



OPEN ACCESS

EDITED BY

Peng Zhenming,
University of Electronic Science and
Technology of China, China

REVIEWED BY

Jiachun You,
Chengdu University of Technology,
China
Wenchao CHEN,
Xi'an Jiaotong University, China
Yatong Zhou,
Hebei University of Technology, China

*CORRESPONDENCE

Yungui Xu,
✉ yungui.xu@outlook.com

RECEIVED 29 January 2023

ACCEPTED 11 April 2023

PUBLISHED 07 July 2023

CITATION

Wang H, Xu Y, Tang S, Wu L, Cao W and
Huang X (2023), Well log prediction while
drilling using seismic impedance with an
improved LSTM artificial neural networks.
Front. Earth Sci. 11:1153619.
doi: 10.3389/feart.2023.1153619

COPYRIGHT

© 2023 Wang, Xu, Tang, Wu, Cao and
Huang. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original author(s)
and the copyright owner(s) are credited
and that the original publication in this
journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Well log prediction while drilling using seismic impedance with an improved LSTM artificial neural networks

Heng Wang^{1,2}, Yungui Xu^{1,2*}, Shuhang Tang^{1,2}, Lei Wu³,
Weiping Cao^{1,2} and Xuri Huang^{1,2}

¹State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Southwest Petroleum University, Chengdu, China, ²School of Geosciences and Technology, Southwest Petroleum University, Chengdu, China, ³Shale Gas Project Management Department of CNPC Chuanqing Drilling Engineering Co., Ltd., Chengdu, China

Well log prediction while drilling estimates the rock properties ahead of drilling bits. A reliable well log prediction is able to assist reservoir engineers in updating the geological models and adjusting the drilling strategy if necessary. This is of great significance in reducing the drilling risk and saving costs. Conventional interactive integration of geophysical data and geological understanding is the primary approach to realize well log prediction while drilling. In this paper, we propose a new artificial intelligence approach to make the well log prediction while drilling by integrating seismic impedance with three neural networks: LSTM, Bidirectional LSTM (Bi-LSTM), and Double Chain LSTM (DC-LSTM). The DC-LSTM is a new LSTM network proposed in this study while the other two are existing ones. These three networks are thoroughly adapted, compared, and tested to fit the artificial intelligent prediction process. The prediction approach can integrate not only seismic information of the sedimentary formation around the drilling bit but also the rock property changing trend through the upper and lower formations. The Bi-LSTM and the DC-LSTM networks achieve higher prediction accuracy than the LSTM network. Additionally, the DC-LSTM approach significantly promotes prediction efficiency by reducing the number of training parameters and saving computational time without compromising prediction accuracy. The field data application of the three networks, LSTM, Bi-LSTM, and DC-LSTM, demonstrates that prediction accuracy based on the Bi-LSTM and DC-LSTM is higher than that of the LSTM, and DC-LSTM has the highest efficiency overall.

KEYWORDS

machine learning, long short-term memory, well log prediction, drilling bit, seismic impedance, neural networks

1 Introduction

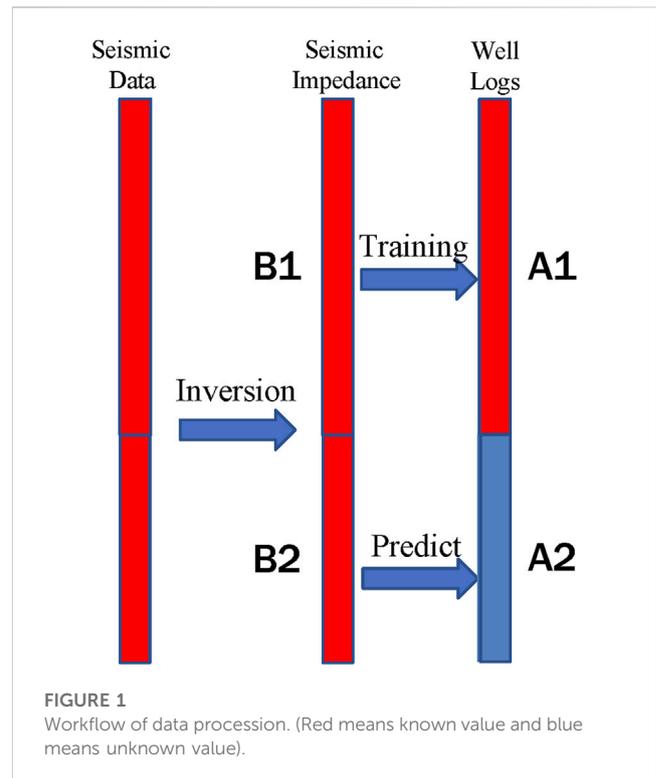
During the drilling process, reliable prediction of the rock properties of the geological formation below the drill bit is of great significance. If well logs can be effectively obtained within a certain depth range below the drill bit, it can surely help to improve the drilling process with lower risk and optimal strategy. Researchers have proposed various methods for well log prediction while drilling (Wang, 2017; Tamim et al., 2019). Wang (2017) established a 3D model by analyzing the drilling core, rock

chip, and seismic data. The well log curve values below the drilling bit along the borehole trajectory are predicted based on the 3D model. Tamim et al. (2019) constructed a Bayesian classifier using well logs as the classification target. With the posterior probability calculated with prior distribution and conditional probability, the spatial distribution of well logs can be predicted. However, the actual geological formations are non-homogeneous and complex. The logging curve values show great fluctuations with depth even when the sampling interval of the well log is small. The mapping relationships between different data points are strongly non-linear, and the traditional methods are not able to effectively predict the actual formation changes.

Predicting rock property with geophysical data is often a non-linear problem in many cases. Describing the intrinsic relationship between properties and their geophysical responses using explicit mathematical or physical equations can be challenging. However, machine learning has shown great potential in recent years, as it can describe these relationships with network parameters by using a large number of training datasets. Examples of relevant machine learning applications include geological parameter estimation (Ahmed Ali Zerrouki and Baddari, 2014; Iturrarán-Viveros and Parra, 2014), lithology discrimination (Wang et al., 2014; Silva et al., 2015), and stratigraphic boundary determination (Singh, 2011; Silversides et al., 2015). In recent years, there has been a surge of research on well log prediction, resulting in significant improvements in performance (Rolon et al., 2009; Alizadeh et al., 2012; Mo et al., 2015; Long et al., 2016; Salehi et al., 2017).

The well log prediction methods discussed above use fully connected neural networks (FCNN) to construct a point-to-point mapping. However, FCNN shows less effectiveness in characterizing the trend of the data because it cannot capture the relations between the point data of well log at different depths. This means that the correlation between the rock properties at a shallower depth and those at a deeper depth is ignored, potentially contradicting the sedimentary principles in the geological sense. To fix this problem with FCNN in well log prediction, many researchers have improved FCNN by coupling non-machine learning methods, such as wavelet transform (Adamowski and Chan, 2011) and singular spectrum analysis (Sahoo et al., 2017). However, these improvements are often complex and cumbersome to implement. An alternative way to utilize recurrent neural networks (RNNs) is able to deal the problems (Schuster and Paliwal, 1997). In the RNN structure, there is a self-looping structure within each neural unit, which allows previous information to be retained and used later. Since the information can flow freely in the RNN, well log prediction with this method integrates the intrinsic connection between different logs and the overall trend with depth, which follows geological principles.

Long Short-Term Memory Neural Networks (LSTM), an advanced form of RNN, have become widely utilized in the deep learning community for various tasks. LSTM incorporates gate structures within each automatic cycle structure to mimic biological neurons' information conduction



patterns, thereby storing more long-term sequential information without additional tuning. This attribute has facilitated LSTM's extensive use in natural language processing (Deng et al., 2019), machine translation (Lokeshkumar et al., 2020), and speech recognition (Graves and Jaitly, 2014). Moreover, RNNs and LSTMs have also found applications in the field of hydrology to deal with time series data-related problems (da Silva and Saggiaro, 2013), and some researchers have used LSTM to generate logs (Jin, 2018).

In the case of well log prediction while drilling, Wang et al. (2020) employed the LSTM model, whereas Shan et al. (2021) utilized the CNN-LSTM hybrid model to achieve the goal. Both methods employed neighboring well logs to predict the logs at the undrilled segment, resulting in successful outcomes. However, when neighboring wells are too distant (as in the case of sparse neighboring wells) and well correlation is weak, the prediction results can become biased. In certain scenarios, prediction may not be possible due to the unavailability of well logs from neighboring wells at the prediction depth.

This study proposes a novel approach for well log prediction while drilling, utilizing seismic impedance and three different artificial intelligence networks: the LSTM network and its two derivative networks, namely, the Bi-LSTM and the newly developed Double Chain LSTM (DC-LSTM). The proposed methodology effectively integrates seismic impedance constraints for well log prediction. The results demonstrate that the two derivative networks outperform the LSTM network in terms of prediction accuracy. Additionally, the DC-LSTM network exhibits superior computational efficiency as compared to the Bi-LSTM network, owing to the reduction in

TABLE 1 The steps of the work.

Steps of well log prediction while drilling
Step 1: Obtaining seismic impedance by the inversion of seismic data.
Step 2: Constructing a dataset with B1 and A1 as the label library.
Step 3: Training the neural network with the label library and establishing the mapping from seismic impedance to well log.
Step 4: Performing the prediction from B2 to A2 based on the trained network.

the number of network parameters and consequently the computing time, while maintaining prediction accuracy. A practical case study is conducted to verify the effectiveness of the proposed methodology.

2 Methodology

2.1 Well log prediction while drilling principle: training and prediction

The acquisition of well logs is traditionally carried out after drilling, while well log prediction while drilling refers to the real-time estimation of well logs at a certain depth below the drill bit during the drilling process. This paper presents a workflow for such prediction, as depicted in Figure 1. The known observed logs are shown in red, whereas the logs yet to be predicted are shown in blue. The seismic impedance obtained through seismic inversion is utilized as input data for the training and prediction of well logs using three different networks, as outlined in Table 1.

2.2 Three networks: LSTM, Bi-LSTM, and DC-LSTM

2.2.1 The LSTM adaption

The LSTM neural network (Hochreiter and Schmidhuber, 1997) is a specialized type of recurrent neural network (RNN) designed to capture long-term dependencies within sequential data. This makes it particularly suitable for processing well logs,

which are a type of sequential data with a relatively small sampling interval. A significant characteristic of well logs is the representation of long-term depth trends within a large depth interval, which can play an essential role in prediction tasks. Unfortunately, these trends are often overlooked in existing models. However, LSTM’s ability to retain information with long-term dependencies from previous, more distant steps allows it to capture and incorporate these trends into its predictions. As a result, LSTM represents an effective tool for well log prediction.

Figure 2 illustrates the architecture of the LSTM network. The inputs to the network are denoted by x_1 to x_n , while h_1 to h_n represent the hidden state vectors. The computation of the LSTM network’s outputs is performed iteratively in a stepwise fashion, from Step 1 to Step n , with the hidden state vectors h_1 to h_n and outputs y_1 to y_n calculated sequentially by the LSTM unit. In the context of well log prediction, only the final output, y_n , is utilized for prediction, and is subsequently fed into a Fully-Connected Neural Network (FCNN).

2.2.2 Bidirectional LSTM network adaption

The Bi-LSTM neural network (Graves and Schmidhuber, 2005) represents a variant of the LSTM architecture that incorporates two parallel LSTMs: one running forward along the input sequence, and another running backward in reverse order. By exploiting both the past and future context of the input data, the Bi-LSTM model can capture more comprehensive and accurate representations of sequential data. Specifically, the forward and backward LSTMs compute hidden state vectors in opposite directions, and the final output of the network is produced by combining the two LSTM’s outputs with appropriate weights in the output layer. A diagram of the Bi-LSTM structure is presented in Figure 3.

In Figure 3, the input is denoted by x , and the hidden state vectors of the forward and backward hidden layers are represented by h^f and h^b , respectively. The forward layer computes its outputs in a sequential manner from Step 1 to n , while the backward layer calculates its outputs from Step n to 1. The LSTM unit calculates the hidden state vectors of the forward and backward layers, i.e., h^f and h^b , respectively. The network generates its outputs based on the following equation,

$$y_n = h^f \otimes h^b \tag{1}$$

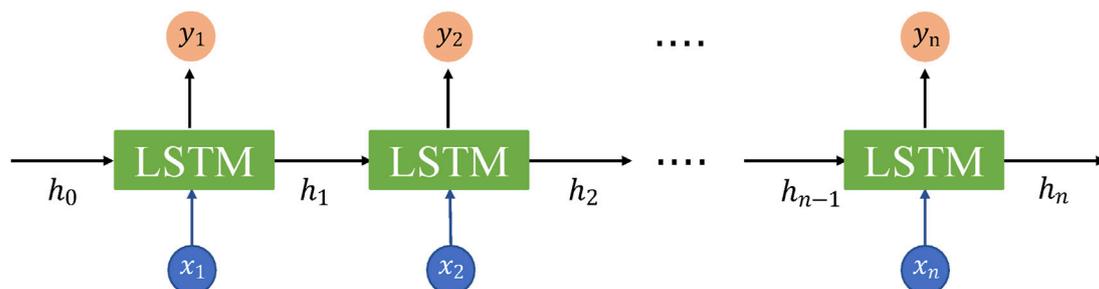
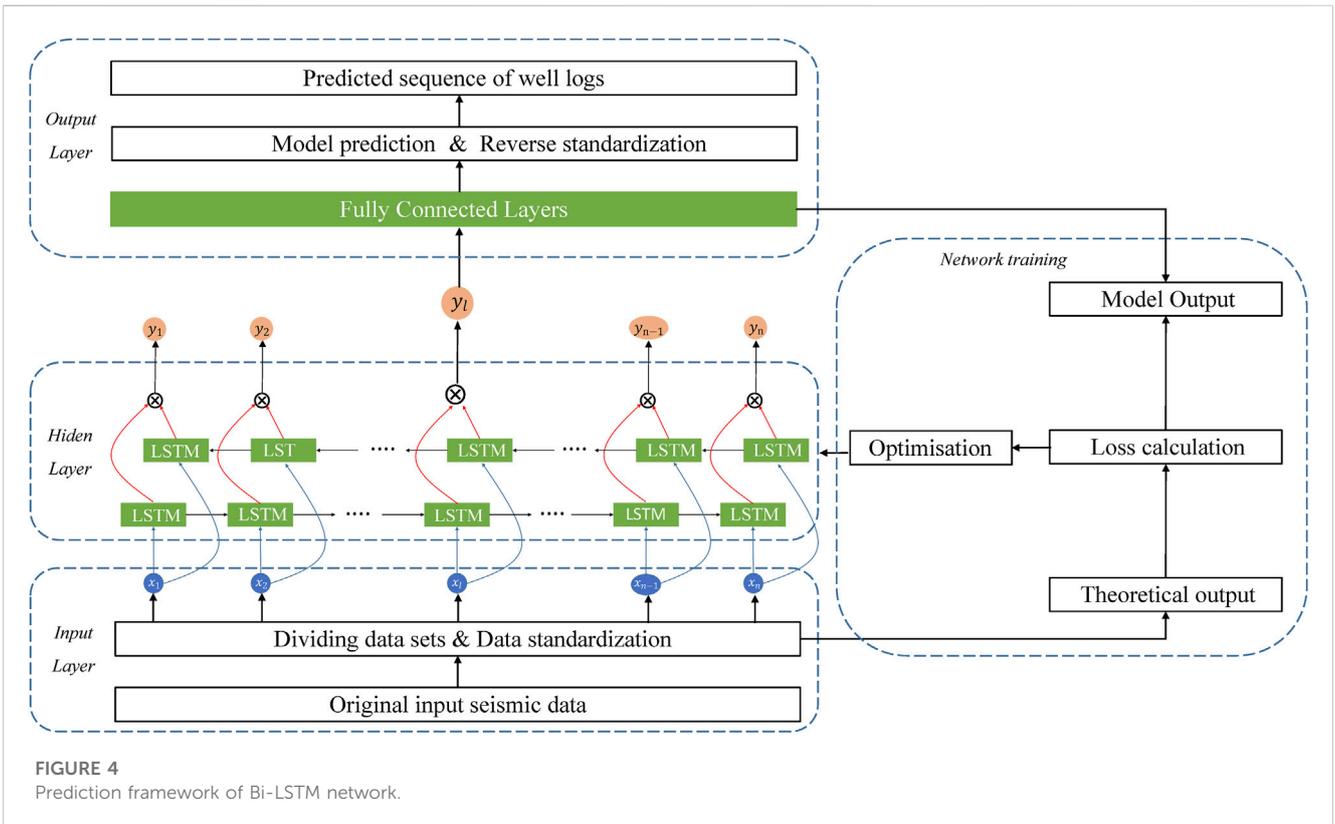
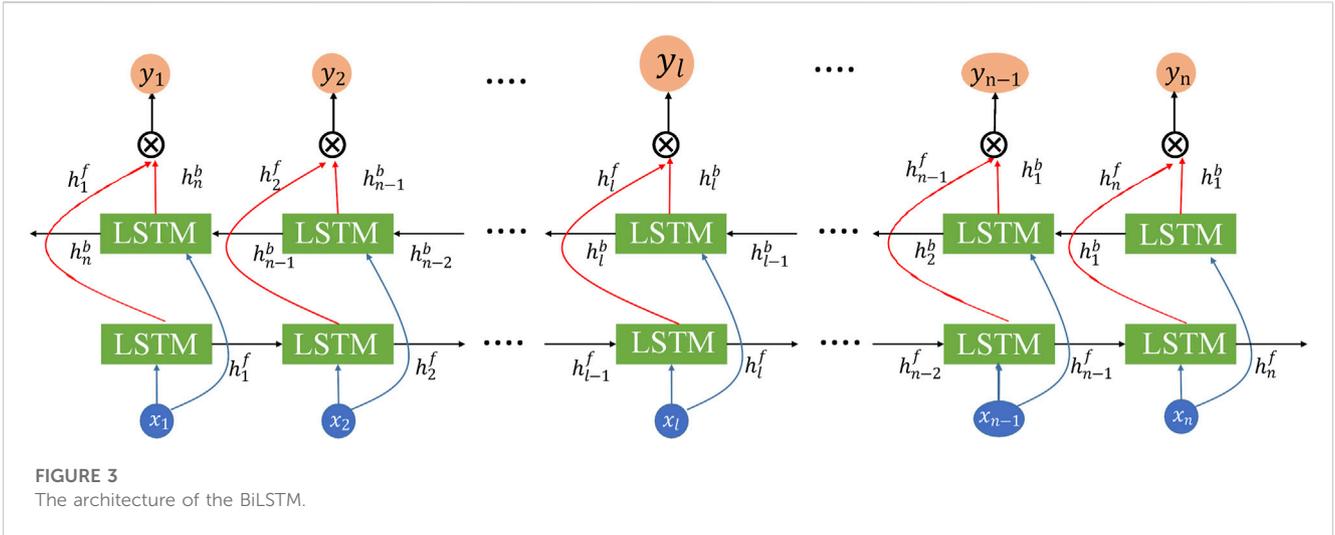


FIGURE 2 The architecture of the LSTM. (Each green box stands for one LSTM unit).



where a \otimes is a matrix operator, which is utilized to couple h^f and h^b sequences. The operator \otimes could be a summation, multiplication, concatenating, or average function. In the experiment of this paper, the operator \otimes is concatenating the horizontal matrix. The final output of a Bi-LSTM network is expressed as vectors, $[y_1, y_2, \dots, y_l, \dots, y_{n-1}, y_n]$. The element in the middle position, y_l , is used to predict the log values.

Figures 4, 5 present the prediction framework using a Bi-LSTM network for mapping the relationship between well logs and seismic impedance. The prediction framework is comprised of three layers, namely, the input layer, the output layer, and the hidden layer. The network is trained using the training set to obtain the mapping relationship between the log segment of the well and seismic

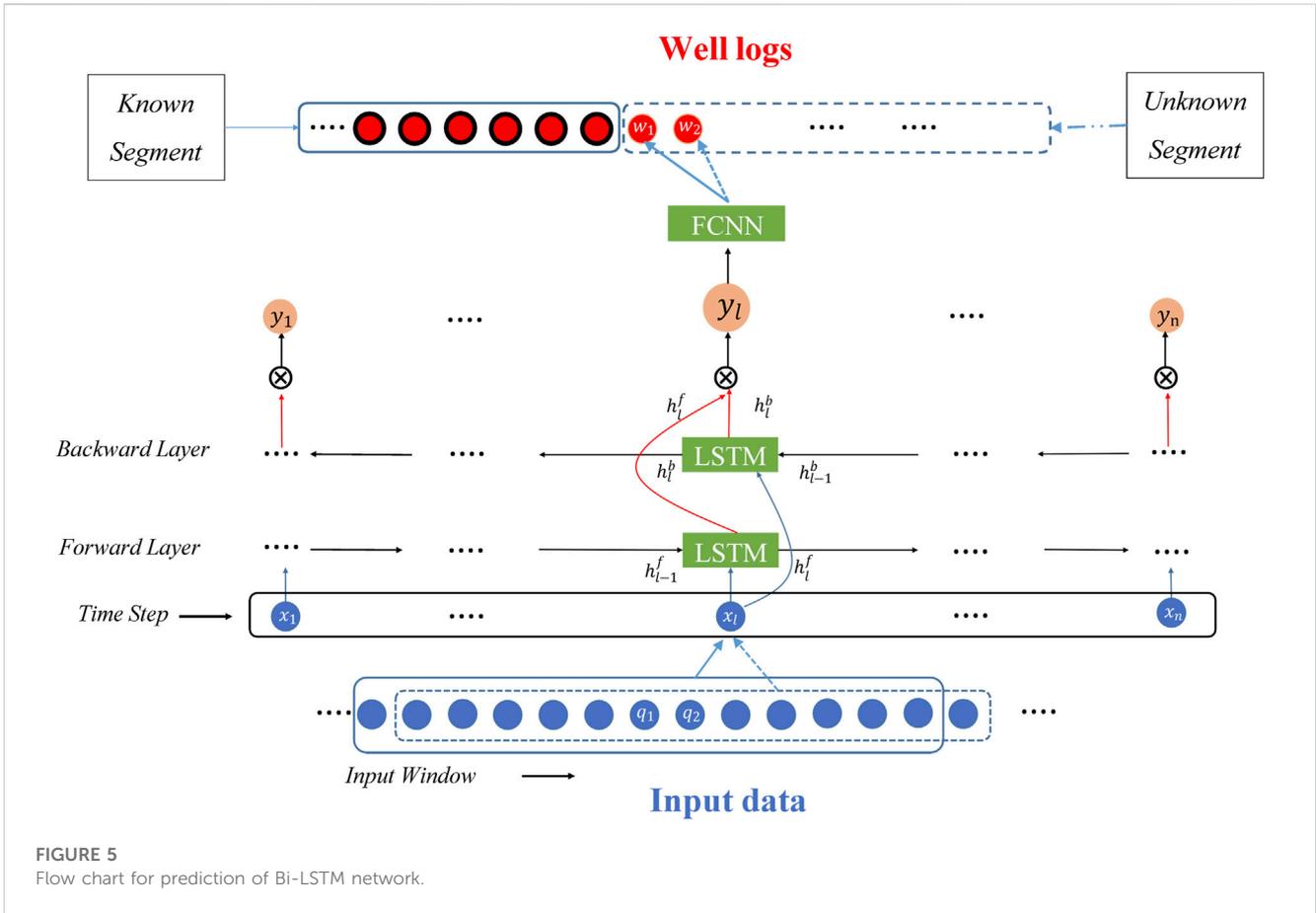


TABLE 2 The specific steps of the prediction phase.

Steps of Bi-LSTM model for well log prediction while drilling
<i>Step 1:</i> Determine the location of the input window
<i>Step 2:</i> As shown in Figure 5, when the data is input to the network (the data in the solid box), q_1 is located in the midpoint of the input window, i.e., q_1 is the input of time step x_l , then the LSTM cell at the position of time step = l can obtain as well as the original seismic data q_1 and the input of two cell states h_{l-1}^f and h_{l-1}^b , outputs h_l^f and h_l^b . After the \otimes operation y_l is obtained and from this, the logging data w_1 for the unknown segment is predicted, w_1 and q_1 being the corresponding data for the same depth.
<i>Step 3:</i> The Input Window is shifted one place to the right, so that when the data is entered into the network (the data inside the dashed box), q_2 is located right in the middle of the Input Window, and it can be predicted that w_2, w_2 , and q_2 are the corresponding data of the same depth.
<i>Step 4:</i> Repeat the step3 to predict all unknown segment data

impedance. The seismic impedance within the undrilled well segment is then used to predict well logs.

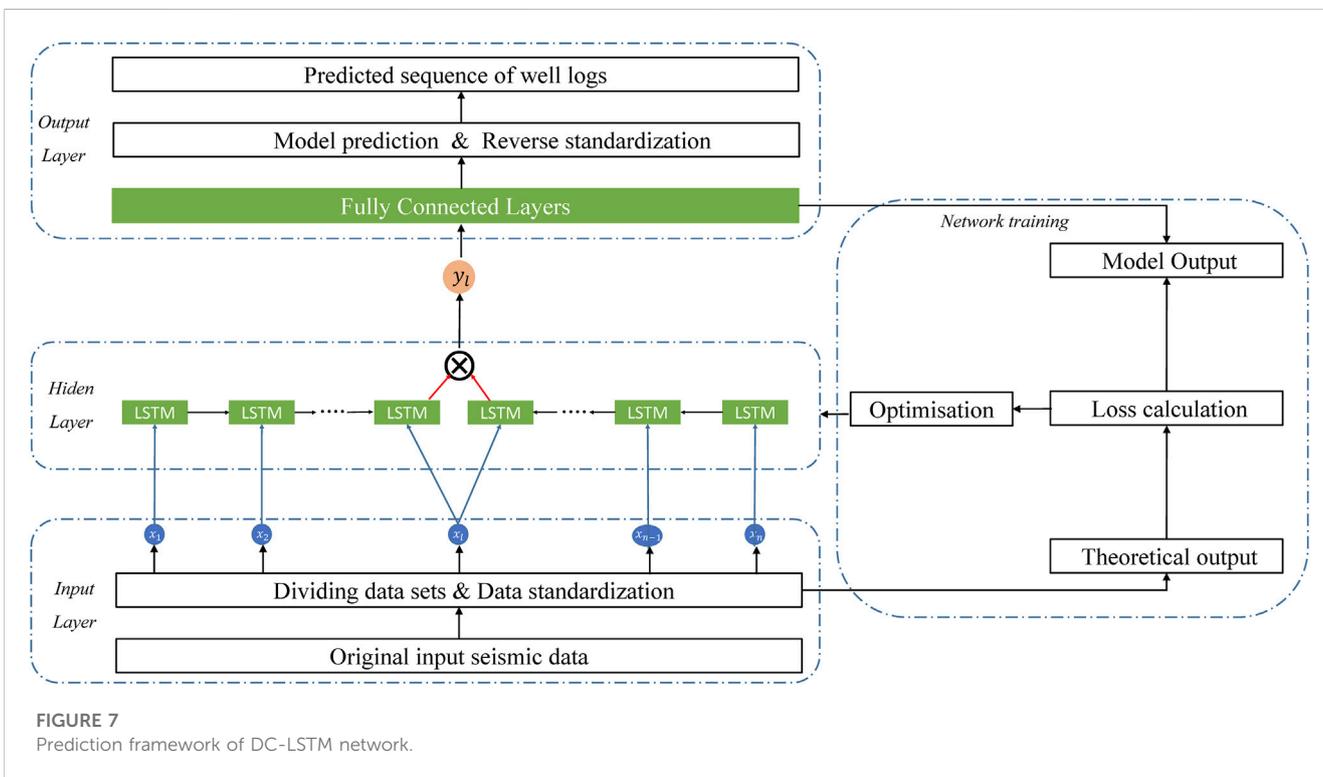
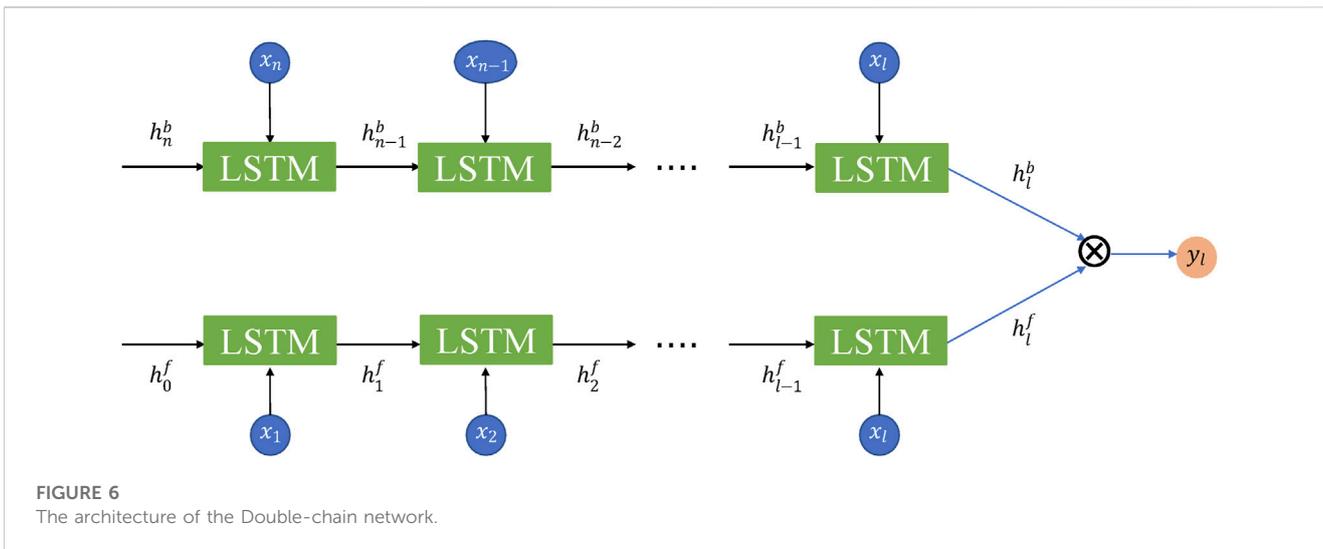
In Figure 5, the model is trained using well logs observed from the drilled segment and the corresponding seismic impedance to optimize the network. The trained model is subsequently used to predict log values of the well segment below the training data along the well trajectory. The input window specifies the length of the input seismic impedance

sequence to the network, and the input window length should be consistent between the training and prediction phases. In the current example, q_1 represents the value at the position x_l , while h_{l-1}^f and h_{l-1}^b denote the hidden state vectors before and after processing q_1 , respectively. Table 2 outlines the prediction steps in detail.

2.2.3 The double-chain LSTM adaption

The Bi-LSTM model utilized for well log prediction while drilling entails numerous prediction parameters. Specifically, the final output of a Bi-LSTM layer is expressed as a set of vectors, denoted as $[y_1, y_2, \dots, y_l, \dots, y_{n-1}, y_n]$. However, only the vector positioned in the middle, y_l , is employed to predict the log values. As a result, other vectors computed during both training and prediction are not utilized, resulting in a strain on computational efficiency. To address this issue, this paper proposes a DC-LSTM network.

The DC-LSTM network is an optimized network that aims to enhance the computational efficiency in well log prediction while drilling in comparison to the Bi-LSTM model. As shown in Figure 6, DC-LSTM divides the input sequence, represented by $[x_1, x_2, \dots, x_l, \dots, x_{n-1}, x_n]$, into two sequences: $[x_1, \dots, x_l]$ and $[x_l, \dots, x_n]$, where x_l is positioned at the middle of the sequence. The two sequences $[x_1, \dots, x_l]$ and $[x_l, \dots, x_n]$ are then separately inputted into two LSTM. Following the \otimes operation, y_l , which is identical to the y_l mentioned in the Bi-LSTM, is obtained. Because

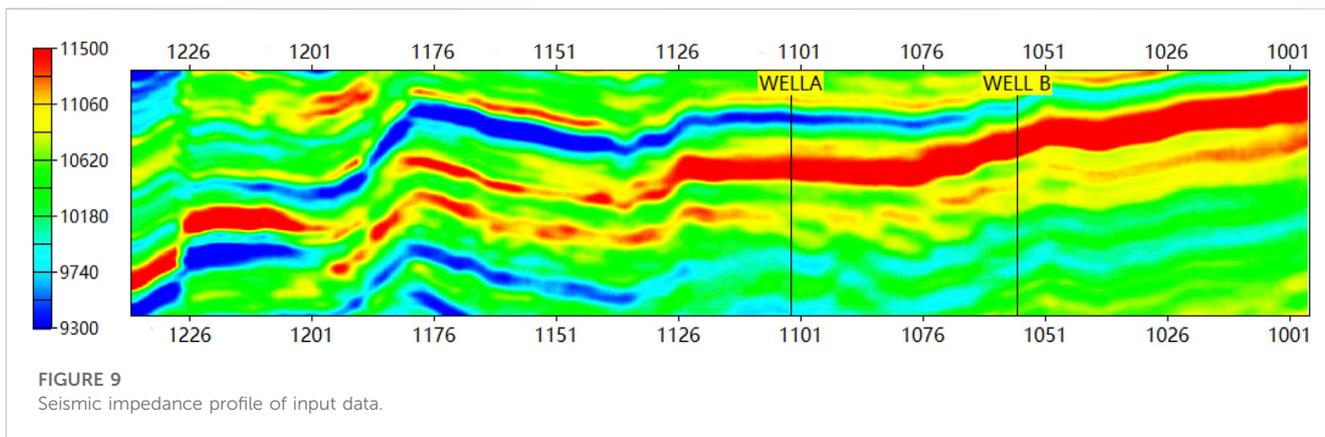
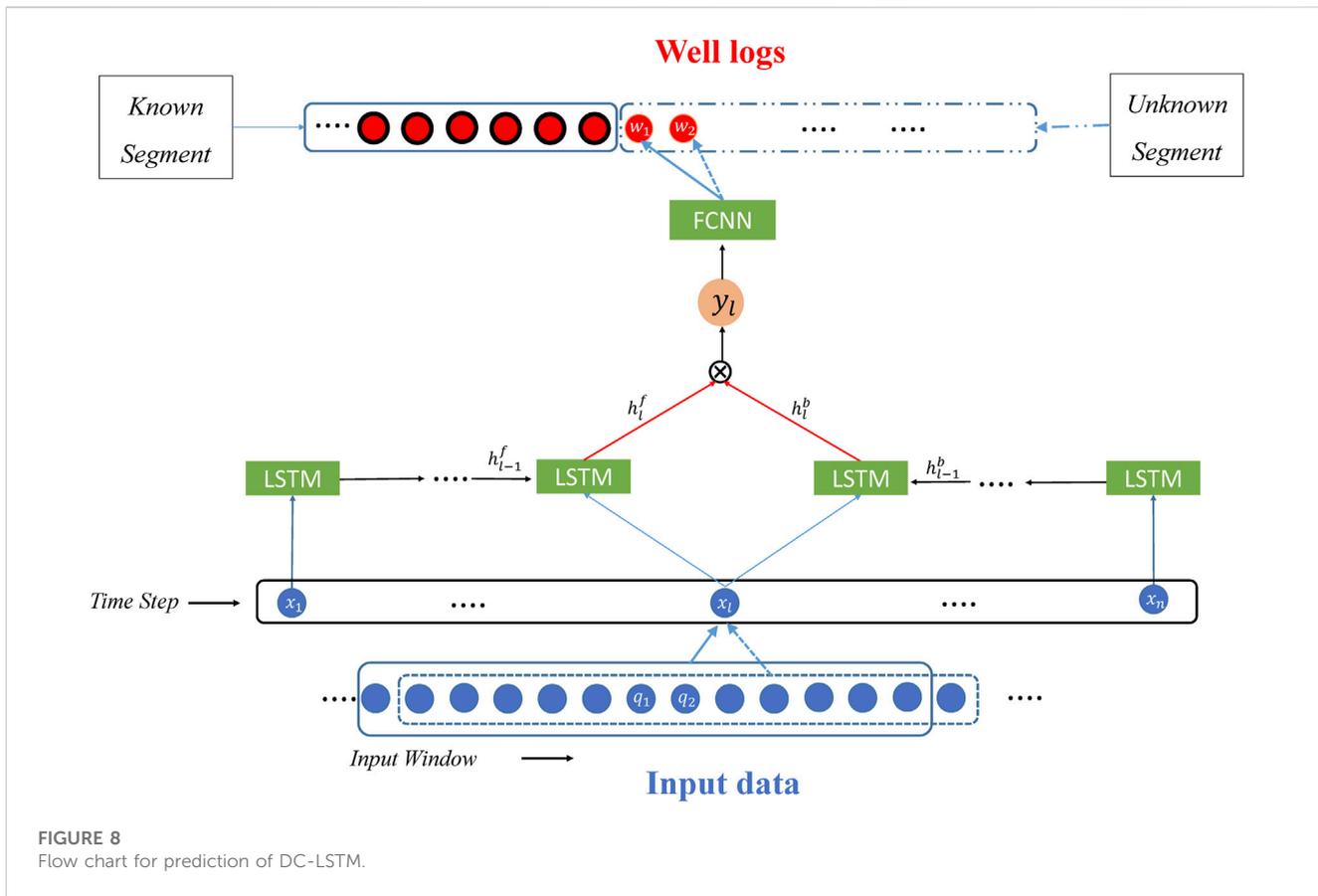


only one vector, y_l , is involved in the prediction process, DC-LSTM significantly reduces the number of network parameters and enhances calculation efficiency relative to Bi-LSTM.

Figures 7, 8 depict the prediction process with DC-LSTM, which is a simplified and more efficient version of Bi-LSTM in well log prediction while drilling. As illustrated in Figure 7, DC-LSTM replaces Bi-LSTM in the hidden layer, resulting in a simpler overall structure. Furthermore, Figure 8 shows that DC-LSTM

requires only $n+1$ LSTM units, whereas Bi-LSTM utilizes $2n$ LSTM units. This reduces the number of parameters and improves computational efficiency. Importantly, y_l calculated in DC-LSTM is identical to that in Bi-LSTM, thereby ensuring the same level of prediction accuracy.

The steps of prediction with DC-LSTM are largely analogous to those outlined in Table 2 for Bi-LSTM. The only difference is that DC-LSTM divides the input sequence $[x_1, x_2, \dots, x_1, \dots, x_{n-1}, x_n]$



into two shorter sequences, namely, $[x_1, \dots, x_l]$ and $[x_l, \dots, x_n]$, which are then fed into two separate LSTMs.

3 Experiment

In this part, we present the results of well log prediction while drilling using seismic impedance with the LSTM, Bi-LSTM, and DC-LSTM.

3.1 Dataset

In this experiment, a pilot area of an oil field in East China is selected for well log prediction. Two wells and a seismic impedance cube in the area are used in this experiment (see Figure 9). The two wells have four logs for testing, density (DEN), natural gamma ray (GR), compensated neutron log (CNL), and induction (CON).

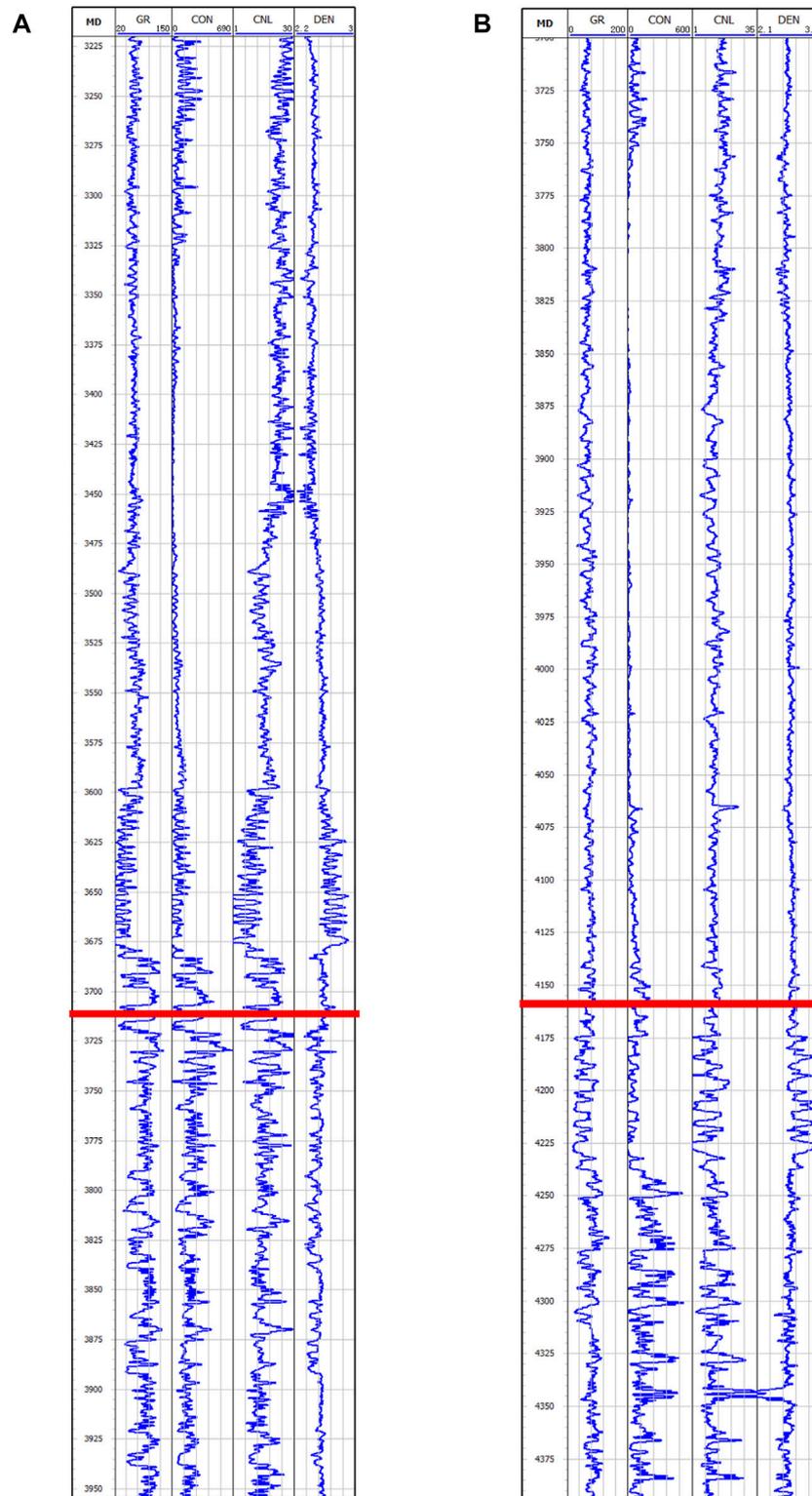


FIGURE 10
Logging curves for well A and well B used in training and validation: (A) well A, (B) well B

Before the prediction with the three networks, two data preprocessing tasks, data cleaning, and standardization are required.

3.1.1 Data cleaning

Data cleaning (Zhang et al., 2015) aims to improve the data quality of input data. Common problems with logs, such as

invalid and missing values, are cleared. Check on the data value range are necessary as well. Any data point, falling outside the normal range, logically unreasonable, or contradictory, is also cleared.

3.1.2 Data standardization

The purpose of data standardization is to centralize data values so that the characteristics of each type of data are balanced. Data standardization also reduces the disturbance of anomalous values that still exist after data cleaning. This paper uses the z-score normalization method (Hrynaszkiewicz, 2010) to standardize input data standardization, expressed in the equations below,

$$x^* = \frac{x_i - \bar{x}}{s} \quad (2)$$

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad (3)$$

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (4)$$

where x^* is the standardized data, x_i is the input data, \bar{x}_i is the mean of the input data, and s is the standard deviation of the input data. The standardized data not only retains the correlations present in the original data but also removes the effects of different magnitudes and ranges of data values.

3.2 Evaluation metrics

The problem discussed in this paper is called the ‘regression task’ in the field of the neural network, and the common loss functions of the regression task are Mean Square Error (MSE), Root Mean Square Error (RMSE), and Mean Absolute Error (MAE). If there are some unusual values (large or small) in the data, these values are more likely to indicate the important geological contrast in rock properties, such as sudden change of lithology, measured acoustic velocity, etc. Therefore, it is necessary to attach more weight on the unusual values. This is achieved by using MSE as the loss function.

MAE and RMSE are used as criteria to evaluate prediction accuracy. MAE is the average of the absolute error between the predicted value and the true value, indicating the true situation of the error, while RMSE is generally used to indicate the deviation between the predicted value and the true value. MAE, MSE, and RMSE are defined as,

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (5)$$

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (6)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (7)$$

TABLE 3 Training parameter setting.

Parameter	Setting
Optimization algorithm	ADAM
Hidden layers	25
Batch size	32
Learning rate	0.002
Dropout	15%
Epochs	150
Input size	1
Output size	1
Layer number	2

where y_i is the predicted value and \hat{y}_i is the true value. The smaller the MAE and RMSE values, the more accurate the model’s prediction of the well logs will be.

In this study, the MAE and RMSE of four logs are calculated and averaged to evaluate the performance of different models. However, the value ranges of these four curves are very different, and the value ranges of the same curve in different wells are also different. For the fair evaluation of the results, the true and predicted values of each curve are put together and normalized before calculating the MAE and RMSE, according to the following equation,

$$y_i = \frac{(y_i - \bar{y})}{(y_{max} - y_{min})} \quad (8)$$

where y_i are the normalized values of a well log, \bar{y} is the mean of the well log, y_{min} refers to the minimum value, and y_{max} denotes the maximum value.

3.3 Experiment results and analyze

In this experiment, two wells, namely, Well A (with measured depth ranging from 3220m to 3960 m) and Well B (ranging from 3700 to 4400 m), along with a seismic impedance segment traversing both wells, were utilized as input data (as depicted in Figure 9). The training data set was constructed by incorporating the first 70% of the well log data (the section above the red line shown in Figure 10), along with the corresponding seismic impedance. The remaining 30% of the logs (the section below the red line in Figure 10) and their corresponding seismic impedance were used for blind test validation. The sampling interval for both wells was set to 0.076 m. The input window length for the networks was fixed at 51 neighboring data points, which corresponds to an actual sampling length of 3.876 m. By using seismic impedance, each input window predicts a single well log point at the depth of the

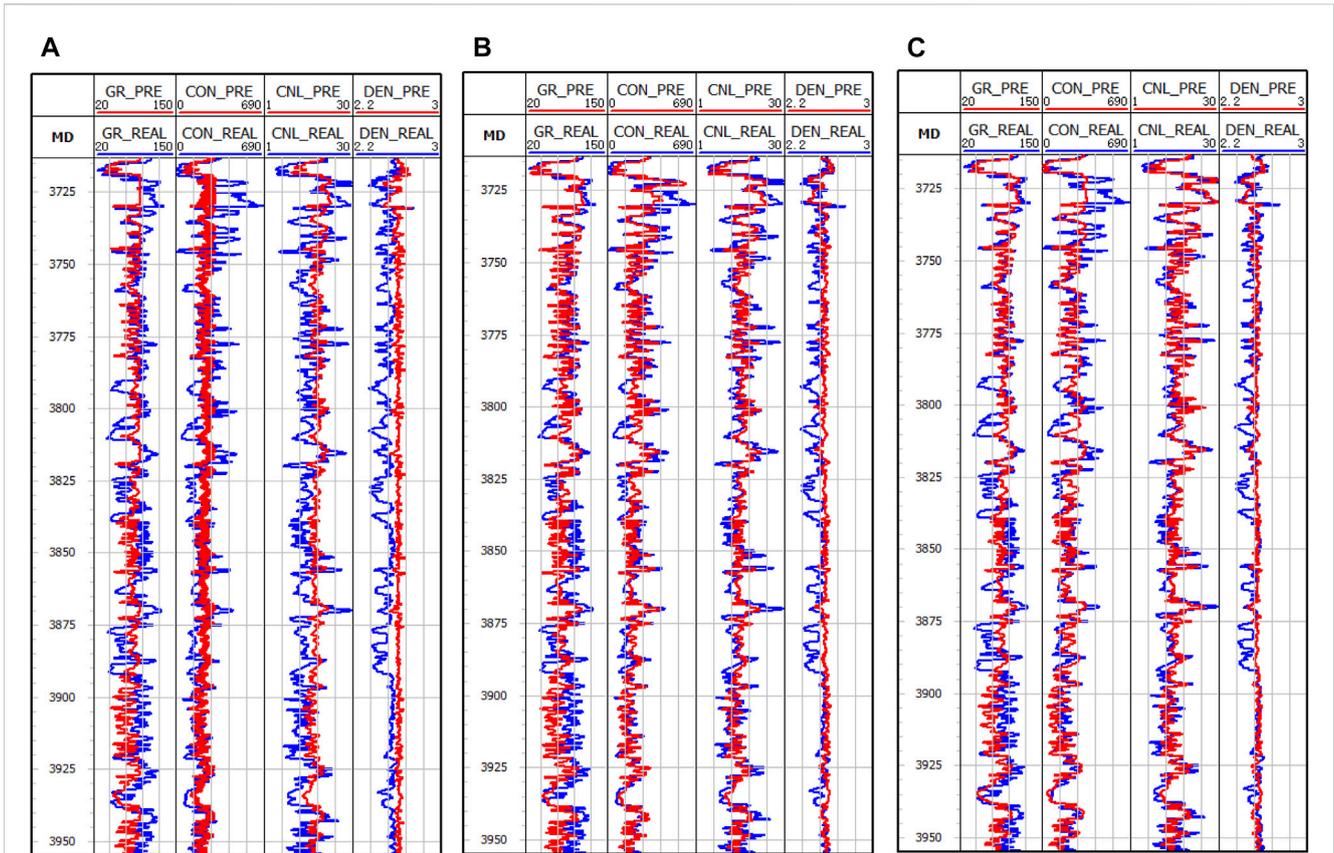


FIGURE 11
Results of the three networks applied to well A: (A) LSTM, (B) Bi-LSTM, (C) DC-LSTM.

TABLE 4 MAE and RMSE of the three methods applied to well A.

	CNL		CON		DEN		GR		Average	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
LSTM	0.154	0.182	0.119	0.153	0.145	0.198	0.176	0.214	0.149	0.187
Bi-LSTM	0.088	0.111	0.067	0.090	0.146	0.183	0.154	0.182	0.114	0.142
DC-LSTM	0.078	0.099	0.076	0.102	0.115	0.156	0.129	0.164	0.099	0.130

TABLE 5 MAE and RMSE of the three methods applied to well B.

	CNL		CON		DEN		GR		Average	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
LSTM	0.062	0.088	0.143	0.207	0.085	0.115	0.167	0.207	0.114	0.154
Bi-LSTM	0.056	0.075	0.119	0.169	0.069	0.110	0.113	0.141	0.089	0.124
DC-LSTM	0.057	0.076	0.110	0.139	0.062	0.089	0.121	0.147	0.086	0.112

middle point of the input window. The window length determines the memory range of the network model during both the training and prediction processes.

In this study, the training data set is obtained by dividing the impedance and well logs into segments with a length of 51. The impedance is utilized as the input, while four well logs, including

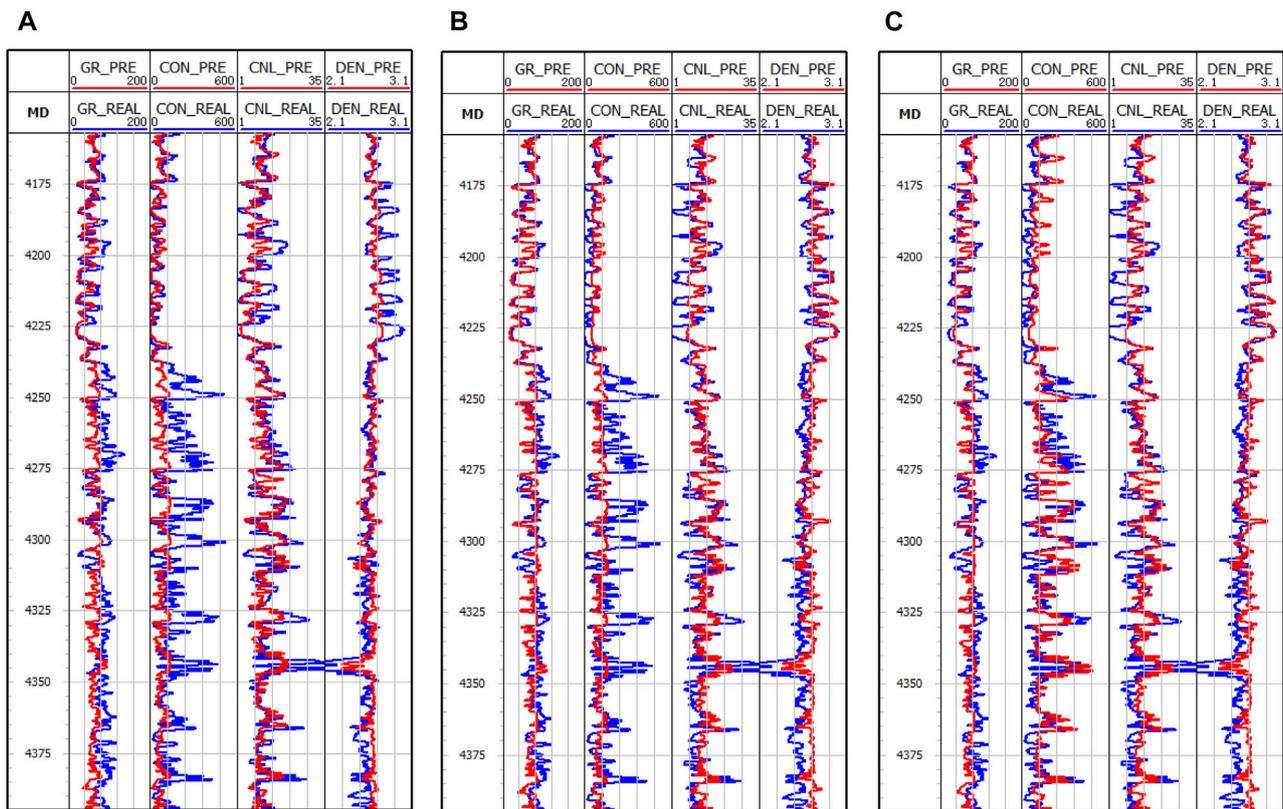


FIGURE 12
Results of the three networks applied to well B: (A) LSTM, (B) Bi-LSTM, (C) DC-LSTM.

TABLE 6 Time spent on three methods (Unit: seconds).

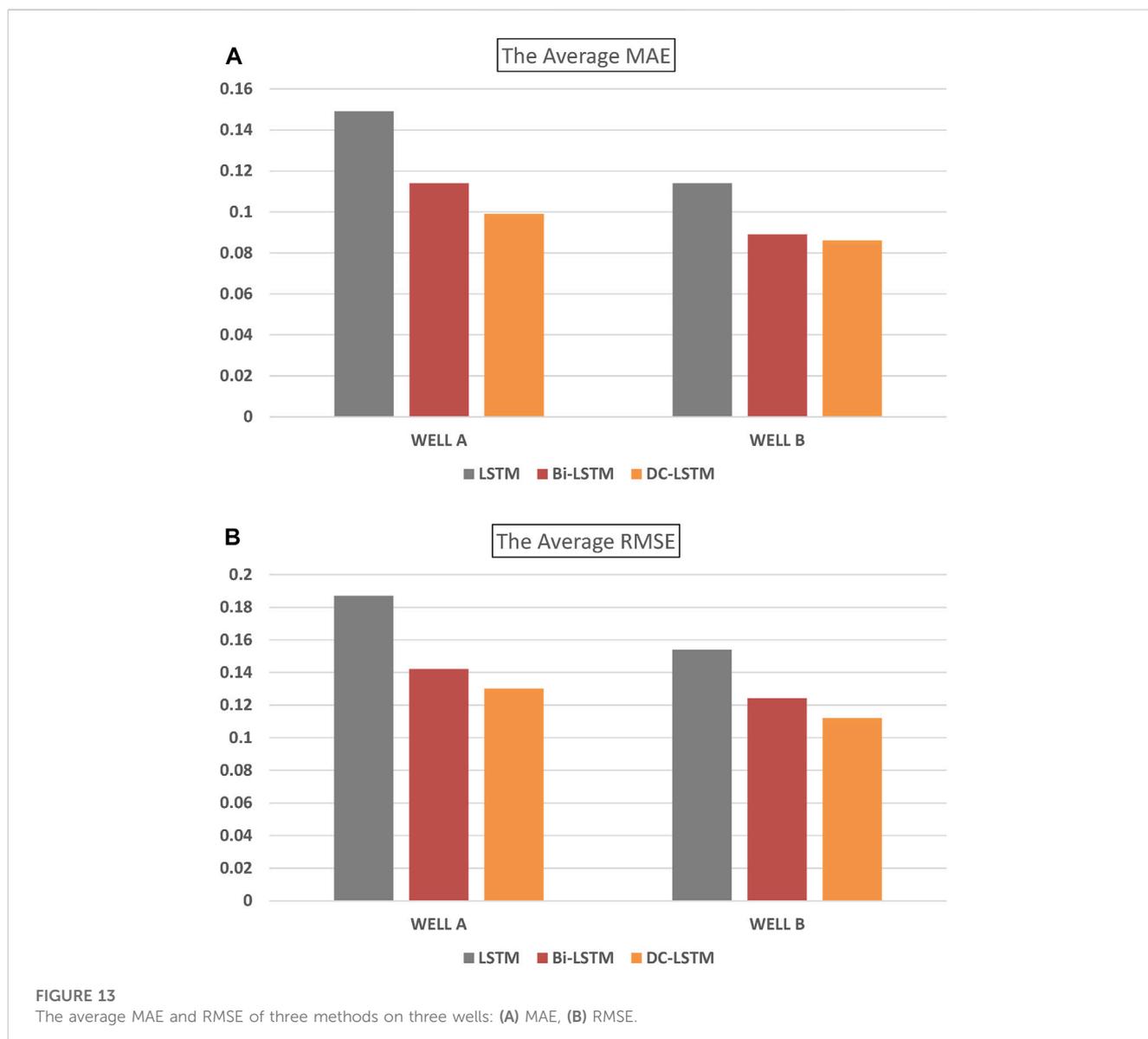
	WELL A		WELL B		Over All		
	Train	Predict	Train	Predict	Train	Predict	Total
LSTM	17.6 ± 1.1	3.4 ± 0.3	18.9 ± 1.2	3.2 ± 0.3	36.5	6.5	43.0
Bi-LSTM	43.5 ± 3.4	8.8 ± 1.1	41.3 ± 3.2	8.1 ± 1.1	84.8	16.9	101.7
DC-LSTM	25.4 ± 2.1	5.1 ± 0.7	24.1 ± 2.1	4.7 ± 0.8	49.5	9.8	59.3

density (DEN), natural gamma ray (GR), compensated neutron log (CNL), and induction (CON) curves, serve as the output. The optimal training parameters, as demonstrated in Table 3, are determined through several rounds of parameter tests.

The training sequences (impedance and well logs) are divided into segments with a length of 51 to produce the training and prediction data set. The impedance is used as the input. Four well logs, including density (DEN), natural gamma ray (GR), compensated neutron log (CNL), and induction (CON) curves,

are used as the output. The best results are obtained by using the training parameters shown in Table 3 after a few times of parameter tests.

Figures 11, 12 depict the outcomes of the three methods employed in the analysis of the two wells, where the red and blue curves indicate the predicted and actual values, respectively. As illustrated in these figures, the performance of the LSTM-based method is suboptimal, as it fails to capture the geological changes adequately, leading to inaccurate overall predictions.

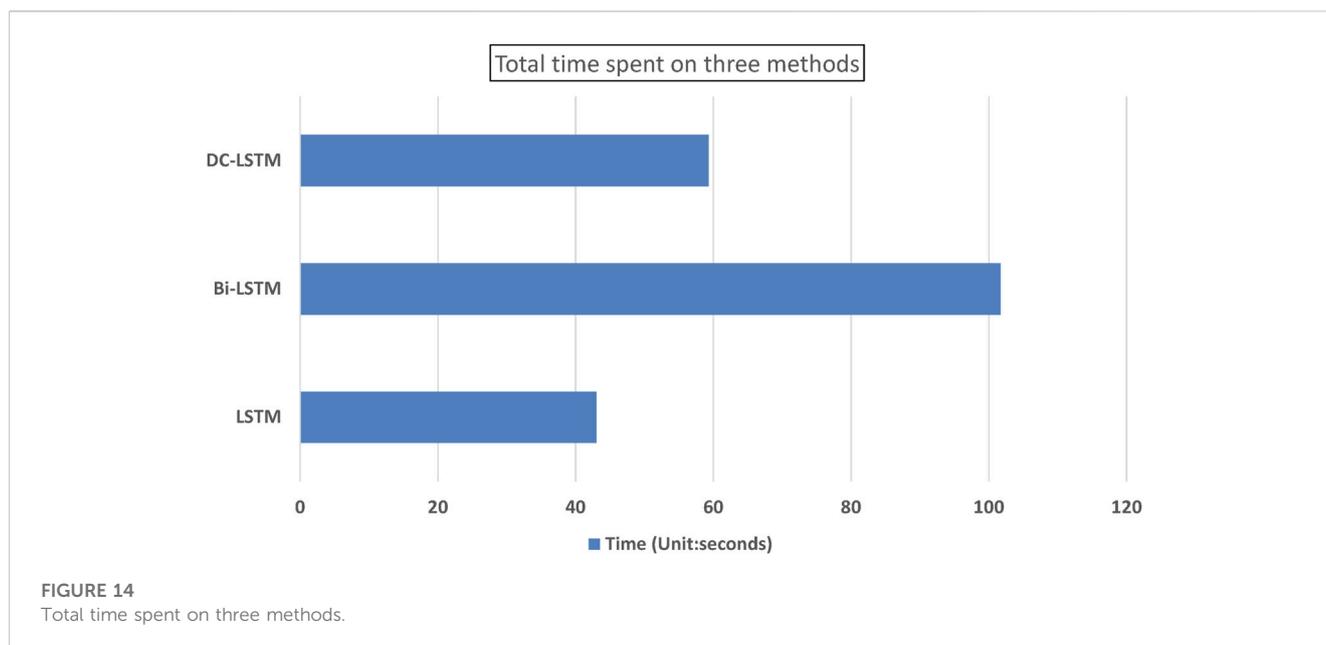


Conversely, the Bi-LSTM and DC-LSTM methods yield favorable results in both detecting the geological variations in detail and predicting the overall trend.

Table 4, 5 present the Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) values of three methods applied to two wells. Figure 13 is a bar chart illustrating the performance of these methods. The results indicate that the MAE and RMSE values of the Bi-LSTM and DC-LSTM models are significantly lower than those of the LSTM model in terms of average value. This finding implies that the predicted values of Bi-LSTM and DC-LSTM are closer to the true values than those of the LSTM model. The MAE and RMSE values of DC-LSTM are similar to those of Bi-LSTM. However,

Table 6 shows that using DC-LSTM significantly reduces the time required for model training and prediction.

In well log prediction while drilling, accuracy is a crucial reference factor. In this study, the models are evaluated based on the time spent, using the same personal computer with NVIDIA GPU (GTX 1650 with 4 GB video storage) and the same set of parameters for training and prediction. Table 6 and Figure 14 show the time spent on the models. Results reveal that the running time of the LSTM model is the shortest among the three methods, but its accuracy is the lowest. Bi-LSTM and DC-LSTM models exhibit comparable higher accuracy, with DC-LSTM being more efficient in terms of computational time.



4 Conclusion

This paper presents a novel approach for predicting well logs using artificial intelligence while drilling, utilizing seismic impedance and three related LSTM networks. The proposed approach incorporates the constraints of sedimentary formation seismic impedance at the drill bit and the changing trend of impedance through the upper and lower formations. This optimization of the prediction process results in improved computational efficiency. Through experimentation, the Bi-LSTM and DC-LSTM derivative networks are demonstrated to have higher prediction accuracy than the base LSTM. Notably, the DC-LSTM network, developed in this paper, is more efficient in reducing the number of training parameters and computational time without compromising prediction accuracy, as compared to the Bi-LSTM. Field data tests of the three networks were conducted using two wells, and the results demonstrate their effectiveness and efficiency.

The data-driven approach utilizing the three networks is effective for predicting well logs in laterally homogeneous target formations, particularly in the case of clastic sedimentation. However, if the formation exhibits significant lateral heterogeneity, such as fractures and faults near the drilling bits, the prediction accuracy may be compromised. Therefore, the challenge of making reliable predictions under such conditions with limited quality data and rapidly changing trends is an important research topic for the future.

Data availability statement

The datasets presented in this article are not readily available because Oilfield sensitive. Requests to access the datasets should be directed to yungui.xu@outlook.com.

Author contributions

HW was responsible for the preliminary literature research, the writing of the whole manuscript, and the drawing of the figure. YX and ST were responsible for the late revision and submission of the manuscript. LW, WC, and XH were responsible for providing inspiration and ideas for the manuscript.

Funding

National Natural Science Foundation of China: U20B2016. Study on high temperature high pressure seismic rock model and integrated intelligent reservoir characterization methods for a mid-deep reservoir in Yingqiong basin.

Conflict of interest

LW was employed by Shale Gas Project Management Department of CNPC Chuanqing Drilling Engineering Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Adamowski, J., and Chan, H. F. (2011). A wavelet neural network conjunction model for groundwater level forecasting. *J. Hydrology* 407 (1-4), 28–40. doi:10.1016/j.jhydrol.2011.06.013
- Ahmed Ali Zerrouki, T. A., and Baddari, K. (2014). Prediction of natural fracture porosity from well log data by means of fuzzy ranking and an artificial neural network in Hassi Messaoud oil field, Algeria. *J. Petroleum Sci. Eng.* 115, 78–89. doi:10.1016/j.petrol.2014.01.011
- Alizadeh, B., Najjari, S., and Kadkhodaie-Ilkhchi, A. (2012). Artificial neural network modeling and cluster analysis for organic facies and burial history estimation using well log data: A case study of the south pars gas field, Persian gulf, Iran. *Comput. Geosciences* 45, 261–269. doi:10.1016/j.cageo.2011.11.024
- da Silva, J. A. C. I., and Saggiaro, N. J. (2013). *Recurrent neural network based approach for solving groundwater hydrology problems*. London, United Kingdom: IntechOpen.
- Deng, H., Zhang, L., and Wang, L. (2019). Global context-dependent recurrent neural network language model with sparse feature learning. *Neural Comput. Appl.* 31 (2), 999–1011. doi:10.1007/s00521-017-3065-x
- Graves, A., and Jaitly, N. (2014). *Towards end-to-end speech recognition with recurrent neural networks*. Beijing, China: International Conference on Machine Learning.
- Graves, A., and Schmidhuber, J. (2005). *Framewise phoneme classification with bidirectional LSTM networks*. Montreal, QC: 2005 IEEE International Joint Conference on.
- Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural Comput.* 9 (8), 1735–1780. doi:10.1162/neco.1997.9.8.1735
- Hrynaskiewicz, I. (2010). A call for BMC Research Notes contributions promoting best practice in data standardization, sharing and publication. *BMC Res. notes* 3, 235. doi:10.1186/1756-0500-3-235
- Iturrarán-Viveros, U., and Parra, J. O. (2014). Artificial Neural Networks applied to estimate permeability, porosity and intrinsic attenuation using seismic attributes and well-log data. *J. Appl. Geophys.* 107, 45–54. doi:10.1016/j.jappgeo.2014.05.010
- Jin, Z. D. C. Y. M. (2018). *Synthetic well logs generation via recurrent neural networks*. Beijing, China: Petroleum Exploration and Development (4), 629–639.
- Lokeshkumar, R., Jayakumar, K., Prem, V., and Nonghuloo, M. S. (2020). Analyses and modeling of deep learning neural networks for sequence-to-sequence translation(Article). *Int. J. Adv. Sci. Technol.* 29 (5), 3152–3159.
- Long, W., Chai, D., and Aminzadeh, F. (2016). *Pseudo density log generation using artificial neural network*. Anchorage, AK: SPE Western Regional Meeting.
- Mo, X., Zhang, Q., and Li, X. (2015). *Well logging curve reconstruction based on genetic neural networks*. Jilin Changchun, China: Jilin University. College of Geoprospection Science and Technology.
- Rolon, L., Mohaghegh, S. D., Ameri, S., Gaskari, R., and McDaniel, B. (2009). Using artificial neural networks to generate synthetic well logs. *J. Nat. Gas Sci. Eng.* 1 (4-5), 118–133. doi:10.1016/j.jngse.2009.08.003
- Sahoo, S., Russo, T. A., Elliott, J., and Foster, I. (2017). Machine learning algorithms for modeling groundwater level changes in agricultural regions of the U.S. *U.S. Water Resour. Res.* 53 (5), 3878–3895. doi:10.1002/2016wr019933
- Salehi, M. M., Rahmati, M., Karimnezhad, M., and Omidvar, P. (2017). Estimation of the non records logs from existing logs using artificial neural networks. *Egypt. J. Petroleum* 26 (4), 957–968. doi:10.1016/j.ejpe.2016.11.002
- Schuster, M., and Paliwal, K. K. (1997). Bidirectional recurrent neural networks. *IEEE Trans.* 45 (11), 2673–2681. doi:10.1109/78.650093
- Shan, L., Liu, Y., Tang, M., Yang, M., and Bai, X. (2021). CNN-BiLSTM hybrid neural networks with attention mechanism for well log prediction. *J. Petroleum Sci. Eng.* 205, 108838. doi:10.1016/j.petrol.2021.108838
- Silva, A. A., Neto, I. A. L., Misságia, R. M., Ceia, M. A., Carrasquilla, A. G., and Archilha, N. L. (2015). Artificial neural networks to support petrographic classification of carbonate-siliciclastic rocks using well logs and textural information. *J. Appl. Geophys.* 117, 118–125. doi:10.1016/j.jappgeo.2015.03.027
- Silversides, K., Melkumyan, A., Wyman, D., and Hatherly, P. (2015). Automated recognition of stratigraphic marker shales from geophysical logs in iron ore deposits. *Comput. Geosciences* 77, 118–125. doi:10.1016/j.cageo.2015.02.002
- Singh, U. K. (2011). Fuzzy inference system for identification of geological stratigraphy off Prydz Bay, East Antarctica. *J. Appl. Geophys.* 75 (4), 687–698. doi:10.1016/j.jappgeo.2011.08.001
- Tamim, N., Laboureur, D. M., Hasan, A. R., and Mannan, M. S. (2019). Developing leading indicators-based decision support algorithms and probabilistic models using Bayesian network to predict kicks while drilling. *Process Saf. Environ. Prot. Trans. Institution Chem. Eng. Part B* 121, 239–246. doi:10.1016/j.psep.2018.10.021
- Wang, G., Carr, T., Ju, Y., and Li, C. (2014). Identifying organic-rich Marcellus Shale lithofacies by support vector machine classifier in the Appalachian basin. *Comput. GEOSCIENCES* 64, 52–60. doi:10.1016/j.cageo.2013.12.002
- Wang, J., Cao, J., Liu, Z., Zhou, X., and Lei, X. (2020). Method of well logging prediction prior to well drilling based on long short-term memory recurrent neural network(Article). *J. Chengdu Univ. Technol. Sci. Technol. Ed.* 47 (2), 227–236. doi:10.3969/j.issn.1671-9727.2020.02.11
- Wang, W., Qiu, Z. S., Zhong, H. Y., Huang, W. A., and Dai, W. H. (2017). Thermo-sensitive polymer nanospheres as a smart plugging agent for shale gas drilling operations. *J. Xi'an Shiyou Univ. Nat. Sci. Ed.* 32 (2), 116–125. doi:10.1007/s12182-016-0140-3
- Zhang, C. J., Chen, L., Tong, Y., and Liu, Z. (2015). *Cleaning uncertain data with a noisy crowd*. Hong Kong, China: Hong Kong University of Science and Technology.