

Regional Risk Prediction Method Based on Deep Learning and Multi-source Data



Yuqiang Xu and Kuan Liu

Abstract Drilling in complex formations in new exploration areas poses the challenge of accurately predicting downhole engineering risks before drilling. To address this issue, researchers are exploring the use of intelligent algorithms to analyze the complex relationship between multi-source data and underground engineering risks. However, due to the limited number of drillings in these areas, there are few risk samples, which can result in insufficient generalization ability of the training model and poor prediction effect. To overcome these challenges, this paper introduces a quantitative evaluation method for drilling well engineering risk, which enables the construction of a complex underground risk probability profile rich in geological-engineering information. This risk profile provides reliable risk samples for subsequent model training. Additionally, the paper proposes the concept of virtual wells and its deployment method. The LSTM deep learning model is used to mine the quantitative relationship between multi-source data, such as seismic interval velocity, well logging, and rock mechanics parameters, and the downhole risk probability profile. To achieve a quantitative prediction of underground engineering risk probability profiles of virtual wells, relevant parameters of virtual wells are calculated using Depth Adjustment and Kriging interpolation method. The example calculation demonstrates that the addition of virtual wells can significantly improve the regional drilling engineering risk understanding compared to the 3D drilling engineering risk body constructed only based on wells. The prediction accuracy of engineering risks in unexplored areas can be increased by up to 23.6%.

Keywords Pre-drill risk assessment · Risk samples · Deep learning · LSTM neural network · 3D model

Y. Xu · K. Liu (✉)

College of Petroleum Engineering, China University of Petroleum (East China), Qingdao 266400, China

e-mail: Z21020060@s.upc.edu.cn

Engineering Research Center of Intelligent Oilfield, Ministry of Education, Qingdao 266400, China

1 Introduction

1.1 A Subsection Sample

Oil and gas exploration and development are increasingly focusing on deep strata and deep water, which present significant challenges such as strong formation information uncertainty, well control risks, and complex underground accidents. To reduce the risk of encountering such challenges during drilling, pre-assessment of downhole engineering risks is crucial. Several methods have been proposed, including Interval Analysis [1], Fault Tree Analysis [2], Tree Naive Bayes Algorithm modeling [3], and Fuzzy Bow-tie Model Method [4], which can qualitatively assess the regional drilling engineering risk and safety of engineering scheme before drilling. However, qualitative or semi-quantitative methods are inadequate for meeting the safety requirements of drilling engineering in complex formations due to real factors and model limitations [5]. To address the uncertainty of deep formation information and the difficulty of quantitative evaluation of downhole engineering risks before drilling, Sheng et al. [5] proposed a quantitative method based on credibility to characterize the uncertainty of formation pressure. This technique builds on the wellbore pressure balance criterion to evaluate underground engineering risks, including lost circulation, kick, borehole collapse, and pipe sticking, to achieve good application results. However, this method requires precise formation pressure profile and drilling construction plan information, limiting its use for pre-drilled wells. Yi et al. [6] proposed a 3D geomechanical model of the entire block using the rock mechanical properties and pressure profile of adjacent wells obtained from logging data and drilling logs. However, this approach relies on well logging data and well history records, making it challenging to achieve quantitative evaluation of the graded risk level. With the development of machine learning and artificial intelligence technologies, there has been a growing interest in applying these techniques to the field of petroleum engineering. Several approaches, including support vector regression machine using particle swarm optimization algorithm [7], MLP algorithm [8], and gradient decision tree algorithm [9], have been proposed to construct complex relationships between various data and downhole risks for quantitative evaluation of drilling risks. However, these methods depend on the number and quality of samples used for data training, and for a well or block, the risk specimens of downhole drilling engineering are limited. Therefore, how to gain sufficient quality and quantity of risk samples is a critical bottleneck for current data-driven risk assessment methods.

In order to solve the above issues, this paper proposes the following solutions. First, using the risk assessment method based on wellbore pressure balance criteria [10], we construct a downhole engineering risk probability profile rich in geological and engineering information, which can greatly improve the quantity and quality of risk samples. Second, based on the machine learning algorithm, the quantitative

relationship between the interval velocity, formation lithology, physical properties, drilling parameters and the downhole engineering risk probability profile were also established, and virtual wells were added in the undrilled areas in the block. Besides, the underground engineering risk probability profile of virtual well was quantitatively predicted by using the quantitative relationship of drilled wells. So as to solve the problem of less drilling and uneven spatial distribution in new exploration areas. Third, use spatial interpolation, depth adjustment, and other methods, based on the underground engineering risk probability profile of drilled wells and virtual wells, a regional 3D (three-dimension) risk model was built. The 3D risk body constructed in this way fully considers the influence of regional geology, engineering design and construction level, which is of great significance to improve the understanding level of drilling engineering risk in deep complex strata before drilling, optimize the engineering design scheme to be drilled, and reduce risks of underground engineering in the drilling process.

2 Quantitative Risk Assessment Method of Drilled Engineering Based on Wellbore Pressure Balance Criterion

2.1 Basic Principle of Method

In view of the problems of strong uncertainty of deep formation information and difficulty in quantitative evaluation of underground engineering risks, Sheng et al. [10, 11] and Sheng and Guan [12] proposed a method to characterize the uncertainty of formation pressure by using reliability quantitative table. On this basis, according to the wellbore pressure balance criterion, a quantitative evaluation method for downhole engineering risks such as lost circulation, kick, borehole collapse and pipe sticking was constructed, the basic principle is shown in Fig. 1, which can quantitatively evaluate the type and specific depth of downhole complex occurrence. Field application has proved that this method has a coincidence rate of more than 92% for the risk assessment of drilled wells.

Since this approach judges the complex underground situations based on the equilibrium relationship between formation pressure and wellbore pressure, its risk assessment results are rich in geological (formation pressure and its uncertainty) and engineering (drilling fluid density, wellbore structure scheme, construction level, etc.) information. Compared with the risk information recorded in well history including complex occurrence, well depth, surface monitoring overflow/lost circulation, etc., the quantity and quality of samples used for data training are greatly enriched.

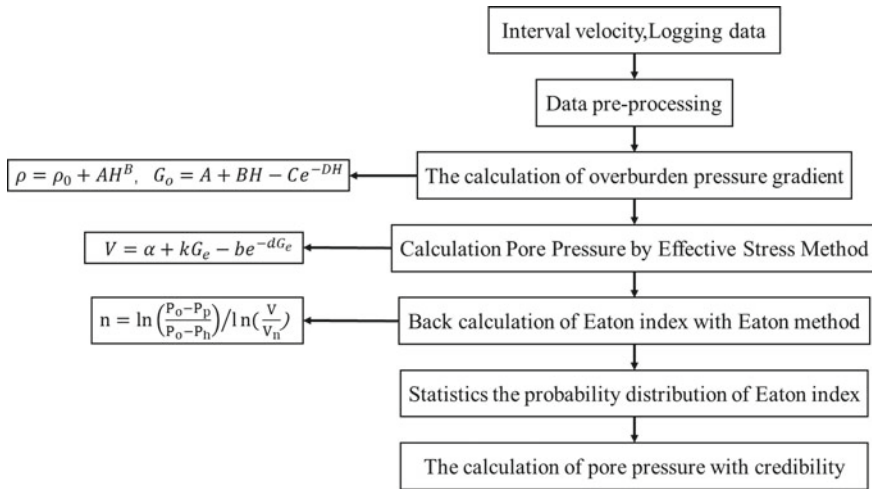
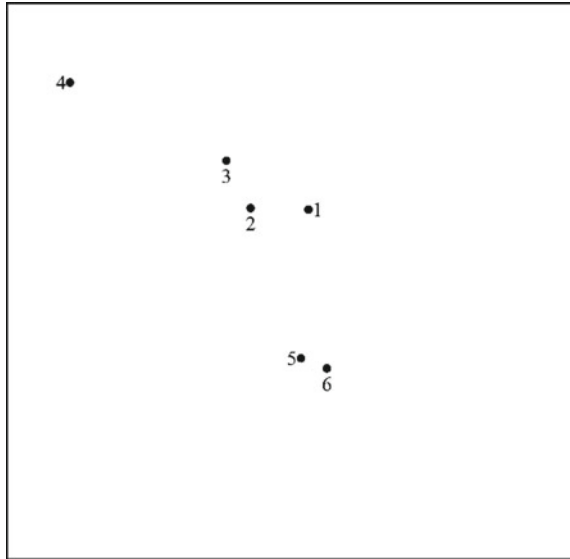


Fig. 1 Basic principle and effect of pore pressure calculation of formation with credibility [13]

2.2 Analysis of Drilling Examples

This paper takes a block in the South China Sea as the instance, which has the characteristics of ultra-high temperature, high pressure and narrow density window. Also, 6 wells have been drilled, kick, lost circulation and other downhole complex multiple, drilling risks are prominent, and the distribution of these wells in the seismic area is shown in Fig. 2. Using the method of 1.1, we can obtain the complex risk probability profiles of 6 drilled wells. In the future, dozens of development wells will be drilled in the seismic area shown in Fig. 2. However, it can be seen from the figure that in this seismic area, the number of wells drilled in this seismic area is small and unevenly distributed. If only the risk probability profile data of 6 drilled wells is used to assess the risk of regional drilling engineering, it is difficult to fully reflect the overall situation of the block. Especially, when the well location of the development well is far away from the drilled well, it is difficult to obtain a relatively accurate risk assessment in the prior art. Therefore, we propose the concept of virtual wells and its deployment method in this paper. By adding virtual wells in the undrilled area of the block, the problem of few wells drilled and uneven spatial distribution can be solved.

Fig. 2 Distribution map of drilled wells in the seismic area



3 Determination Method of Virtual Well Location and Related Attributes

3.1 Determination Method of Virtual Well Location

It can be seen from Fig. 2 that the distribution of drilled wells in the block is relatively dispersed, and there is a large area of undrilled areas. It is difficult to predict the risk characteristics of the whole block with high precision if only relying on the well-drilled data. Therefore, it is necessary to insert virtual wells in the undrilled area to enrich the block information and improve the prediction accuracy of drilling risks in the whole block. In order to ensure that the selected virtual wells are randomly and evenly distributed in the undrilled area, a random function [14] is used to generate virtual wells. Figure 3 shows the well location of the virtual wells in the seismic area.

3.2 Determination Method of Virtual Well Depth

According to the well history data, all six drilled wells are vertical wells, and the virtual wells are also set as vertical wells. Meanwhile, in order to ensure that the data that has been drilled can be fully considered when using the interpolation method to obtain the relevant attributes of the virtual wells, the depth of the virtual well is uniformly set to the bottom boundary of the deepest formation encountered by the six wells.

Fig. 3 Distribution of drilled wells and virtual wells in the block

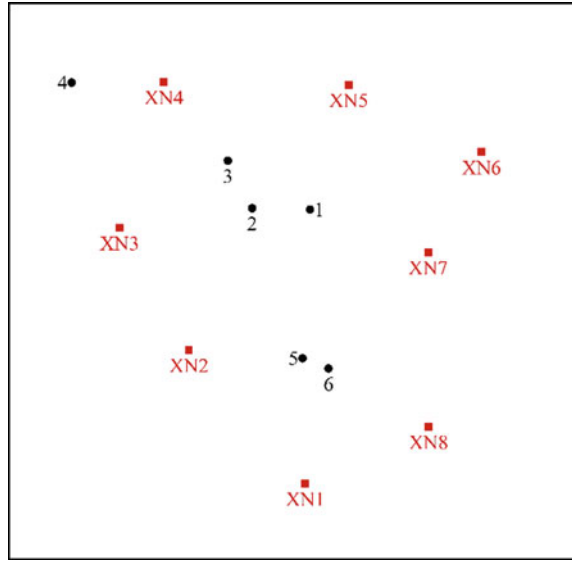


Table 1 Depth of virtual wells

Well	Depth (m)	Well	Depth (m)
XN1	4382.36	XN5	3890.62
XN2	4279.54	XN6	3953.33
XN3	4266.8	XN7	3920.19
XN4	3942.59	XN8	4251.6

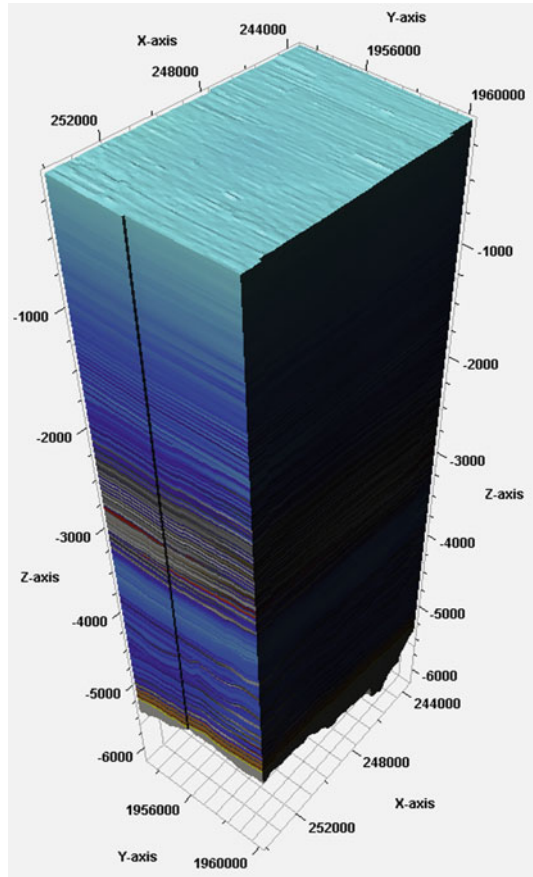
The bottom boundary of the deepest stratum that has been drilled in this area is T41. Based on this stratum level, the well depths of eight virtual wells in Fig. 3 are determined. Table 1 shows the depth of the virtual wells.

3.3 Calculation Method of Seismic Interval Velocity Virtual Well

The virtual well information involves various data such as seismic interval velocity and rock mechanics parameters, and the seismic interval velocity can be extracted from 3D seismic interval velocity volume. Based on the geodetic coordinates of the virtual wells, combined with the SEG Y seismic body of the block, a 3D seismic interval velocity model is established, as shown in Fig. 4. The single-well seismic interval velocity of the drilled wells and virtual wells were extracted from the model according to the well trajectory.

The interval velocity of drilled and virtual wells are shown in Fig. 5.

Fig. 4 3D seismic interval velocity model



3.4 Calculation Method of Rock Mechanical Parameters in Virtual Well

Virtual well's rock mechanics parameters can be quantitatively calculated using logging data, depth adjustment method and Kriging interpolation method [15]. The depth adjustment method corrects the depth of each well to establish a correspondence between the data points before and after processing. The Kriging interpolation method considers the distance between drilled wells and the location of virtual wells to calculate a weight coefficient that reflects their spatial relationship more reasonably [16].

By applying the depth adjustment and Kriging interpolation methods to the rock mechanics parameters of drilled wells, comprehensive calculations can be made to determine the rock mechanics parameters of virtual wells. Figure 6 is the rock mechanical parameters of both drilled and virtual wells.

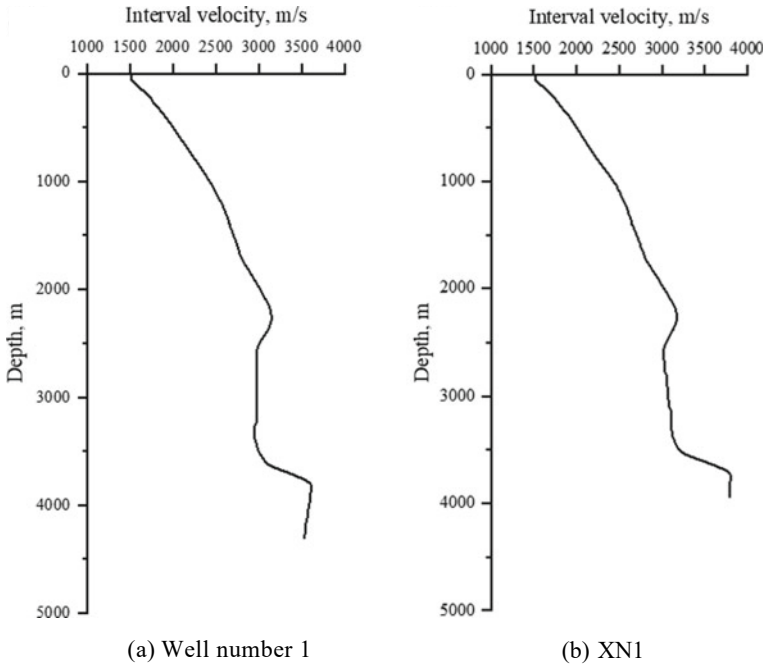


Fig. 5 Interval velocity of drilled wells and virtual wells

4 Regional Drilling Risk Prediction Method Based on Deep Learning

To predict drilling risk in a region, deep learning can be employed using the following steps. Firstly, a LSTM model is utilized to evaluate the correlation between rock physical parameters, seismic interval velocity, and downhole risk. LSTM, a variant of RNN, has a unique characteristic where neurons can be re-fed as inputs after output, maintaining data dependencies. This model is capable of processing data with sequence changes. Subsequently, the obtained correlation is employed to predict the risk of virtual wells, based on the rock physical parameters and seismic interval velocity of the virtual wells. Figure 7 is the basic process of regional drilling risk assessment process.

4.1 Construction Method of LSTM Model

The first step in training an LSTM model for prediction is to establish correlations between input and output attributes. In this paper, the input x_t consists of rock mechanical parameters of a drilled well, along with seismic interval velocity,

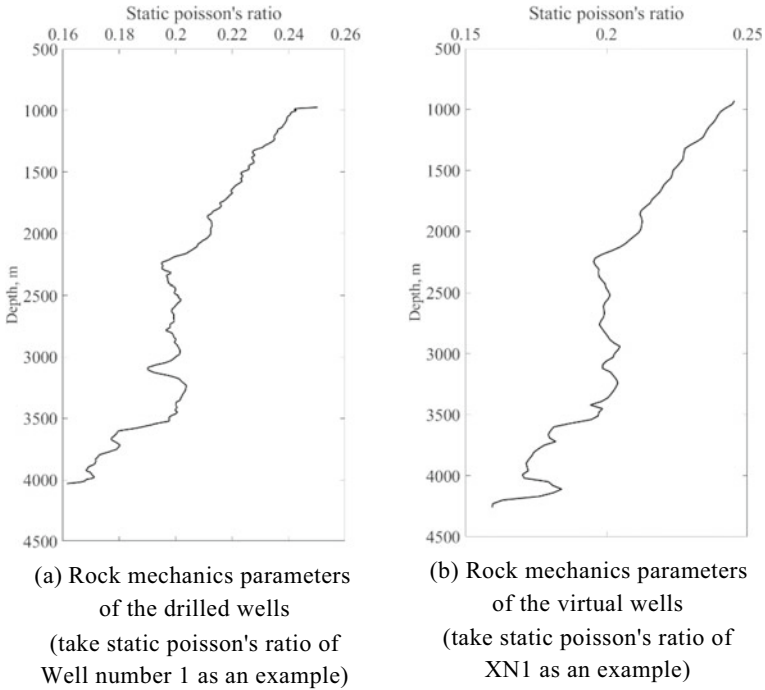


Fig. 6 Rock mechanics parameters

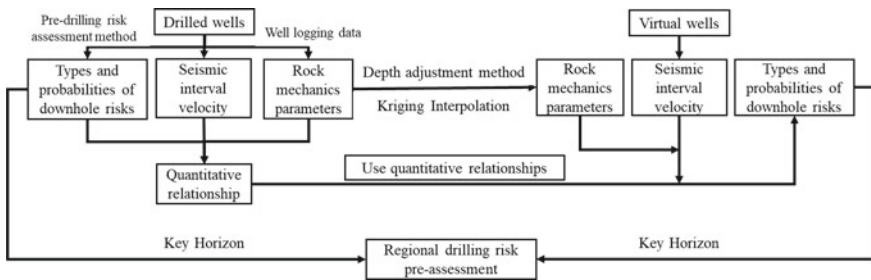


Fig. 7 Basic process of regional drilling risk assessment

with each attribute corresponding to a specific depth. The output h_t represents the corresponding risk level. However, the risk section of most wells is relatively short, resulting in a majority of the risk probability data being '0'. To reduce the amount of calculation and shorten the training time, some of the '0' data is eliminated, so that the final non-'0' risk data accounts for around 90–95% of the total data. The data is deleted from both ends of the dataset to ensure continuous depth data in the training section with a consistent step size of 0.1 m for both the training and prediction sets. If a well has multiple risks, each risk is predicted separately. To minimize the influence

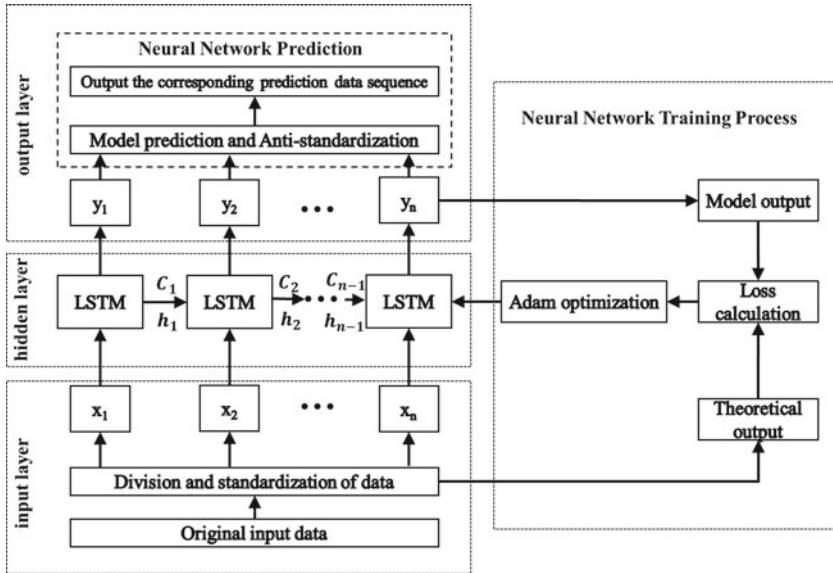


Fig. 8 LSTM network prediction framework

of unknown factors, the data set is randomly sorted, and the first 90% of the new sorted data set is used as the training section, while the last 10% is applied as the verification section.

To build a risk prediction model, it's crucial to standardize the data to avoid the dimension's influence on model learning, leading to a significant improvement in the model's training speed [17]. The input feature has seven dimensions, and the output feature has one dimension. Consequently, the input layer has seven cell units, the output layer has one cell unit, and the number of cell units in the hidden layer is determined using an empirical formula [18]. To speed up the network learning, the Adam optimization algorithm is utilized, and the model undergoes 500 rounds of training. Moreover, to prevent gradient explosion, the gradient threshold is set to one, and the initial learning rate is 0.001. After the first 20 rounds of training, the learning rate is reduced by a factor of 0.9. The computation of the hidden layer is the core of the network's structure [19], as illustrated in Fig. 8.

4.2 Prediction Effect Analysis

Taking the prediction of lost circulation as an example, five drilled wells (Well number 2, Well number 3, Well number 4, Well number 5 and Well number 6) were used as the training set, and the Well number 1 was used as the test set to analyze the

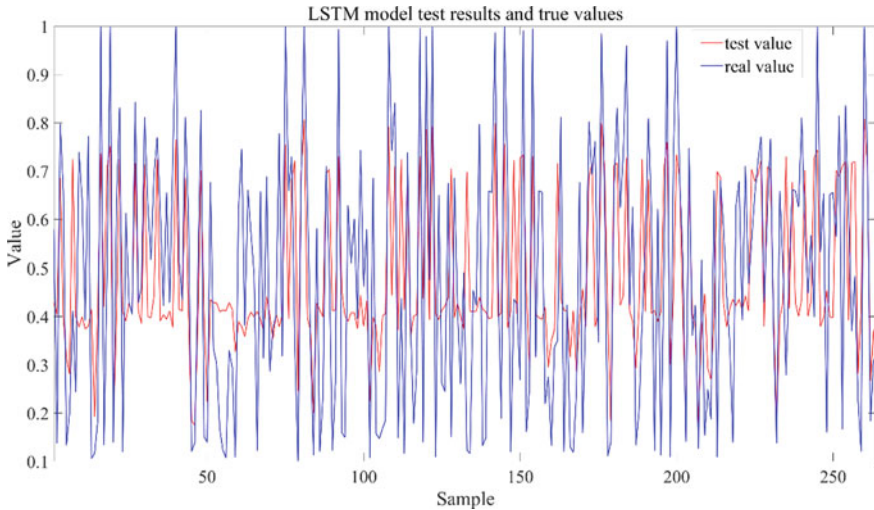


Fig. 9 Comparison of risk test results

prediction accuracy of the constructed LSTM model. The training results of the five drilled wells are as shown in Fig. 9.

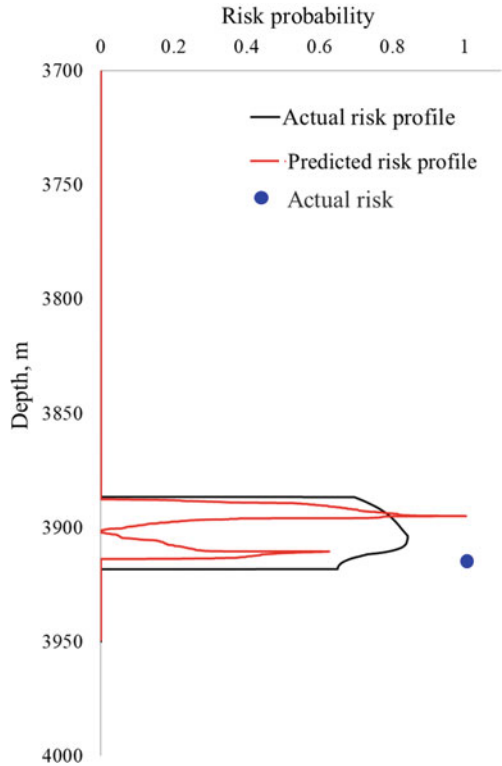
It can be seen that the test data curves of LSTM is not only similar to the actual data curves, but also have similar values, indicating that LSTM has high reliability in the prediction of 10% of the training data segment. The quantitative relationship is used to predict the Well number 1, and the predicted risk profile of the Well number 1 is compared with the actual risk profile that is shown in Fig. 10.

Based on Fig. 10, the forecasted occurrence depth and probability of each risk in Well number 1 aligns well with the actual risk profile and complex situation. By computing the prediction error using the maximum of the actual and predicted risk profile as the benchmark, the model's relative error is determined to be 11.04%.

4.3 Virtual Well Risk Probability Profile Prediction

Taking the prediction of lost circulation as an example, and using the trained model to foresee the risk of virtual wells, the risk probability profile of the virtual well is shown in Fig. 11.

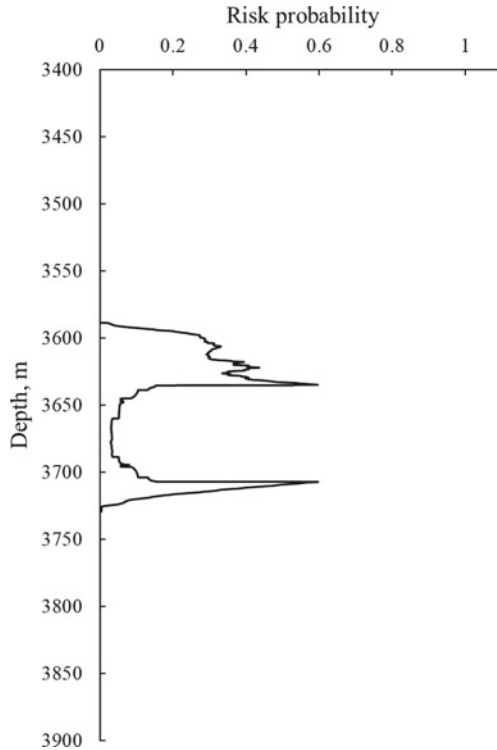
Fig. 10 The lost circulation of Well number 1



5 Construction Method of Regional 3D Risk Body

In order to more intuitively analyze the impact of adding virtual wells on the accuracy of block risk prediction, one of the six drilled wells was selected as the control well. Based on the sequential Gaussian simulation method, the block risk model was established by using the remaining five wells, and the block risk model after adding eight virtual wells is established by using the geological modeling software. The risk data of the control well are extracted from the two models and compared with the actual risk location obtained from the well historical data to determine the impact of the virtual well on the risk prediction accuracy. Sequential Gaussian simulation is the most widely used stochastic modeling method for continuous geological variables, and the basic idea is to perform sequential simulation on conditional data to normal distribution [20].

Fig. 11 Virtual well risk probability profile (take XN1 as an example)

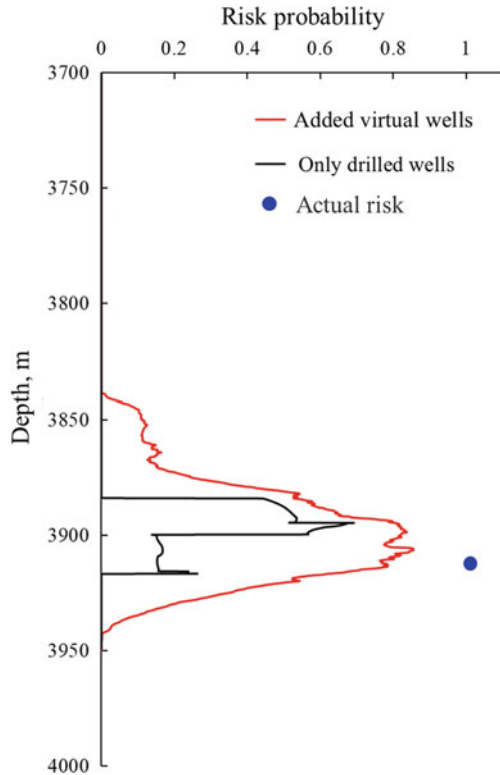


5.1 Comparative Analysis

A 3D (three dimension) risk model was established for a block with six drilled wells using the sequential Gauss simulation method. Well number 1 was selected as the control well, and the risk data of the other five wells were used to establish the model layer by layer. The risk data of Well number 1 was extracted from this model. Next, eight virtual wells were added to the five drilled wells, and the risk model of the entire block was established using the hierarchical group of sequential Gaussian simulation method. The risk data of Well number 1 was again extracted from this model. The extracted risk data of Well number 1 was then compared with the actual risk location based on well history data. Taking the lost circulation as an example, the findings were presented in Fig. 12.

It can be seen from Fig. 12 that the location interval of the risk of Well number 1 extracted by only five drilled wells for the first time is obviously far from the location of the actual risk, while the risk interval of Well number 1 after the second increase of eight virtual wells is not much different from the location of the actual risk. Using the maximum value of the risk distribution extracted from the two times to calculate the accuracy, the second highest prediction accuracy increased by 23.6%.

Fig. 12 Risk comparison of Well number 1



6 Conclusion

In this paper, by introducing the quantitative evaluation method of underground engineering risk before drilling, a complex underground risk probability profile rich in geological-engineering information is achieved, which solves the issue of insufficient generalization ability of the data-driven model due to less risk samples. Also, the concept and deployment method of virtual wells are proposed, and the quantitative calculation of rock mechanics parameters of virtual wells is realized by means of depth adjustment and Kriging interpolation, which provides data basis for model training.

Moreover, according to the timing series characteristics of multi-source data, the confusion matrix is applied to optimize the LSTM model, which is employed to excavate the quantitative relationship between the multi-source data such as well-drilling seismic, logging, rock mechanics parameters and the risk probability profile of downhole. The accuracy of the model is verified by drilled wells, and the relative error is 11.04%. On this basis, using the quantitative relationship constructed by LSTM, the risk probability profile of virtual wells is quantitatively predicted.

The computed findings of the case indicate that compared with the three-dimensional drilling engineering risk body built only by drilling, adding virtual wells can significantly improve the awareness of regional drilling engineering risks, and the prediction accuracy of unexplored regional engineering risks can be increased by 23.6%.

References

1. Zhichuan, G., Kai, W., Shenglin, F., Tingfeng, Z.: Risk evaluation method for drilling engineering based on interval analysis. *Petrol. Drill. Tech.* **41**(04), 15–18 (2013)
2. Dengchuan, Y., Jiajie, Y., Wei, C.: Study on risk assessment of drilling blowout accident based on fault tree analysis. *Drill. Prod. Technol.* **38**(05), 12–14+6 (2015)
3. Adedigba, S.A., Oloruntobi, O., Khan, F., Butt, S.: Data-driven dynamic risk analysis of offshore drilling operations. *J. Petrol. Sci. Eng.* **165**, 444–452 (2018)
4. Baozhan, L., Weiwei, J., Wei, A., Jianping, Z.: Oil spill risk assessment for blowout of deepwater drilling platform based on fuzzy bow-tie model. *Ship Ocean Eng.* **49**(02), 1–5 (2020)
5. Sheng, Y., Guan, Z., Luo, M.: A quantitative evaluation method of drilling risks based on uncertainty analysis theory. *J. China Univ. Petrol. (Ed. Natl. Sci.)* **43**(2), 91–96 (2019)
6. Yi, M., Huang, Z., Zhang, J., Li, B., Xu, Y., Zhang, S.: Three-dimensional geomechanical modeling of Gaoquan structure along the southern margin of the Junggar Basin and its application to the risk evaluation of deep exploration wells. *Oil Drill. Prod. Technol.* **43**(01), 21–28 (2021)
7. Liang, H., Zou, J., Li, Z., Khan, M.J., Lu, Y.: Dynamic evaluation of drilling leakage risk based on fuzzy theory and PSO-SVR algorithm. *Future Gener. Comput. Syst.* **95**, 454–466 (2019)
8. Jahanbakhshi, R., Keshavarzi, R., Jalili, S.: Artificial neural network-based prediction and geomechanical analysis of lost circulation in naturally fractured reservoirs: a case study. *Eur. J. Environ. Civ. Eng.* **18**(3), 320–335 (2014)
9. Zhang, X.: Research on lost circulation prediction and diagnosis theoretical model based on machine learning algorithm. *Drill. Eng.* **49**(2), 58–66 (2022)
10. Sheng, Y., Guan, Z., Zhang, G., Ye, L., Liu, S., Feng, H.: Borehole structure optimization based on pre-drill risk assessment. *Oil Drill. Prod. Technol.* **38**(4), 415–421 (2016)
11. Sheng, Y. N.: Research on Risk Assessment and Control Technology of Drilling Engineering. China University of Petroleum (East China) (2019)
12. Sheng, Y.N., Guan, Z.: Design of drilling engineering risk evaluation and control system and its software development. *Oil Drill. Prod. Technol.* **41**(2), 191–196 (2019)
13. Ke, K.: Casing level and depth in deepwater drilling. Ph.D. dissertation, China University of Petroleum (East China) (2019)
14. Xiaogang, W., Decai, L.: A method of accessing point randomly in a bounded closed region and its realization with mathematica software. *J. Yangzhou Polytech. Coll.* **18**(02), 48–50 (2014)
15. Zhang, D., Chen, Y., Meng, J.: Synthetic well logs generation via recurrent neural networks. *Pet. Explor. Dev.* **45**(4), 629–639 (2018)
16. Jiaqi, H.: Logging curve prediction and reservoir identification method based on deep learning. Ph.D. dissertation, Shaanxi University of Science & Technology (2020)
17. Heng, Z., Zhongyuan, W., Xin, Z., Chunlei, Z., Qiaoyu, M.: Shear wave prediction method based on LSTM recurrent neural network. *Fault-Block Oil Gas Field* **28**(06), 829–834 (2021)
18. Kai, W., Zhichuan, G., Jiehong, W.: Drilling engineering risk evaluation method based on neural network and Monte-Carlo simulation. *China Saf. Sci. J.* **23**(2), 125–130 (2013)

19. Wang, J., Cao, J., Liu, Z., Zhou, X., Lei, X.: Method of well logging prediction prior to well drilling based on long short-term memory recurrent neural network. *J. Chengdu Univ. Technol.* **47**, 227–236 (2020)
20. Yuan, G.: Study on geological modeling method of oil and gas reservoir. Ph.D. dissertation, Xi'an University of Science and Technology (2014)