# Prediction of Safe Mud Window Based on Seismic Data in Carbonate Formation

Huiwen Pang
*State Key Laboratory of Petroleum Resources and Prospecting, China University of Petroleum, Beijing, China*
*College of Science, China University of Petroleum, Beijing, China*

Hanqing Wang
*Petroleum Exploration and Production Research Institute, SINOPEC, Beijing, China*

Yan Jin
*State Key Laboratory of Petroleum Resources and Prospecting, China University of Petroleum, Beijing, China*
*College of Petroleum Engineering, China University of Petroleum, Beijing, China*

ABSTRACT: To address the problem of predicting the safe mud window in China's TH oil field, we present a feasible seismic-based workflow that employs machine learning. Initially, multiple drilling fluid and mud loss engineering records were used to establish secure mud density windows for eight distinct wells with differing depths. Then, the well logs served as the link between drilling fluid density and through-well seismic data, and the relationship between drilling fluid density and seismic data was constructed using machine learning techniques involving ensemble learning. Finally, a 3D distribution model of safe drilling fluid density is generated, and its dependability is evaluated. The results of one validation well indicate that the model's complete blind test accuracy exceeds 75%. The model has a transverse resolution of 25 meters and a longitudinal resolution of 15 meters. It may offer theoretical guidance for devising drilling fluid density and wellbore construction.

*Keywords: Safe mud window, Seismic data, Machine learning, Carbonate formation.*

## 1 INTRODUCTION

The safe mud window (SMW) is the permissible drilling mud weight range. Keeping the mud weight below the SMW limits would assist in preventing a number of significant wellbore issues, such as wellbore instability, decreased circulation, pipe sticking, etc. The SMW is defined by the minimal mud weight below which shear failure (breakout) is possible (MWBO) and the maximum mud weight beyond which tensile failure (breakdown) is possible (MWBD) (Gowida et al., 2022). Most carbonate formations have a restricted SMW due to the presence of multiple leakage zones, such as vugs, fractures, and caverns, as well as locally developed complex shale, sandstone, and mudstone. Maintaining mud circulation control requires balancing wellbore and formation pressure by regulating mud weight (Tan et al., 2020; Tan et al., 2021). Before drilling, it is crucial to accurately anticipate the SMW in carbonate formations.

Nowadays, SMW prediction methods primarily fell into three categories: The first approach is the empirical formula method, which is utilized most frequently. This method involves calculating and expressing various pressure profiles, such as pore pressure, collapse pressure, fracture pressure, and leakage pressure, in terms of equivalent mud weight, with pore pressure or collapse pressure typically serving as the lower limit of SMW and fracture pressure or leakage pressure serving as the upper limit of SMW. The representative methods are the Hubbert-Willis formula (Hubbert and Willis, 1957), the Matthews and Kelly formula (Matthews, 1967), the Eaton formula (1969), the Zoback formula (1984), and the Huang equation (1984). The second method is numerical simulation, which is primarily used to predict the SMW after the formation structure type has been determined. Combination modes employing continuous and discrete coupling models are extensively used in this method for a variety of structures. Representative continuity models include Lavrov's model (Lavrov et al. 2006), Majidi's model (Majidi et al. 2010), and Gulbransen's model (Gulbransen et al. 2010). In addition, the representative discrete model comprises the models of Yao (Yao et al., 2010), Wang (Wang et al., 2020), and Wei (Wei et al., 2022). In recent years, with the development of artificial intelligence and machine learning, a new technique for SWM prediction has emerged, which is gradually becoming the third most prominent method. In contrast to the model of the mechanism established by conventional methods, this method primarily employs data-driven models. (Noshi and Schuster, 2018) The method has unique advantages for coping with the uncertainty of drilling complex problems, identifying hidden patterns, and revealing useful information. Geng et al. (2019) used machine learning to predict the risk of soil loss based on seismic attributes. Ding et al. (2021) predicted sediment loss in a fractured formation using post-stack seismic data. Using machine learning, Pang et al. (2022) predicted sediment loss rates and evaluated their dependability based on seismic data. Using machine learning, mud recording data are also used to predict or diagnose mud loss (Pang et al., 2021).

In this paper, we propose a practicable method for predicting the SMW in China's TH oil field. Using machine learning, the SMW of drilled wells and post-stack seismic data are linked. Using the post-stack seismic data, the SMW of pre-drill wells can then be predicted, providing theoretical guidance for mud loss prevention and control.

## 2 METHODLOGY

The TH Oilfield, located in the northern Tarim Basin of Xinjiang, is China's first 100-million-ton Palaeozoic marine oil field. Drilling reveals that the carbonate stratigraphic structure of the TH oilfield is complex, resulting in additional downhole complexity. In each layer's pressure system, there are significant issues with elevated in-situ stress and radial stress imbalances. The formation's integrity is compromised, the condition of the stress distribution is ambiguous, and the SWM is difficult to determine.

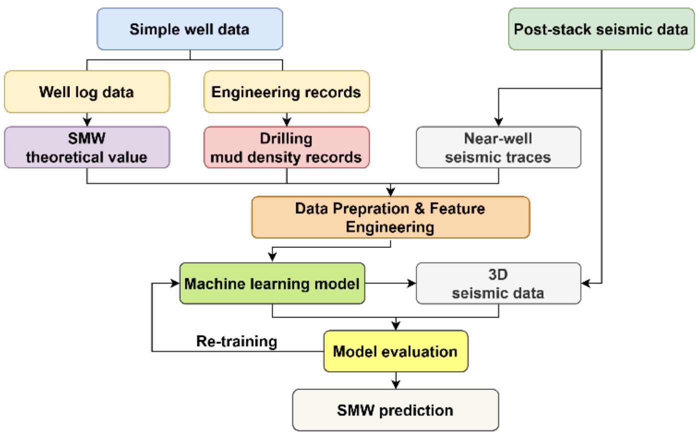


Figure 1. The workflow of SMW prediction.

In undrilled areas, only seismic data can be used to anticipate the SMW, as opposed to well logging data. However, the poor longitudinal resolution and time domain of post-stack seismic data makes it challenging to establish a one-to-one correspondence with drilling engineering records. The seismic track data were extracted and matched with SWM (comprehensive logging calculation results and actual drilling engineering documents) based on time-depth relationships. The relationship between seismic data and SWM was determined using machine learning, and the original post-stack seismic data was replaced with sediment density to achieve an accurate prediction of SWM. Figure 1 depicts the workflow of the technical procedure.

## 2.1 Data Preparation

Post-stack seismic data primarily consists of amplitude information. Figure 1 depicts the distribution of seismic amplitude data in the study area. Figure 2 demonstrates that the amplitude value falls within the range of -2500 to 2500 and follows a normal distribution.

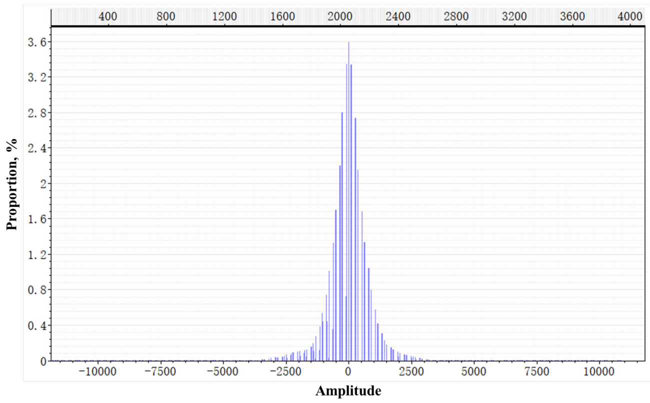


Figure 2. Distribution of post-stack seismic data.

Pore pressure is typically used as the lower limit of the sediment density window for carbonate reservoirs, while rupture pressure is the upper limit. During new drilling, the aforementioned data are typically derived from well logging calculations of neighboring wells and adjusted to account for the actual situation. Table 1 presents the pore pressure, fracture pressure equivalent mud density, and actual mud density of various strata in the study area.

Table 1. Drilled well pressure profile information.

| Formation | MD, m | Pore pressure equivalent density, g/cm$^3$ | Fracture pressure equivalent density, g/cm$^3$ | Actual drilling fluid density, g/cm$^3$ |
|---|---|---|---|---|
| $Q\sim N_2k$ | 10.5-1998 | 1.07~1.09 | 1.83~1.94 | 1.09~1.10 |
| $N_1k$ | -2973 | 1.10 | 1.88~1.96 | 1.10~1.12 |
| $N_1j$ | -3391 | 1.13 | 1.89~1.96 | 1.12~1.14 |
| $E_3s\sim E_{1\text{-}2}km$ | -3538 | 1.13 | 1.89~1.98 | 1.12~1.14 |
| $K_1\sim T$ | -5067 | 1.09~1.14 | 1.87~1.98 | 1.14~1.28 |
| $C_1kl\sim C_1b$ | -5500 | 1.10~1.15 | 1.87~1.98 | 1.28~1.30 |
| $O_{1\text{-}2}y$ | -5620 | 1.05~1.07 | 1.88~1.98 | 1.13 |

## 2.2 Machine Learning Model

Ensemble learning improves machine learning outcomes by combining multiple models, which facilitates greater predictive performance than a single model. Friedman (2001) describes Gradient Boosted Decision Trees (GBDT) as an extension of boosting to loss functions with indeterminate differentiation. GBDT is a precise and efficient off-the-shelf technique applicable to regression and classification issues in a variety of disciplines. Safe mud window prediction is a regression problem, and GBDT is also known as Gradient-Boosted Regression Trees (GBRT) for regression problems. The GBRT diagram is displayed in Figure 3.

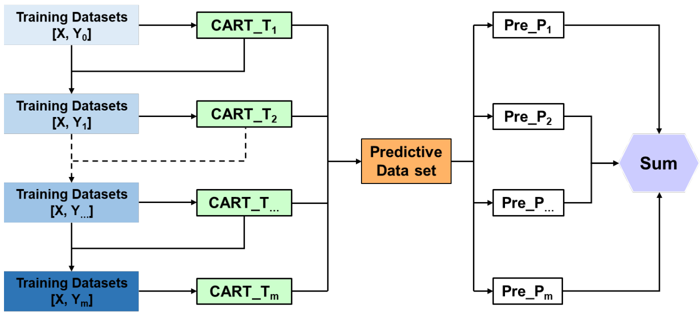


Figure 3. Diagram of GBRT.

GBRT regressors are additive models with the following form of prediction for a given input:

$$\hat{y}_i = F_M\left(x_i\right) = \sum_{m=1}^{M} h_m\left(x_i\right) \tag{1}$$

Where, the $h_m$ are estimators called weak learners in the context of boosting. Gradient Tree Boosting uses decision regressors of fixed size as weak leaners. The constant M corresponds to the number of estimators.

Like other boosting algorithms, a GBRT is built in a greedy fashion:

$$F_m\left(x\right) = F_{m-1}\left(x\right) + h_m(x) \tag{2}$$

Where, the newly added tree $h_m$ is fitted to minimize a sum of losses, given the previous ensemble:

$$h_m = \arg \underset{h}{m} L_m = \arg \min_{h} \sum_{i=1}^{n} l\left(y_i, F_{m-1}(x_i) + h\left(z_i\right)\right) \tag{3}$$

Where, $l\left(y_i, F(x_i)\right)$ is the loss function.

By default, the initial model is chosen $F_0$ as the constant that minimizes the loss: for a least-squares loss, this is the empirical mean of the target values. The initial model can also be specified via the initial argument.

Using a first-order Taylor approximation, the value of $l$ can be approximated as follows:

$$l\left(y_i, F_{m-1}(x_i) + h_m(x_i)\right) \approx l\left(y_i, F_{m-1}(x_i) + h_m(x_i)\right)\left[\frac{\partial l\left(y_i, F(x_i)\right)}{\partial F\left(x_i\right)}\right]_{F=F_{m-1}} \tag{4}$$

Where, the quantity $\left[\frac{\partial l\left(y_i, F(x_i)\right)}{\partial F\left(x_i\right)}\right]_{F=F_{m-1}}$ is the derivative of the loss with respect to its second parameter, evaluated at $F_{m-1}(x_i)$.

It is easy to compute for any given $F_{m-1}(x_i)$ in a closed form since the loss is differentiable. We will denote it by $g_i$.

Taking out the constant terms, we get:

$$h_m \approx \arg \min_{h} \sum_{i=1}^{n} h\left(x_i\right) g_i \tag{5}$$

This is minimized if $h\left(x_i\right)$ is fitted to predict a value that is proportional to the negative gradient $g_i$. Therefore, at each iteration, the estimator $h_m$ is fitted to predict the negative gradients of the samples. Every iteration updates the gradients. This can be thought of as a gradient descent in a functional space.

In accordance with the aforementioned formulations, the sample set is generated by extracting the amplitude data of the seismic track of a single feature well and matching the amplitude data with the pressure profile data of a single well using the time-depth relation.

## *2.3 Model Evaluation*

The evaluation of model deviations and predictive ability to determine the model's performance is depicted in Figure 4. Figure 4 (a) depicts the train and test error for each iteration. Therefore, 50 is the optimal number of boosting interactions. Figure 4 (b) depicts the validation outcomes of a single random well with an overall blind test accuracy of over 75%.

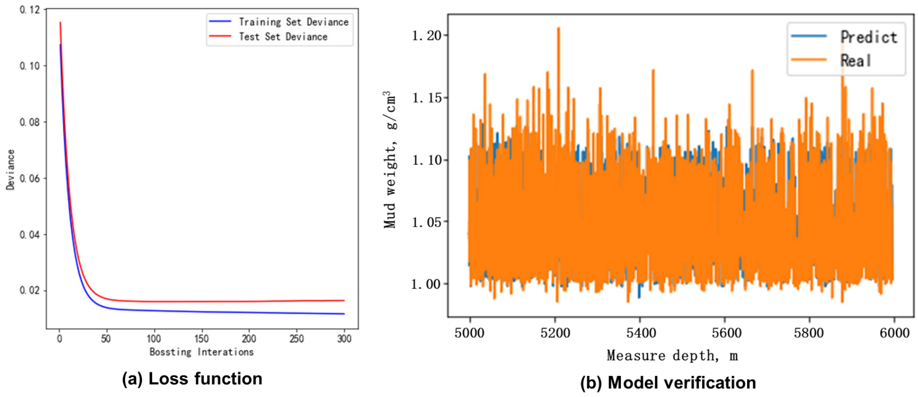


Figure 4. Model evaluation.

## 3 RESULT AND DISCUSSION

Figure 5 illustrates the final forecast results, where 5 (a) is a three-dimensional representation of the safe mud density and 5 (b) is a representation of the safe mud density along the crossline profile and the depth profile, allowing for the optimization of well placement and wellbore trajectory. The calculation and prediction results of a single well are depicted in Figure 5 (c). The figure demonstrates that the method can be used to predict the presence of safe mud density in undrilled or drilled areas.

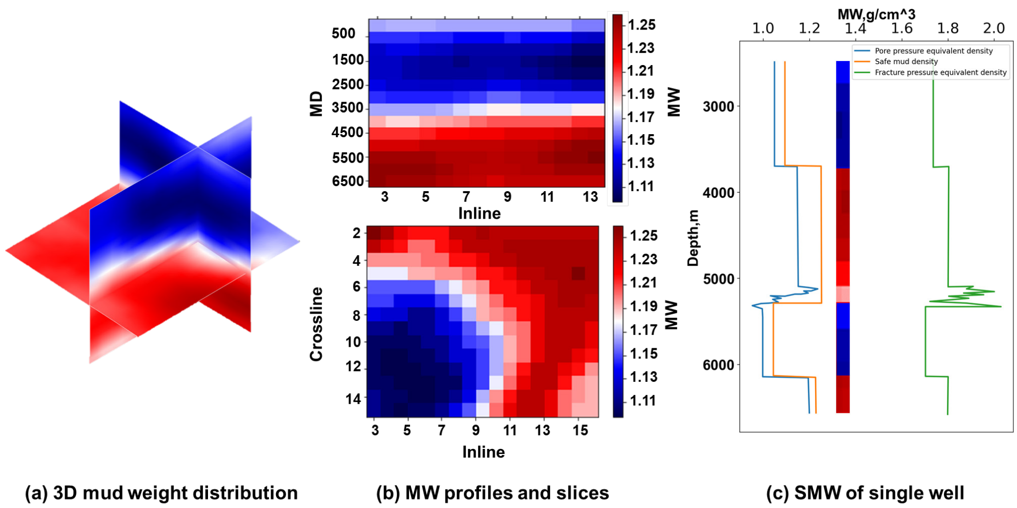


Figure 5. SMW prediction results.

## 4 CONCLUTION

(1) SWM can be predicted by comparing amplitude data with mud density data, as amplitudes disclose stratigraphic characteristics.

(2) The GBDT is superior at predicting SWM, and its ability to accurately characterize data makes it valuable for predicting mud density.

(3) The limitation of the prevalent seismic frequency renders the resolution of prediction results inadequate. However, seismic data remain the only viable approach for predicting SWM.

## ACKNOWLEDGEMENTS

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

Ding, Y., Cui, M., Wang, H., Zhao, F., Shi, X., & Huang, K. 2021. Predicting seismic-based anisotropy for prevent pre-drill risk using a novel type neural network. In *Proc. SPE/IATMI Asia Pacific Oil & Gas Conference and Exhibition*, Virtual, October 12–14, 2021.

Eaton, B. A. 1969. Fracture gradient prediction and its application in oilfield operations. *Journal of petroleum technology*, 21(10), 1353-1360. DOI: doi.org/10.2118/2163-PA.

Friedman, J. H. 2002. Stochastic gradient boosting. *Computational statistics & data analysis*, 38(4), 367-378. DOI: doi.org/10.1016/S0167-9473(01)00065-2.

Geng, Z., Wang, H., Fan, M., Lu, Y., Nie, Z., Ding, Y., & Chen, M. 2019. Predicting seismic-based risk of lost circulation using machine learning. *Journal of Petroleum Science and Engineering*, 176, 679-688. DOI: doi.org/10.1016/j.petrol.2019.01.089

Gowida, A., Ibrahim, A. F., & Elkatatny, S. 2022. A hybrid data-driven solution to facilitate safe mud window prediction. *Scientific Reports*, 12(1), 15773. DOI: 10.1038/s41598-022-20195-7

Gulbransen, A. F., Hauge, V. L., & Lie, K. A. 2010. A multiscale mixed finite-element method for vuggy and naturally fractured reservoirs. *Spe Journal*, 15(02), 395-403. DOI: doi.org/10.2118/119104-PA.

Huang Rongzun. 1984. A model for prediction formation fracture pressure. *Journal of China University of Petroleum (Edition of Natural Science)*., 4: 335-347 (In Chinese).

Hubbert, M. K., & Willis, D. G. 1957. Mechanics of hydraulic fracturing. *Transactions of the AIME*, 210(01), 153-168. DOI: doi.org/10.2118/686-G.

Lavrov, A. 2006. Newtonian fluid flow from an arbitrarily-oriented fracture into a single sink. *Acta mechanica*, 186(1-4), 55-74. DOI: doi.org/10.1007/s00707-006-0324-9.

Majidi, R., Miska, S. Z., Yu, M., Thompson, L. G., & Zhang, J. 2010. Quantitative analysis of mud losses in naturally fractured reservoirs: the effect of rheology. *SPE Drilling & Completion*, 25(04), 509-517. DOI: doi.org/10.2118/114130-PA

Noshi, C. I., & Schubert, J. J. 2018. The role of machine learning in drilling operations: a review. *In SPE/AAPG Eastern regional meeting*, Pittsburgh, Pennsylvania, USA, October 7-11, 2018.

Pang, H., Wang, H., Jin, Y., Lu, Y., & Fan, Y. 2022. Prediction of Mud Loss Type Based on Seismic Data Using Machine Learning. *In 56th US Rock Mechanics/Geomechanics Symposium*, Santa Fe, New Mexico, USA, June 26–29, 2022.

Tan, P., Pang, H., Zhang, R., Jin, Y., Zhou, Y., Kao, J., & Fan, M. 2020. Experimental investigation into hydraulic fracture geometry and proppant migration characteristics for southeastern Sichuan deep shale reservoirs. *Journal of Petroleum Science and Engineering*, 184, 106517. DOI: doi.org/10.1016/j.petrol.2019.106517

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. *Engineering Fracture Mechanics*, 248, 107707. DOI: 10.1016/j.engfracmech.2021.107707.

Wang, H., Chen, M., Wei, S., Lu, Y., Nie, Z., & Li, Y. 2020. The influence of barrier coastal sedimentary system lost circulation in sandstone. *Journal of Petroleum Science and Engineering*, 185, 106654. DOI: 10.1016/j.petrol.2019.106654.

Zoback, M. D., & Healy, J. H. 1984. Friction, faulting and «in situ» stress. 1984. *Annales geophysicae*, 2(6), 689-698.