Petroleum Science 18 (2021) 1619–1629

Contents lists available at ScienceDirect

Petroleum Science

journal homepage: www.keaipublishing.com/en/journals/petroleum-science

Original Paper

The application of machine learning under supervision in identification of shale lamina combination types — A case study of Chang 7_3 sub-member organic-rich shales in the Triassic Yanchang Formation, Ordos Basin, NW China

Yuan-Yuan Zhang [a], Ke-Lai Xi [a, *], Ying-Chang Cao [a], Bao-Hai Yu [b], Hao Wang [a], Mi-Ruo Lin [a], Ke Li [a], Yang-Yang Zhang [c]

[a] Deep Oil and Gas Key Laboratory, China University of Petroleum (East China), Qingdao, Shandong 266580, China
[b] No.7 Oil Recovery Plant of PetroChina Changqing Oilfield Company, Xi'an, Shaanxi 710000, China
[c] Fengcheng Oil Recovery Plant of PetroChina Xinjiang Oilfield Company, Karamay, Xinjiang 834000, China

ARTICLE INFO

Article history:
Received 22 December 2020
Accepted 29 March 2021
Available online 21 September 2021

Edited by Jie Hao and Teng Zhu

Keywords:
Organic-rich shale
Laminae combination
Conventional logs
Machine learning
Ordos Basin

ABSTRACT

Organic rich laminated shale is one type of favorable reservoirs for exploration and development of continental shale oil in China. However, with limited geological data, it is difficult to predict the spatial distribution of laminated shale with great vertical heterogeneity. To solve this problem, taking Chang 7_3 sub-member in Yanchang Formation of Ordos Basin as an example, an idea of predicting lamina combinations by combining 'conventional log data — mineral composition prediction — lamina combination type identification' has been worked out based on machine learning under supervision on the premise of adequate knowledge of characteristics of lamina mineral components. First, the main mineral components of the work area were figured out by analyzing core data, and the log data sensitive to changes of the mineral components was extracted; then machine learning was used to construct the mapping relationship between the two; based on the variations in mineral composition, the lamina combination types in typical wells of the research area were identified to verify the method. The results show the approach of 'conventional log data — mineral composition prediction — lamina combination type identification' works well in identifying the types of shale lamina combinations. The approach was applied to Chang 7_3 sub-member in Yanchang Formation of Ordos Basin to find out planar distribution characteristics of the laminae.

1. Introduction

Continental shale oil in China has huge resource potential. Continental shale strata have laminated structure commonly, and diverse types of laminae in frequent variation vertically (Du et al., 2019; Zhao et al., 2020; Liu et al., 2021a,b). Studies on shale oil's enrichment regularity conducted by many researchers show the laminated shale reservoir with developed lamellation, high organic matter content, and high brittle mineral content, has advantages over other types of shale reservoirs for shale oil enrichment, making it the most preferred reservoir type for shale oil exploration and development (Du et al., 2019; Hu et al., 2020; Liu et al., 2021a,b). Shale lamina combinations of different structures make different contributions to oil and gas accumulation, among which the lamina combination with alternate organic-rich shale and thin reservoir layers has the highest hydrocarbon expulsion efficiency and storage capability (Wang et al., 2016). Therefore, the research on types and combinations of shale laminae is of great significance for finding out the shale oil enrichment regularity.

Continental shale strata are often influenced by factors such as climate changes, hydrodynamic conditions, and source material feeding modes, so laminae in shale formations of different mineral composition often show feature of sedimentary cycle (Zeng et al., 2017; Fu et al., 2020) with wide differences in contents of mineral components (Zhang et al., 2020; Xi et al., 2020). The differences in

* Corresponding author.
E-mail address: xikelai@upc.edu.cn (K.-L. Xi).

https://doi.org/10.1016/j.petsci.2021.09.033

mineral component contents can be used to identify the types of shale lamina combinations. Whole-rock analysis technique by X-ray diffraction (XRD) is a common testing method of mineral composition (Yan et al., 2015; Zhao et al., 2018a,b). Since coring cannot be done on whole sections of all wells, this technique can only provide mineral composition data of scattered points in shale strata vertically (Alnahwi and Loucks, 2019). Elemental capture spectroscopy log and natural gamma-ray spectroscopy log are special logging technologies that can indirectly obtain the vertical continuous data of whole-rock minerals and clay minerals from well log data respectively (Wang et al., 2007; Shi et al., 2019). But these technologies are limited in application due to the high cost. So far, the basic field data universally available is conventional well log data, which is commonly used to identify the reservoir lithology. This kind of data can identify the reservoir lithology both vertically and laterally (Zhao et al., 2017), but can't meet the demand of lamina type identification. If the relationships between conventional well log data and mineral components are established, the spatial distribution of shale laminae can be predicted based on mineral differences of lamina combinations.

There have been successful trials of high precision quantitative prediction of minerals based on conventional well log data by using machine learning under supervision (Li et al. 2019a,b, 2020). Machine learning under supervision makes use of massive experience and instances, or sample sets composed of multiple sets of data (inputs) and corresponding known objective function values (outputs), to train under self-monitoring to find out the relationship between two sets of values and construct models. The models then can be used to predict the values of unknown objectives with new data. Machine learning under supervision can realize the vertical continuous prediction of mineral composition through sample set composed of limited XRD data and corresponding conventional well log data. In view of the current geological data richness and economic perspective, machine learning under supervision is more widely used than elemental capture spectroscopy log, which is a method of quantifying mineral components by analyzing the relationship between main elements and minerals in formation.

Organic-rich shale in Chang 7_3 sub-member of Ordos Basin has laminated structure generally, and the different types of lamina combinations have distinct features and wide differences in mineral composition (Xi et al., 2020). Taking the work area as an example, machine learning under supervision was used to set up the mode of 'conventional log data — mineral composition prediction — lamina combination type identification'. Then, the spatial distribution of laminae was analyzed to verify the reliability of this mode. This approach can provide certain methodological guidance for research on continuity distribution characteristics of continental shale laminae with limited core data and strong heterogeneity.

2. Geological background

Located in the west of the North China Platform, the Ordos Basin is divided into 6 secondary tectonic units, namely, western margin thrust belt, Tianhuan depression, Yishan slope, Jinxi flexure zone, Yimeng uplift, and Weibei uplift (Yang et al., 2016; Fu et al., 2020). In the Late Triassic epoch, it was a large inland depression lacustrine basin. The Yanchang Formation is a set of fluvial — delta — lacustrine facies strata formed in this period and is divided into 10 members from top to bottom, namely Chang 1 to Chang 10 (Fu et al., 2013). Among them, Chang 7 is a set of thick organic-rich mud shale that is mainly semi-deep lacustrine and deep lacustrine sub-facies deposit. Chang 7 is subdivided into 3 sub-members from bottom to top, Chang 7_3, Chang 7_2, and Chang 7_1 respectively. Chang 7_3 sub-member depositing in the peak epoch of the lacustrine basin is the major source rock of the Ordos Basin (Zhao et al., 2018a,b; Yuan et al., 2018; Li et al. 2019a,b; Zhu et al., 2019). Chang 7_3 sub-member shale strata contain intercalated sandy layers (Fig. 1a). In the center area of the lacustrine basin there develops deep-water gravity-flow deposit caused by slumps while in the southwest area of the basin center there develops deep-water gravity flow deposit caused by floods, like hyperpycnal flow sandy layer and mud shale interbeds (Xi et al., 2020).

Chang 7_3 sub-member shale in the research area has two kinds of lamellation characteristics, rich white volcanic tuff strips and black parallel laminae (Fig. 1b) (Xi et al., 2020). The shale has 4 types of laminae, tuff-rich, organic-rich, silty felsic, and clay laminae. They combine into 3 types of lamina combinations, namely massive mudstone, 'organic-rich and silt-grained feldspar-quartz', and 'organic-rich and tuff-rich' lamina combinations. The massive mudstone mainly consists of clay minerals and a small amount of fine silts — very fine silts, with low organic carbon content and undeveloped reservoir space. Whereas, the 'organic-rich and silt-grained feldspar-quartz' lamina combination and 'organic-rich and tuff-rich' lamina combination (Fig. 1c) have higher average organic carbon content and are the main reservoirs of shale oil (Xi et al., 2020). At present, though there have some understandings on Chang 7_3 sub-member's lamina combinations and shale oil enrichment models, the spatial distribution features of the lamina combinations, especially the two types of lamina combination rich in shale oil are still unclear, hindering the prediction of favorable shale oil reservoirs.

3. Theory and methodology

3.1. Machine learning theory

Machine learning under supervision includes 4 elements, namely DATA, MODEL, OBJECTIVE FUNCTIONG and OPTIMIZATION ALGORITHM. It is a multi-layered feedforward neural network trained by error feedback propagation algorithm (Fig. 2). A neural network is connected and computed by massive neurons and its model consists of three layers: the input layer directly receives data; the hidden layer that is the intermediate layer of the network, single or multiple layers, and responsible for analyzing and balancing input data to perform related weight calculation ($\sum$) and non-linearly transferring data to next layer of the network through activating the function $f(x)$; the output layer presents the results after learning and training. The optimization algorithm compares the predicted values of the model with the objective function and reports errors to the model's starting point. By successive iteration and training of each sample, the weight of neurons at the hidden layer is adjusted toward the right direction to reach the goal of minimizing errors eventually.

3.2. Methodology

Based on machine learning under supervision, mineral components of Chang 7_3 sub-member in Ordos Basin in vertical direction were predicted quantitatively by using conventional well log data. First, the main mineral types in the research area and well log data sensitive to changes in mineral components were sorted out and taken as the predicted objective values and input parameters, respectively. Then, the model was developed through machine learning under supervision and used (Fig. 3).

The specified workflow of mineral component quantitative analysis will be presented in detail in the case study section.

3.2.1. Data Pre-processing

(1) Determination of Predicted Objective Values

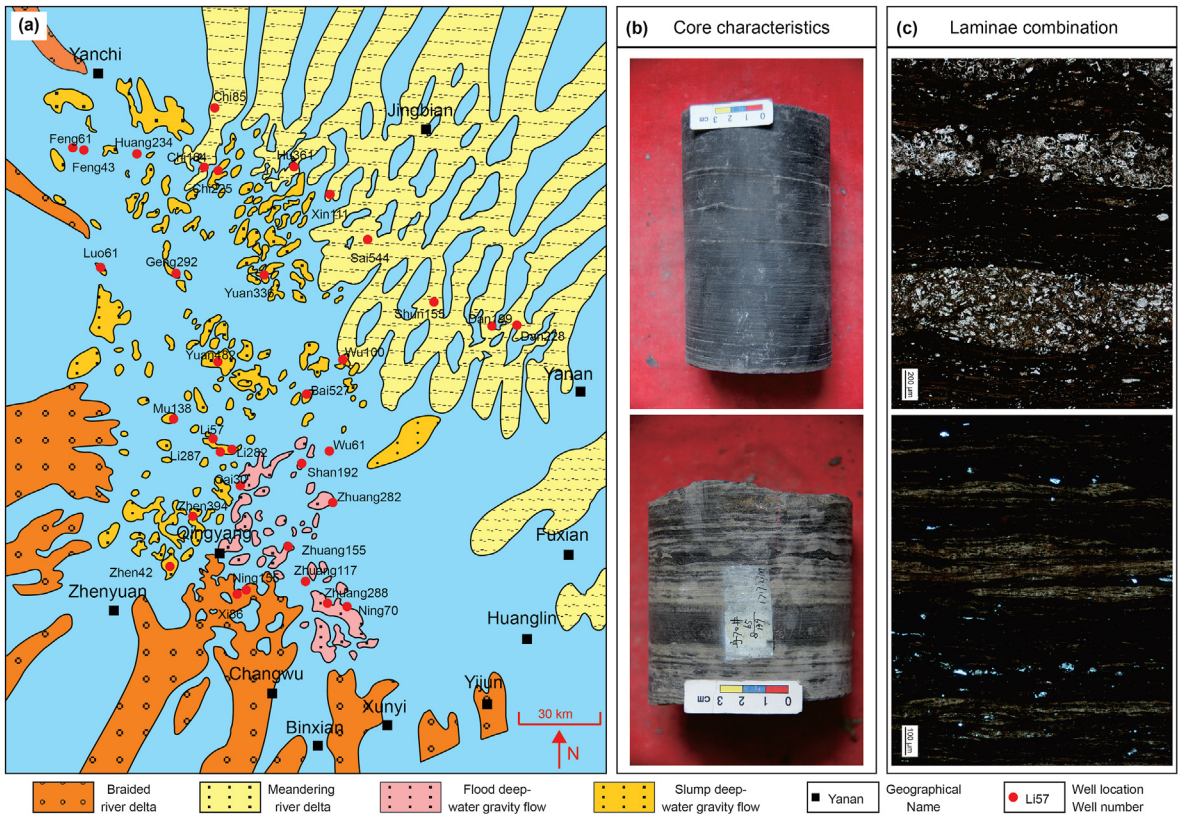

Fig. 1. (a) Areal distribution map of sedimentary facies of Chang 7_3 sub-member of the Triassic Yanchang Formation in the Ordos Basin, modified from Xi et al. (2020), (b) The core characteristics of the Chang 7_3 sub-member, (c) The characteristics of laminae combination of Chang 7_3 sub-member.

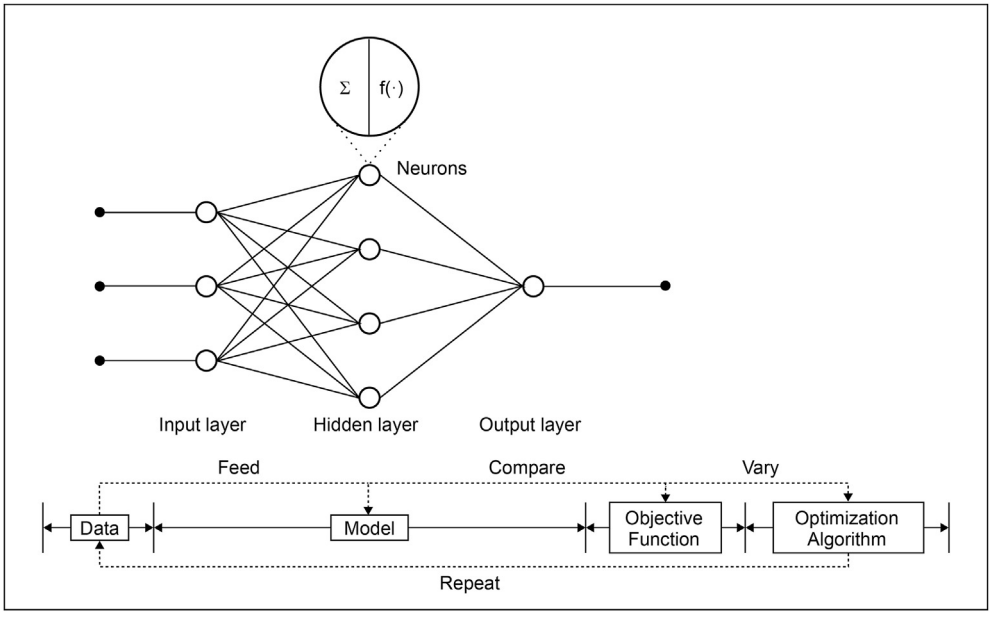

Fig. 2. Schematic diagram of supervised machine learning.

The whole-rock mineral analysis data needed for model construction comes from samples of 6 typical wells with Chang 7_3 sub-member encountered in the Ordos Basin. The samples, 75 in total, cover all lamina types. According to the XRD's quantitative analysis results, the main minerals in the research area are quartz, potassium feldspar, pyrite, and clay minerals (Fig. 4a), with average contents of 35.0%, 12.8%, 17.8%, and 31.2%, respectively (Fig. 4b). Among the four types of minerals, feldspar has the lowest content, and the four types of minerals take up 96.8% of the total mineral composition, while the other minerals only account for an average

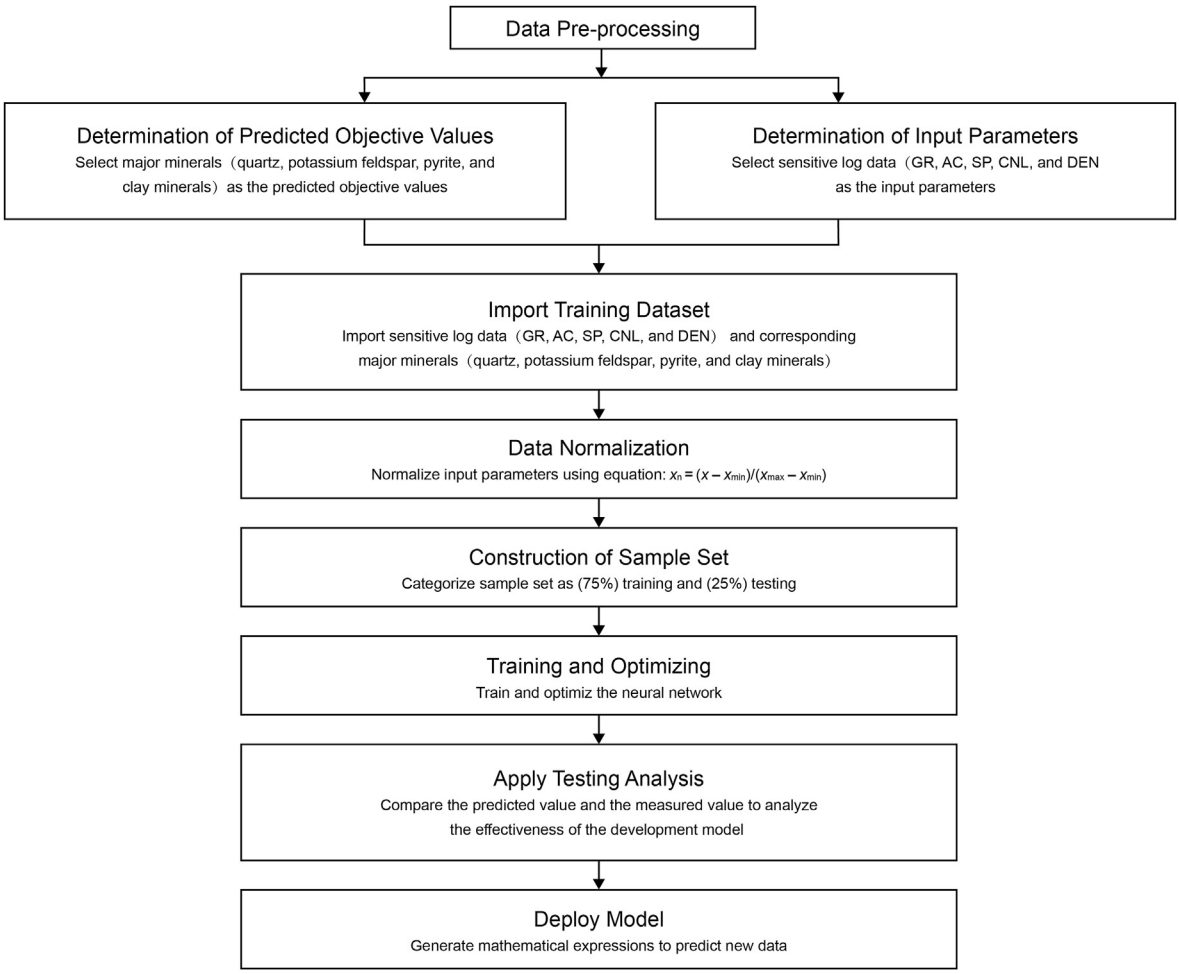

Fig. 3. A proposed workflow showing the steps followed in this study to quantify the mineralogy.

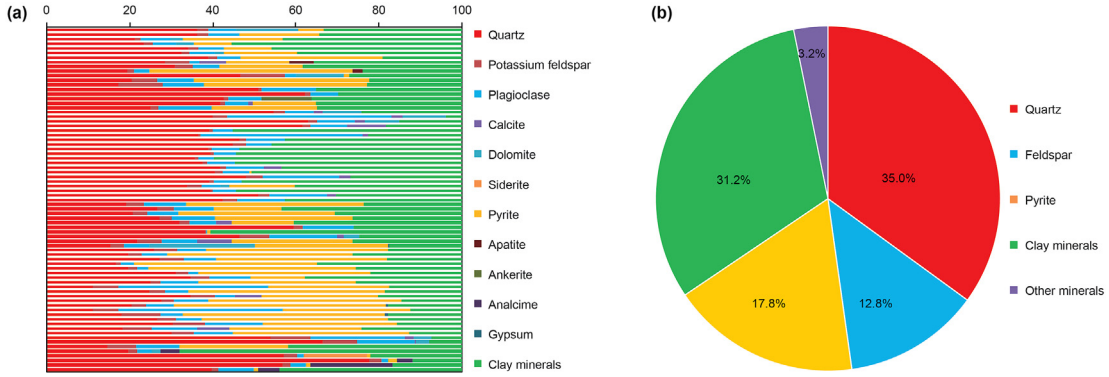

Fig. 4. (a) The mineralogical compositions of samples from the Chang 7_3 sub-member, (b) Average mineral composition of shale.

of 3.2% of the total composition. Therefore, the quartz, potassium feldspar, pyrite, and clay minerals were taken as predicted objective values.

(2) Correlation Analysis

In terms of input parameters, too many types of parameters and relatively low correlations will have negative impact on the model's accuracy. Therefore, well log data with higher correlation indexes should be taken as the main controlling factors to reduce complexity and enhance accuracy of the model. The correlation analysis between the 4 types of predicted objective minerals and well log data shows (Table 1) that the types of well log data that are sensitive to the above minerals are natural gamma ray log (GR), acoustic log (AC), spontaneous potential log (SP), neutron log (CNL), and density log (DEN). Therefore, the above types of 5 well log data were taken as input parameters for model construction.

Through correlation analysis, it is found quartz, pyrite, and clay

Table 1
Correlation analysis between the minerals and log data.

Mineral type	GR, API	AC, μs/ft	CNL, %	DEN, g/cm^3	SP, mV	CAL, IN	RT, Ω·m
Quartz	−0.547	−0.512	−0.621	0.290	0.222	0.169	0.121
Feldspar	0.122	0.023	0.045	0.070	−0.195	−0.168	−0.104
Clay	−0.447	−0.303	−0.386	−0.007	0.474	0.468	0.139
Pyrite	0.721	0.594	0.729	−0.553	−0.530	−0.377	−0.298

GR = natural gamma ray log; AC = acoustic log; CNL = neutron log; DEN = density log; SP = spontaneous potential log; CAL = calliper log; RT = true resistivity log.

minerals show obvious correlations with the above well log data. Feldspar has lower correlations with the log data, mainly due to the low mineral content, which leads to the responses on the log data is not obvious (Li et al. 2019a,b).

3.2.2. Train of machine learning

The model in this study was constructed by using Tensorflow online open source code software library, the program for machine learning was written by python 3.0 design language, machine learning was used to quantitatively predict individual minerals, and Keras was adopted to build, evaluate and supervise the machine learning model. The detailed process is as follows.

(1) Import Training Dataset

75 sets of sample data composed of predicted objectives mineral contents and corresponding conventional well log data were imported into the online neural network design software to conduct self-monitored learning. The 5 types of logging data (GR, AC, SP, CNL and DEN) were the input data, while the 4 types of minerals to be predicted (quartz, feldspar, pyrite and clay minerals) were output objective values.

(2) Data Normalization

Different types of logging data have a great difference in their dimensions, for instance, in the research area, the natural gamma ray log has data ranging from 0 API to 700 API, while the density log ranges from 1.5 g/cm^3 to 3.0 g/cm^3. Input parameters with greater values may attenuate weights of the ones with smaller values in the neural network. To eliminate the deviation caused by largely different dimensions of input parameters, all input parameters need to be normalized (Lai et al., 2015). In this work, the data was normalized through linear transformation into the range [0, 1]. The normalization formula is:

$$x_n = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

Where x_n is the normalized value; x is the original sample data; x_{min} and x_{max} are the maximum and minimum values of the main controlling factors of sample sets, respectively.

(3) Construction of Sample Set

The samples were divided randomly into two sets at a ratio of 3:1. The set with 75% of the samples was used for training to find the optimal model. The set with the other 25% of the samples was used to verify the effectiveness of the model. To ensure the objectivity of the test, the latter set of the data didn't take part in the training process. During the machine learning under supervision of this research area, the number of actual samples for training and test were 56 and 19, respectively.

(4) Training and Optimizing

Neural network optimizes by looking for a minimum error. The estimated value with minimum error is considered the optimum. The error minimization can be achieved by adding hidden layers and neurons to complicate the neural network (Alnahwi and Loucks, 2019).

Note that the final training goal of machine learning is to obtain the minimum error by a relatively simple neural network and fewer iteration times. Meanwhile, previous experience shows that a neural network with a single layer can build a model with high accuracy (Alnahwi and Loucks, 2019; Li et al. 2019a,b, 2020). Therefore, the number of hidden layers of neural network in this study was adjusted to be 1 or 2, the number of neurons in hidden layers based on Kolmogorov theorem (2) (Liu et al., 2019) and the formula (3) (Gao, 1998) was set between 3 and 9. With other parameters remaining the same, the parameters were adjusted one by one within the set range.

$$s = 2m + 1 \quad (2)$$

$$s = \sqrt{0.43mn + 0.12nn + 2.54m + 0.77n + 0.35} + 0.51 \quad (3)$$

where s represents the number of neurons in the hidden layer, m represents the number of neurons in the input layer, and n represents the number of neurons in the output layer.

Activation function of the neural network in this research used 'rule', the activation level was dependent upon the weight of neurons, and the weight of neuron can decide whether the data is exported to neurons in the next layer. In the process, the neurons construct complicated nonlinear models, which usually are more suitable for predicting nonlinear data's trend than a linear relationship. The optimizer used 'adam' to calculate self-adaptive learning rate to speed up convergence rate and shorten the massive time spent on repeating training by machine learning; loss function used the mean square error 'MSE'.

(5) Model Performance Analysis

The predicted mineral values of the training and test sample sets from the models and the actual measured values were plotted, while straight lines from fitting were evaluated, the model with the highest correlation index (ideal condition is a slope of 1) was the optimal model.

(6) Model Applications

The model developed from training the set with 75 samples from Chang 7_3 sub-member in Ordos Basin was used to predict mineral contents of target organic rich shale layer in vertical direction continuously with conventional well log data.

4. Results and discussion

4.1. Analysis of mineral composition differences in lamina combinations

Through analysis of massive XRD data of the shale layers in Chang 7_3 sub-member, Ordos Basin, it is found that different lamina combinations in the research area differ significantly in mineral composition (Fig. 5).

(1) Mineral composition of 'organic-rich and silt-grained feldspar-quartz' lamina combination

This lamina combination features high feldspar content, medium quartz and pyrite contents, and low clay mineral content. This kind of lamina combination has relative quartz contents from 11% to 38%, mainly 15% to 35%, on average 27.3%; relative feldspar contents from 7% to 46%, mainly from 10% to 25%, and 18.2% on average; relative pyrite contents from 13% to 42%, mainly 10% to 35%, and on average 28.5%; relative clay mineral contents from 13% to 36%, mainly 10% to 35%, and on average 24.5%.

(2) Mineral composition of 'organic-rich and tuff-rich' lamina combination

This kind of lamina combination features high pyrite content, medium contents of quartz and clay minerals, and low feldspar content. They have quartz relative contents from 15% to 29%, mainly 15% to 30%, on average 21.7%; feldspar relative contents from 4% to 15%, mainly 0% to 15%, and on average 8.0%; pyrite relative contents from 9% to 51%, mainly 25% to 45%, and on average 34.7%.

(3) Mineral composition of massive mudstone

This kind of lamina combination features high contents of quartz and clay minerals, low feldspar content, and very low pyrite content. They have quartz relative contents from 34% to 62%, mainly 35% to 55%, and on average 44.2%; feldspar relative contents from 2% to 40%, mainly 0% to 10%, and on average 8.8%; pyrite relative contents from 0% to 16%, mainly 0% to 5%, and on average 3.8%; relative contents of clay minerals from 18% to 55%, mainly 35% to 55%, and on average 43.9%.

4.2. Analysis of effect of the developed Model's quantitative prediction

Based on machine learning under supervision, the best model was developed by machine learning of the training sample set of 75 samples XRD analysis results and corresponding depth's conventional well log data with high correlation indexes and validated on the testing sample set. Through repeated iteration and optimization, the optimal structural parameters of the neural network were: 1 hidden layer, 9 neurons in the hidden layer, learning rate of 0.01, and circulated training times of 2000. The fitting result between XRD values from the developed model and actual measured values is shown in Fig. 6. The mineral contents of the training sample set and test sample set predicted by machine learning show a good linear relationship with the actual measured values, with linear coefficient close to 1, and overall above 0.83. Compared with prediction results of other minerals, the predicted results of feldspar of the training and test sample sets are lower in accuracy, as feldspar has lower correlations with conventional logging curves and lower

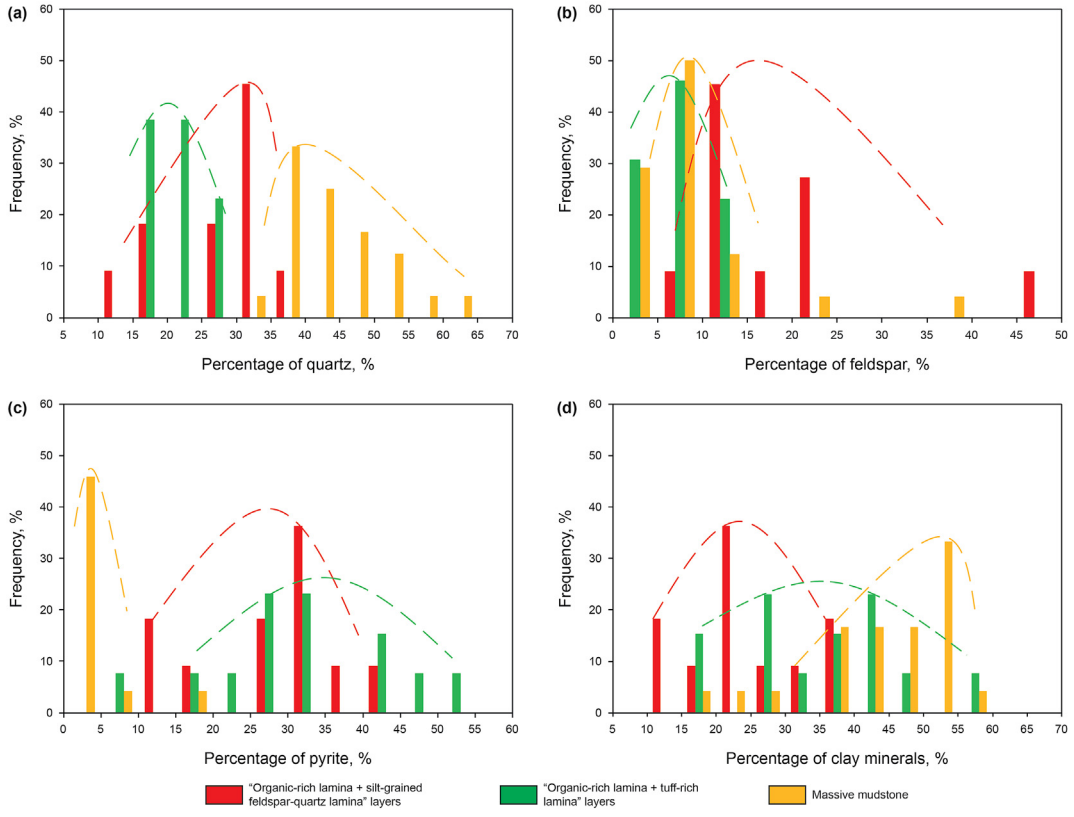

Fig. 5. The difference of mineral content of three types of laminae combination in the Chang 7_3 sub-member. (a) Quartz, (b) Feldspar, (c) Pyrite, (d) Clay minerals.

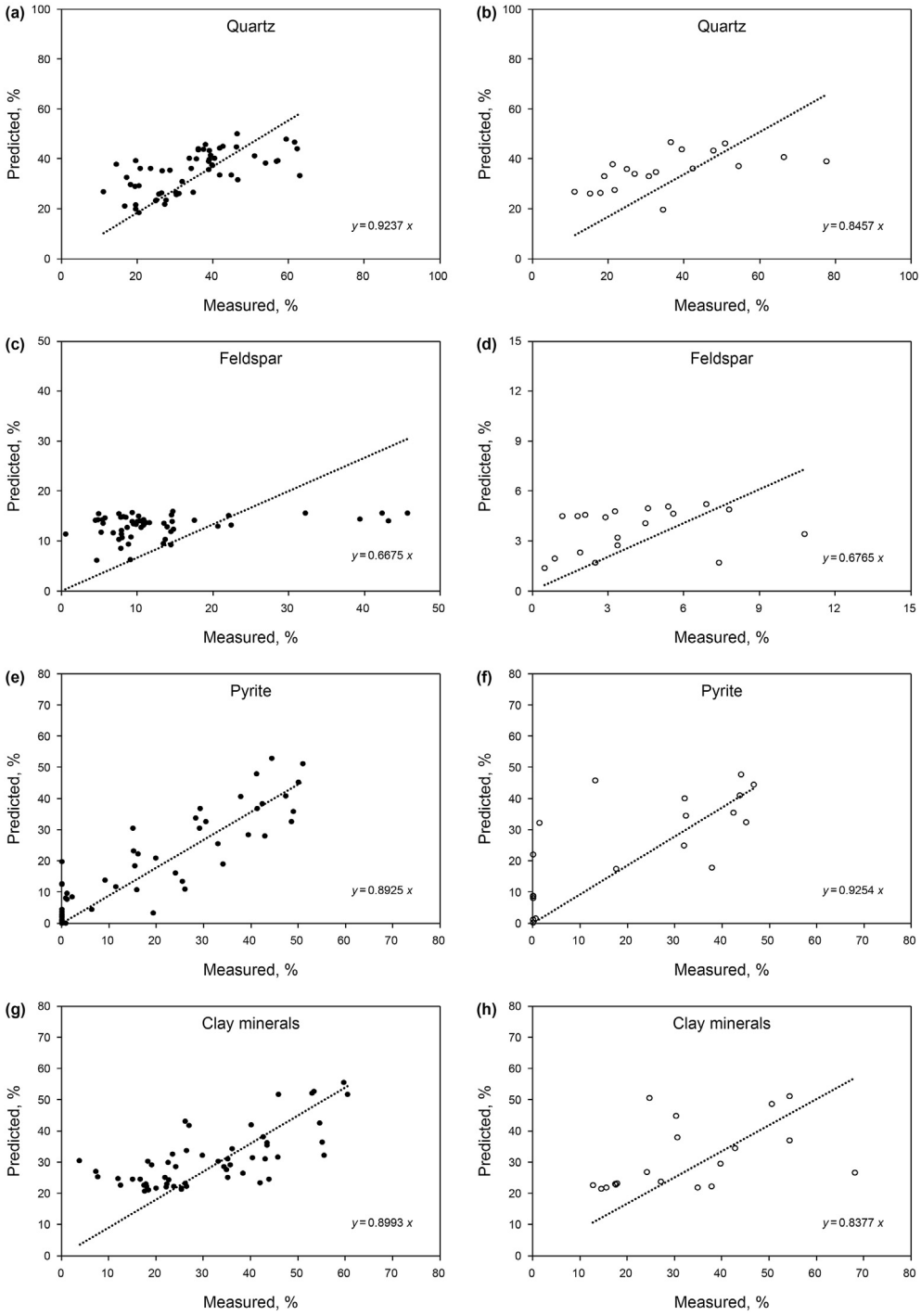

Fig. 6. Cross plot of measured values vs. predicted values. (a) Training samples of quartz, (b) Testing samples of quartz, (c) Training samples of feldspar, (d) Testing samples of feldspar, (e) Training samples of pyrite, (f) Testing samples of pyrite, (g) Training samples of clay minerals, (h) Testing samples of clay minerals.

content, very small prediction deviation could result in significant fluctuation in its correlation. But the predicted feldspar contents are in the ideal range, with a linear coefficient of around 0.6.

4.3. Lamina combination identification and lamina distribution analysis

The model built from machine learning under supervision was used to predict mineral contents of an organic-rich shale section (1709.7 m - 1721.7 m, 12 m in total) in Well Ning 70 of the research area, and then the lamina combinations were identified based on the differences in mineral composition. The type of lamina combination was identified following the principle of identification according to the accuracy of predicted mineral contents from high to low (pyrite, quartz, clay minerals, feldspar): a) The predicted value of pyrite was taken as the first evidence for identification, the lamina combination with low pyrite content and a relatively high content of quartz was identified as Massive structure shale; while the lamina combination with high pyrite content and low quartz content was identified as combination of two kinds of laminae. b)

Then the shale layers with high pyrite contents were further divided, among which the shale layers with relatively low pyrite content, high feldspar content and low clay mineral content were deemed 'organic-rich and silt-grained feldspar-quartz' lamina combination; while the shale layers with relatively high pyrite content, low feldspar content and higher clay mineral content were identified as 'organic-rich and tuff-rich' lamina combination. The analysis results are shown in Fig. 7. Track 1 and 2 are the conventional well log data used as inputs during the model construction; Track 3 is the logging depth; Track 4 is the lithological profile; Track 5 to track 8 are percentages of different minerals predicted by the model, in which the solid points are actual measured values of core samples by XRD that are taken as the benchmark values to validate the model's accuracy; Track 9 shows the identification results of lamina combinations; Track 10 shows the microscopic characteristics of thin sections at the corresponding depths.

It can be seen the predicted mineral contents of Well Ning 70 show good agreement with the actual measured values. The mineral components with higher contents have clear log responses and therefore much smaller deviations in predicted contents than the mineral component with low content (feldspar), but the vertical variation tendency of the feldspar with low content still coincides with that of the actual measured values. Compared with the adjacent relatively thick continuous sand layer, the organic-rich shale section has great differences in mineral composition (1711.3 m - 1721.2 m), featuring a rapid increase in pyrite content and drop in contents of quartz, feldspar, and clay minerals. Through the mode 'conventional log data—mineral composition prediction—lamina combination identification', the prediction results of lamina combination types in the organic-rich shale section are similar to the observation results of core samples under microscope. The organic-rich shale section mainly has two types of lamina combinations, 'organic-rich and silt-grained feldspar-quartz' lamina combination and 'organic-rich and tuff-rich' lamina combination. Massive mudstone mainly occurs between relatively thick continuous sand layers stacking vertically. With the model, the mud shale section in Well Ning 70 was estimated at around 8.5 m thick in total. In this section, 'organic-rich and silt-grained feldspar-quartz' lamina combination was estimated at 4.4 m thick, accounting for the largest proportion of 51.8%; 'organic-rich and tuff-rich' lamina combination was estimated at 3.7 m thick, accounting for 43.5%; massive mudstone was estimated at 0.4 m thick, accounting for 4.7%.

Based on the mode 'conventional log data—mineral composition prediction—lamina combination type identification', lamina combinations in the organic-rich shale section of Well Ning 70 were

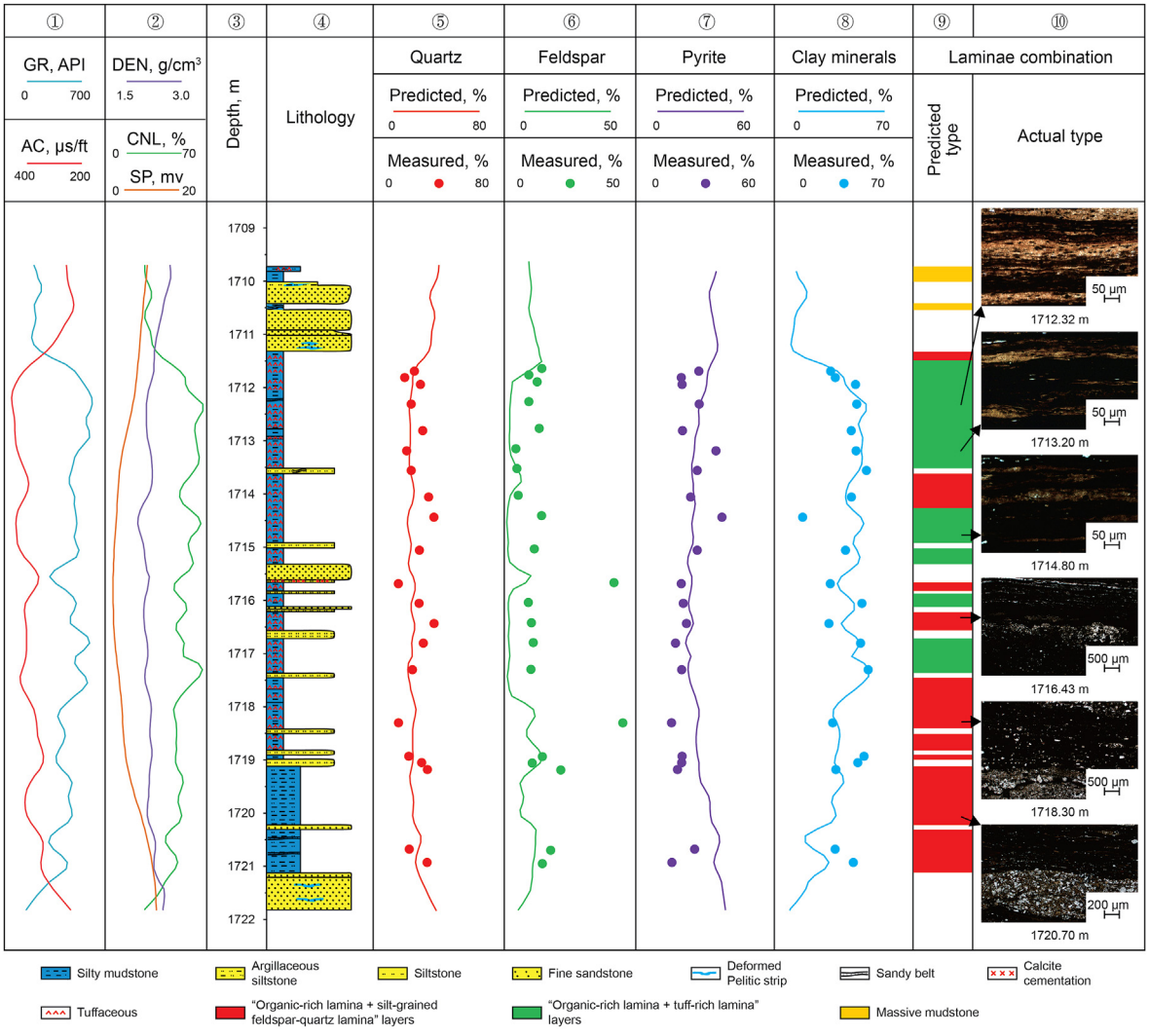

Fig. 7. Prediction of the mineralogical abundance of well Ning 70 and the vertical distribution of laminae combination in the Chang 7_3 sub-member.

identified, and the results show the model developed has high reliability and can be popularized to other oil wells in the research area. This approach was also applied to analyze the planar distributions of favorable lamina combinations in wells Cai 30, Zhuang 282, Huang 234, and Hu 361. The results are shown in Fig. 8. In the research area, the 'organic-rich and tuff-rich' lamina combination mainly concentrates in the center area of the lacustrine basin; while the 'organic-rich and silt-grained feldspar-quartz' lamina

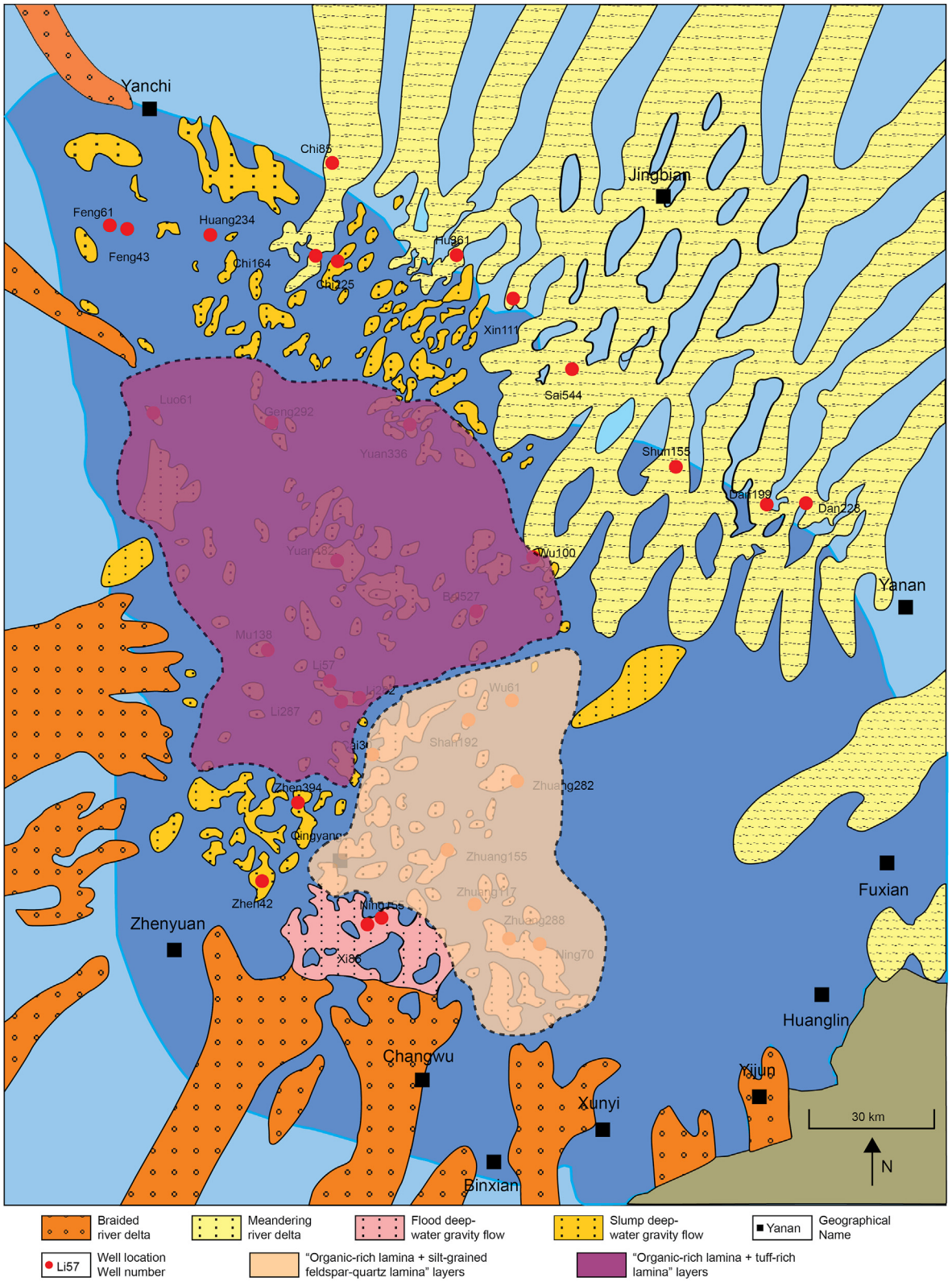

Fig. 8. The planar distribution characteristics of laminae combination in the Chang 7_3 sub-member of Yanchang Formation in the Ordos Basin.

combination is mainly distributed in the southwest of the lacustrine basin center.

The sedimentation of Chang 7_3 sub-member in Yanchang Formation of Ordos Basin was influenced by eruptions of Qinling Volcano located in the southwest of Late Triassic lacustrine basin (Zhang et al., 2009; Dong et al., 2015). Massive volcanic ash entered the lacustrine basin carried by air or water (Zou et al., 2008; Qiu et al., 2014). In the basin's center area, volcanic ash drifted down and directly settled into tuffaceous bedding, interbedding with organic-rich laminae (Zhang et al., 2009). In contrast, in the southwest region of the lacustrine basin, affected by deep-water hyperpycnal flow, coarse volcanic materials previously settled on land were carried into the deep water zone under strong water currents (Liu et al., 2014; Xi et al., 2020); and alternating water currents created silty felsic laminae interbedding with organic-rich laminae in this region.

5. Conclusions

By using machine learning under supervision and neural network analysis technique, the mineral component contents of organic-rich shale were quantitatively evaluated continuously in vertical direction from conventional well log data. Then, on this basis, a mode of 'conventional log data—mineral composition prediction—lamina combination type identification' was proposed to identify the types of lamina combinations based on the mineral composition differences of different types of lamina combinations. This mode has been used successfully in predicting mineral composition quantitatively and identifying lamina combination types of the organic-rich shale section in Chang 7_3 sub-member of the typical well Ning 70 in Ordos Basin. The predicted results coincide with the XRD measured values and thin section observation results, proving the effectiveness and reliability of the mode. This mode was then further applied to Chang 7_3 sub-member organic-rich shale strata across the research area, to analyze the spatial distribution of lamina combinations and predict the plane distribution characteristics of the favorable lamina combinations.

The approach of "conventional log data—mineral composition prediction—lamina combination type identification" takes the advantage of abundant conventional well log data in the field and overcomes the issues of scattered XRD measured points and insufficient special well log data, providing a new methodology for the identification of shale lamina combination with great heterogeneity and frequent vertical variations.

Acknowledgement

This study was co-supported by the National Natural Science Foundation of China (Grant Nos. U1762217, 42072161).

References

Alnahwi, A., Loucks, R.G., 2019. Mineralogical composition and total organic carbon quantification using x-ray fluorescence data from the Upper Cretaceous Eagle Ford Group in southern Texas. AAPG (Am. Assoc. Pet. Geol.) Bull. 103 (12), 2891–2907. https://doi.org/10.1306/04151918090.

Dong, Y.P., Zhang, X.N., Liu, X.M., et al., 2015. Propagation tectonics and multiple accretionary processes of the Qinling Orogen. J. Asian Earth Sci. 104, 84–98. https://doi.org/10.1016/j.jseaes.2014.10.007.

Du, J.H., Hu, S.Y., Pang, Z.L., et al., 2019. The types, potentials and prospects of continental shale oil in China. China Petroleum Exploration 24 (5), 560–568. https://doi.org/10.3969/j.issn.1672-7703.2019.05.003 (in Chinese).

Fu, J.H., Li, S.X., Liu, X.Y., et al., 2013. Geological theory and practice of petroleum exploration in the Ordos Basin. Natural Gas Geoscience 24 (6), 1091–1101. CNKI: SUN:TDKX.0.2013-06-001. (in Chinese).

Fu, J.H., Li, S.X., Niu, X.B., et al., 2020. Geological characteristics and exploration practice of the self-sourced reservoir in Chang 7 member of Triassic Yanchang Formation, Ordos Basin, NW China. Petrol. Explor. Dev. 47 (5), 1–14 (in Chinese). https://kns.cnki.net/kcms/detail/11.2360.TE.20200729.0920.008.html.

Gao, D.Q., 1998. On structures of supervised linear basis function feed forward three-layered neural networks. Chin. J. Comput. 21 (1), 80–86 (in Chinese).

Hu, S.Y., Zhao, W.Z., Hou, L.H., et al., 2020. Development potential and technical strategy of continental shale oil in China. Petrol. Explor.Dev. Online 47 (4). https://doi.org/10.1016/S1876-3804(20)60103-3.

Lai, J., Wang, G., Huang, L., et al., 2015. Brittleness index estimation in a tight shaly sandstone reservoir using well logs. J. Nat. Gas Sci. Eng. 27, 1536–1545. https://doi.org/10.1016/j.jngse.2015.10.020.

Li, J.B., Lu, S.F., Wang, M., et al., 2019a. A novel approach to the quantitative evaluation of the mineral composition, porosity, and kerogen content of shale using conventional logs: a case study of the Damintun Sag in the Bohai Bay Basin. China. Interpretation. 7 (1), 83–95. https://doi.org/10.1190/INT-2018-0088.1.

Li, J.B., Wang, M., Lu, S.F., et al., 2020. A new method for predicting sweet spots of shale oil using conventional well logs. Mar. Petrol. Geol. 113, 104097. https://doi.org/10.1016/j.marpetgeo.2019.104097.

Li, S., Zhu, R.K., Cui, J.W., et al., 2019b. The petrological characteristics and significance of organic-rich shale in the Chang 7 member of the Yanchang Formation, south margin of the Ordos basin, central China. Petrol. Sci. 16 (6), 1255–1269. https://doi.org/10.1007/s12182-019-00386-0.

Liu, B., Gao, Y.F., Liu, K.Q., et al., 2021a. Pore structure and adsorption hysteresis of the middle Jurassic Xishanyao shale formation in the Southern Junggar Basin, northwest China. Energy Explor. Exploit. 39 (3). https://doi.org/10.1177/0144598720985136.

Liu, B., Sun, J.H., Zhang, Y.Q., et al., 2021b. Reservoir space and enrichment model of shale oil in the first member of Cretaceous Qingshankou Formation in the Changling sag, southern Songliao Basin, NE China. Petrol. Explor. Dev. 48 (2), 1–16 (in Chinese). https://kns.cnki.net/kcms/detail/11.2360.TE.20201202.1032.004.html.

Liu, X., Gui, X.J., Ding, X.Q., et al., 2014. Coupling research of event deposition and Qinling orogenesis of late Triassic, South Ordos Basin. J. NE . Pet.Univ. 38 (4), 59–67. https://doi.org/10.3969/j.issn.2095-4107.2014.04.009 (in Chinese).

Liu, Y.J., Liu, S.Q., Ma, Q., et al., 2019. Application of BP neutral network method to identification of shale lithofacies of Lucaogou Formation in Santanghu Basin. Lithologic Reservoirs 31 (4), 101–111. https://doi.org/10.12108/yxyqc.20190411 (in Chinese).

Qiu, X.W., Liu, C.Y., Mao, G.Z., et al., 2014. Late Triassic tuff intervals in the Ordos basin, Central China: their depositional, petrographic, geochemical characteristics and regional implication. J. Asian Earth Sci. 80, 148–160. https://doi.org/10.1016/j.jseaes.2013.11.004.

Shi, P.Y., Wang, C.Y., Yan, W.L., et al., 2019. A method for calculating mineral content of metamorphic rock with ECS logging data. Well Logging Technol. 43 (6), 597–600. https://doi.org/10.16489/j.issn.1004-1338.2019.06-009 (in Chinese).

Wang, Y., Wang, X.J., Song, G.Q., et al., 2016. Genetic connection between mud shale lithofacies and shale oil enrichment in Jiyang Depression, Bohai Bay Basin. Petrol. Explor. Dev. 43 (5), 759–768. https://doi.org/10.1016/S1876-3804(16)30091-X.

Wang, Z.W., Liu, J.H., Huang, Q., 2007. Using natural gamma-ray spectroscopy log method to determine clay mineral content. Lithologic Reservoirs 19 (2), 108–111+116. CNKI:SUN:YANX.0.2007-02-022 (in Chinese).

Xi, K.L., Li, K., Cao, Y.C., et al., 2020. Laminae combination and shale oil enrichment patterns of Chang 7 organic-rich shales in the Triassic Yanchang Formation, Ordos Basin, NW China. Petrol. Explor. Dev. 47 (6), 1–12. https://doi.org/10.11698/PED.2020.06.18 (in Chinese).

Yan, J.H., Pu, X.G., Zhou, L.H., et al., 2015. Naming method of fine-grained Sedimentary rocks on basis of X-ray diffraction data. China Petroleum Exploration 20 (1), 48–54. https://doi.org/10.3969/j.issn.1672-7703.2015.01.005 (in Chinese).

Yang, H., Niu, X.B., Xu, L.M., et al., 2016. Exploration potential of shale oil in Chang7 member, Upper Triassic Yanchang Formation, Ordos Basin, NW China. Petrol. Explor. Dev. 43 (4), 511–520. https://doi.org/10.11698/PED.2016.04.02 (in Chinese).

Yuan, W., Liu, G.D., Xu, L.M., et al., 2018. Petrographic and geochemical characteristics of organic-rich shale and tuff of the Upper Triassic Yanchang Formation, Ordos Basin, China: implications for lacustrine fertilization by volcanic ash[J]. Can. J. Earth Sci. 56 (1). https://doi.org/10.1139/cjes-2018-0123.

Zeng, X., Cai, J.G., Dong, Z., et al., 2017. Sedimentary characteristics and hydrocarbon generation potential of mudstone and shale: a case study of middle Sub-member of member 3 and Upper Sub-member of member 4 in Shahejie formation in Dongying sag. Acta Pet. Sin. 38 (1), 31–43. https://doi.org/10.7623/syxb201701004 (in Chinese).

Zhang, J.F., Xu, X.Y., Bai, J., et al., 2020. Enrichment and exploration of deep lacustrine shale oil in the first Member of Cretaceous Qingshankou Formation, southern Songliao Basin, NE China. Petrol. Explor. Dev. 47 (4), 637–652. https://doi.org/10.11698/PED.2020.04.01 (in Chinese).

Zhang, W.Z., Yang, H., Peng, P.A., et al., 2009. The influence of Late Triassic volcanism on the development of Chang 7 high grade hydrocarbon source rock in Ordos Basin. Geochimica 38 (6), 573–582. https://doi.org/10.19700/j.0379-1726.2009.06.007 (in Chinese).

Zhao, W.Z., Hu, S.Y., Hou, L.H., et al., 2020. Types and resource potential of continental shale oil in China and its boundary with tight oil. Petrol. Explor. Dev. 47 (1), 1–10. https://doi.org/10.1016/S1876-3804(20)60001-5 (in Chinese).

Zhao, W.Z., Hu, S.Y., Hou, L.H., 2018a. Connotation and strategic role of in-situ conversion processing of shale oil underground in the onshore China. Petrol. Explor. Dev. 45 (4), 563–572. https://doi.org/10.1016/S1876-3804(18)30063-6.

Zhao, X.Z., Pu, X.G., Han, W.Z., et al., 2017. A new method for lithology identification

of fine grained deposits and reservoir sweet spot analysis: a case study of Kong 2 Member in Cangdong sag, Bohai Bay Basin, China. Petrol. Explor. Dev. 44 (4), 492–502. https://doi.org/10.11698/PED.2017.04.02 (in Chinese).

Zhao, X.Z., Zhou, L.H., Pu, X.G., et al., 2018b. Geological characteristics of shale rock system and shale oil exploration breakthrough in a lacustrine basin: a case study from the Paleogene 1st sub-member of Kong 2 Member in Cangdong sag, Bohai Bay Basin, China. Petrol. Explor. Dev. 45 (3), 361–372. https://doi.org/10.11698/PED.2018.03.01 (in Chinese).

Zhu, R.K., Cui, J.W., Deng, S.H., et al., 2019. High-precision dating and geological Significance of Chang 7 Tuff Zircon of the Triassic Yanchang Formation, Ordos Basin in Central China. Acta Geol. Sin. 93 (6), 1823–1834. https://doi.org/10.1111/1755-6724.14329.

Zou, Z.F., Qi, Y., Ge, X.R., et al., 2008. Effect of late Triassic volcanic sediment event on hydrocarbon accumulation conditions in Ordos Basin. J. Lanzhou Univ. 44 (3), 12–15. https://doi.org/10.13885/j.issn.0455-2059.2008.03.015 (in Chinese).