Research Article

Acoustic Log Prediction on the Basis of Kernel Extreme Learning Machine for Wells in GJH Survey, Erdos Basin

Jianhua Cao, Yancui Shi, Dan Wang, and Xiankun Zhang

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

Correspondence should be addressed to Jianhua Cao; caojh@tust.edu.cn

In petroleum exploration, the acoustic log (DT) is popularly used as an estimator to calculate formation porosity, to carry out petrophysical studies, or to participate in geological analysis and research (e.g., to map abnormal pore-fluid pressure). But sometime it does not exist in those old wells drilled 20 years ago, either because of data loss or because of just being not recorded at that time. Thus synthesizing the DT log becomes the necessary task for the researchers. In this paper we propose using kernel extreme learning machine (KELM) to predict missing sonic (DT) logs when only common logs (e.g., natural gamma ray: GR, deep resistivity: REID, and bulk density: DEN) are available. The common logs are set as predictors and the DT log is the target. By using KELM, a prediction model is firstly created based on the experimental data and then confirmed and validated by blind-testing the results in wells containing both the predictors and the target (DT) values used in the supervised training. Finally the optimal model is set up as a predictor. A case study for wells in GJH survey from the Erdos Basin, about velocity inversion using the KELM-estimated DT values, is presented. The results are promising and encouraging.

1. Introduction

Oil and gas exploration in sedimentary basins is very complicated, since all the targets are buried underground and they cannot be viewed or touched directly. So all the properties for the buried targets have to be predicted or estimated by using modern electrical or magnetic tools. The physical properties of the geologic formations include pore-fluid pressure, rock lithology, porosity, permeability, and oil or water saturation. Nowadays the conventional tool for characterizing these geophysical properties is well logging, and some logs such as gamma ray (GR), dual induction log, formation density (DEN) compensated, deep resistivity (REID), self-potential (SP), and sonic log (DT) are usually recorded. Among them, the sonic log (DT) has largely been used to predict rock porosity, to perform petrophysical analysis, or to carry out well-to-seismic inversion.

Owing to historical operation mistakes or recording loss, the sonic log may not be available in well logging suites. The traditional way solving this problem is to transform the DEN or REID log to DT log based on some experimental formula built between these logs. It might be feasible for some area, but sometimes the errors are unacceptable.

Artificial intelligence techniques have the advantage in connecting unrelated parameters and solving nonlinear problems. Such techniques, including BP neural network, fuzzy reasoning, or evolutionary computing for data analysis and interpretation have become effective tools in the workflow for well drilling and reservoir characterization [1–10]. However, traditional neural networks have many known drawbacks in the learning process, such as multiple local minima, slow learning speed, and poor generalization performances [11].

Extreme learning machine (ELM) is a single-hidden layer feed-forward neural network (SLFN) proposed by Huang et al. [12, 13]. 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 can

learn thousands of times faster than conventional popular learning algorithms for feed-forward neural networks [14]. A lot of real-life applications [15–18] have already demonstrated advantages of using basic ELM. A kernel-based ELM (KELM) has also been developed lately [19], where the hidden layer feature mapping is determined by the kernel matrix. In this version, only the kernel function and its parameters are needed to be defined; the number of hidden nodes is not required. With the use of kernel function, KELM is expected to achieve better generalization performance than basic ELM. Furthermore, as randomness does not occur in KELM, the chance of result variations could be reduced [20].

In this paper, kernel-based extreme learning machine is used to predict missing sonic (DT) logs when only common logs (e.g., natural gamma ray—GR, bulk density—DEN, or deep resistivity—REID) are available. By using KELM, we first create and train a supervised network model based on experimental data and then confirm and validate the model by blind-testing the results. The optimal model is at last applied to wells containing the predictor data but with lack of DT log. We use this workflow in GJH survey from Erdos Basin and the KELM-estimated DT logs are then integrated in the seismic inversion to identify the sandstone reservoir.

The rest of this paper proceeds as follows. Section 2 gives a short review of ELM and KELM. Section 3 describes the experiments using KELM, including the data preparation, parameter selection, and model validation. Section 4 gives the prediction application in GJH survey. Finally, Section 5 gives the conclusion of this work.

2. Methodology

In this study, the kernel extreme learning machine (KELM) is employed to predict the DT logs for the wells in GJH survey. So we present an overview of the ELM and kernel-based ELM as follows.

2.1. ELM. The classical ELM was proposed for SLFNs by Huang et al. [12, 13]. Different from BP network, the input weights and biases of ELM are randomly assigned and need not be fine-tuned within the training phase, and the output weights can be determined analytically by finding the least-square solution. The prediction of ELM is given by

$$f_L(\mathbf{x}) = \sum_{i=1}^{L} \boldsymbol{\beta}_i \mathbf{h}_i(\mathbf{x}) = \mathbf{h}(\mathbf{x})\boldsymbol{\beta}, \quad (1)$$

where $\boldsymbol{\beta} = [\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_L]^T$ is the weight vector connecting the hidden node and the output nodes and $\mathbf{h}(\mathbf{x}) = [\mathbf{h}_1(\mathbf{x}), \ldots, \mathbf{h}_L(\mathbf{x})]^T$ is the output of the hidden layer with respect to the sample $\mathbf{x}$. Since the weights and biases are initially assigned for the hidden layer, when the activation function is set, $\mathbf{h}(\mathbf{x})$ is determined and need not be tuned. And the only unknown parameter is $\boldsymbol{\beta}$, which can be solved as constrained optimization problem:

$$\text{Minimize:} \quad \|\mathbf{H}\boldsymbol{\beta} - T\|_p^{\alpha 1} + \frac{C}{2}\|\boldsymbol{\beta}\|_q^{\alpha 2}, \quad (2)$$

where C is control parameter for a tradeoff between structural risk and empirical risk, T is the target output for the network.

And when $p, q = F$ and $\alpha 1, \alpha 2 = 2$, a popular and efficient closed-form solution for $\boldsymbol{\beta}$ is

$$\boldsymbol{\beta} = \begin{cases} \mathbf{H}^T\left(C\mathbf{I} + \mathbf{H}\mathbf{H}^T\right)^{-1}T & N \geq L \\ \left(C\mathbf{I} + \mathbf{H}\mathbf{H}^T\right)^{-1}\mathbf{H}^T T & N \leq L \end{cases} \quad (3)$$

2.2. KELM. As proposed in Huang et al. [19], if $\mathbf{h}(\cdot)$ is unknown, that is, an implicit function, one can apply Mercer's conditions on ELM and define a kernel matrix for ELM that takes the form

$$\mathbf{K}_{\text{ELM}} = \mathbf{H}\mathbf{H}^T :$$
$$\mathbf{K}_{\text{ELM}_{i,j}} = \mathbf{h}(\mathbf{x}_i) \cdot \mathbf{h}(\mathbf{x}_j) = \mathbf{k}(\mathbf{x}_i, \mathbf{x}_j), \quad (4)$$

where $\mathbf{k}(\mathbf{x}_i, \mathbf{x}_j)$ is a kernel function. Many kernel functions can be used in kernel-based ELM, such as linear, polynomial, and radial basis function, so that we can obtain the kernel form of the output function as follows:

$$f_L(\mathbf{x}) = \begin{bmatrix} \mathbf{k}(\mathbf{x}, \mathbf{x}_1) \\ \vdots \\ \mathbf{k}(\mathbf{x}, \mathbf{x}_N) \end{bmatrix} (C\mathbf{I} + \mathbf{K}_{\text{ELM}})^{-1} T \quad (5)$$

Similar to the SVM, $\mathbf{h}(\mathbf{x})$ need not be known; instead, its kernel can be provided (e.g., Gaussian kernel $\mathbf{k}(\mathbf{u}, \mathbf{v}) = \exp(\|\mathbf{u} - \mathbf{v}\|^2/\sigma)$). The optimal penalty parameter C and kernel width s are determined by try and error way. Node number of the hidden layer L need not be available beforehand either. The experimental and theoretical analysis of Huang et al. showed that KELM produces improved generalization performance over the SVM/LS-SVM [21].

For the given type of the kernel function, the training dataset, and the initial parameters of the network, the following steps are considered.

Step 1. Initiate the population based on the kernel function.

Step 2. Evaluate the fitness function of each parameter.

Step 3. The optimal parameters of kernel function can be determined. Then, based on the optimized parameters, the hidden layer kernel matrix is computed.

Step 4. Determine the final output weights.

3. Experimental Study

3.1. Problem Description and Related Work. Well logging is the practice of making a detailed record of the geologic formations penetrated by a borehole. Normally the log is based on the physical measurements made by instruments lowered into the borehole. According to the geophysical properties of the rocks, the logs are always classified as follows: electrical logs, porosity logs, lithology logs, and miscellaneous logs. Sonic log (DT) belongs to the porosity logs, and it provides a formation interval transit time, which

typically varies lithology and rock texture, especially porosity for the rocks. Gamma ray log is a log of the natural radioactivity of the formation along the borehole, measured in API units, particularly useful for distinguishing between sands and shales in a siliciclastic environment. This is because sandstones are usually nonradioactive quartz, whereas shales are naturally radioactive due to potassium isotopes in clays and adsorbed uranium and thorium.

The main datasets used in this study include acoustic log (DT), the gamma ray (GR), the resistivity log (REID), which represents the variation of the electric resistivity, the density (DEN), which records the density variation with depth in the borehole, and the self-Potential (SP), a measurement of natural electric potential. These geophysical parameters DT, GR, REID, DEN, and SP are intrinsically linked, since each of them reflects some physical property of the same rock layer. Take sandstone as an example. Pores are sure to exist at the sandstone interval, and if the pores are not filled with other types of tight materials, fluid is the only also important stuffing. There might be oil or gas and water as well. Since the fluid has different physical parameters than the surrounding sandstone, obvious differences will be recorded on the measuring logs: lower GR, lower DT, higher REID, lower DEN, and abnormal change on SP. Thus just observing the characters of the logs, especially those abnormal changes, the experienced researchers have confidence to tell the geological information along the borehole. And then some researchers try to build theoretical relationships between the logs. Thousands of experiments result in empirical equations. For example, DEN could be transformed using DT log when DEN is missing and the relation is defined as Gardener formula [6]:

$$DEN = \alpha DT^{-\beta}, \tag{6}$$

where, α, β are the coefficients and their values are up to the core tests for the studied area.

In this study, the key we focus on is the DT log, and we want to find the optimal way to get the DT log when it is missing.

The sonic log (DT) is very important in petroleum exploration phase. One way for using DT is to estimate rock porosity, which is the critical parameter for the reservoir evaluation, and identify the fluid information along the borehole. Additionally, since DT log has both time and velocity information, it becomes the reliable key for the time-depth conversion when using seismic data to interpret structures and geological mapping. In one word, the DT log is indispensable for the geophysical and geological study.

But there has always been imperfection, and sometimes, owing to operation mistake or recording loss, DT log may not be available in some wells. One solution for obtaining the DT log is to carry out empirical transformation from other logs, and the model is built by experiment analysis. The formula is just for specific field condition, and it can not be used for all the formation conditions. For instance, Faust formula is just for DT calculation using REID log, and cases [7] have shown that the formula is not suitable when fluid exits in the formation. So another study to synthesize missing DT is to use soft-computing methods, such as artificial neural network, gene expressing programming, and fuzzy reasoning. ANN (artificial neural network) has been frequently used in petrophysical properties estimation, and results show satisfied performances when choosing proper models and parameters [8, 9, 16]. The most important property of ANNs is their ability to approximate virtually any function in a stable and efficient way. By using ANNs, it is possible to create a platform on which different models can be constructed. Baziar et al. [22] tested coactive neurofuzzy inference system which combines fuzzy model and neural network in permeability prediction in a tight gas reservoir and gained convincing results.

Since DT has intrinsic links with the other geophysical logs, researchers often use logs like GR, REID, and so forth as the original inputs and the DT as outputs. Linear and nonlinear relationships have been set up using the soft-computing methods. But the results are not always satisfied. Thus our purpose is to build an optimal and reliable relationship between those geophysical logs and DT log.

In this paper, we investigate the capability of a kernel extreme learning machine in building the nonlinear mathematical model that best explains DT (target) as a function of GR, REID, DEN, and SP (predictors).

3.2. Data Preparation. In order to validate the use of KELM in the context of log data recorded in oil and gas wells, we employed datasets obtained from seven wells drilled in the GJH survey in Erdos Basin.

The study involves the following well logging parameters: gamma ray (GR), deep resistivity (REID), self-potential (SP), formation density (DEN), and sonic log (DT). Among the wells, wells of YQ2, Y209, S211, S212, and S215 have full suites of well logs, while DT log is not available in the other two wells (S219 and S205). According to the evaluation conclusion for the logging process, we choose the farther four wells as training dataset sources and well S215 as the testing dataset. Shanxi group of the Permian formation is set as the analysis interval. Logs of GR, REID, SP, DEN, and DT in the interval from the mentioned four wells are collected and grouped as training dataset, while logs of well S215 as the validation target.

Figure 1 is the example of logs showing of well YQ2 in the Shanxi group of Permian formation ranging from 2700 to 2798 meters. The lithology includes sandstone, mudstone, and thin coal layer, and it is easy to differentiate them from the GR log. Coal layer has very low GR and DEN response and abnormal high DT and REID. Thus, for the same type of rocks, these logs have close geophysical link, which is the foundation for DT prediction using these logs.

We select data in the same interval from the four wells of YQ2, Y209, S210, and S212 as the training samples. To ensure the quality of the logs, we use caliper log (CAL) as the reference. Constant diameter of the wellbore (described by CAL) means good environment for the other suite of logs. Totally about 40,000 data items are available for the training process.

To speed up the convergence of the gradient descent algorithm, data normalization is mandatory for the performance. And the above-mentioned logs have different measurement units. All of the logs are normalized before formally inputting
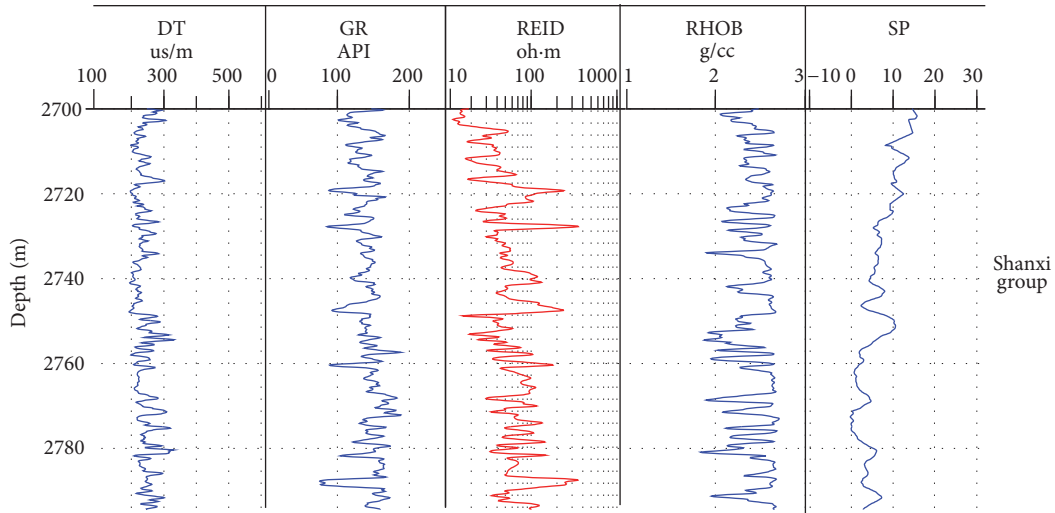
FIGURE 1: Logs showing of well YQ2 in the Shanxi group of Permian formation.
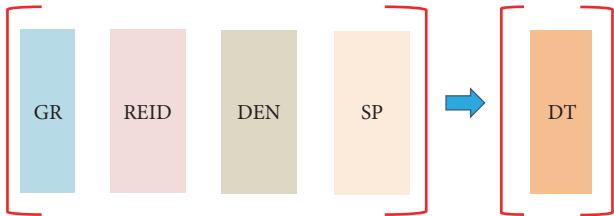

FIGURE 2: Example of multi-input versus single-output sonic log prediction using KELM. Details about the well logging parameters depicted in the figure are given in text.

into the network. The normalized variable has the following form:

$$X_{\text{new}} = \frac{X_{\text{old}} - \min X}{\max X - \min X}, \tag{7}$$

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

In KELM network learning, the output model is created by learning patterns from the training examples provided. Therefore, the training dataset should be carefully chosen in order to provide correct examples. And noise should be removed from the samples; otherwise errors may affect the final performance.

3.3. KELM Model Training. For the KELM network model, there are totally four input neurons and one target at the output layer. The four inputs include GR, REID, SP, and DEN logs, and the main task is to build reliable prediction model between these inputs logs and DT log (shown as Figure 2). Gaussian radial basis kernel function is used because it usually produces good results and outperforms other functions for regression.

In the algorithms of KELM, two hyperparameters, namely, the regularization factor (C) and the basis width

parameter of the kernel function ($s2$), are necessary. To select the best values for these hyperparameters, leave-one-out cross-validation (LOOCV) is usually applied [9]. In the preliminary experiment, the KELM model achieves the best performance when the values of C and s are set to (10, 1), so these values are finally chosen in our experiment.

The quality of the trained model is evaluated based on the prediction accuracy. The Mean Squared Error (MSE) is computed as the average over all squared deviations of the predictions from the real values.

After training, the model could be presented in the following form:

$$Y = \begin{bmatrix} K(X, x_1) \\ \cdots \\ K(X, x_n) \end{bmatrix} \boldsymbol{\beta}, \tag{8}$$

where $K(*)$ is the Gaussian radial basis kernel function, n is the number of training data, and $\boldsymbol{\beta}$ is the trained weight matrix of the model based on the training data. By providing unseen input data X to the model, the corresponding model output Y can be predicted.

Furthermore, in order to testify the advantages of KELM, BP network algorithm is used in the model training and testing process to compare with KELM. Backpropagation (BP) feed-forward network is the most commonly used ANN approach, and it is also criticized on its difficulty to decide learning rates, being easy to be stuck on local minimums, overfit problems, and being time-consuming [11].

Table 1 shows the results on testing data. Accuracy, MSE, and training time are three factors in comparison, and the values are obtained by averaging estimations of the samples in well YQ2. The table shows the accuracy, Mean Squared Error (MSE), and total time in seconds for the two processing approaches, respectively. Best results are achieved by KELM with an accuracy of 0.906, mean absolute error of 0.423%, and fast learning speed (23 seconds).

Table 1: Comparison of porosity prediction performance results on KELM against BP methodology for well YQ2. The comparison strata belong to the Shanxi group of the Permian formation.

Algorithm	Accuracy	MSE (%)	Training time (s)
BP	0.752	1.812	912
KELM	0.906	0.423	23

3.4. KELM Model Validation and Prediction. Through the above-mentioned training process, the KELM model for predicting DT is established finally. Although the training dataset has almost 40,000 data points, the training task costs very short time and the performance is satisfying. To validate the KELM model, we use well S215 as blind well. The four logs are collected and processed for the well, and then we input them into the model and keep the network parameters. Since the data for validation is small group with nearly 6000 samples, the process only costs 6 seconds and one predicted DT is generated. In well S215, there has been DT log, so that the predicted DT can be used in comparison with the real DT. Figure 3 shows the comparison result. The curve with the red color is the predicted one from KELM model, and the curve with the blue color stands for the recorded DT log. It is easy to see that the total changing trend and the finest part are almost the same; thus the model is qualified in this study and is reliable to be a predictor.

In this study, DT log is missing in the two wells of S219 and S205. Here the KELM model is then recommended to do the prediction task for the two wells. Luckily, the four input logs (GR, REID, DEN, and SP) are guaranteed in both of the wells. Using the same noise-filtering and normalization step in the training and validating step, we firstly input the four predictor logs of well S219 into the model and generate DT log for this well. And then we repeat the steps for the well S205 and also get the DT log. Figure 4 shows the predicted DT log for well S219 in the Shanxi group of Permian formation.

4. KELM-Estimated DT Application

The above analysis has shown the reliability and accuracy of the KELM-based prediction model. All of the 7 wells in the studied area have DT logs now, although two of them are generated using KELM model.

In reservoir description phase, seismic profiles are just wiggle-based and not so convenient for researchers to understand and identify the potential fluid zone. Thus transforming the wiggle shape of seismic sections into velocity or lithological profiles are the necessary step in seismic interpretation. That goal of transformation in geophysical process is the seismic inversion. Since DT log has time unit and velocity information, while seismic data is just in time unit, in the inversion task, DT can be used to do the well-to-seismic calibration and mark the reservoir interval. Here we just focus on the KELM-estimated DT application in the seismic inversion other than discussing the complex inversion technique.

Figure 5 shows part section of the seismic inversion result for line 400 using the predicted DT log of well S205.
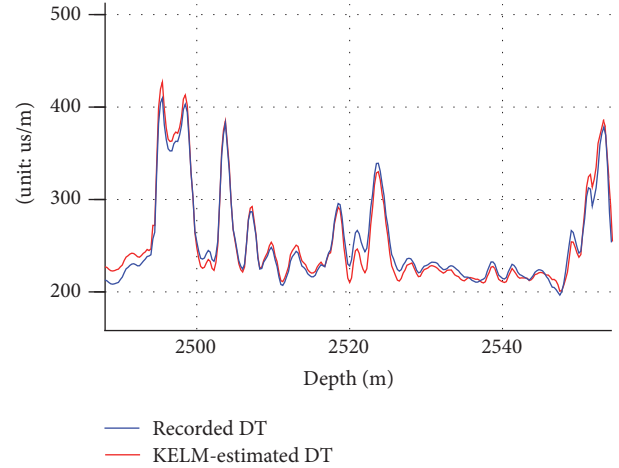

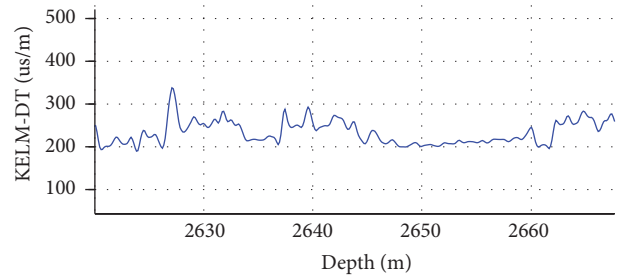Figure 3: Logs comparison in Shanxi group of Permian formation in well S215.

Figure 4: Estimated DT log using KELM model (colored in blue) for S219.

The inversion result is colored, and the color stands for the velocity change within the Permian formation. Warm color of red and yellow is the high velocity area, while the cold color of green and blue is the relatively low velocity area. Since the rocks within the interval have the difference in velocity reference, the color changes can be viewed as the lithology component difference. Normally sandstone has higher velocity than mudstone, and coal layer has the lowest velocity character. Therefore warm color in the section represents the sandstone area, while the pure blue color is the index of coal layer. So when interpreting the inversion result with the geological reference, we may divide the interval into three parts: the upper part-I, which is mainly composed of sandstone and mudstone and the farther is richer, the middle part-II, with upper half-dominant coal layer and lower half-dominant sandstone, and the lower part-III, which has almost the same bedding principal as the middle part, with thinner sandstone and coal layer. The estimated DT log is inserted as color plot and the meaning of color ranges is the same as the inversion section. It almost matches the section in color resolution, and that is the normal phenomenon. DT log has finer sample interval than the section, and, for the section, more focus will be directed to the horizontal color difference interpretation. The continuous horizontal color zones mean a lot for the geologists and engineers.
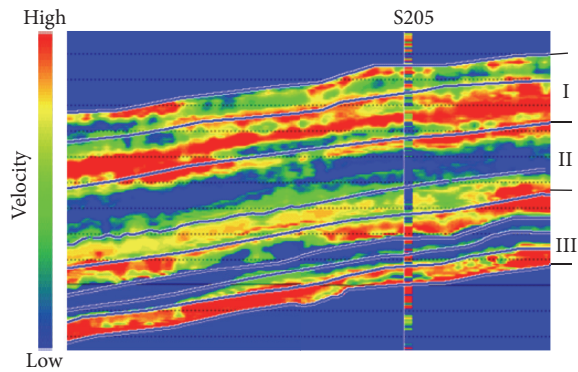
FIGURE 5: KELM-based inverted velocity section crossing well S205.

5. Conclusions

This paper discusses kernel extreme learning machine as a tool for predicting the sonic log in gas/oil wells based on other available common logs. Strict steps including data normalization, training set selection, and optimization of the ELM parameters are very important for deciding the prediction power, the generalization capability, and the complexity of the derived regression model. Extensive applications are carried on to investigate the prediction power of model-predicted DT log use for seismic inversion.

The method presented here is not limited to modeling DT logs only. It can be extended, with appropriate modifications of the algorithm, in any area of well logging studies, where missing log values are needed. Thus, we offer a blueprint for future similar applications.

Competing Interests

The authors declare that they have no conflict of interests.

Acknowledgments

References

[1] A. F. Al-Anazi and I. D. Gates, "Support vector regression to predict porosity and permeability: effect of sample size," *Computers and Geosciences*, vol. 39, pp. 64–76, 2012.

[2] J. Asadisaghandi and P. Tahmasebi, "Comparative evaluation of back-propagation neural network learning algorithms and empirical correlations for prediction of oil PVT properties in Iran oilfields," *Journal of Petroleum Science and Engineering*, vol. 78, no. 2, pp. 464–475, 2011.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] J. R. Hearst, P. H. Nelson, and F. L. Paillet, *Well Logging for Physical Properties*, John Wiley & Sons, 2000.

[8] A. Kouider El Ouahed, D. Tiab, A. Mazouzi, and A. J. Safraz, "Application of artificial intelligence to characterize naturally fractured reservoirs," Paper SPE 84870, 2013.

[9] A. Ouenes, "Practical application of fuzzy logic and neural networks to fractured reservoir characterization," *Computers & Geosciences*, vol. 26, no. 8, pp. 953–962, 2000.

[10] R. S. Zazoun, "Fracture density estimation from core and conventional well logs data using artificial neural networks: the Cambro-Ordovician reservoir of Mesdar oil field, Algeria," *Journal of African Earth Sciences*, vol. 83, pp. 55–73, 2013.

[11] 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.

[12] 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*, pp. 985–990, Budapest, Hungary, July 2004.

[13] 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.

[14] 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.

[15] J. B. Butcher, D. Verstraeten, B. Schrauwen, C. R. Day, and P. W. Haycock, "Reservoir computing and extreme learning machines for non-linear time-series data analysis," *Neural Networks*, vol. 38, pp. 76–89, 2013.

[16] J. Cao, J. Yang, Y. Wang, D. Wang, and Y. Shi, "Extreme learning machine for reservoir parameter estimation in heterogeneous sandstone reservoir," *Mathematical Problems in Engineering*, vol. 2015, Article ID 287816, 10 pages, 2015.

[17] R. Minhas, A. Baradarani, S. Seifzadeh, and Q. M. Jonathan Wu, "Human action recognition using extreme learning machine based on visual vocabularies," *Neurocomputing*, vol. 73, no. 10–12, pp. 1906–1917, 2010.

[18] 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.

[19] 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.

[20] S. Shamshirband, K. Mohammadi, H.-L. Chen, G. N. Samy, D. Petković, and C. Ma, "Daily global solar radiation prediction from air temperatures using kernel extreme learning machine: a case study for Iran," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 134, pp. 109–117, 2015.

[21] 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.

[22] S. Baziar, M. Tadayoni, M. Nabi-Bidhendi, and M. Khalili, "Prediction of permeability in a tight gas reservoir by using three soft computing approaches: a comparative study," *Journal of Natural Gas Science and Engineering*, vol. 21, pp. 718–724, 2014.