HIERARCHICAL NEURAL NETWORKS IN PREDICTING THE PROPERTIES OF OIL AND GAS RESERVOIRS BASED ON WELL AND SEISMIC DATA

I.I. Priezzhev[1,2], D.A. Danko[1], A.N. Onishchenko[2]

[1]*Gubkin Russian State University of Oil and Gas, Leninsky pr. 65, Moscow, 119991, Russia*

[2]*OOO Priezzhev Laboratory, Leninskiye Gory 1, 119234, Russia*

Received: 27 May 2024
Accepted: 2 August 2024
Published online: September 2024
DOI: 10.2113/RGG20244760

A reference:
Priezzhev I.I., Danko D.A., Onishchenko A.N. Hierarchical neural networks in predicting the properties of oil and gas reservoirs based on well and seismic data // Russian Geology and Geophysics, 2024, DOI: 10.2113/RGG20244760

✉*E-mail*: danko.d@gubkin.ru

This paper describes a technique for hierarchical neural networks based on the nearest neighbor method with preliminary clustering of the original training dataset and construction of a search cluster decision tree. This method is a promising alternative to neural network technologies with deep learning and has quite a few advantages: high learning rate, identification of objects with a low degree of similarity, and the ability to generalize and retrain. As shown by testing the hierarchical neural network method on real data from the West Siberian oil and gas province, predicting the oil saturation in the Vikulov suite interval is much faster and more efficient than inversion approaches to quantitative interpretation of seismic data while achieving fairly similar geological results. This characterizes the proposed method of hierarchical neural networks as an effective tool for the quantitative interpretation of seismic data to solve geological problems.

Neural networks, seismic exploration, interpretation, Vikulov suite, oil and gas, Western Siberia

INTRODUCTION

The application of machine learning methods in various scientific fields has recently shown stunning results. The use of a large amount of information has contributed not only to the ability to classify, recognize, and generate speech and texts, but also to the advent of smart chatbots, virtual assistants, and more. The relative universality of algorithms and the similarity of tasks make it possible to apply machine learning methods in related fields of research, including the interpretation of well and seismic data for the purpose of forecasting promising oil and gas objects.

Such problems are often solved using machine learning methods, in which the forecast is based on a training dataset. The latter consists of a set of (X, Y) pairs, where X is a vector of values characterizing the properties or characteristics of objects (vector of object features) and Y is a vector of forecast parameter values.

The most common way to obtain a forecast for the dataset being analyzed is to use deep machine learning algorithms based on multilayer neural networks. When setting the task of directly forecasting reservoir properties based on well data and seismic exploration, the training dataset can be enormous. This often happens when a considerable number of wells are involved in the project and seismic data are presented over broad areas or have high multiplicity. In this case, a large multilayer neural network is required to forecast reservoir properties, which is quite difficult to train. The training process itself can take a long time, which is not always feasible under tight project deadlines. In addition, neural networks are very sensitive to the quality of well-to-seismic ties if the forecast is performed in the time domain, as well as to the quality of seismic data processing results. This can lead to noisy and laterally intermittent forecast results and a decrease in the accuracy of forecast parameters at well locations.

Another similar and common approach is the nearest neighbor method, which seems to be quite a simple and popular algorithm in various fields of machine learning. It requires directly comparing all the features of the analyzed object with the features of objects from the training set. This method does not have a training stage. The forecast result is calculated on the basis of selecting the most similar elements in the training dataset pairs. Due to the high degree of freedom and smoothness of the proximity field, the nearest neighbor method is less sensitive to the quality of the well tie and to areas where the seismic data quality is reduced. However, performing the operation of comparing the objects under study with each element of a large training dataset can take an unacceptably long time, which is a disadvantage of this method.

There are many approaches to accelerating the nearest neighbor search, such the method of hierarchical navigable small world graphs (Malkov and Yashunin, 2018). It is suggested in this paper that one should use a similar method, which is related to hierarchical neural networks and based on sequential clustering of the original training dataset and the construction of a search cluster decision tree.

HIERARCHICAL NEURAL NETWORKS

The method proposed in this study is based on the idea of accelerating the nearest neighbor method by pre-creating a search cluster decision tree based on the training dataset.

In the case of predicting the properties of oil and gas reservoirs based on well measurements and seismic data, the proposed method can be represented as a scheme consisting of the following blocks (Fig. 1):

Block 0: Input Data Conditioning. This block is not directly related to the method under consideration, but is necessary and important for all types of scientific and production work related to the quantitative interpretation of seismic data.

It includes data gathering, analysis, loading, and quality control of input geological and geophysical information. Particular attention is paid to the preparation of well data and wave fields for quantitative interpretation. Seismic data should be processed to preserve true amplitude variation with offset (AVO) effects and should have a high signal-to-noise ratio. Well measurements are used to calculate target parameters that are consistent with well logging (WL) data and laboratory results of core data analysis, such as volumetric models of rock mineral composition, fluid saturation, discrete distributions (lithotypes), porosity, permeability, and geomechanical properties. The quality of well-to-seismic ties in the time or depth scale domains is also emphasized.

Block 1: Training Set Formation. The training set is formed by grouping input seismic data and target parameters in wells to create a set of *seismic response – forecast parameter* pairs (for convenience, "response" and "parameter" are italicized to describe the training set pairs in this paper). Full stack and partial (angle or offset) stacks, or prepared pre-stack seismic gathers after migration can be used as input seismic data (Fig. 2).

In the given example, the target parameter is a porosity distribution obtained from the quantitative interpretation of WL data. When forming a training dataset pair, the *forecast parameter* corresponds to the average porosity value falling within a cell of a regular seismic grid and is calculated along the wellbore trajectory. The seismic response is an m x n seismic wavefield window, where m is the number of vertical samples in the time or depth domain of the input seismic data, and n is the number of the nearest traces. The window center corresponds to the location of the *forecast parameter* cell. The values of m and n are user-defined and selected during quantitative prediction.

Block 2: Construction of a Search Cluster Decision Tree. A search cluster decision tree is constructed on the basis of the idea of multilevel clustering of seismic responses depending on their similarity or differences. As a result, clusters of seismic responses with a high degree of similarity are formed at the end nodes of the tree (tree leaves), which have the corresponding values of the target parameters. This approach significantly accelerates the decision-making process. As compared to simple enumeration, this approach makes it easier and quicker to obtain target parameter distributions between wells based on seismic data using the nearest neighbor or several nearest neighbors.

The construction of a search cluster decision tree is performed based on a set of *seismic responses* from the pairs of the training sample formed in Block 1 (Fig. 2). In the initial («root») node of the decision tree, the input set of seismic responses is divided into a given number of clusters. In the next internal tree node, each upper-level cluster is divided into lower-level clusters down to the lowest level – the "leaves" of the decision tree (Fig. 3).

The number of clusters and the tree depth are determined by the user during the parameter analysis when performing quantitative prediction. If there is division into two clusters for each node, then a binary search tree is created. Dividing into a larger number of clusters allows for more complex search trees, with the formation of leaves at different levels. If the minimum number of objects in the leaves is set, the tree depth is determined automatically.

The clustering process is based on the unsupervised classification algorithm known as Kohonen's method of self-organizing maps (Kohonen, 2001) using the Kohonen 1D/2D/3D multidimensional space projection options (Priezzhev et al., 2019). With this approach, two seismic responses are assigned to the same cluster if their shapes are similar. The similarity measure is defined by the distance between response shapes in a multidimensional

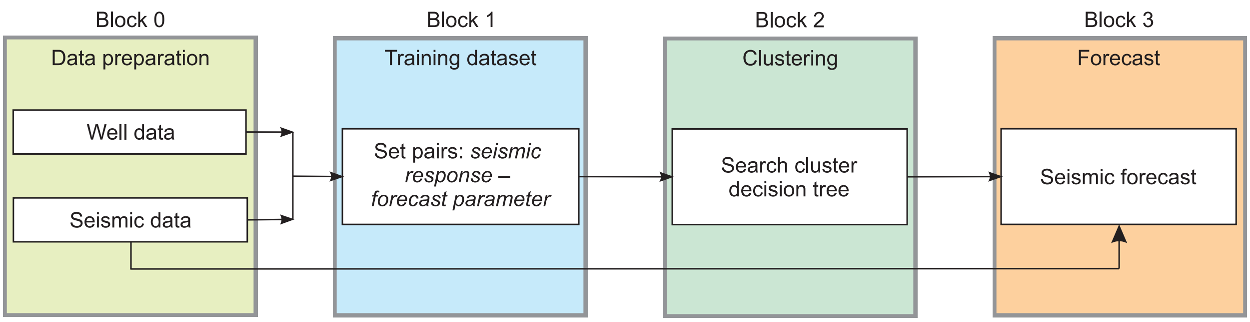


Fig. 1. Scheme of the hierarchical neural networks applied.

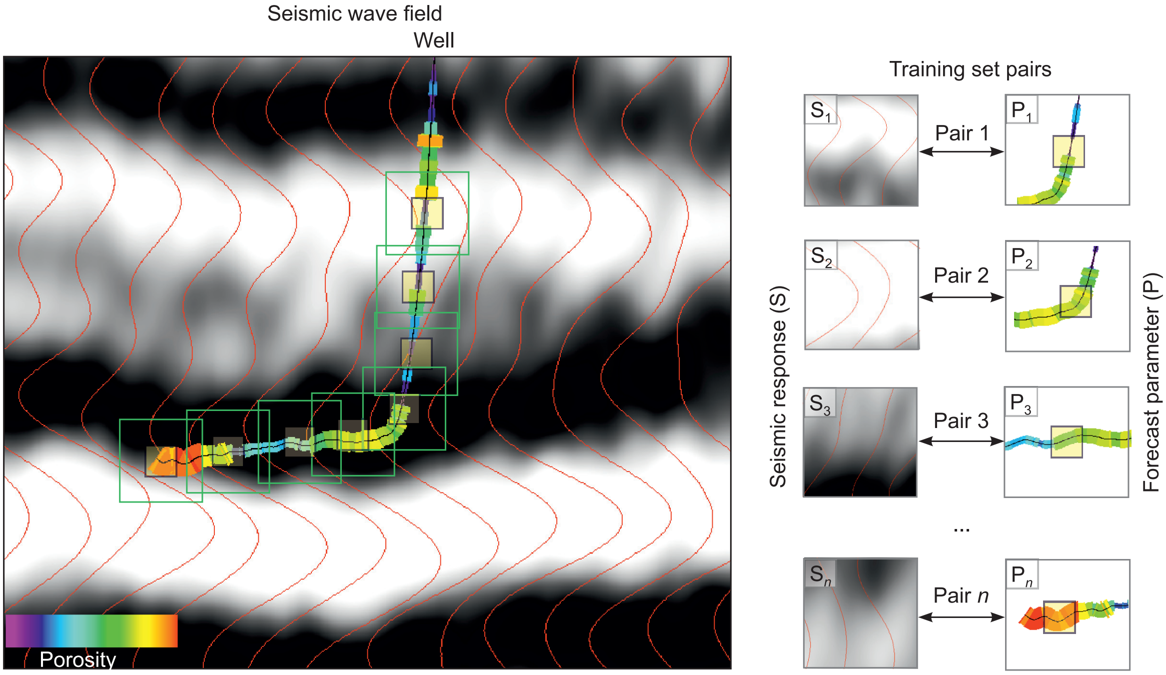


Fig. 2. Principle of forming a training dataset using well and seismic data. Here and in Fig. 4, the porosity colour scale shows the change in values from low (cold tones) to high (warm tones), and porosity is chosen as an example of a predictive geological parameter. The green square indicates the window in which the seismic response is determined. Together with the value of the forecast parameter (yellow square), pairs of training samples are formed.

space. In a simple case, the distance can be defined as the difference in squares between two responses (L2 norm).

Mathematically, the classification procedure for one tree level can be described as follows. For simplicity, the Kohonen 1D clustering variant is considered. Let the training sample of Block 1 include the element S_i as the i-th seismic response, which contains n seismic samples. The use of the Kohonen 1D algorithm assumes creating a

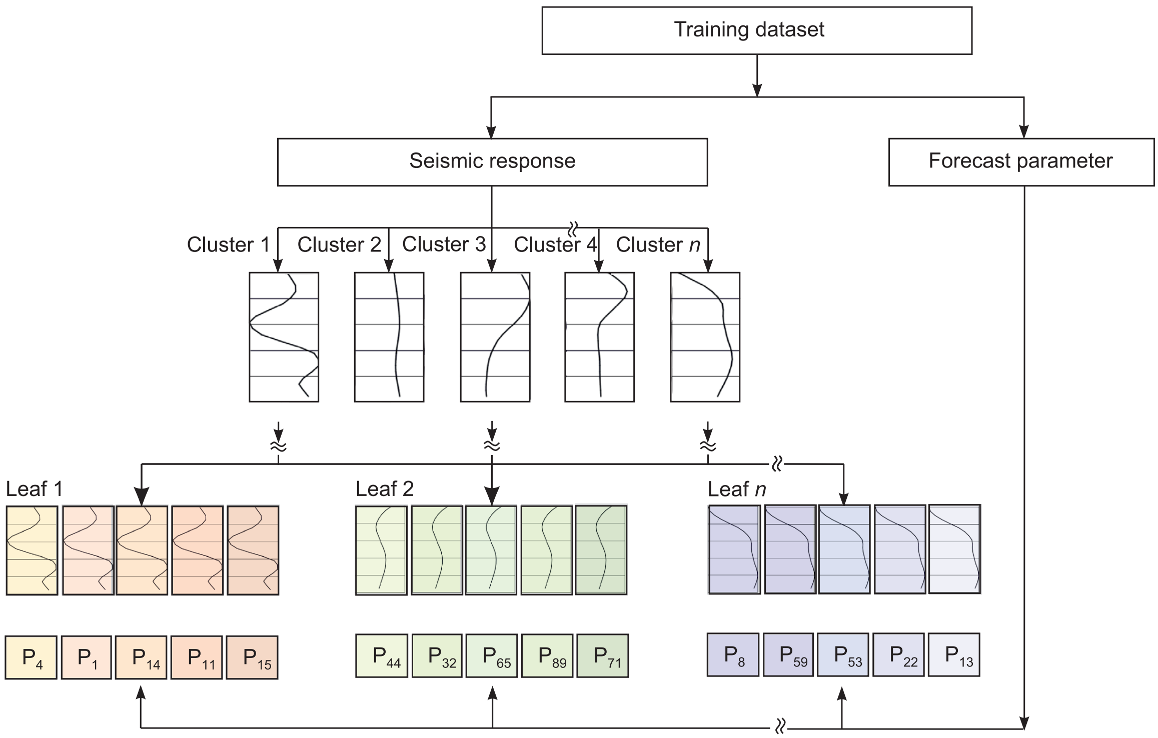


Fig. 3. Cluster decision tree construction scheme.

neural network of a predetermined size k, where each j-th neuron ($j = \overline{1, k}$) is described by vectors of the same size

$$W_j = \left\{w_{1j}, w_{2j}, \ldots, w_{nj}\right\}.$$

Initially, neurons are initialized with randomly selected seismic responses from the training dataset. Further cycles are organized over all elements of seismic responses S_i from the training sample, and the closest "winning" neuron W_j from the neural network is obtained using various measures of proximity, for example, in the form of L2 norm:

$$L_2\left(S_i, W_j\right) = \sqrt{\frac{1}{n}\sum_{l=1}^{n}\left(s_{li} - w_{lj}\right)^2}$$

The values of the "winning" neuron are then adjusted with weights a_{ij} inversely proportional to the number of seismic responses previously included in it (corrections performed):

$$W_j^{new} = W_j^{old} + a_{ij} S_i$$

The values of neurons closest to the "winning" neuron by index j are also adjusted. The adjustment is performed using weight coefficients, which increase in value as they approach the "winning" neuron.

For the Kohonen 2D and Kohonen 3D clustering variants, one can use a structure of neuron arrangement either with two indices (j^1 and j^2) and dimension $k \cdot k$ (class map) or with three indices (j^1, j^2, and j^3) and dimension $k \cdot k \cdot k$ (class cube), respectively. In this case, the values of the neurons close to the "winning" neuron in all indices are adjusted. The resulting indices can be interpreted as new variables, and the Kohonen 1D/2D/3D clustering algorithm can be considered not only as a method for grouping similar objects (seismic responses), but also as a method for reducing the feature space from size n to size k (1D), $k \cdot k$ (2D), or $k \cdot k \cdot k$ (3D). Increasing the clustering dimension to 2D or 3D reduces the possibility of clustering errors, thereby improving the quality of search cluster trees when constructing a hierarchical neural network.

After creating a cluster tree, the *seismic response – forecast parameter* pairs are analyzed in each tree leaf. It is logical to assume that, if the seismic responses in each tree leaf have a high similarity measure, then the *forecast parameters* should also be similar to each other. In this case, the use of the k-nearest neighbors method allows one to average pairs by accounting for the weight coefficients of their similarity measure and/or spatial proximity. In the case of predicting discrete property values based on well data, such as reservoir/nonreservoir or lithofacies, the averaging operation is performed by selecting the maximum number of identical values in a list of the nearest neighbors. As a result, the most probable *forecast parameter* among the k-nearest neighbors is obtained. If the *forecast parameters* differ significantly from each other within one leaf, then the *seismic response – forecast parameter* connectivity in the training set is either low or absent.

Block 3: Prediction Based on Seismic Data. In this block, target parameters are predicted for the entire seismic dataset using the "nearest neighbors" method based on the search cluster decision tree created in the previous block.

The entire seismic dataset used to form the training sample is fed to the cluster decision tree. The *seismic response* in each trace of the input set is compared within a user-specified time window (target interval) to seismic responses in each cluster of the search decision tree. The comparison process begins with the root node and ends with one of the tree leaves. Thus, the seismic response is assigned the value of the target *forecast parameter* in accordance with the pairs of the training dataset (Fig. 4).

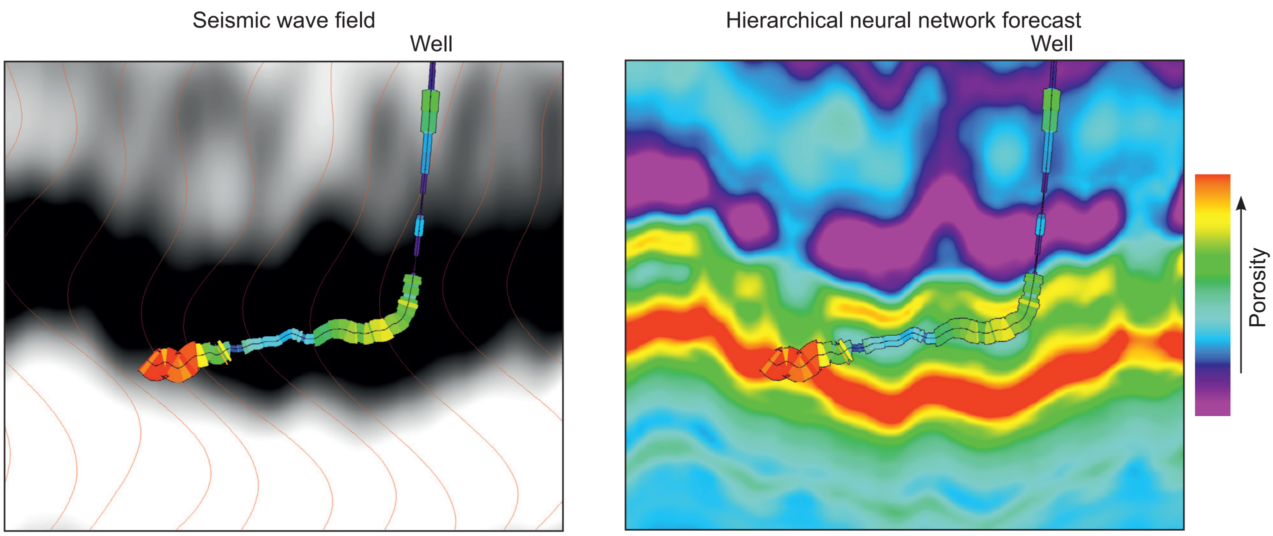


Fig. 4. Example of hierarchical neural network porosity forecast based on seismic data.

If the input seismic data contains a response that has a low degree of similarity with the responses from the training sample and, accordingly, is not used in clustering to create a decision tree, then such a response is placed in a separate unlabeled cluster. All subsequent responses that are similar to the response from the unlabeled cluster are added to this cluster. Otherwise, a new unlabeled cluster with the next ordinal number is created. In this way, additional training occurs and these new clusters can be added to the process of pattern recognition and prediction of target parameters for the entire seismic dataset.

TESTING ON REAL DATA

The proposed hierarchical neural network algorithm is tested at one of the oil and gas fields of the Krasnoleninsky arch in Western Siberia. The target object is the terrigenous deposits of the Vikulov suite from the Upper Aptian to the Lower Albian of the Lower Cretaceous. The oil potential of the suite is associated with shallow-water and coastal-marine sandstones in reservoirs VK_1 and $VK_{2\text{-}3}$ of lenticular structure. The oil deposits are undersaturated and characterized by extensive transitional water-oil zones and increased water content in the rocks, leading to high water cut during field operations and low confirmability of the initial geological and recoverable oil reserves (Medvedev, 2010; Shubin et al., 2020).

The Vikulov suite has been developed for over 30 years, resulting in the accumulation of a vast amount of geological and geophysical information, including 2D/3D seismic surveys and data from thousands of drilled wells. Despite this, there are few examples of quantitative interpretation of seismic data being used because the main criterion for drilling wells is a structural plan (Fig. 5*a*). Therefore, the presence of vast volumes of seismic and well information within the Krasnoleninsky arch is a real challenge not only for standard approaches to quantita-

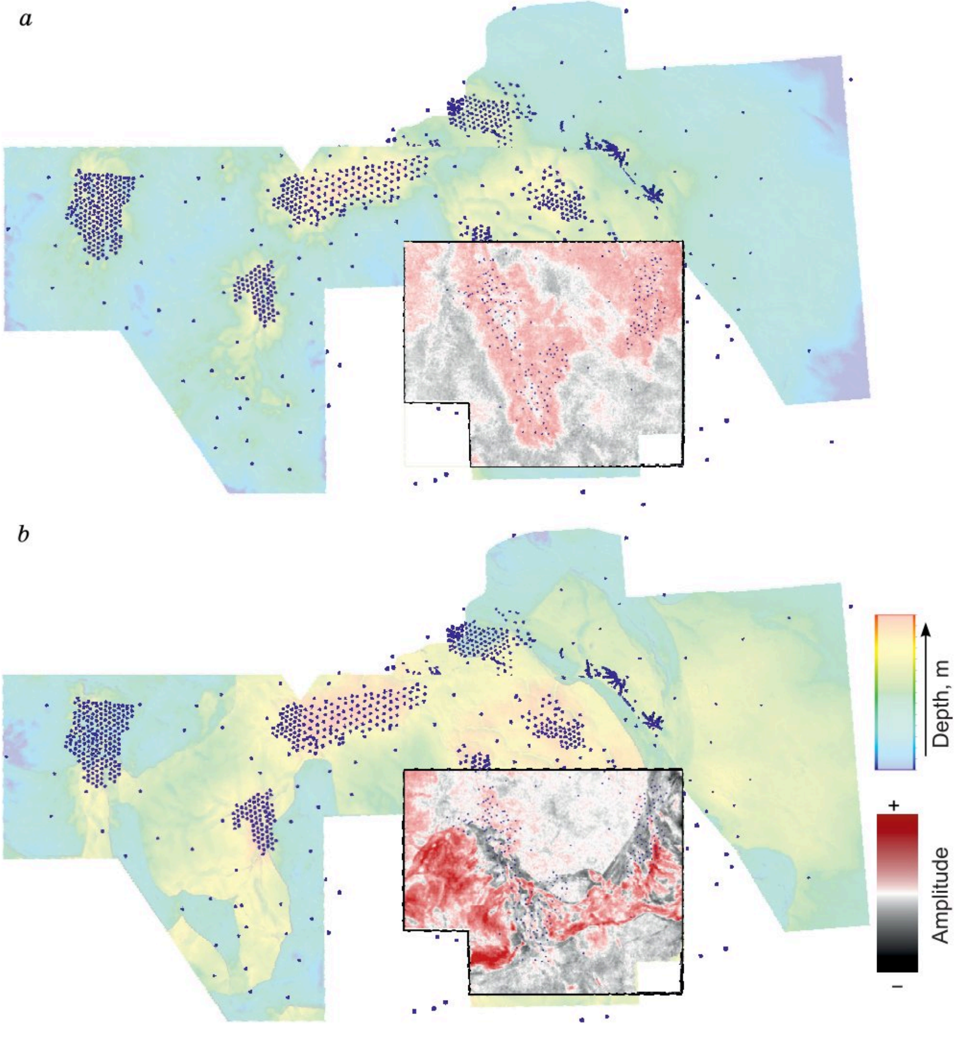


Fig. 5. Structural maps of the Vikulov suite beds of the studied deposit with the outline of the reference region of seismic exploration data: *a* – along the top of reservoir VK_1; *b* – along the top of reservoir $VK_{2–3}$, complicated by deposits of the incised river valley. Here and in Fig. 7, the blue dots show the position of the wells in the area.

tive interpretation of seismic data, but also for the case of search for promising nonstructural oil and gas targets using neural networks.

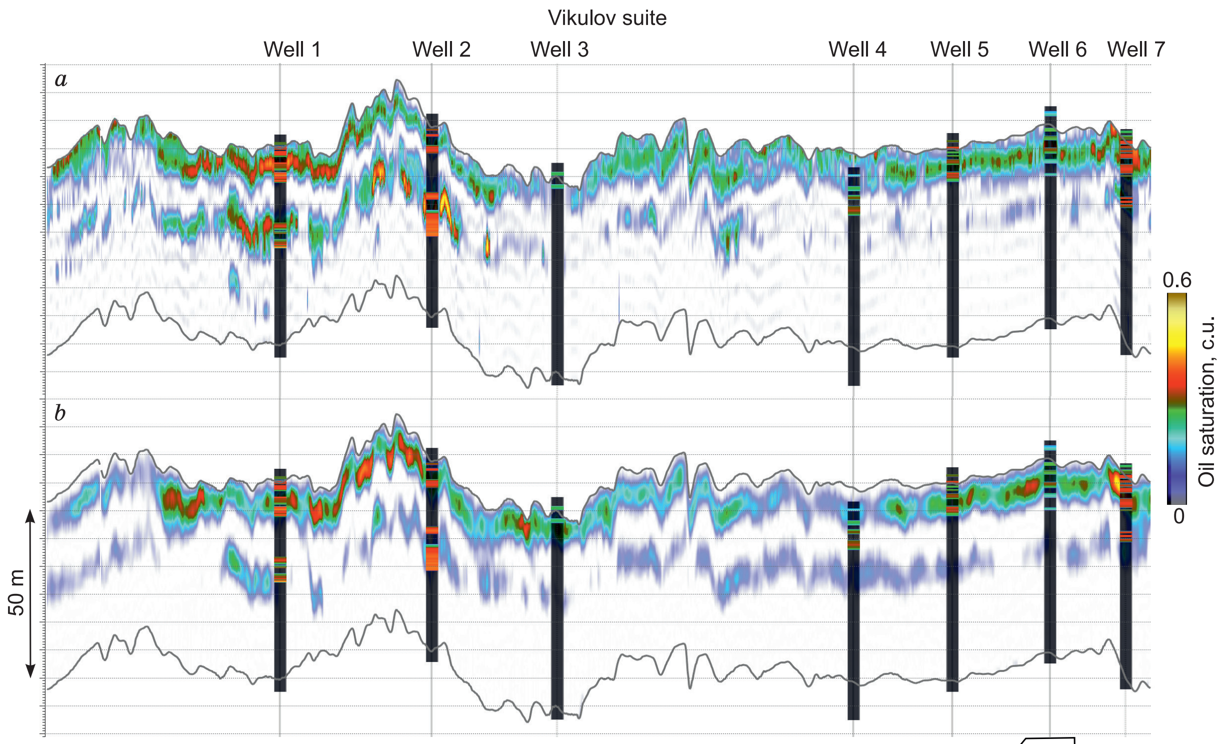
The field under study has 3D seismic data in both the time and depth domains, processed using the well-driven approach. The total area of the seismic survey is about 1700 km^2 with a seismic fold of 72. Additionally, more than 1500 wells have been drilled in the field, and the effective oil-saturated thicknesses have been calculated in 1250 of them. Moreover, the detailed interpretation of WL data with continuous assessment of oil saturation in the studied interval has been performed in 37 wells (Fig. 5). Part of the Vikulov suite in the field is complicated by an incised valley ($VK_{2–3\ inset}$), whose deposits are also oil-saturated (Fig. 5*b*).

It should be clarified that the hierarchical neural network method described in this study has been tested over the entire seismic survey area. However, for reasons of confidentiality, the research results are shown in a limited region highlighted by a polygon in Fig. 5 with a modified survey geometry.

The main target parameter for neural network prediction in the studied field is the oil saturation of the pore space in the Vikulov suit reservoirs. It is noteworthy that it is often impossible to forecast fluid saturation using seismic data due to many factors, including the rigidity of the mineral matrix, the low porosity of reservoir rocks, and the similar elastic properties of oil and water. However, as shown in (Shubin et al., 2023), the effect of oil saturation in the Vikulov suit interval on the seismic response is expressed by the presence of a regression relationship between effective oil-saturated thicknesses and the far offset (angle) amplitude values. This creates the prerequisites for predicting oil saturation based on seismic data over the entire region under study, including the use of neural networks.

The training set for the hierarchical neural network algorithm described in this paper includes oil saturation curves obtained from 30 wells after the detailed interpretation of WL data (seven wells were used as blind wells) as the target parameter, as well as sets of angle stacks with average angles (equal to 5°, 15°, 25°, 35°, and 45°), their amplitude/frequency attributes, and calculated AVO (Amplitude Variation Offset – the dependence of the change in the amplitude of reflected waves on the distance) attributes (intercept and gradient) as the *seismic response* (Fig. 1, Block 1). Hierarchical neural networks are trained using the Kohonen 1D classification algorithm with 30 nearest neighbors. The number of hidden layers corresponds to five, with 15 neurons in each hidden layer.

One of the main problems in using neural networks in the quantitative interpretation of seismic data is the geo-



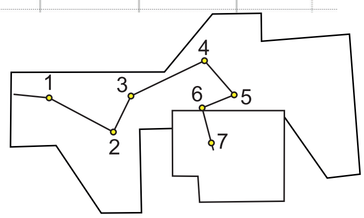


Fig. 6. Oil saturation forecast sections passing through blind wells: *a* – based on hierarchical neural networks; *b* – based on petroelastic inversion. Here and in Fig. 8, the well numbers in the diagram correspond to the well numbers in the section. The contour of the polygon is indicated around well 7, according to which the publication of results is allowed. For more information, see the text.

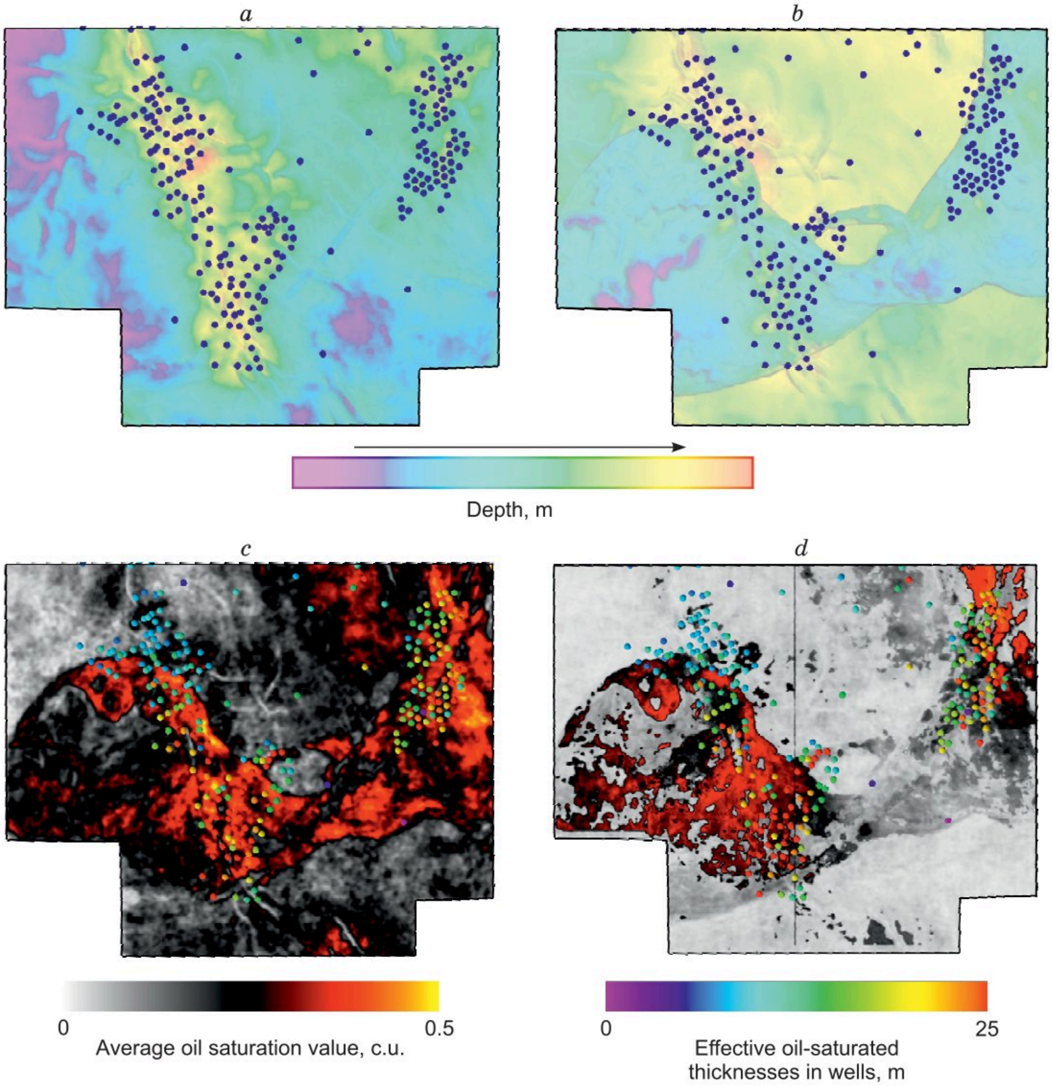


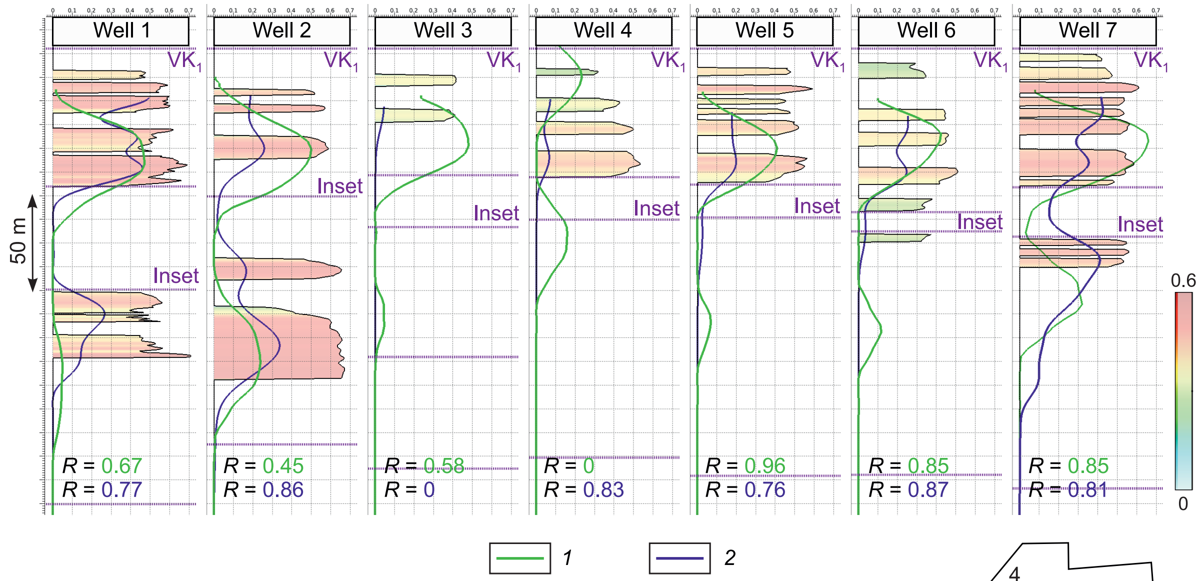
Fig. 7. Structural maps of the Vikulov suite formations: along the top of the VK_1 formation *a*; along the top of the VK_{2-3} formation complicated by deposits of the incised river valley *b*. Maps of average oil saturation values in the VK_1–VK_{2-3} intervals: according to the result of petroelastic inversion *c*; according to the result of hierarchical neural networks *d*.

logical justification of the forecast obtained. Often, mathematical estimates of the neural network results show a high degree of reliability in the forecast. However, it is quite difficult to geologically comprehend and integrate the obtained result into the existing geological model. Therefore, the proposed hierarchical neural network forecast is compared and verified using the oil saturation estimate obtained at the studied field on the basis of the inverse rock physics (IRP) technique (Shubin et al., 2020). It is noteworthy that IRP was previously used with the same set of seismic and well data on which the hierarchical neural networks were tested. The obtained results were reported and approved in subsoil user companies and were also presented in scientific publications (Shubin et al., 2020, 2022, 2023).

In Fig. 6, the oil saturation sections calculated using hierarchical neural networks (Fig. 6*a*) are compared to the posterior average oil saturation obtained using the IRP technique (Fig. 6*b*). The wells shown in the sections are not included in the predictions of both approaches and are used only as blind wells to verify the results obtained.

Figure 7 shows the maps of average oil saturation values in the VK_1-VK_{2-3} intervals also calculated using the two methods under consideration.

The comparisons shown in Figs. 6 and 7 indicate the geological similarity of the results obtained, thereby confirming the effectiveness of the proposed hierarchical neural network algorithm for solving the geological problems set using seismic data. Figure 8 shows a detailed comparison of oil saturation forecasts obtained using both methods at the locations of the blind wells. The results are presented on a single scale for visual convenience. As shown by the analysis, both methods yield forecast results comparable in errors and accuracy: 0.66



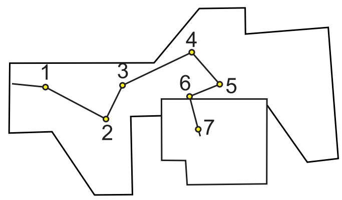


Fig. 8. Comparison of oil saturation forecasts obtained using petroelastic inversion and hierarchical neural networks (‘blind’) wells on a single general scale. *1* – the result of the petroelastic inversion, *2* – the result of the hierarchical neural networks application. *R* – the correlation coefficient, c.u. Inset – the designation of a stratigraphic reference corresponding to an incised river valley.

for the overall correlation coefficient of IRP and 0.71 for the hierarchical neural networks.

RESULTS AND DISCUSSION

As shown by testing the proposed algorithm on real data, hierarchical neural networks are a fairly effective and fast tool for quantitative interpretation of seismic data.

The use of hierarchical neural networks in one of the fields of the Krasnoleninsky arch in Western Siberia indicates that the obtained results are quite similar to the IRP results for quantitative estimates of oil saturation (Fig. 8) and lateral geological details (Fig. 6). However, a significant advantage of the developed algorithm is the quickness of forecast calculation. For example, the calculation time for the entire study area (1700 km^2) in the Vikulov suite interval is approximately 1.5 h, whereas the calculation using only IRP based on the available acoustic impedance cubes and the v_P/v_S ratio requires about two days. The calculations are performed using a workstation with the following specifications: Intel Xeon CPU E5-2667 3.20GHz, 192 GB RAM, and Nvidia Tesla M60x2.

For the purpose of better visual comparison, the stages of predicting the proposed hierarchical neural network algorithm and the IRP technique are presented in the table below (Table 1):

It is also worth noting that the number of stages required for performing a forecast using hierarchical neural networks is significantly smaller than in classical methods of dynamic interpretation of seismic data. This largely reduces the possibility of subjective human error in the neural network forecasting process as compared to multistage methods, such as seismic inversion and IRP (Table 1).

Table 1. Comparison of the stages of the process of obtaining a forecast using IRP and hierarchical neural networks

	Forecast stage	Petroelastic inversion	Hierarchical neural networks
1	Construction of a volumetric lithological-petrophysical model on the basis of WL data	+	+
2	Analysis of petroelastic bonds	+	+
3	Petroelastic modeling	+	↓
4	Seismic synchronous inversion	+	↓
5	Inversion of petroelastic model parameters	+	↓
6	IRP/neural network forecast	+	+
7	Interpretation of results	+	+

Note. The arrow indicates skipping the stages. «+» – the stage must be completed.

Aside from its application in the direction of quantitative interpretation of seismic data, the proposed method of hierarchical neural networks can be used in scientific fields beyond the scope of the oil and gas industry. For example, the method has also been tested in handwritten character recognition tasks using a well-known dataset from the Modified National Institute of Standards and Technology (MNIST). The MNIST database contains 60,000 images of handwritten characters with a size of 28 × 28 pixels for training and 10,000 images for testing. The method involves constructing a search cluster decision tree based on the training set of characters, which is then used to find similar images from the test set. In the case of using the full training dataset without dividing it into clusters, the search time for the nearest neighbors is 80.41 min with a recognition error of 369 characters (3.69%) across all tested characters. After constructing a search cluster decision tree consisting of ten clusters at each level and with a tree depth of up to five levels, the prediction time for the test set is reduced to 0.12 min, making it 670 times faster. However, the recognition error also increases to 564 characters (5.64%) due to the ambiguity of cluster division.

CONCLUSIONS

The hierarchical neural network algorithm presented in this paper is an effective tool for the quantitative interpretation of seismic data, which can be used as an alternative or supplement to classical interpretation technologies for solving complex geological problems. The main advantages of the algorithm include a high learning rate, the ability to identify seismogeological objects with a low degree of similarity, and the capacity for generalization and retraining.

As compared to multistage methods, such as seismic inversion and inverse rock physics, the use of hierarchical neural networks allows for performing one-step calculations without relying on a number of complex software packages to obtain intermediate results. This significantly reduces the time required to calculate target geological and geophysical parameters, decreases the probability of error accumulation when transitioning from one stage of quantitative interpretation to another, and increases the efficiency of the resulting forecasts of oil and gas reservoir properties based on well and seismic data.

The authors express their gratitude to colleagues from the National University of Oil and Gas (Gubkin University), the Lomonosov Moscow State University, PAO NK Rosneft, and PAO Lukoil for the opportunity to test the developed hierarchical neural network algorithm on real data and to publish the results obtained.

REFERENCES

Kohonen, T., 2001. Self-Organizing Maps. Third Ed., Heidelberg, Springer Berlin, doi: 10.1007/978-3-642-56927-2.

Malkov, Y.A., Yashunin, D.A., 2018. Efficient and robust approximate nearest neighbor search using hierarchical navigable small world graphs. IEEE Trans. Pattern Anal. Mach. Intell. 42 (4), 824–836, doi: 10.1109/TPAMI.2018.2889473.

Medvedev, A.L., 2010. Aptian incised river valleys of the Kamennoye field, Western Siberia: Regional aspects of petroleum potential. Neftegazovaya Geologiya. Teoriya i Praktika, vol. 5 (3), 1–27.

Priezzhev, I.I., Veeken, P.C.H., Egorov, S.V., Nikiforov, A.N., Strecker, U., 2019. Seismic waveform classification based on Kohonen 3D neural networks with RGB visualization. First Break 37 (2), 37–43, doi: 10.3997/1365-2397.2019012.

Shubin, A.V, Klyazhnikov, D.V., Ryzhkov, V.I., 2020. Inverse rock physics: predicting porosity and fluid saturation of the Vikulov Formation. Geophysics 6, 31–41.

Shubin, A.V., Klyazhnikov, D.V., Ryzhkov, V.I., 2022. Rock physics inversion for fluid saturation and porosity prediction, in: Proceedings of the Second EAGE Conference on Seismic Inversion. European Association of Geoscientists and Engineers, pp. 1–5, doi: 10.3997/2214-4609.202229011.

Shubin, A.V, Klyazhnikov, D.V., Ryzhkov, V.I., 2023. On the possibility of forecasting oil saturation of a terrigenous reservoir based on seismic data, in: Proceedings of the Geomodel Conference [in Russian]. Geomodel, Gelendzhik, pp. 243–246.