-Gaussian, N is the Gaussian probability density function, ω_j , μ_j , $\sigma_j^2(x)$ are the weight, mean, and variance of j -th sub-Gaussian, respectively. The 3 parameters must be met: $\sum_{j=1}^D \omega_j = 1$, $\mu_j(x) > 0$, $\sigma_j^2(x) > 0$.

Considering that MDN is based on maximum likelihood estimation, the training process loss function is defined as the negative log-likelihood function:

$$L(y|x) = -\log \sum_{j=1}^D \omega_j N[y|\mu_j(x), \sigma_j^2(x)] \quad (6)$$

Once an MDN is trained, the predictive mean and variance can be computed by selecting the mean and variance of the mixture of the highest mixture weight.

The prediction error of the MDN can be expressed as follows (Choi et al., 2018):

$$\begin{aligned} E\|y - \hat{f}(x)\|^2 &= E\|y - f(x) + f(x) - \hat{f}(x)\|^2 \\ &= E\|y - f(x)\|^2 + E\|f(x) - \hat{f}(x)\|^2 = \sigma_a^2 + \sigma_e^2 \end{aligned} \quad (7)$$

where, $\hat{f}(x)$ is the value calculated from the machine learning model, $f(x)$ is the value calculated from the ideal condition. σ_a^2 is the aleatoric error, σ_e^2 is the epistemic error, both error means the error between the trained model and the actual model.

Eq. (7) indicates the total predictive variance is the sum of aleatoric uncertainty and epistemic uncertainty.

The mean value of the prediction result can be expressed as follows:

$$E(y|x) = \int y \sum_{j=1}^D \alpha_j N[y|\mu_j(x), \sigma_j^2(x)] dy = \sum_{j=1}^D \alpha_j \mu_j(x) \quad (8)$$

The variance of model prediction can be expressed as follows:

$$\begin{aligned} Var(y|x) &= \int \|y - E(y|x)\|^2 p(y|x) dy \\ &= \sum_{j=1}^D \alpha_j \sigma_j^2(x) + \sum_{j=1}^D \alpha_j \left\| \mu_j - \sum_{j=1}^D \alpha_j \mu_j(x) \right\|^2 \end{aligned} \quad (9)$$

where, α_j is the weight corresponds to the j -th individual Gaussian distribution in Eqs. (8) and (9).

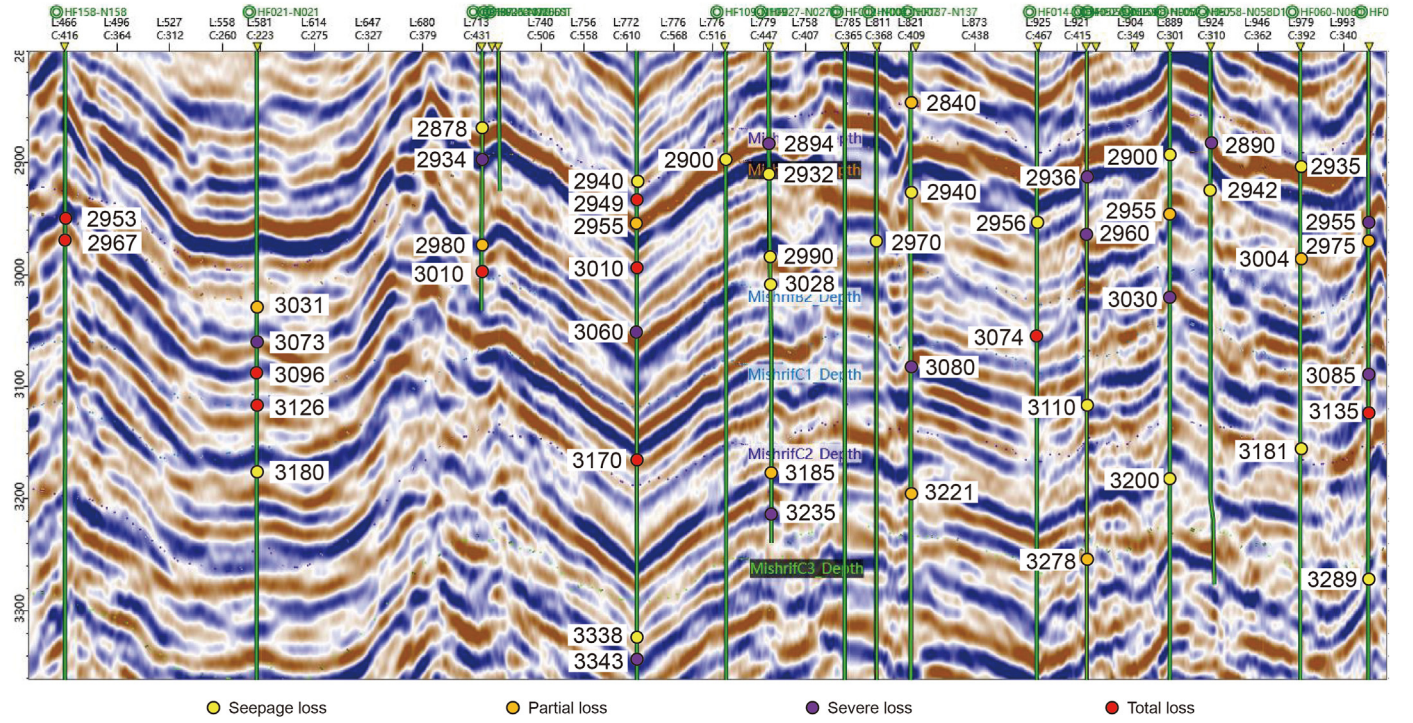


Fig. 9. Calibration of mud loss rate in a single well.

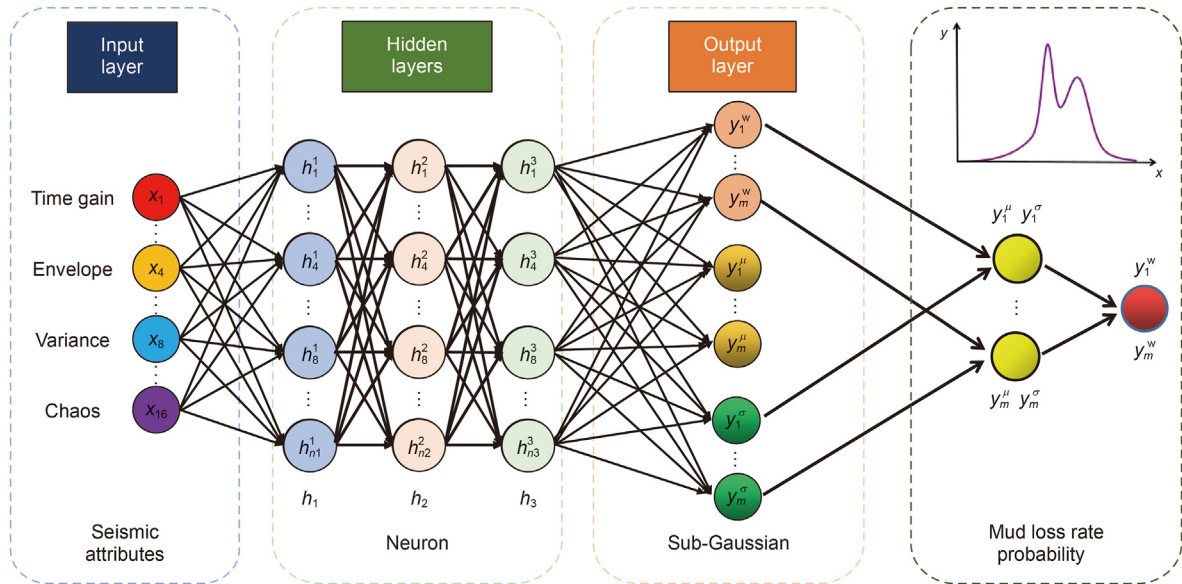


Fig. 10. The schematic diagram of the MDN (modify from Pang et al., 2022).

In Eq. (9), $\sum_{j=1}^D \alpha_j \sigma_j^2(x)$ is equal to σ_a^2 , which means the weighted average of the variance of each sub-Gaussian distribution.

$\sum_{j=1}^D \alpha_j \left\| \mu_j - \sum_{j=1}^D \alpha_j \mu_j(x) \right\|^2$ is equal to σ_e^2 , which describes the credibility of the results predicted by the model. The $\sum_{j=1}^D \alpha_j \sigma_j^2(x)$ expresses the uncertainty of the data itself caused by systematic errors such as noise, and it also can be called random uncertainty.

The $\sum_{j=1}^D \alpha_j \left\| \mu_j - \sum_{j=1}^D \alpha_j \mu_j(x) \right\|^2$ expresses the uncertainty of the model chosen, and it also can be called cognitive uncertainty.

Even if random uncertainty cannot be totally eliminated, cognitive uncertainty can be reduced by gathering more sample data. More sample features can be covered by more training samples, producing more accurate prediction results. These two different types of uncertainty describe the uncertainty of the prediction outcomes. It increases the objectivity of the appraisal of the outcomes.

The Gaussian density function is typically selected as the distribution function, but it should be noted that the MDN with a Gaussian distribution differs from the Gaussian mixture model (GMM). Each MDN model's mean, variance, and weight are produced by the neural network. The weight of the network is

calculated using maximum likelihood estimation as the loss function for backpropagation (to identify a more precise Gaussian distribution parameter), but the mean, variance, and weight for each GMM model are estimated. The typical method for determining this through continuous iteration is the Expectation-Maximum (EM) algorithm. MDN usually does well on regression problems since it is robust and easy to apply.

2.2.2. Model evaluation

For GMN, the number of sub-Gaussians significantly affects the performance of the model. To further evaluate the influence of the number of sub-Gaussians on the model, the mean relative error, time consuming, and R^2 were calculated for the model with varying number of sub-Gaussians. Fig. 11 shows the relationship between the mean relative error, time consuming, and R^2 . As the number of sub-Gaussians increases, the mean relative error decreases significantly. At the same time, the time consuming and R^2 also start to increase significantly. It can be found that when the number of sub-Gaussians is 10, the mean relative error is stable at about 0.075, while R^2 is stable at approximately 0.88. Thereafter, there are no significant changes. However, when the sub-Gaussian is more than 10, the time efficiency considerably increases. It can be concluded that the number of sub-Gaussians is 10, which is a reasonable choice.

In order to further evaluate the prediction effect of MDN model, it is compared with the other two models. Considering that MDN model is an improved model on DNN model and BP model, in order to better highlight the advantages of MDN model, we compare the prediction results of DNN and BP with this model. The parameters used by the three models are shown in Table 3.

Fig. 12 demonstrates the efficacy of the three models. The relative error of the MDN model is 7.5% and the correlation coefficient is 90%, which is 1.5% and 4.5% less than the relative error of the DNN model and the BP model, respectively, but the correlation coefficient is 5% and 10% higher than the other two models. It demonstrates that the MDN model has superior fitting precision and generalizability.

To further evaluate the generalization effect of the model, blind well H1 was selected for verification. Fig. 13 illustrated the relative error in well H1 at various depths. The synthesis error is about 20%, and there are 4 high error zones in each specific depth interval. The relative error in these 4 zones is larger than 30%, and the relative errors in zones 1 and 4 are relatively discrete, but those in zones 2 and zone 3 are centralized. Due to the limitation of the resolution of seismic data, when the accuracy is assessed in meters, relatively high and concentrated relative errors will be presented in areas prone to mud loss, as shown in zones 2 and 3 in the figure.

To evaluate the reliability of the prediction results of the model,

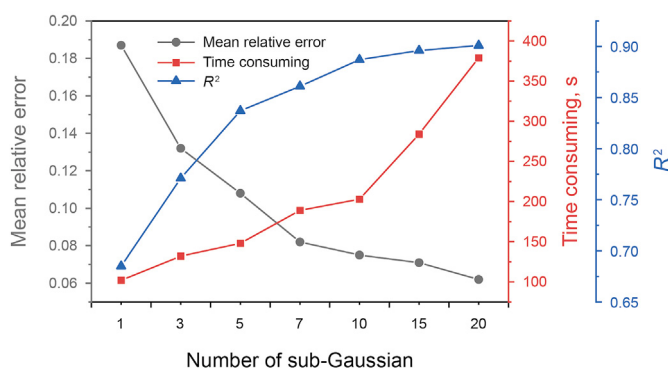


Fig. 11. The evaluation of sub-Gaussian number.

it is necessary to determine the threshold values of the parameters of random uncertainty and cognitive uncertainty. Ten blind wells were selected to evaluate the relative errors of prediction results under different uncertain parameters. As shown in Fig. 14, combined with the previous comprehensive prediction results of the single well, it is found that when the random uncertainty is greater than 0.2 and the cognitive uncertainty is greater than 0.07, the relative error will exceed 20%. When the stochastic uncertainty exceeds 0.20 and the cognitive uncertainty exceeds 0.07, it is considered that the prediction results will no longer be credible.

3. Result

3.1. Mud loss rate prediction in different layers

The M formation can be divided into three sets of layers from top to bottom, namely MA, MB, and MC. In further compare and analyze the difference between the mud loss rate prediction results of different small layers, depth slices were created in the middle of each set of small layers, and the original seismic amplitude data, leakage rate prediction results, corresponding random uncertainty and cognitive uncertainty were compared. At the same time, a well that was not included in the modeling was also drilled. Fig. 15(a), (b) and (c) show the depth slices of the original post-stack seismic data in the middle of the three small layers. There is little difference in the general trend of the three maps, and they all have obvious anticlinal structures, but it is difficult to determine the locations with high loss risk. In contrast, Fig. 15(d), (e) and (f) predict the distribution of mud loss rate per unit footage in different regions. Compared to the MA formation, the MB and MC formations have a broader area of high mud loss rate distribution. Combined with the evaluation results of random uncertainty in Fig. 15(g), (h) and (i) and cognitive uncertainty in Fig. 15(j), (k) and (l), the wells H₂, H₃ and H₇ have higher mud loss risk especially in the MB and MC formation.

Wells located in high loss zones were extracted and compared to actual drilling data, as shown in Table 4. The mud loss rate per unit footage can be used to evaluate the level of mud loss rate, and random uncertainty and cognitive uncertainty can be used to evaluate the confidence coefficient of mud loss rate. As can be seen in the table, the predicted results are in relatively good agreement with the actual data, and the risk of mud loss should be considered while designing well location in the above regions.

3.2. Mud loss rate prediction in longitudinal profile

The mud loss risk along the wellbore trajectory was predicted by a blind well's mud loss rate per unit footage profile, cognitive uncertainty section and random uncertainty. The M formation is the typical vug-fractured limestone reservoirs almost without caves and faults. As a result, the original seismic data show no significant differences in each small layer of the M formation. The mud loss rate per unit footage profile, similar to the plane distribution diagram, displays changes in the distribution of mud loss rate at different depths and layers. Fig. 16 depicts the mud loss rate per unit footage section, random uncertainty section, and cognitive uncertainty section for a directional well. It demonstrates that the MB and MC layers have a relatively high risk of mud loss. According to the mud loss rate per unit footage section, the well will encounter severe mud loss in 3000 m of MB and 3100 m of MC. The random uncertainty is 0.15 and 0.1, and the cognitive uncertainty is 0.03 and 0.02, respectively. The predictions are trustworthy since the values of the two types of uncertainty are low. The actual drilling data shows that different mud loss occurred in MB and MC layers. Conventional well logging interpretation reveals anomalies

Table 3
Parameters of different models.

Algorithms	Parameters
MDN	Laraning_rate = 0.003, Activation_function = 'Relu', Hidden_layers = [64,128,64]
DNN	Laraning_rate = 0.003, Activation_function = 'Relu', Hidden_layers = [64,128,64]
BP	Laraning_rate = 0.01, Activation_function = 'Relu', Hidden_layers = [128]

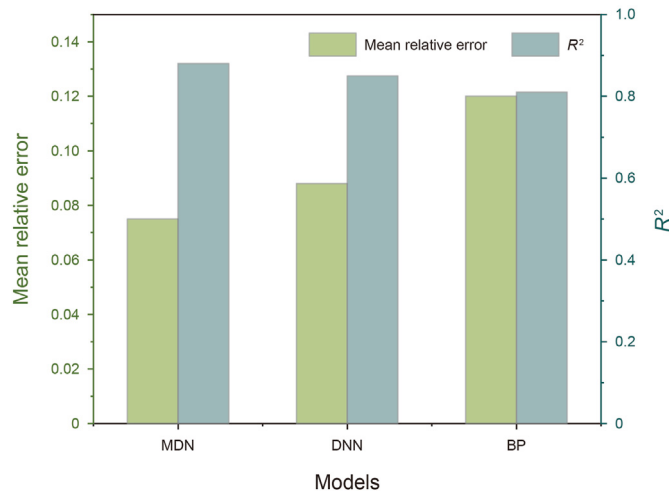


Fig. 12. Comparison of training and testing results of different models.

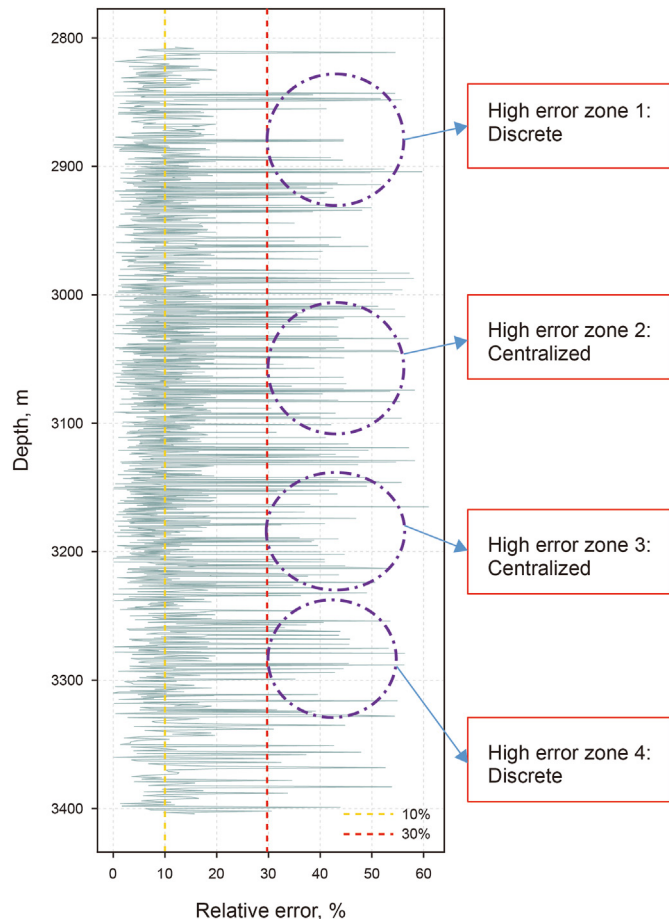


Fig. 13. The evaluation of sub-Gaussian number.

in these two layers, implying the presence of natural fractures. Combined with the case study, the well geological design will inform the formation of the possible risk of mud loss when designing a well trajectory, but it will frequently fail to estimate the probability of mud loss. The MDN model is used to calculate the mud loss rate per unit footage, random uncertainty, and cognitive uncertainty, which can provide guidance for the prediction of well design.

4. Discussion

4.1. The relationship between well location and prediction reliability

When calibrating the mud loss of typical well, the vertical well trajectory usually does not completely coincide with the seismic trace. Because there is no data between two seismic traces, it is customary to match the seismic trace data of the nearest borehole in general (Fig. 17(a)). This leads to some uncertainty in the final prediction results, so a central question to focus on how reliable the prediction results are. In the pre-drilling design of a well, multiple offset wells are usually used as references. On the premise that there is no obvious anomaly in the seismic section, it is considered that the closer the well is to be drilled, the higher the reference value. Considering the corresponding relationship between the well and seismic trace, the reference drilled well and the planned well can be divided into a fan-shaped area composed of four seismic measuring trace to discuss the credibility problems caused by the distance between them (Fig. 17(b)). After the locations of the drilled well and the planned well are determined, the seismic section along wells is extracted for further analysis (Fig. 17(c)).

Based on the above assumptions, the problem can be simplified into two cases.

4.1.1. Case 1: the planned well is closer to the seismic trace than the drilled well

The waveforms and predictions of the four seismic traces closest to the planned well, as well as the predictions of the planned well and drilled well, are extracted, and compared to the actual data. As seen from Fig. 18, the mud loss prediction of the seismic trace closest to the planned well, the planned well, and the drilled well are similar and consistent with the actual drilling results, however, the mud loss prediction of the other three seismic traces differs in various ways. The predictions between seismic traces are displayed by interpolation, and their values are mainly affected by the variation trend of seismic trace data. The trend of the four seismic traces closest to the planned well is similar, so the predictions are relatively stable. When the distance between the planned well and the seismic trace is close, the prediction can be used for drilling design.

4.1.2. Case 2: the planned well is closer to the drilled well than the seismic trace

As in case 1, another blind well was selected for analysis in conjunction with seismic trace data. The waveforms and predictions of the four seismic traces closest to the planned well, as

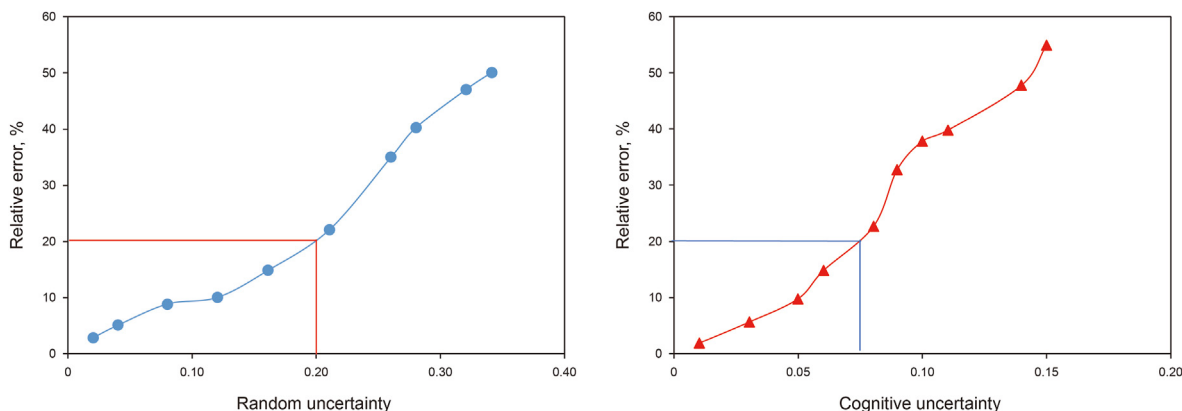


Fig. 14. Evaluation of uncertainty parameters.

well as the predictions of the planned well and drilled well are extracted and compared to the actual data. Fig. 19 shows that the seismic trace prediction closest to the planned well is basically consistent with the planned well prediction. The prediction is more accurate and covers a wider area than the reference well (drilled well), even though the planned well is closer to the drilled well.

According to the above analysis, none of the above cases have a significant impact on mud loss prediction. The reason is that the object region studied in this paper is a gentle anticline structure with a dip angle of less than 5° and with no large faults or fracture zones, and there is basically no stratigraphic unconformity in the M formation. In addition, in all drilled well data, the distance between the two nearest wells is about 20 m, while the distance between the inlines and the crosslines are both 25 m. The particularity of the above conditions greatly reduces the uncertainty of the prediction. However, if the formation has a complex structure and is sedimentary and the distance between two survey line is large, further discussion is needed to obtain more general conclusions.

4.2. The effect of time-depth relation on prediction accuracy in the depth domain

Mud loss engineering records are depth-domain data, and time-depth conversion during calibration is problematic. Ideal time-depth conversion would employ a high-precision velocity volume from geophysical interpreters, which would greatly eliminate the bias generated during the conversion. In most cases, it is impossible to obtain high-precision velocity volume, which is usually generated by the interpolation method through the time-depth relationship established by several single wells. The prediction may produce relatively server errors in local areas, especially in formations with complex tectonic processes. Fig. 20 shows the effect of synthetic records of well H_2 and the prediction of a blind well H_9 . Due to the differences between the synthetic record and the actual seismic record (which cannot be one-to-one correspondences), errors will be generated during deep calibration. Even one-to-one correspondences may also be unreasonable and incorrect. In addition to logging and seismic data quality issues, the extent to which errors can be reduced depends on the experience of geophysical/geological interpreters, which is also a kind of cognitive uncertainty. The blind well H_9 of Fig. 20 shows the deviation of the prediction caused by such an error. The prediction is about 10 m away from the actual deviation. Currently, this problem is solved by constantly checking the time-depth relationship. For example, when the first drilling hole is completed, the new logging data is utilized to correct the combined time-depth relationship in

conjunction with the comprehensive logging calibration results. There is a lot of discussion about whether it is more reasonable to interpret and process seismic data in the time domain or the depth domain, but no conclusion has yet been reached.

4.3. Uncertainty analysis of lost circulation records

For an ideal training set of mud loss record samples, the best case is that each seismic data and each mud loss record can be matched one by one to a certain extent. This one-to-one mapping relationship will greatly improve the prediction accuracy, but in fact, this is not the case. Since the distribution range and characteristics of the mud loss record in this paper are not ideal (Fig. 6), relatively concentrated, and the degree of differentiation is not obvious, we are unable to use the original seismic amplitude data directly. Instead, 16 seismic attributes are used to achieve the purpose of expansion of seismic data characteristics, which helps to partially solve the problem. The 16 seismic attributes selected in this paper are mainly based on the seismic interpretation experience of the H oilfield and the correlation analysis with leakage. However, due to the wide variety of seismic attribute bodies and the huge difference between different oilfield strata, how many seismic attribute bodies should be selected is a problem that needs further study and discussion. We believe that a reasonable method needs to balance the relationship between correlation, computational efficiency, and accuracy, and form a complete set of technical processes.

5. Conclusions

- (1) The accuracy and impact of mud loss prediction based on seismic data may be significantly increased by integrating multiple seismic attributes with optimized mud loss engineering records and applying machine learning methods.
- (2) In comparison to the general mud loss prediction model, the MDN model can not only predict the mud loss rate but also evaluate the results from the two aspects of data quality and model quality, namely random uncertainty, and cognitive uncertainty, to predict the mud loss situation more objectively.
- (3) Due to different stratigraphic conditions, the major method to increase prediction accuracy is to improve the quality and matching relationship of seismic data, mud loss records, and other multi-source data, including time-depth relationship, and select a reasonable model.

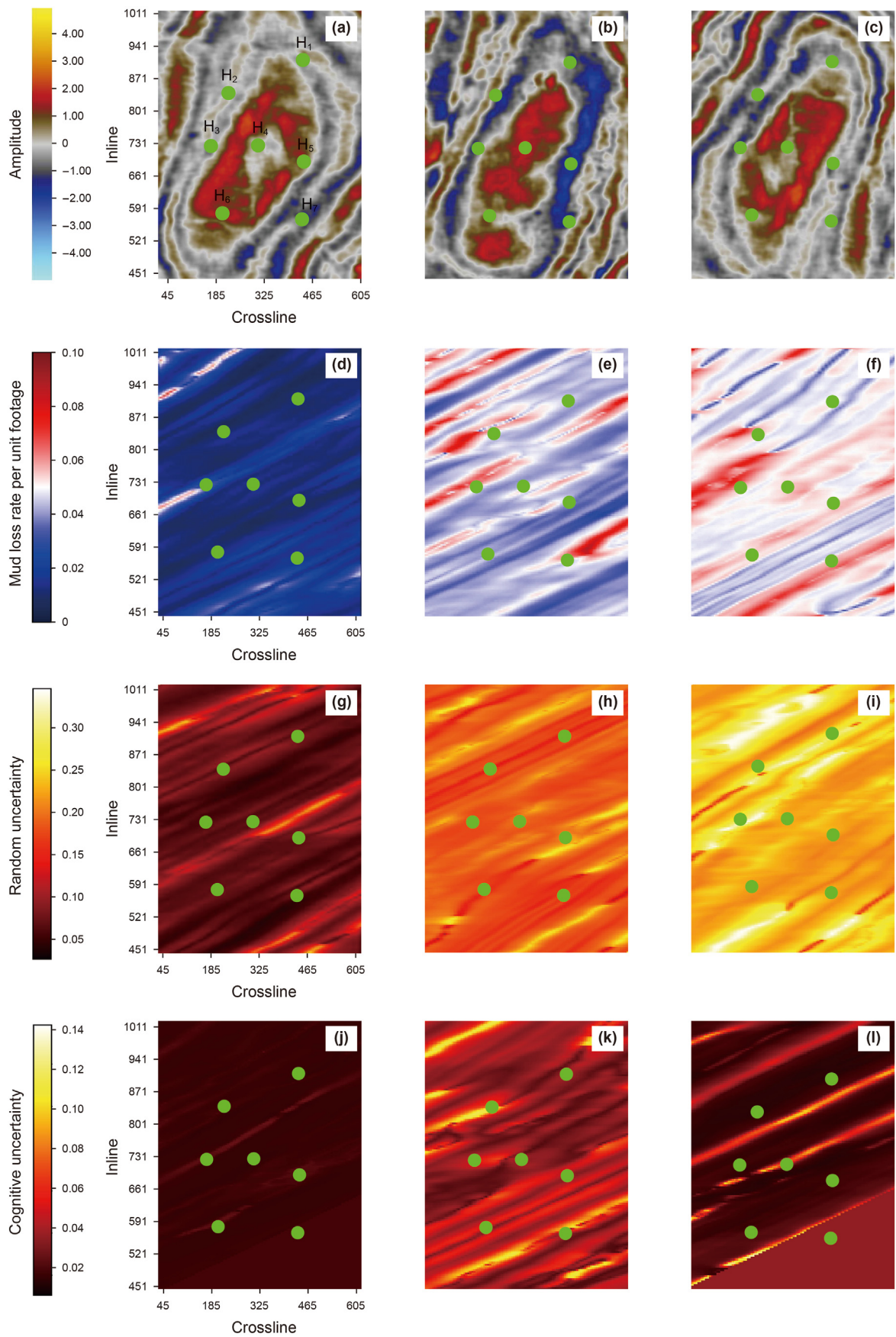


Fig. 15. Original seismic amplitude slice ((a), (b) and (c)) and corresponding prediction of mud loss rate distribution per unit footage((d), (e) and (f)), random uncertainty ((g), (h), (i)), and cognitive uncertainty((j), (k), (l)).

Table 4
Statistics analysis of 16 attributes from seismic data.

Well Name	Prediction (MA/MB/MC)			Result
	Mud loss rate per unit footage	Random uncertainty	Cognitive uncertainty	
H ₂	0.03/0.10/0.08	0.02/0.15/0.20	0.01/0.12/0.01	The mud loss rate of 15 m ³ /h occurs in the MB formation, with a total loss of 30 m ³ . The mud loss rate of 20 m ³ /h occurs in the MC formation, with a total loss of 10 m ³ .
H ₃	0.05/0.08/0.08	0.05/0.15/0.20	0.03/0.06/0.03	The mud loss rate of 25 m ³ /h occurs in the MB formation, with a total loss of 35 m ³ .
H ₇	0.01/0.1/0.08	0.02/0.15/0.20	0.00/0.03/0.04	The mud loss rate of 10 m ³ /h occurs in the MC formation, with a total loss of 40 m ³ .

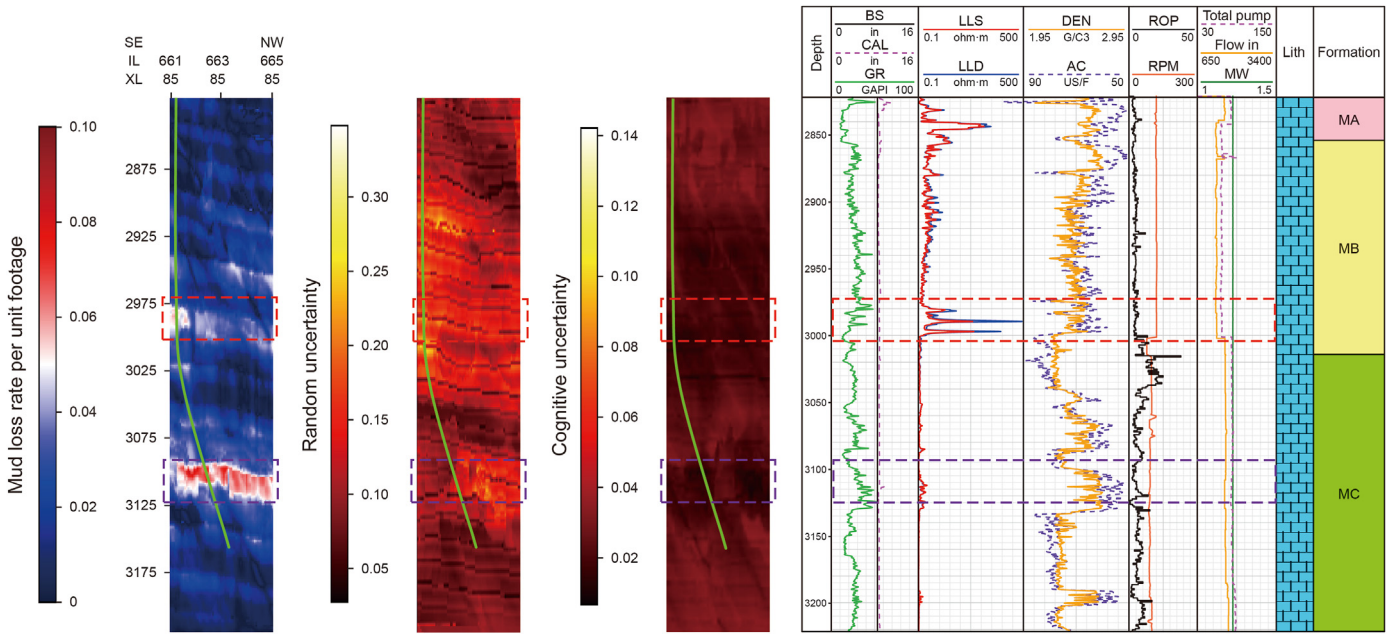


Fig. 16. Single well mud loss risk assessment and validation.

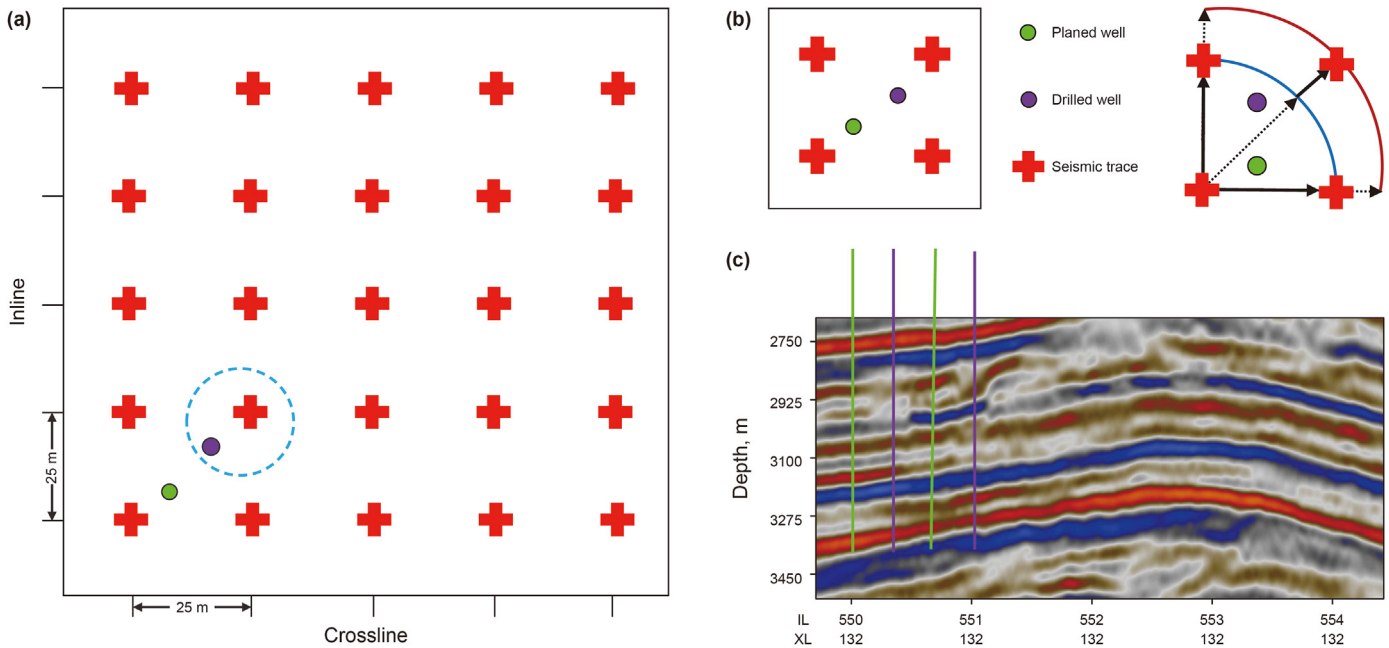


Fig. 17. Single well lost circulation risk assessment and design.

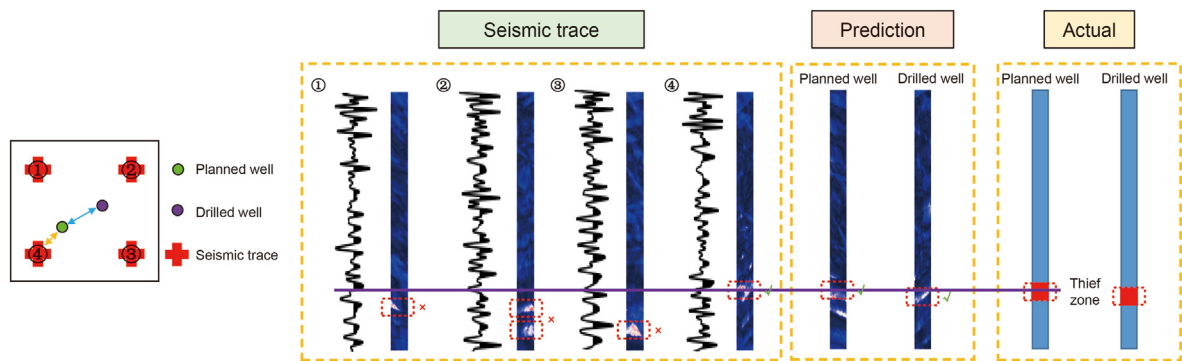


Fig. 18. Reliability analysis of single well prediction when plan well closer than seismic trace.

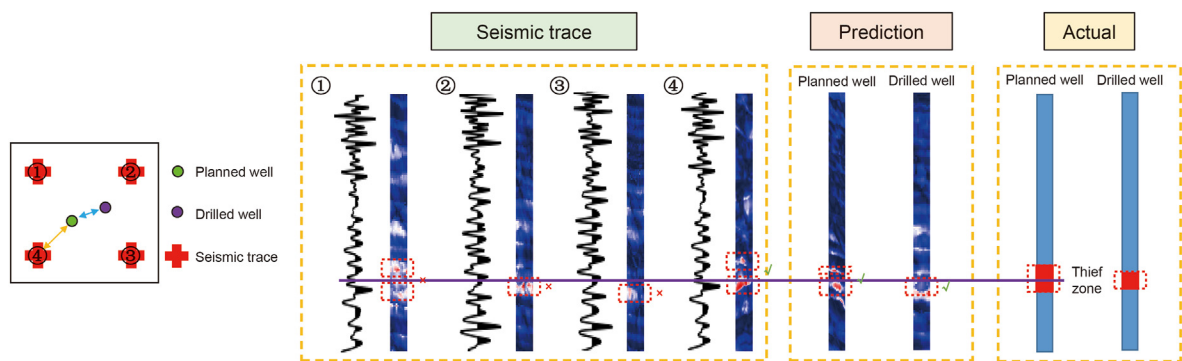


Fig. 19. Reliability analysis of single well prediction when plan well closer than drilled well.

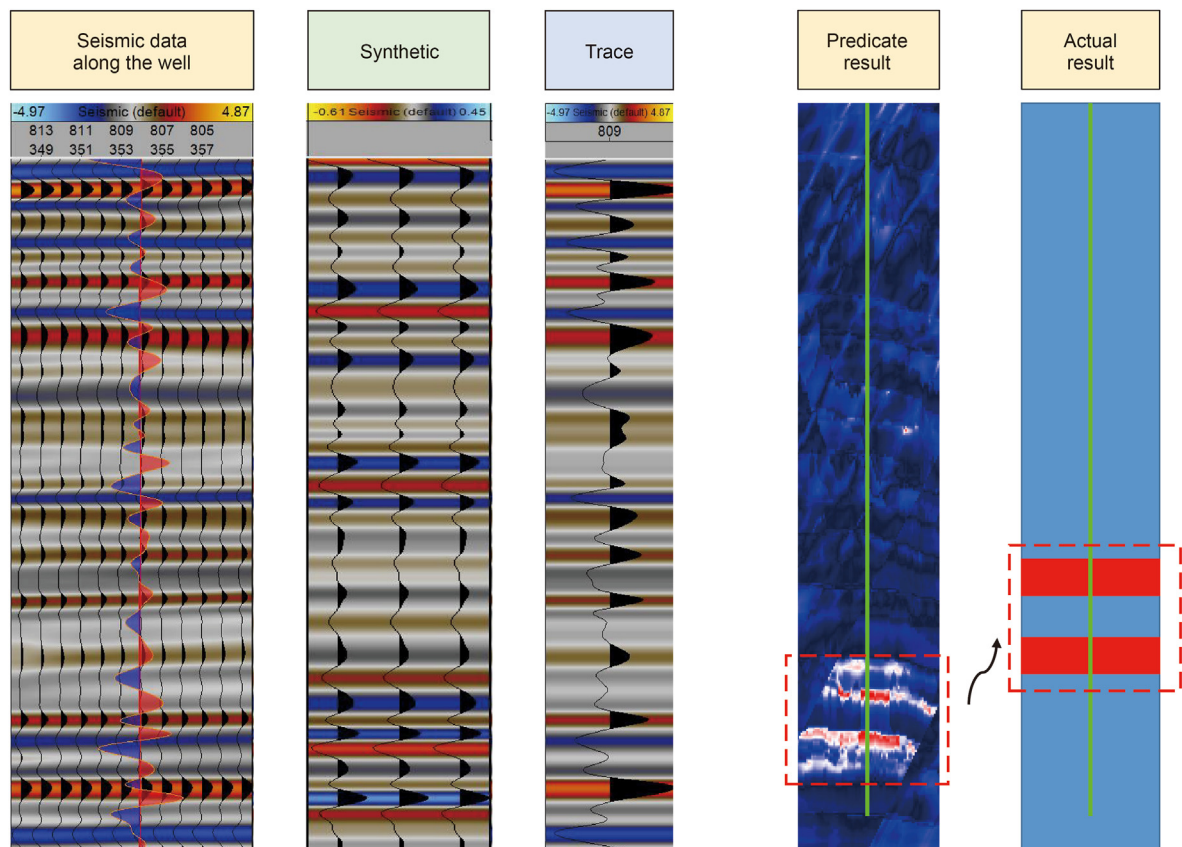


Fig. 20. The effect of synthetic records on prediction.

CRediT authorship contribution statement

Hui-Wen Pang: Conceptualization, Formal analysis, Funding acquisition, Writing – original draft, Writing – review & editing. **Han-Qing Wang:** Conceptualization, Data curation, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Yi-Tian Xiao:** Supervision. **Yan Jin:** Project administration, Resources, Supervision, Validation. **Yun-Hu Lu:** Project administration, Resources. **Yong-Dong Fan:** Data curation, Methodology, Software, Visualization. **Zhen Nie:** Project administration, Resources.

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.

Acknowledgments

The authors would like to acknowledge gratefully the financially supported by the National Natural Science Foundation of China (Grant No. 52104013) and the China Postdoctoral Science Foundation (Grant No. 2022T150724).

References

- Abbas, A.K., Bashikh, A.A., Abbas, H., et al., 2019. Intelligent decisions to stop or mitigate lost circulation based on machine learning. *Energy* 183, 1104–1113. <https://doi.org/10.1016/j.energy.2019.07.020>.
- Ali, P.J.M., Faraj, R.H., Koya, E., et al., 2014. Data normalization and standardization: a technical report. *Mach Learn Tech Rep* 1 (1), 1–6.
- Anderson, R.A., Ingram, D.S., Zanier, A.M., 1973. Determining fracture pressure gradients from well logs. *J. Petrol. Technol.* 25 (11), 1259–1268. <https://doi.org/10.2118/4135-PA>.
- Ahdaya, M., Al Ibrahim, A., Bai, B., et al., 2022. Low-temperature recrosslinkable preformed particle gel as a material for lost circulation control. *SPE J.* 1–11. <https://doi.org/10.2118/209783-PA>.
- Bishop, C.M., 1994. Mixture Density Networks. Technical Report NCRG/97/004, Neural Computing Research Group, Aston University.
- Brown, A.R., 1996. Seismic attributes and their classification. *Lead. Edge* 15 (10), 1090. <https://doi.org/10.1190/1.1437208>.
- Chan, A.W., Hauser, M., Couzens-Schultz, B.A., et al., 2013. An alternative interpretation of leakoff and lost circulation pressure measurements. In: The 47th U.S. Rock Mechanics/Geomechanics Symposium, 23–26 June. ARMA-2013-302.
- Chen, Q., Sidney, S., 1997. Seismic attribute technology for reservoir forecasting and monitoring. *Lead. Edge* 16 (5), 445–448. <https://doi.org/10.1190/1.1437657>.
- Choi, S., Lee, K., Lim, S., et al., 2018. Uncertainty-aware learning from demonstration using mixture density networks with sampling-free variance modeling. In: 2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE, pp. 6915–6922.
- Chopra, S., Marfurt, K.J., 2005. Seismic attributes—a historical perspective. *Geophysics* 70 (5), 350–2850. <https://doi.org/10.1190/1.2098670>.
- Civan, F., Rasmussen, M.L., 2002. Further discussion of fracture width logging while drilling and drilling mud/loss-circulation-material selection guidelines in naturally fractured reservoirs. *SPE Drill. Complet.* 17 (4), 249–250.
- Cook, J., Growcock, F., Guo, Q., et al., 2011. Stabilizing the wellbore to prevent lost circulation. *Oilfield Rev.* 23 (4), 26–35.
- Camacho-Velázquez, R., Vásquez-Cruz, M., Castrejón-Aívar, R., et al., 2005. Pressure-transient and decline-curve behavior in naturally fractured vuggy carbonate reservoirs. *SPE Reservoir Eval. Eng.* 8 (2), 95–112. <https://doi.org/10.2118/77689-PA>.
- Daines, S.R., 1982. Prediction of fracture pressures for wildcat wells. *J. Petrol. Technol.* 34 (4), 863–872. <https://doi.org/10.2118/9254-PA>.
- Ding, Y., Cui, M., Wang, H., et al., 2021. Predicting seismic-based anisotropy for prevent pre-drill risk using a novel type neural network. In: The SPE/IATMI Asia Pacific Oil & Gas Conference and Exhibition, Virtual, 12–14 October. <https://doi.org/10.2118/205710-MS>.
- Dyke, C.G., Wu, B., Milton-Taylor, D., 1995. Advances in characterising natural fracture permeability from mud log data. *SPE Form. Eval.* 10 (3), 160–166. <https://doi.org/10.2118/25022-PA>.
- Eaton, B.A., 1969. Fracture gradient prediction and its application in oilfield operations. *J. Petrol. Technol.* 21 (10), 1353–1360. <https://doi.org/10.2118/2163-PA>.
- Fomel, S., 2007. Local seismic attributes. *Geophysics* 72 (3), A29–A33. <https://doi.org/10.1190/1.2437573>.
- Gao, J., Lin, H., Sun, J., et al., 2021. A porothermoelastic model considering the dynamic temperature-perturbation boundary effect for borehole stability in fractured porous rocks. In: The SPE/AAPG/SEG Asia Pacific Unconventional Resources Technology Conference, Virtual, 16–18 November. <https://doi.org/10.15530/AP-URTEC-2021-208266>.
- Geng, Z., Wang, H., Fan, M., et al., 2019. Predicting seismic-based risk of lost circulation using machine learning. *J. Pet. Sci. Eng.* 176, 679–688. <https://doi.org/10.1016/j.petrol.2019.01.089>.
- Gulbransen, A.F., Hauge, V.L., Lie, K.A., 2010. A multiscale mixed finite-element method for vuggy and naturally fractured reservoirs. *SPE J.* 15 (2), 395–403. <https://doi.org/10.2118/119104-PA>.
- Hauser, M.R., 2021. Evaluating an integrated physical model for borehole-leakoff pressures. *SPE J.* 26 (6), 3389–3409. <https://doi.org/10.2118/204453-PA>.
- Huang, R., 1984. A model for prediction formation fracture pressure. *Journal of East China Petroleum Institute* 4, 335–347 (in Chinese).
- Hubbert, M.K., Willis, D.G., 1957. Mechanics of hydraulic fracturing. *Transactions of the AIME* 210 (1), 153–168. <https://doi.org/10.2118/686-G>.
- Hou, X., Yang, J., Yin, Q., et al., 2020. Lost circulation prediction in South China Sea using machine learning and big data technology. In: The Offshore Technology Conference, 4–7 May. <https://doi.org/10.4043/30653-MS>.
- Jiang, X., 2006. Drilling Accidents and Complex Problems. Petroleum Industry Press, Beijing, p. 439 (in Chinese).
- Jin, J., Jin, Y., Lu, Y., et al., 2021. Image processing and machine learning based cavings characterization and classification. *J. Pet. Sci. Eng.*, 109525. <https://doi.org/10.1016/j.petrol.2021.109525>.
- Lavrov, A., Tronvoll, J., 2004. Modeling mud loss in fractured formations. In: The Abu Dhabi International Conference and Exhibition, Abu Dhabi, United Arab Emirates, 10–13 October. <https://doi.org/10.2118/88700-MS>.
- Li, X., Zhao, Q., Zhang, J., et al., 2019. A prediction model for extension limit of horizontal drilling based on leakage pressure in carbonate formations. In: The 53rd U.S. Rock Mechanics/Geomechanics Symposium, 23–26 June. ARMA-2019-1872.
- Li, Z., Chen, M., Jin, Y., et al., 2018. Study on intelligent prediction for risk level of lost circulation while drilling based on machine learning. In: The 52nd U.S. Rock Mechanics/Geomechanics Symposium, Seattle, Washington, USA, 17–20 June. ARMA-2018-105.
- Liétard, O., Unwin, T., Guillot, D.J., Hodder, M.H., 1999. Fracture width logging while drilling and drilling mud/loss-circulation-material selection guidelines in naturally fractured reservoirs. *SPE Drill. Complet.* 14 (3), 168–177. <https://doi.org/10.2118/57713-PA>.
- Majidi, R., Miska, S.Z., Yu, M., et al., 2010. Quantitative analysis of mud losses in naturally fractured reservoirs: the effect of rheology. *SPE Drill. Complet.* 25 (4), 509–517. <https://doi.org/10.2118/114130-PA>.
- Matthews, W.R., Kelly, J., 1967. How to predict formation pressure and fracture gradient from electric and sonic logs. *Oil Gas J.* 92–106.
- Morita, N., Black, A.D., Guh, G.F., 1990. Theory of lost circulation pressure. In: The SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, 23–26 September. <https://doi.org/10.2118/20409-MS>.
- Noshi, C.I., Schubert, J.J., 2018. The Role of Machine Learning in Drilling Operations: A Review. In: The Role of Machine Learning in Drilling Operations: A Review. The SPE/AAPG Eastern Regional Meeting, Pittsburgh, Pennsylvania, USA 7–11 October. <https://doi.org/10.2118/191823-18ERM-MS>.
- Özkaya, S.I., 2003. Fracture length estimation from borehole image logs. *Math. Geol.* 35 (6), 737–753. <https://doi.org/10.1023/B:MATG.0000002987.69549.ba>.
- Pang, H., Meng, H., Wang, H., et al., 2022. Lost circulation prediction based on machine learning. *J. Pet. Sci. Eng.* 208, 109364. <https://doi.org/10.1016/j.petrol.2021.109364>.
- Rabia, H., 2002. Well Engineering & Construction. Entrac Consulting Limited, London, pp. 288–289.
- Sabah, M., Mehrad, M., Ashrafi, S.B., et al., 2021. Hybrid machine learning algorithms to enhance lost circulation prediction and management in the Marun Oil Field. *J. Pet. Sci. Eng.* 198, 108125. <https://doi.org/10.1016/j.petrol.2020.108125>.
- Sawaryn, S.J., 2001. Discussion of fracture width logging while drilling and drilling mud/loss-circulation-material selection guidelines in naturally fractured reservoirs. *SPE Drill. Complet.* 16 (4), 268–269.
- Schmidhuber, J., 2015. Deep learning in neural networks: an overview. *Neural Network*. 61, 85–117. <https://doi.org/10.1016/j.neunet.2014.09.003>.
- Shi, L., Jiang, H., Guo, Q., et al., 2012. Investigation on the differential pressure leakage of thief formation. In: The IADC/SPE Asia Pacific Drilling Technology Conference and Exhibition, 9–11 July. <https://doi.org/10.2118/155699-MS>.
- Tan, P., Jin, Y., Pang, H., 2021. Hydraulic fracture vertical propagation behavior in transversely isotropic layered shale formation with transition zone using XFEM-based CZM method. *Eng. Fract. Mech.* 248, 107707. <https://doi.org/10.1016/j.engfracmech.2021.107707>.
- Tan, P., Pang, H., Zhang, R., et al., 2020. Experimental investigation into hydraulic fracture geometry and proppant migration characteristics for southeastern Sichuan deep shale reservoirs. *J. Pet. Sci. Eng.* 184, 106517. <https://doi.org/10.1016/j.petrol.2019.106517>.
- Tian, M., Xu, H., Cai, J., et al., 2019. Artificial neural network assisted prediction of dissolution spatial distribution in the volcanic weathered crust: a case study from Chepaizi Bulge of Junggar Basin, northwestern China. *Mar. Petrol. Geol.* 110, 928–940. <https://doi.org/10.1016/j.marpetgeo.2019.08.045>.
- Wang, H., Chen, M., Wei, S., et al., 2020a. Breathing effect of sandstone reservoir formed in barrier coastal sedimentary system. In: The ISRM International Symposium - EUROCK 2020, Physical Event Not Held, 14–19 June. ISRM-EUROCK-2020-031.

- Wang, H., Chen, M., Wei, S., et al., 2020b. The influence of barrier coastal sedimentary system lost circulation in sandstone. *J. Pet. Sci. Eng.* 185, 106654. <https://doi.org/10.1016/j.petrol.2019.106654>.
- Wei, S., Jin, Y., Xia, Y., 2020. Predict the mud loss in natural fractured vuggy reservoir using discrete fracture and discrete vug network model. *J. Pet. Sci. Eng.* 195, 107626. <https://doi.org/10.1016/j.petrol.2020.107626>.
- Wu, H., Pollard, D.D., 2002. Imaging 3-D fracture networks around boreholes. *AAPG Bull.* 86 (4), 593–604. <https://doi.org/10.1306/61EEDB52-173E-11D7-8645000102C1865D>.
- Wu, Y., Di, Y., Kang, Z., Fakcharoenphol, P., 2011. A multiple-continuum model for simulating single-phase and multiphase flow in naturally fractured vuggy reservoirs. *J. Pet. Sci. Eng.* 78 (1), 13–22. <https://doi.org/10.1016/j.petrol.2011.05.004>.
- Xia, Y., Jin, Y., Chen, M., 2015. Comprehensive methodology for detecting fracture aperture in naturally fractured formations using mud loss data. *J. Pet. Sci. Eng.* 135, 515–530. <https://doi.org/10.1016/j.petrol.2015.10.017>.
- Zheng, S., Yang, M., Kang, Z., et al., 2019. Controlling factors of remaining oil distribution after water flooding and enhanced oil recovery methods for fracture-cavity carbonate reservoirs in Tahe Oilfield. *Petrol. Explor. Dev.* 46 (4), 786–795. [https://doi.org/10.1016/S1876-3804\(19\)60236-3](https://doi.org/10.1016/S1876-3804(19)60236-3).
- Zoback, M.D., Healy, J.H., 1984. Friction, faulting and “in situ” stress. *Ann. Geophys.* 2, 689–698.