

Research Article

Extreme Learning Machine for Reservoir Parameter Estimation in Heterogeneous Sandstone Reservoir

Jianhua Cao, Jucheng Yang, Yan Wang, Dan Wang, and Yancui Shi

College of Computer Science and Information Engineering, Tianjin University of Science and Technology, Tianjin 300222, China

Correspondence should be addressed to Jucheng Yang; jcyang@tust.edu.cn

Received 21 August 2014; Revised 8 November 2014; Accepted 10 November 2014

Academic Editor: Zhan-li Sun

Copyright © 2015 Jianhua Cao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study focuses on reservoir parameter estimation using extreme learning machine in heterogeneous sandstone reservoir. The specific aim of work is to obtain accurate porosity and permeability which has proven to be difficult by conventional petrophysical methods in wells without core data. 4950 samples from 8 wells with core data have been used to train and validate the neural network, and robust ELM algorithm provides fast and accurate prediction results, which is also testified by comparison with BP (back propagation) network and SVM (support vector machine) approaches. The network model is then applied to estimate porosity and permeability for the remaining wells. The predicted attributes match well with the oil test conclusions. Based on the estimations, reservoir porosity and permeability have been mapped and analyzed. Two favorable zones have been suggested for further research in the survey.

1. Introduction

In geosciences, reservoir is defined as the underground accumulation of oil or natural gas in sedimentary basins, and it is of great importance for petroleum exploration and development. Among the steps for well planning decisions, reservoir characterization is the essential one, and physical parameters estimation, including porosity and permeability, is the basic requirement in the characterization workflow.

As for the two geophysical parameters, porosity describes the fraction of void space in the sedimentary rocks, where the void may contain fluids, such as oil or natural gas. The more porous the rock is, the more the oil or gas may be preserved in the void spaces. And permeability describes the ability of rocks to transmit fluids. The more permeable the rock is, the easier the oil or gas could flow through. These two types of reservoir parameters are to some extent determining factors for reserve estimation and oil or gas production.

Practically, it is very complex and difficult for porosity and permeability estimation since lots of factors could affect the estimation accuracy, such as depositional formations, lithologic mineral components, measurement tools, data quality, and computational method.

There are mainly two types of approaches that have been used to acquire porosity and permeability data in reservoir research workflow. The first one is laboratory core analysis. Cores are obtained from drilled wells. Porosity and permeability can be determined precisely under strict core test principles. The results are reliable and are often used as reference for further estimation using mathematical ways. Due to the expensive cost, cores are often few in numbers for most of the oilfields. The second one is borehole log interpretation. The logs data are physical measurements performed by electric instruments lowered into the borehole. Specific physical characteristics of the rocks surrounding the borehole are recorded by logs with depth variations. Conventional logs include gamma ray (GR), acoustic slowness (AC), density (DEN), compensated neutron logs (CNL), and deep resistivity (RT). Among these logs, GR is often used to predict rock lithology, and the three logs of AC, DEN, and CNL have largely been used to estimate rock porosity. Permeability is estimated by combination of RT log and the former-estimated porosity. Empirical mathematical equations are often used when carrying out log interpretations. These equations are regression models built based on the correlation between geophysical logs and core-measured reservoir parameters. Since logs are

run for all wells in oilfield and the mathematical empirical equations are feasible to be used, log interpretation becomes the most used method in porosity and permeability estimation. But the estimation results rely greatly on the equations or the correlation models. Meanwhile, the relations between logs and geophysical parameters of rocks are nonlinear and very complicated. It is hard to get a universal solution for all wells in one survey or for all oilfields. So some nonlinear numerical method and artificial intelligence are brought into the log interpretation process and proposed as supplementary approach, so that more reliable and precise estimation data could be obtained for further reservoir evaluation.

Artificial neural networks have been proved to be capable of approximating any nonlinear function to any degree of accuracy provided that there are sufficient number of samples for network training and learning and have some successful applications in petroleum engineering, such as sedimentary microfacies prediction [1], lithology classification [2, 3], and reservoir prediction [4–6].

In petrophysical analysis, the neural network models have always acted as a predictor or estimator of deriving geophysical parameters, such as porosity and permeability where no core data is available [7–12]. Among the neural networks, BP network and SVM are the two commonly used learning algorithms in porosity and permeability estimation. BP neural network is a typical full-connected neural network with forward and error backpropagation part. The error could be backpropagated by adjusting weights in the learning process until it converges to a targeted value, which is very effective in solving nonlinear problems [13]. Support vector machine (SVM) is a network based on statistical learning theory and is especially designed for classification problems with different convolution kernel functions [14]. Satisfactory accuracy of estimations has been achieved when the networks are optimized with appropriate model parameters [15–19], although there still are some shortcomings in the applications, such as time-consuming and overfitting problems [20].

Extreme learning machine (ELM) is a single-hidden layer feedforward neural network (SLFN) proposed by Huang et al. [21, 22]. The ELM approach to training SLFN consists in the random generation of the hidden layer weights, followed by solving a linear system of equations by least squares for the estimation of the output layer weights. This learning strategy is very fast and gives good prediction accuracy. Theoretically and practically, this algorithm can produce good generalization performance in most cases and the speed has been proved to be much faster than conventional popular learning algorithms for feedforward neural networks. Till now, ELM has been widely studied and accepted by researchers and has demonstrated good generalization and prediction performance in many real-life applications [23–26]. But in petroleum reservoir prediction, there are still few applications.

In this paper, we examine the potential of ELM to predict porosity and permeability parameters in a heterogeneous sandstone reservoir in Permian formation, Yanqi survey of Ordos basin, China. Prediction models are established by SLFN trained with ELM and optimally pruned ELM (OP-ELM). OP-ELM is a variation of the ELM introducing

an optimal selection of the number of hidden units and variables modeling the problem [27, 28]. It is more robust and generic than conventional ELM [26]. In the study, reservoir parameters measured from cores and logs value at the same depth are paired as samples. Optimal prediction models for porosity and permeability estimation are established, which are finally used to interpret reservoir porosity and permeability for all wells in the survey.

The outline of this paper is as follows: Section 2 is the geologic background of the survey and brief introduction about the sandstone reservoir. Section 3 gives a short review of ELM and OP-ELM. Section 4 describes the preparation for the network model establishment. Section 5 gives the prediction results. Finally, Section 6 gives the conclusion of this work.

2. Geological Background

Yanqi survey is located in eastern Ordos basin, China. In the survey, there are 15 wells that encountered Permian sandstone reservoir. Oil has been discovered in 6 wells, which are marked with red color-filled circles in Figure 1.

According to the comprehensive study of Permian formation, the structure is monoclinical and west-dipping with few faults developed. The Permian formation is dominated by calcareous quartz sandstone, interbedded with organic-rich mudstone and thin-layer coal.

Owing to the complex depositional process and variable diagenesis, the sandstone reservoir is heterogeneous and changes fast spatially, which is controlled by the fluvial sedimentary microfacies. Core analysis shows that the main pore type is intragranular pore (Figure 2(b)), with proportion more than 55%, and the rest include secondary intergranular pore (Figure 2(c)), matrix pore, and microcrack. And the measured permeability varies greatly both horizontally and vertically.

Figure 3 is the cross-plot of actual porosity and permeability measured from core samples. It shows that porosity ranges from 2% to 14% and permeability value has large span from 0.0001 md to 10 md. The two parameters have highly positive correlation. Exponential regression has been executed for porosity and permeability, which is

$$\text{Permeability} = 0.0009e^{0.6567\text{Porosity}} \quad (1)$$

and the correlation coefficient (R) is 77.9%. In conventional log interpretation, this statistical regression model can be recommended for permeability estimation if porosity is determined, but practically porosity is hard to be accurately calculated using current empirical formula. What is more, the correlation coefficient is not good enough, and the error of porosity calculation can also be brought into the permeability estimation process and affect its accuracy. So the model may not be suitable for all wells of the survey.

According to the petroleum industry criteria, the sandstone reservoir in Yanqi survey has medium porosity and permeability. In the survey, 6% is set as the threshold of porosity for sandstone reservoir, which means if porosity is lower than

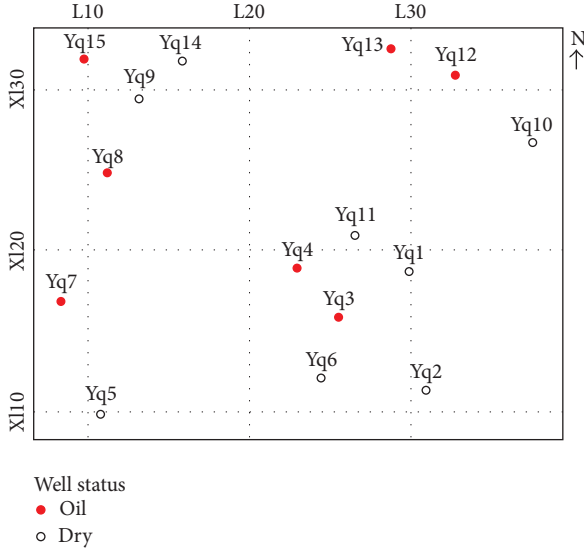


FIGURE 1: Well location and base map of the Yanqi 3D survey. There are totally 15 wells, with 7 wells encountering oil in Permian sandstone reservoir, while the rest are dry.

the threshold, the sandstone reservoir is considered tight and uneconomic.

3. Extreme Learning Machine

In this section we review the fundamental definitions of the two learning algorithms applied in the experiments reported below to predict porosity and permeability parameters in Permian sandstone: the ELM and the OP-ELM.

3.1. ELM. Extreme Learning Machine (ELM) is a simple supervised learning algorithm proposed by Huang et al. [21] for Single-hidden Layer Feedforward Neural Network (SLFN). Comparing with the conventional neural networks, ELM has better performance in learning efficiency and universal approximation capability. Different from BP network, the input weights and biases of ELM are randomly assigned and need not be adjusted within the training phase, and the output weights can be determined analytically by finding the least squares solution. Therefore the neural network is obtained after a few steps with very low computation cost.

Given a dataset containing N training samples $(\mathbf{x}_i, \mathbf{t}_i)$, $i = 1, 2, \dots, N$, where \mathbf{x}_i is a $n \times 1$ input vector $\mathbf{x}_i = [\mathbf{x}_i^1, \mathbf{x}_i^2, \dots, \mathbf{x}_i^n]^T \in \mathbf{R}^n$ and \mathbf{t}_i is a $m \times 1$ target vector $\mathbf{t}_i = [\mathbf{t}_i^1, \mathbf{t}_i^2, \dots, \mathbf{t}_i^m]^T \in \mathbf{R}^m$, the output of a SLFN with M hidden nodes can be expressed mathematically as

$$\sum_{i=1}^M \beta_i g_i(\mathbf{x}_j) = \sum_{i=1}^M \beta_i g(\mathbf{w}_i \cdot \mathbf{x}_j + \mathbf{b}_i) = \mathbf{t}_j, \quad (2)$$

$$j = 1, 2, \dots, N,$$

where \mathbf{w}_i is the weight vector connecting the i th hidden node with the input nodes, β_i is the weight vector connecting the i th hidden node with the output nodes, and \mathbf{b}_i is the threshold

of the i th hidden node. $g(\cdot)$ denotes the nonlinear activation function of the hidden node. It can be the identity sigmoid or Gaussian function, among a large collection of polynomial functions.

Equation (2) can be written in a more compact format as follows:

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{T}, \quad (3)$$

where \mathbf{H} is the hidden layer output matrix of the network:

$$\mathbf{H} = \begin{bmatrix} g(w_1 x_1 + b_1) & \dots & g(w_N x_1 + b_N) \\ \vdots & \ddots & \vdots \\ g(w_1 x_M + b_1) & \dots & g(w_N x_M + b_N) \end{bmatrix}, \quad (4)$$

$\boldsymbol{\beta}$ is the matrix of hidden-to-output weights, and \mathbf{T} is the target matrix.

In (4), weights (\mathbf{w}_i) and biases (\mathbf{b}_i) are randomly assigned and $g(\cdot)$ is known to be selected as sigmoid function, so the output of hidden nodes could be determined, which is \mathbf{H} in (3). The remaining problem becomes a set of linear equations and can be solved by minimum square error estimation:

$$\min_{\boldsymbol{\beta}} \|\mathbf{H}\boldsymbol{\beta} - \mathbf{T}\|. \quad (5)$$

According to the definition of the Moore-Penrose generalized inverse, the smallest norm least squares solution of (3) is given as

$$\hat{\boldsymbol{\beta}} = \mathbf{H}^{-1}\mathbf{T}, \quad (6)$$

where \mathbf{H}^{-1} is the Moore-Penrose generalized inverse of matrix \mathbf{H} .

For (6), once the \mathbf{H} and \mathbf{T} are set, it is not difficult to get the $\boldsymbol{\beta}$ matrix. The process has proved several advantages: (1) the training error is minimized; (2) the generalization performance is optimal; (3) the solution is unique.

Totally, the ELM algorithm can be summarized as follows [24].

Given a training set $\{(x_i, t_i) \mid x_i \in \mathbf{R}^n, t_i \in \mathbf{R}^m, i = 1, \dots, N\}$, an activation function g , and the number of hidden nodes M ,

- (1) randomly set input weights \mathbf{w}_j and biases \mathbf{b}_j ;
- (2) calculate the hidden layer output matrix \mathbf{H} ;
- (3) calculate the output weight matrix $\hat{\boldsymbol{\beta}}$.

3.2. OP-ELM. The optimally pruned extreme learning machine (OP-ELM) is a variation of ELM algorithm for SLFN. The OP-ELM algorithm is made of three main steps summarized as follows [27, 28].

Given a training set $\{(x_i, t_i) \mid x_i \in \mathbf{R}^n, t_i \in \mathbf{R}^m, i = 1, \dots, N\}$.

(1) Build a regular ELM model with initially large number of neurons.

(2) Rank the hidden layer neurons by their contribution to the linear explanation of the ELM output by the multiresponse sparse regression (MRSR), which was proposed by Similä and Tikka in 2005 [29].

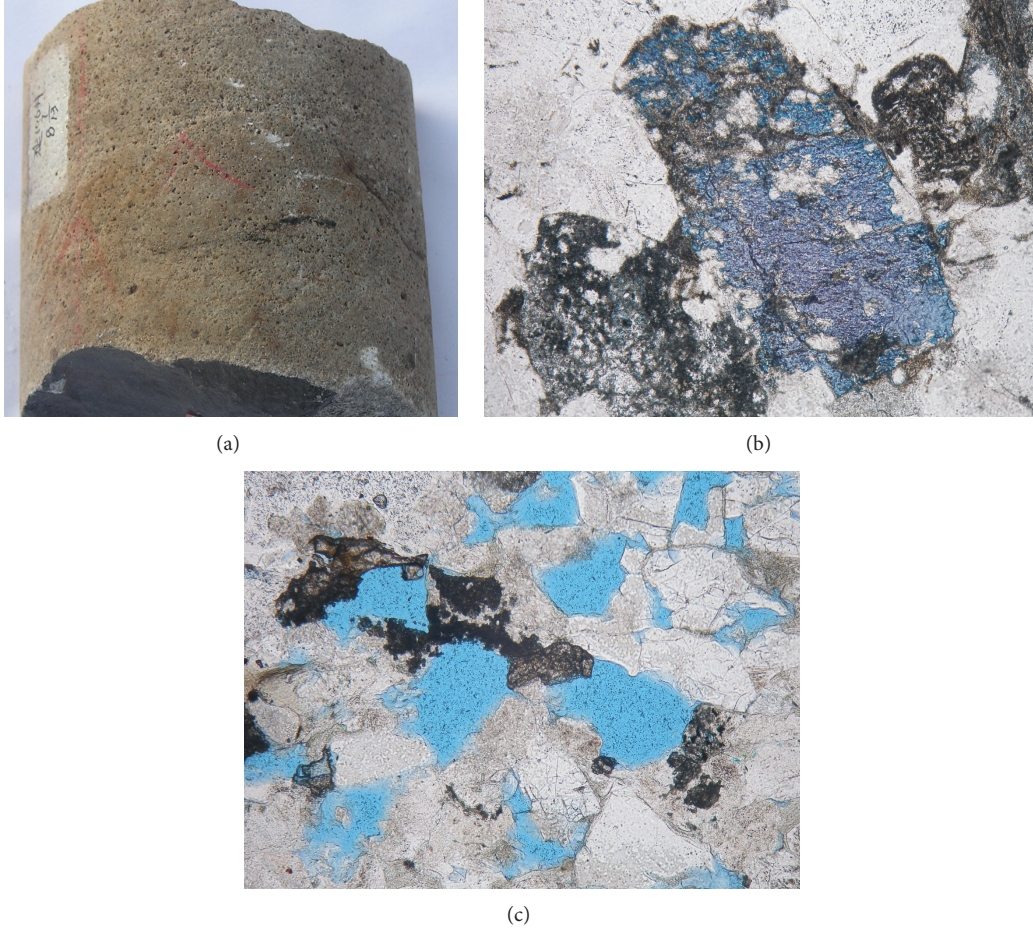


FIGURE 2: Photo and microscope slice of core samples from Permian formation of well Yq3 in Yanqi survey: (a) actual core photo, (b) intragranular pore, and (c) intergranular pore. Small pores can be recognized on the side of core. In the microscope slices, pore spaces are colored and the grey part is the matrix.

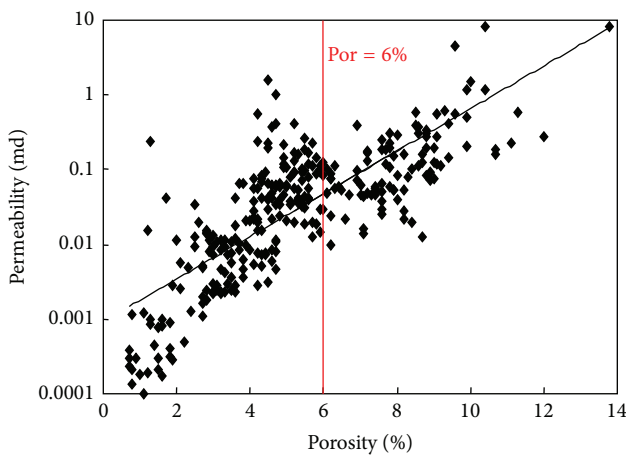


FIGURE 3: Cross-plot of core-measured porosity and permeability.

MRSR is used to get rid of the useless neurons of the hidden layer. Its main idea is as follows: add columns of the regressor matrix \mathbf{H} in $\mathbf{H}\hat{\beta} = \mathbf{T}$ and corresponding nonzero

rows in $\hat{\beta}$ and thus obtain a series of approximations $\mathbf{H}^k \hat{\beta} = \mathbf{T}^k$. Then hidden nodes are ordered by the corresponding decrease in the prediction error $\|\mathbf{T}^k - \mathbf{T}\|$ obtained in the model. More specific details of the MRSR algorithm can be found in [29].

(3) Decide the optimal number of neurons by the leave-one-out (LOO) validation method. Compute the LOO using the PRESS (prediction sum of squares) statistic in the linear case:

$$\epsilon^{\text{PRESS}} = \frac{t_i - \mathbf{h}_i \mathbf{b}_i}{1 - \mathbf{h}_i \mathbf{P} \mathbf{h}_i^T}, \quad (7)$$

where \mathbf{h}_i and \mathbf{b}_i are the i th column and i th row of \mathbf{H} and $\hat{\beta}$, respectively. The process is greedily incremental, and units are added in order until the LOO method decreases below a preset threshold.

For robustness and more generality, the OP-ELM algorithm has the suggestion of using a combination of three types of kernels: linear, sigmoid, and Gaussian kernel, while the original ELM proposed to use only sigmoid kernels [21]. Problems discussed in this paper are not linear, and

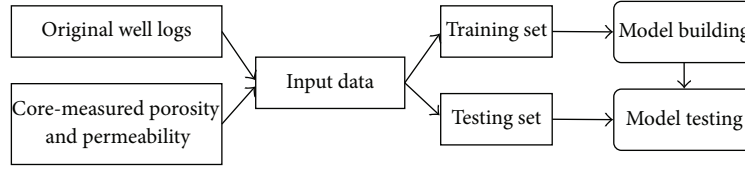


FIGURE 4: Experimental flow diagram of the estimation experiments.

the experiments conducted below will compare the efficiency using different kernel types and choose the optimal one for both ELM and OP-ELM.

4. Experimental Design

The experimental analysis is conducted to find the best model for predicting porosity and permeability from log data. Figure 4 shows the flow diagram of the computational experiments carried out. Rectangle boxes correspond to data, while rounded boxes correspond to process. For the experiment, the first step is to prepare the input data, including the original log data and core-measured porosity and permeability data. The datasets are divided into training set and testing set. The second step is to set up the network model with the training dataset. Model parameters, including kernel types and neuron numbers of hidden layers, are determined. The third step is to validate the model using the testing dataset. If the error is below the threshold, the model is feasibly and appropriately built; thereafter it can be used in reservoir parameters prediction for the remaining wells.

4.1. Data Preparation. In the survey, all of the wells in the survey have run the conventional logging, and the five types of logs for reservoir parameters estimation are well prepared. About 8 wells have cores from Permian sandstone formation. Most of the cores have laboratory analysis results, including the measured porosity and permeability values. Unlike the continuously recorded well logs data, the core data are sparsely sampled (as shown in Figure 5). So when collecting input data for the network, each core-measured data matches with the log data at the same depth.

Five types of logs including AC, DEN, GR, CNL, and RT are used as input for the networks, and porosity (POR) and permeability (PERM) measured from cores are the two output targets. Two parts are prepared using the datasets, including training part and test part. The training part is used to train the model, while the test part is used to compute predictions and compare them with the measured values. Mean square error (MSE) is computed as evaluation of the prediction accuracy and quality of the trained model.

The total number of samples in Permian sandstone of 8 wells of the survey is summed up to 4950, 90% of which will be used as the training data for the ELM network, while the remaining 10% will be used as test samples.

Data normalization is a necessary preprocessing step for network data analysis. So all of the logs and core-measured

parameters are normalized before formally inputting into the network. The normalized variable has the following form:

$$X_{\text{new}} = \frac{X_{\text{old}} - \min X}{\max X - \min X}, \quad (8)$$

where X stands for logs of GR, AC, DEN, CNL, and RT. The new normalized variable X_{new} takes the range from 0 to 1 for all the parameters.

4.2. Network Architecture. For porosity and permeability estimation, five logs including GR, AC, DEN, CNL, and RT are physically related to petrophysical properties. So these five logs are fed to the ELM network as input with each node denoting one log. Porosity and permeability are to be taken as the two network neurons at the output layer. The network architecture is shown as Figure 6.

4.3. Network Parameter Selection. For ELM network, appropriate kernel and number of hidden nodes are the two critical parameters to be determined.

In the study, four types of kernels have been tested using training dataset, and they are sigmoid function, radial basis function, hardlim function, and triangular function. At the same time, numbers of hidden nodes are tested accordingly. When selecting one of the kernels, number of hidden nodes will start from 5, with 10 as incremental step. Figure 7 is the accuracy comparison by using different kernels and node numbers. The MSE between prediction results and measured properties decreases rapidly as node number of the hidden layer gradually increases. Among the four kernels, sigmoid-based model comes to the threshold first when node number is set as 55, while overfitting problem appears as the node number is bigger than 65. Triangular-based and radial-basis-based models have the same trend, and it seems that the node number might exceed 100 when the minimum errors are close to the threshold. For the hardlim-based model, the MSE reaches the lowest point at 3.26% when node number is set as 75.

Based on the experimental tests, sigmoid kernel is optimal and when node number of the hidden layer is set as 55, the network model can obtain the best prediction accuracy.

For the ELM network model, node number of hidden layer can also be optimally determined by OP-ELM training process instead of time-consuming and arbitrary testing. The very first step of OP-ELM methodology is to construct the SLFN network using the original ELM structure with 100 neurons at the hidden layer and sigmoid kernel. Both of training dataset and testing dataset are input to the model,

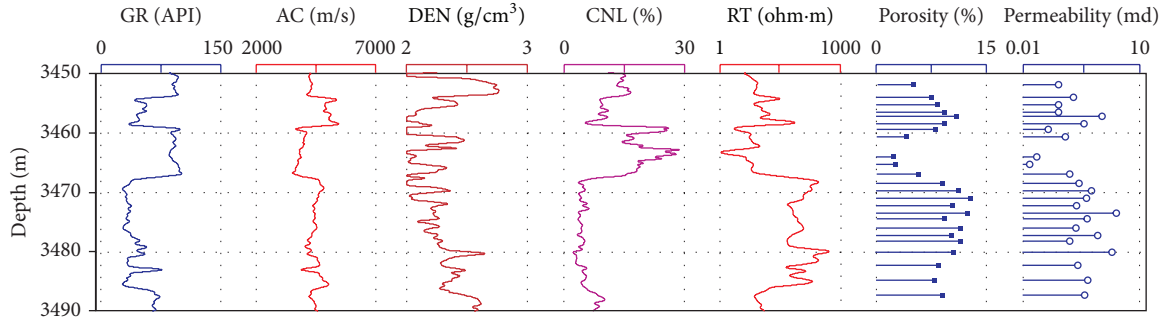


FIGURE 5: Part of logs and corresponding core-measured porosity and permeability in Permian formation of well Yq12. The measured parameters are sparsely distributed along the borehole, while logs are continuously sampled, and the sampling interval of logs is 0.5 m.

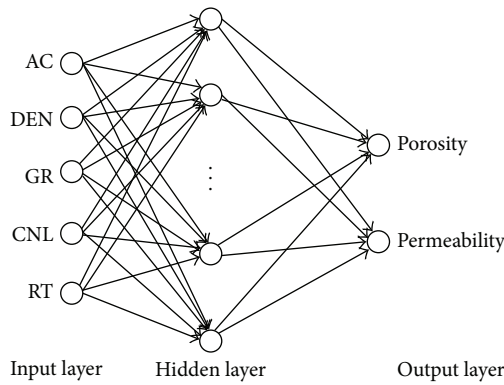


FIGURE 6: Schematic of architecture of the ELM network model. Node number of the input layer is 5, and porosity and permeability are the two nodes at the output layer. Number of hidden layer nodes is to be settled by OP-ELM method.

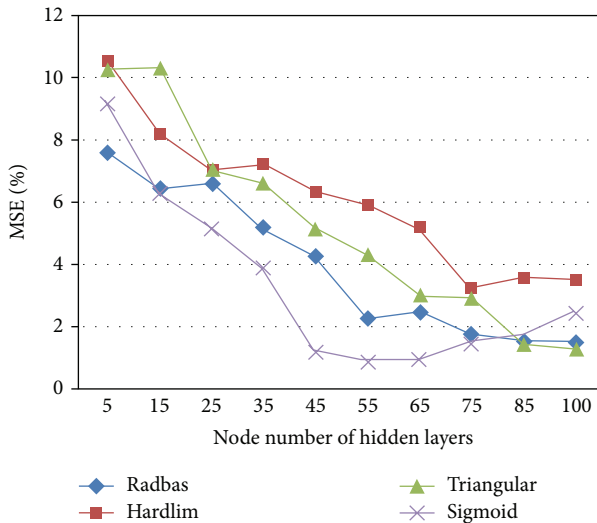


FIGURE 7: Comparison of mean squared error (MSE) obtained from different kernel-based ELM models with increasing node numbers at the hidden layers.

and MSEs are computed for looking for the best accuracy. Throughneuron contribution sorted by MRSR and optimal

nodes selected by LOO validation, the optimal neuron at the hidden layer is finally determined as 62, which is close to the original ELM test.

4.4. Accuracy Analysis. When the parameters for the network model are finally settled, the following step is to validate the model using testing dataset. Core porosity and permeability are set as targets to the network predictor. Figure 8 is the validation result for the well Yq12 with depth ranging from 3450 m to 3490 m (mentioned in Figure 5). The model outputs are superimposed on the core data. Regression plots in Figures 9(a) and 9(b) reflect the accuracy of the OP-ELM network estimator. Coefficients of 0.9932 and 0.9917 are obtained for porosity and permeability estimation, respectively. The accuracy is satisfactory. Comparison analysis of the results from all of the testing datasets demonstrates good performance using the network predictor for the two geophysical parameters, especially for permeability, which is more sensitive to the rocks heterogeneity [30].

Furthermore, in order to testify the advantages of OP-ELM, BP network and support vector machine (SVM) algorithm are used in the model training and testing process for comparison with OP-ELM. Backpropagation feedforward network (BP) is the most commonly used ANN approach, and it is also criticized for having difficulty to decide learning rates, being easy to be stuck on local minimums, having overfit problems, and being time consuming [6]. SVM is a competitive technique which has been intensively used for nonlinear modeling. It has two advantages over traditional deterministic methods: strong nonlinear approximation capabilities and good generalization effectiveness. Experiments have been performed to porosity and permeability prediction in [15–17], which shows better prediction performance than multilayer perceptron (MLP). References [20, 31] have compared performances of BP, ELM, and SVM models, and the experimental results show that ELMs outperform SVMs on reliabilities, while SVMs are better on output distributions. So, it is interesting to compare them in this research.

In the BP network model, the typical structure is used, with three layers including one input layer, one hidden layer, and one output layer. At input layer, 5 neurons stand for the five input logs, and 2 neurons for the output layer. Sigmoid

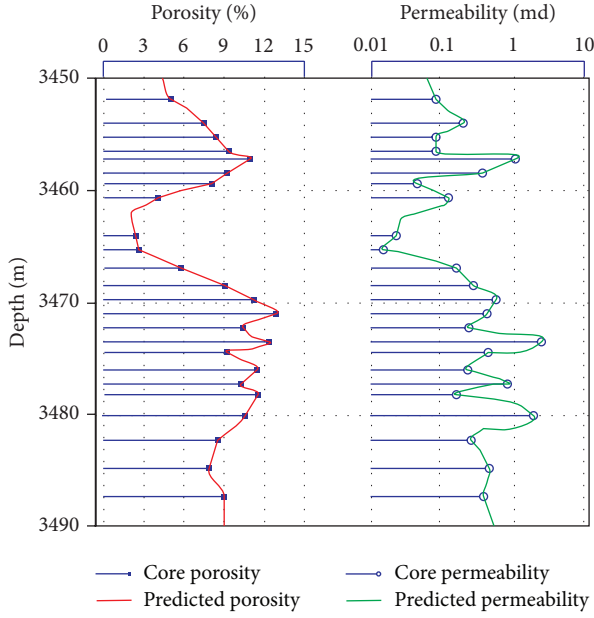


FIGURE 8: Prediction porosity and permeability from OP-ELM network and core values with depth.

kernel function is also adopted for the network model. 33 neurons at the hidden layer are finally determined after dozens of testing in the training process.

As for the SVM predictors, when using the same dataset for model training and testing, minor difference could be created for the SVM models with different kernel function, especially for such numerical approximation problems [31]. But accuracy always needs to be guaranteed, so three types of kernel function have been tested and compared using the dataset of the survey. The three common kernel functions include (1) polynomial function, (2) radial basis function (RBF), and (3) sigmoid function. Here, grid and pattern search methods are used to determine the optimal set of SVM input parameters. Table 1 shows the final results. Gaussian RBF function seems better than the others in performing the estimation, so the final SVM model uses RBF as the kernel function.

Table 2 shows the comparison result using the three types of network predictors with the same testing dataset. Accuracy, MAE, and training time are three factors in comparisons, and the values are obtained by averaging estimations of the samples in well Yq8. The table shows the accuracy, mean absolute error (MAE), and total time in seconds for the three processing approaches, respectively. Good performances have been done using the three optimized network estimators, and best results are achieved by OP-ELM with an accuracy of 95.6%, mean absolute error of 0.205, and fast learning speed of 23 seconds.

Figure 10 shows the prediction porosity and permeability in Permian sandstone of Yq8 by using the three algorithms: BP, SVM, and OP-ELM. All three approaches have conducted good prediction performance for the sandstone reservoir parameters estimation, and the model predictions are very

TABLE 1: Comparison of capabilities of SVM with different kernel functions.

Kernel function	Kernel parameters	Training accuracy (%)	Validation accuracy (%)
Gaussian RBF	$c = 0.125$, $r = 0.125$	100	97.93
Sigmoid	$c = 1$, $r = 0.0625$	100	97.86
Polynomial	$c = 0.125$, $r = 0.125$, $d = 1$	100	96.85

TABLE 2: Comparison of porosity prediction performance results on OP-ELM against BP and SVM methodology for well Yq8. The comparison strata belong to the Permian reservoir.

Algorithm	Accuracy	MAE	Training time (s)
BP	0.901	0.827	72
SVM	0.935	0.639	51
OP-ELM	0.956	0.205	23

close to the core-measured parameters, which shows the advantages in generalization and prediction of the neural network method. But for the three network predictors, OP-ELM has better accuracy than the other two methods. In the interval of 3375–3390 m, the lithology is dominated by porous and permeable sandstone, with average porosity of about 10% and permeability of more than 1 md.

5. Reservoir Parameters Prediction

The above analysis has shown the reliability and accuracy of the OP-ELM prediction model. Therefore the model is then used to estimate porosity and permeability for the remaining wells in the survey. Log data of 7 wells without core data have been input into the model, and porosity and permeability of the Permian reservoir have been estimated. Figure 11 is the plot of original logs and predicted reservoir parameters of well Yq4 from 3560 to 3600 m. This interval belongs to the Permian formation, and oil show exists at the depth of 3591–3598 m. According to the predicted result, porosity of the oil-bearing interval is about 10.78% and permeability is about 3 md, which means the sandstone reservoir from 3590 m to 3598 m has good petrophysical properties and can be recommended for further evaluation.

Since all wells have been processed using the ELM-based model, statistical reservoir analysis is then to be carried out in the survey. According to the geological correlation analysis, tops of target sandstone reservoir in the wells are determined. Thereafter under the constraints of tops boundaries of the reservoir, average porosity and permeability of the target reservoir interval have been calculated for all of 15 wells in the survey. Then flex gridding algorithm is used to interpolate the attributes between wells. Figure 12 is the average porosity map for target reservoir. In the map, red-yellow color stands for high porosity with values bigger than 9.5%. It is obvious

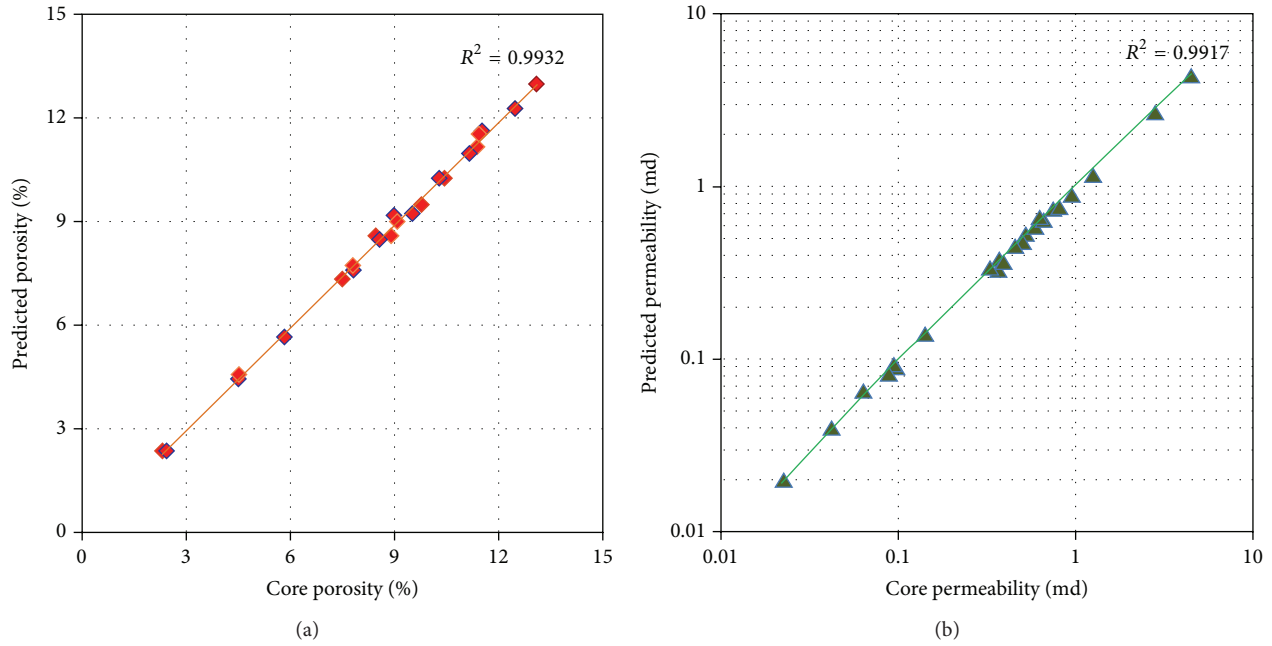


FIGURE 9: Regression analysis of the predicted and core-measured parameters: (a) porosity plot, with correlation coefficient about 0.9932, and (b) permeability plot, with correlation coefficient about 0.9917.

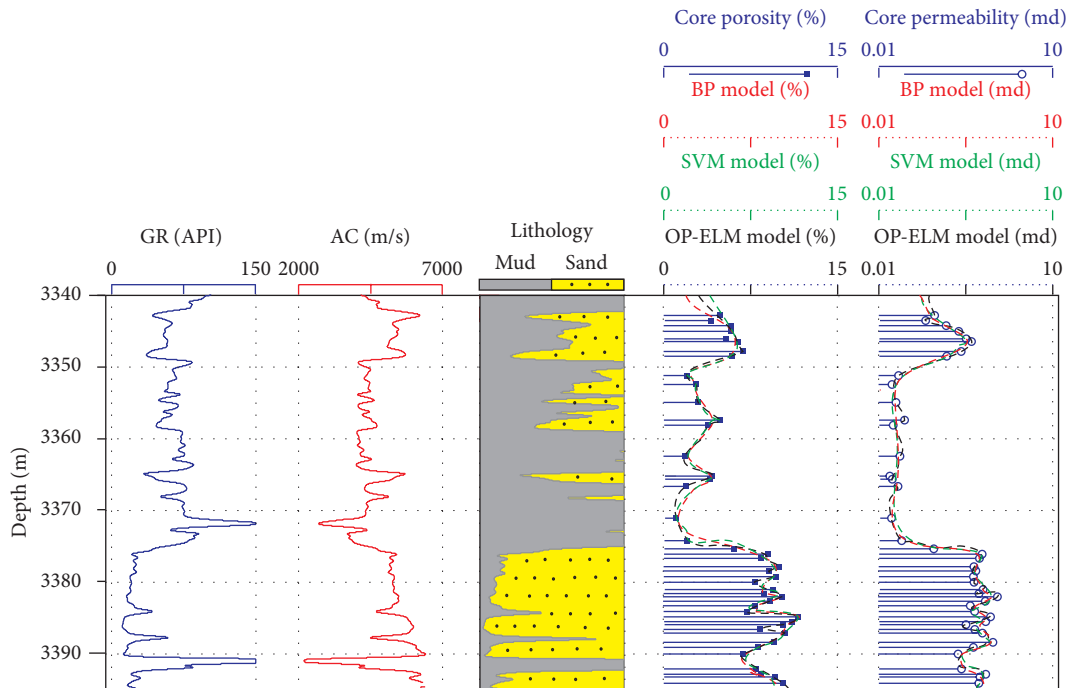


FIGURE 10: Porosity and permeability prediction comparison by three neural network models, BP, SVM, and OP-ELM, for Permian sandstone in Yq8.

that approximate semicircular zones of wells Yq12 and Yq13 at the northeast part of the survey have porosity higher than 10% and also for the triangle zone including wells Yq15, Yq8, Yq4, Yq3, and Yq7 at the western part in the survey. These two areas have better porous and permeable sandstone reservoir, and the six wells included in the two zones have

encountered oil in Permian sandstone reservoir. Since there is no direct structural trap in the survey, stratigraphic trap is the dominant type of traps. Fluvial reservoir is the critical factor for well planning and economic assessment. Therefore the above-mentioned two favorable areas are proposed as potential for next-round well selections in the Yanqi survey.

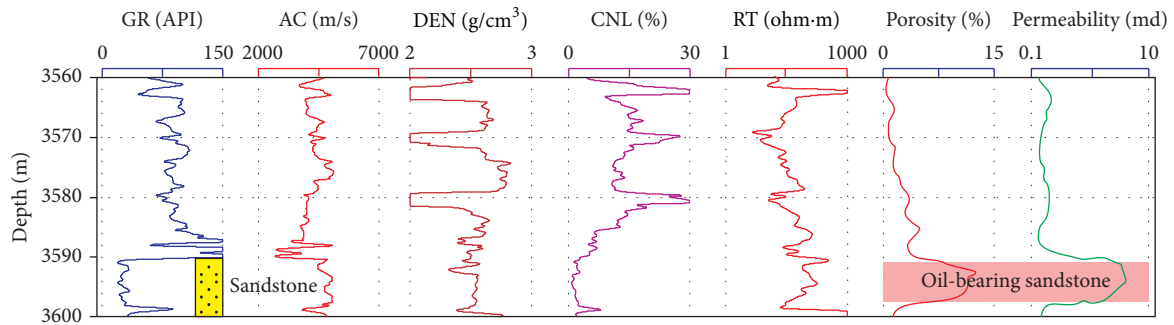


FIGURE 11: Reservoir parameters prediction for well Yq4 in Yanqi survey. Curves of porosity and permeability with blue-dashed line are predicted from the network model. Sandstone reservoir from 3591 m to 3598 m is oil-bearing and has median porosity but high permeability on the basis of the prediction results.

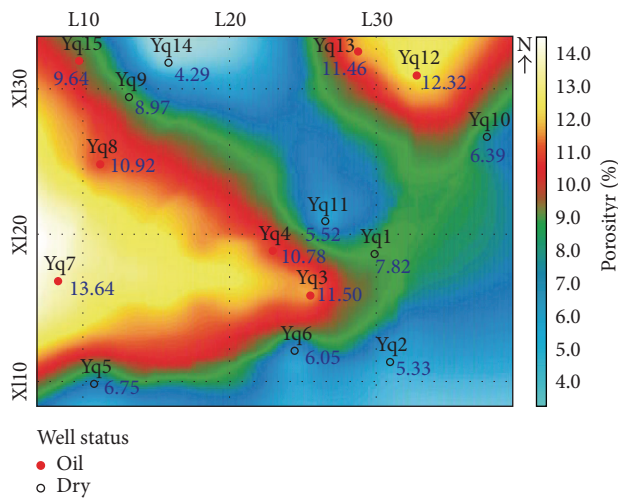


FIGURE 12: Average porosity map for the target reservoir of Yanqi survey. Values near well symbols are the average porosity of Permian sandstone based on the ELM network output.

6. Conclusions

The prediction of porosity and permeability is an essential but complex research problem in reservoir characterization of petroleum exploration. In this paper, the authors adopt the ELM-based predictors for solving such crucial problem in heterogeneous sandstone reservoir of Permian formation in Yanqi survey. The proposed ELM and OP-ELM approaches have been reviewed and applied to build estimation models to predict reservoir parameters. Logs data and core-measured porosity and permeability are input into the network model. Sigmoid kernel is used in the network, and the node number of hidden layer is determined by OP-ELM. The advantages of estimation accuracy and learning speed have been testified for ELM methodology in the research. Reliable network prediction models have been established in the study, and porosity and permeability are estimated for Permian sandstone reservoir of all wells in the survey. Potential areas are suggested finally, including two favorable zones: the northeast approximate semicircular zone of wells Yq12 and Yq13 and

western triangle zone including wells Yq15, Yq8, Yq4, Yq3, and Yq7.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is supported by National Natural Science Foundation of China (no. 61402331). The Foundation of Educational Commission of Tianjin City, China (Grant no. 20140803), also funded this research. The authors are also grateful to the three reviewers of the original paper. Jianhua Cao is the first author who wrote the whole paper including text, tables, and figures. Questions can also be sent to the author at the following email: caojh@tust.edu.cn.

References

- [1] A. K. El Ouahed, D. Tiab, A. Mazouzi, and A. J. Safraz, "Application of artificial intelligence to characterize naturally fractured reservoirs," Paper SPE 84870, 2013.
- [2] G. Wang and T. R. Carr, "Methodology of organic-rich shale lithofacies identification and prediction: a case study from Marcellus Shale in the Appalachian basin," *Computers & Geosciences*, vol. 49, pp. 151–163, 2012.
- [3] L. Qi and T. R. Carr, "Neural network prediction of carbonate lithofacies from well logs, Big Bow and Sand Arroyo Creek fields, Southwest Kansas," *Computers & Geosciences*, vol. 32, no. 7, pp. 947–964, 2006.
- [4] S. Chikhi and M. Batouche, "Probabilistic neural method combined with radial-bias functions applied to reservoir characterization in the Algerian Triassic province," *Journal of Geophysics and Engineering*, vol. 1, no. 2, pp. 134–142, 2004.
- [5] A. K. El Ouahed, D. Tiab, and A. Mazouzi, "Application of artificial intelligence to characterize naturally fractured zones in Hassi Messaoud Oil Field, Algeria," *Journal of Petroleum Science and Engineering*, vol. 49, no. 3–4, pp. 122–141, 2005.
- [6] R. B. C. Gharbi and G. A. Mansoori, "An introduction to artificial intelligence applications in petroleum exploration and

- production," *Journal of Petroleum Science and Engineering*, vol. 49, no. 3-4, pp. 93-96, 2005.
- [7] A. K. Verma, B. A. Cheadle, A. Routray, W. K. Mohanty, and L. Mansinha, "Porosity and Permeability estimation using neural network approach from well log data," CSPG/CSEG/CWLS GeoConvention 2012, Search and Discovery Article #41276, 2012.
 - [8] M. Ali and A. Chawathé, "Using artificial intelligence to predict permeability from petrographic data," *Computers & Geosciences*, vol. 26, no. 8, pp. 915-925, 2000.
 - [9] M. Baneshi, M. Behzadijo, M. Schaffie, and H. Nezamabadi-Pour, "Predicting log data by using artificial neural networks to approximate petrophysical parameters of formation," *Petroleum Science and Technology*, vol. 31, no. 12, pp. 1238-1248, 2013.
 - [10] H. Khoshdel and M. A. Riahi, "Multi attribute transform and neural network in porosity estimation of an offshore oil field—a case study," *Journal of Petroleum Science and Engineering*, vol. 78, no. 3-4, pp. 740-747, 2011.
 - [11] M. A. Ahmadi, M. Reza, S. M. Hosseini, and M. Ebadi, "Connectionist model predicts the porosity and permeability of petroleum reservoirs by means of petro-physical logs: application of artificial intelligence," *Journal of Petroleum Science and Engineering*, vol. 123, pp. 643-656, 2014.
 - [12] M. Saemi, M. Ahmadi, and A. Y. Varjani, "Design of neural networks using genetic algorithm for the permeability estimation of the reservoir," *Journal of Petroleum Science and Engineering*, vol. 59, no. 1-2, pp. 97-105, 2007.
 - [13] G. L. Jing, W. Du, and Y. Y. Guo, "Studies on prediction of separation percent in electrodialysis process via BP neural networks and improved BP algorithms," *Desalination*, vol. 291, pp. 78-93, 2012.
 - [14] C. Nello and S. T. John, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, Publishing House of Electronics Industry, 2004.
 - [15] A. F. Al-Anazi and I. D. Gates, "Support vector regression to predict porosity and permeability: effect of sample size," *Computers & Geosciences*, vol. 39, pp. 64-76, 2012.
 - [16] A. F. Al-Anazi and I. D. Gates, "Support vector regression for porosity prediction in a heterogeneous reservoir: a comparative study," *Computers & Geosciences*, vol. 36, no. 12, pp. 1494-1503, 2010.
 - [17] C. Cranganu and M. Breaban, "Using support vector regression to estimate sonic log distributions: a case study from the Anadarko Basin, Oklahoma," *Journal of Petroleum Science and Engineering*, vol. 103, pp. 1-13, 2013.
 - [18] R. Gholami, A. Moradzadeh, M. Shahoo, A. Saman, and J. Hanachi, "Applications of artificial intelligence methods in prediction of permeability in hydrocarbon reservoirs," *Journal of Petroleum Science and Engineering*, vol. 122, pp. 643-656, 2014.
 - [19] U. Iturrarán-Viveros and J. O. Parra, "Artificial Neural Networks applied to estimate permeability, porosity and intrinsic attenuation using seismic attributes and well-log data," *Journal of Applied Geophysics*, vol. 107, pp. 45-54, 2014.
 - [20] H. Zhong, C. Miao, Z. Shen, and Y. Feng, "Comparing the learning effectiveness of BP, ELM, I-ELM, and SVM for corporate credit ratings," *Neurocomputing*, vol. 128, pp. 285-295, 2014.
 - [21] G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: a new learning scheme of feedforward neural networks," in *Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN '04)*, vol. 2, pp. 985-990, Budapest, Hungary, July 2004.
 - [22] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1-3, pp. 489-501, 2006.
 - [23] G.-B. Huang, D. H. Wang, and Y. Lan, "Extreme learning machines: a survey," *International Journal of Machine Learning and Cybernetics*, vol. 2, no. 2, pp. 107-122, 2011.
 - [24] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics*, vol. 42, no. 2, pp. 513-529, 2012.
 - [25] R. Minhas, A. Baradarani, S. Seifzadeh, and Q. M. J. Wu, "Human action recognition using extreme learning machine based on visual vocabularies," *Neurocomputing*, vol. 73, no. 10-12, pp. 1906-1917, 2010.
 - [26] R. Moreno, F. Corona, A. Lendasse, M. Graña, and L. S. Galvão, "Extreme learning machines for soybean classification in remote sensing hyperspectral images," *Neurocomputing*, vol. 128, pp. 207-216, 2014.
 - [27] A. Mozaffari and N. L. Azad, "Optimally pruned extreme learning machine with ensemble of regularization techniques and negative correlation penalty applied to automotive engine cold-start hydrocarbon emission identification," *Neurocomputing*, vol. 131, pp. 143-156, 2014.
 - [28] Y. Miche, A. Sorjamaa, P. Bas, O. Simula, C. Jutten, and A. Lendasse, "OP-ELM: optimally pruned extreme learning machine," *IEEE Transactions on Neural Networks*, vol. 21, no. 1, pp. 158-162, 2010.
 - [29] T. Similä and J. Tikka, "Multiresponse sparse regression with application to multidimensional scaling," in *Artificial Neural Networks: Formal Models and Their Applications—ICANN 2005*, vol. 3697, pp. 97-102, 2005.
 - [30] P. Bagheripour, "Committee neural network model for rock permeability prediction," *Journal of Applied Geophysics*, vol. 104, pp. 142-148, 2014.
 - [31] J. Chorowski, J. Wang, and J. M. Zurada, "Review and performance comparison of SVM- and ELM-based classifiers," *Neurocomputing*, vol. 128, pp. 507-516, 2014.

