ACS OMEGA

http://pubs.acs.org/journal/acsodf

Article

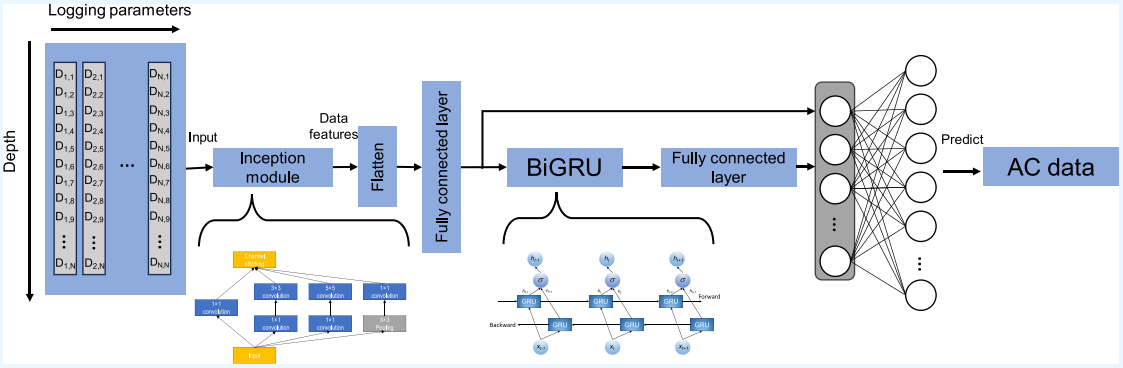
The Bidirectional Gated Recurrent Unit Network Based on the Inception Module (Inception-BiGRU) Predicts the Missing Data by Well Logging Data

Youzhuang Sun, Junhua Zhang,* Zhengjun Yu, Yongan Zhang, and Zhen Liu

Cite This: *ACS Omega* 2023, 8, 27710−27724

Read Online

ACCESS | Metrics & More | Article Recommendations



ABSTRACT: As a key bridge between logging and seismic data, acoustic (AC) logging data is of great significance for reservoir lithology, physical property analysis, and quantitative evaluation, and completing AC logging data can help to obtain high-resolution inversion profiles, which can provide a reliable basis for reservoir geological interpretation. However, in the actual mining process, the AC logging data is always missing due to instrument failure and borehole collapse in many areas, and re-logging is not only expensive but also difficult to achieve. However, the AC data can be completed by other obtained logging parameters. In this paper, a bidirectional gated recurrent unit network based on the Inception module is developed to complete the AC logging data. The Inception module extracts the logging data features and inputs the extracted logging data features into the bidirectional gated recurrent unit network, which can fully consider the characteristics of the current data and the data before and after the logging sequence to complete the missing AC logging data. Experimental results show that the hybrid model (Inception-BiGRU) has higher accuracy than traditional and widely used series forecasting models (gated recurrent unit network and long short-term memory network), and this method also provides a new idea for the completion of AC logging data.

1. INTRODUCTION

The logging data records the parameters of the geological structure, and obtaining a complete logging curve is crucial for lithology identification, reservoir analysis, and 3D geological modeling.[1] However, when obtaining logging data, due to unavoidable factors such as instrument quality limitations[2] and borehole problems,[3] some logging data may be missing, and it is impossible to measure a certain required logging curve in each well, and re-measuring these logging curves will cost a lot and is not advisable in actual production.[4] Initially, researchers used intersection plots,[5] multiple regression,[6] and other methods to predict the missing data, but because these methods require strong mathematical ability and a lot of professional knowledge, the prediction results are greatly affected by human factors and the stability is not high. Therefore, it is crucial to study an efficient and intelligent algorithm for logging curve reconstruction.

In recent years, machine learning[7] has developed rapidly, and it can intelligently and quickly reflect the complex nonlinear correlation between the original data and the target data, opening a new chapter for logging data. Srisutthiyakorn[8] used a Bayesian regularized neural network, which is more robust than other empirical methods in 2012 to predict missing AC logging data. Mo et al.[9] proposed a logging curve reconstruction technique based on the genetic neural network (GNN) in 2015 and used genetic algorithms to optimize the topology, weights, and thresholds of traditional neural networks. The results

Received: May 26, 2023
Accepted: June 23, 2023
Published: July 23, 2023

Table 1. Research Status of Missing Logging Data Prediction

year	authors	method	predicted logging data
2012	Srisutthiyakorn et al.	Bayesian regularized neural network	missing AC data
2015	Mo et al.	genetic neural network	missing AC data
2015	Ayoub et al.	grouping method data handling	missing AC data
2018	Zhang et al.	long short-term memory (LSTM) network	missing AC data
2019	Pham et al.	bidirectional convolutional long short-term memory network	missing AC data
2020	Pham et al.	bidirectional convolutional long short-term memory cascaded with fully connected neural network	missing AC data
2020	Han et al.	ensemble deep neural network	missing DEN and AC data
2021	Fan et al.	kernel ridge regression	missing AC data
2022	Cheng et al.	bidirectional long short-term memory network	missing AC data

showed that the GNN well reconstruction results were better than the traditional neural network, and the calculation error was smaller than that of the traditional neural network. Ayoub and Mohamed[10] used the grouping method data handling (GMDH) in 2015 to predict the log loss curve, and the experimental results showed that the method had good generalizability. Han et al.[11] used multiple linear regression, Bayesian learning, and traditional machine learning methods in 2020 to predict missing logging curves, but these methods often fall into local optimum. Therefore, an ensemble deep neural network was established, and experimental results showed that the proposed method was more accurate than the traditional machine learning algorithm. Kwon et al.[12] proposed a method to generate logging data using cutting-edge machine learning techniques in 2020. Unlike existing methods, it is possible to process various missing patterns with a single model and combine it with a supervised model. Experimental results show that the model can accurately generate the missing logging data and outperform the existing supervised models. Fan et al.[13] predicted the missing acoustic data in 2021 using the kernel ridge regression method. Experimental results showed that the proposed method had high prediction accuracy.

In related research, series forecasting models are widely used in logging acoustic data reconstruction and achieve good results. For example, in 2018, Zhang et al.[14] established a logging data reconstruction method using a long short-term memory (LSTM) network that can combine depth change trends and contextual information. Compared to FCNNs, the LSTM method had higher accuracy. In 2019, Pham and Whu[15] predicted missing acoustic logging data by using a bidirectional convolutional long short-term memory network, and experimental results showed that the method had high accuracy. In 2020, Pham et al.[16] used bidirectional convolutional long short-term memory cascaded with fully connected neural networks (FCNNs) to predict missing logging data using the ConvLSTM architecture to consider the depth trend and local characteristics of the logging data. This method was tested in acoustic logging prediction and could accurately predict acoustic logging. In 2022, Cheng et al.[17] proposed a recurrent neural network-based method (bidirectional long short-term memory network) to complete the missing logging curve. The experimental results showed that this method could improve the prediction accuracy. Table 1 shows the research status of prediction of missing logging data.

There are two main improvements proposed by predecessors to the sequence models. The first is to use the swarm optimization algorithm to adaptively adjust the hyperparameters of the sequence models, such as using the particle swarm algorithm or genetic algorithm to optimize the hyperparameters of the LSMT model. The second is to use variations of sequence models, such as variants of LSMT (ConvLSTM or BiLSMT).

Based on the previous research on series forecasting models, this paper innovatively combines a sequence model (bidirectional gated recurrent unit network) with a deep learning module (Inception module) to predict AC data. The Inception module is mainly used for the extraction of data features,[18] and the gated recurrent unit network has high accuracy for logging data prediction,[19] so it is input into the gated recurrent unit network for the prediction of missing acoustic curves after extracting logging data features.[20] The advantage of using the feature extraction method (Inception module) is that it can reduce data redundancy and data dimension, make the model pay more attention to valuable features, reduce the risk of overfitting, and improve the effect of the model. The disadvantage is that adding modules will lead to longer computing time and consume more time. In addition, the prediction model BiGRU in this paper has two main advantages. On the one hand, the before and after information can be

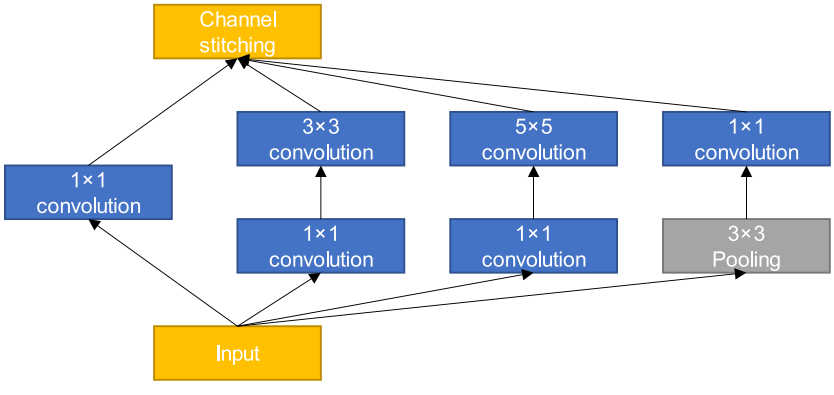


Figure 1. Inception basic structure diagram.

https://doi.org/10.1021/acsomega.3c03677
ACS Omega 2023, 8, 27710−27724

considered at the same time so that the characteristic information of sequence data can be fully captured. On the other hand, the gradient vanishing problem that often occurs in traditional models can be effectively avoided. The disadvantage is that the reverse information is introduced, and the complexity of the calculation becomes greater.

2. METHODS

2.1. Inception Module. The convolutional neural network has been widely used in many fields and can extract data features[21] layer by layer through the slipping operation of convolution kernels on the feature map. Increasing the depth and width of the network is the main way to improve network performance, but it leads to overfitting and difficult network training.

The Inception module[22] is an effective way to solve these problems, which uses convolution kernels of different sizes for the same layer feature map for feature extraction, and then passes through 1 × 1 convolution kernel, which is used for channel dimensionality reduction, and finally, the channel splicing is summarized to extract feature information. Based on the structure of the Inception module, the number of parameters remains relatively unchanged while expanding the network width. Figure 1 shows the basic structure of the Inception module.

Batch standardization is generally used between the convolutional layer and the activation function, which normalizes the data in the channel dimension, which can effectively solve the problem of gradient vanishing, reduce overfitting phenomenon, and improve training speed. The batch normalization formula is as follows:

$$\mu_B = \frac{1}{m}\sum_{i=1}^{m} x_i \tag{1}$$

$$\sigma_B^2 = \frac{1}{m}\sum_{i=1}^{m}(x_i - \mu_B)^2 \tag{2}$$

$$\widehat{x_i} = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 - \varepsilon}} \tag{3}$$

$$y_i = \gamma\widehat{x_i + \beta} \tag{4}$$

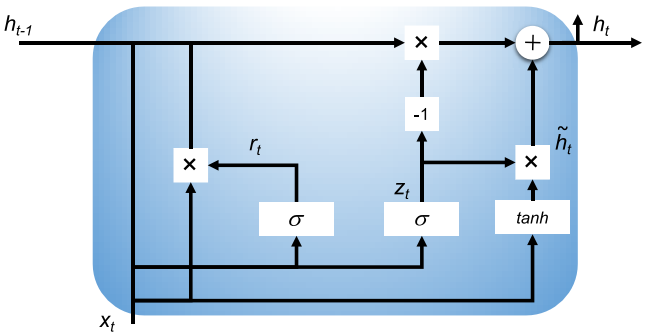


Figure 2. Structure diagram of the GRU network.

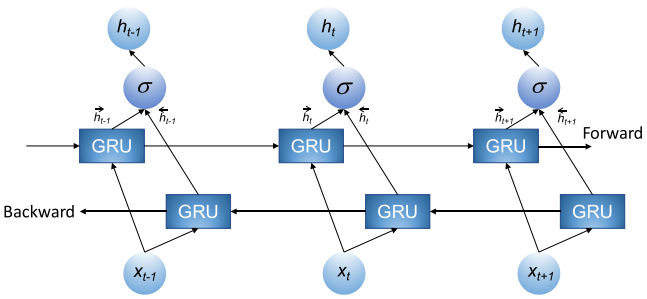


Figure 3. Bidirectional GRU network structure diagram.

Table 2. Model Parameters

layer	type	structural parameters	output size
1	input layer	(10,5)	(10,5,1)
2	(Inception module)		
	convolutional layer	(1,1,32)	(10,5,32)
	convolutional layer	(1,1,32)	(10,5,32)
		(3,3,32)	
	convolutional layer	(1,1,32)	(10,5,32)
		(5,5,32)	
	pooling layer/convolutional layer	(3,3)	(10,5,32)
		(1,1,32)	
3	fully connected layer	50	50
4	BiGRU layer	50	50
5	fully connected layer	50	50
6	jump connection layer		100
7	fully connected layer	10	10

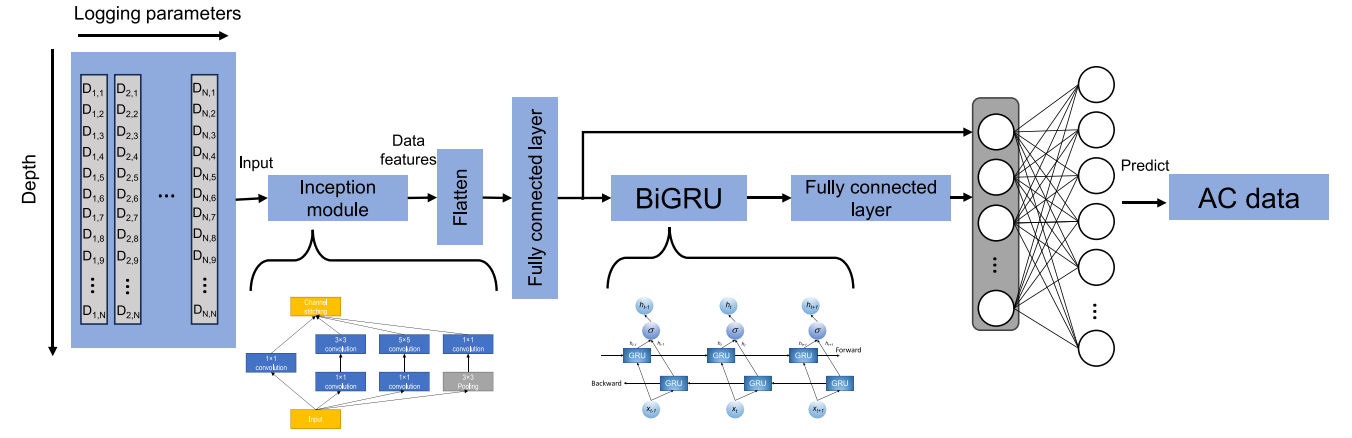


Figure 4. Inception-BiGRU network structure diagram.

https://doi.org/10.1021/acsomega.3c03677
ACS Omega 2023, 8, 27710−27724

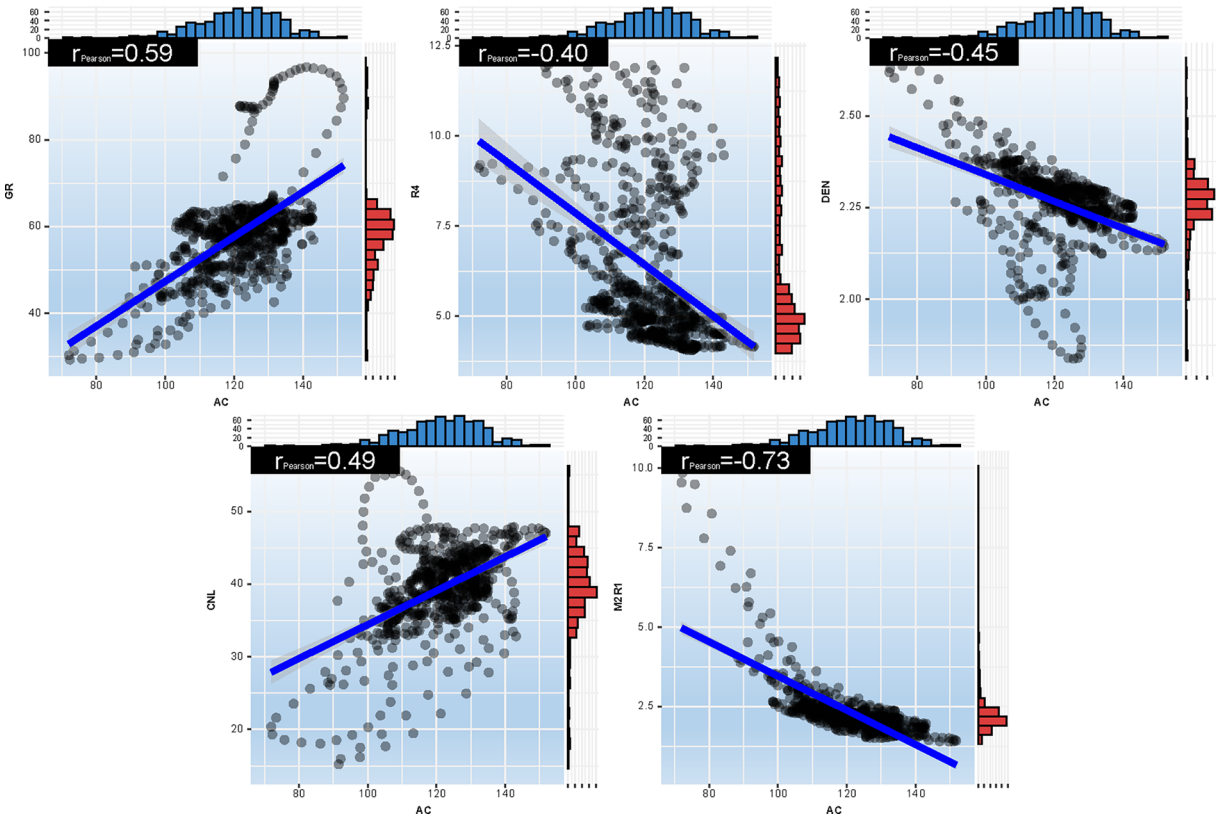


Figure 5. Well 1 logging parameters and AC scatter plot distribution plot.

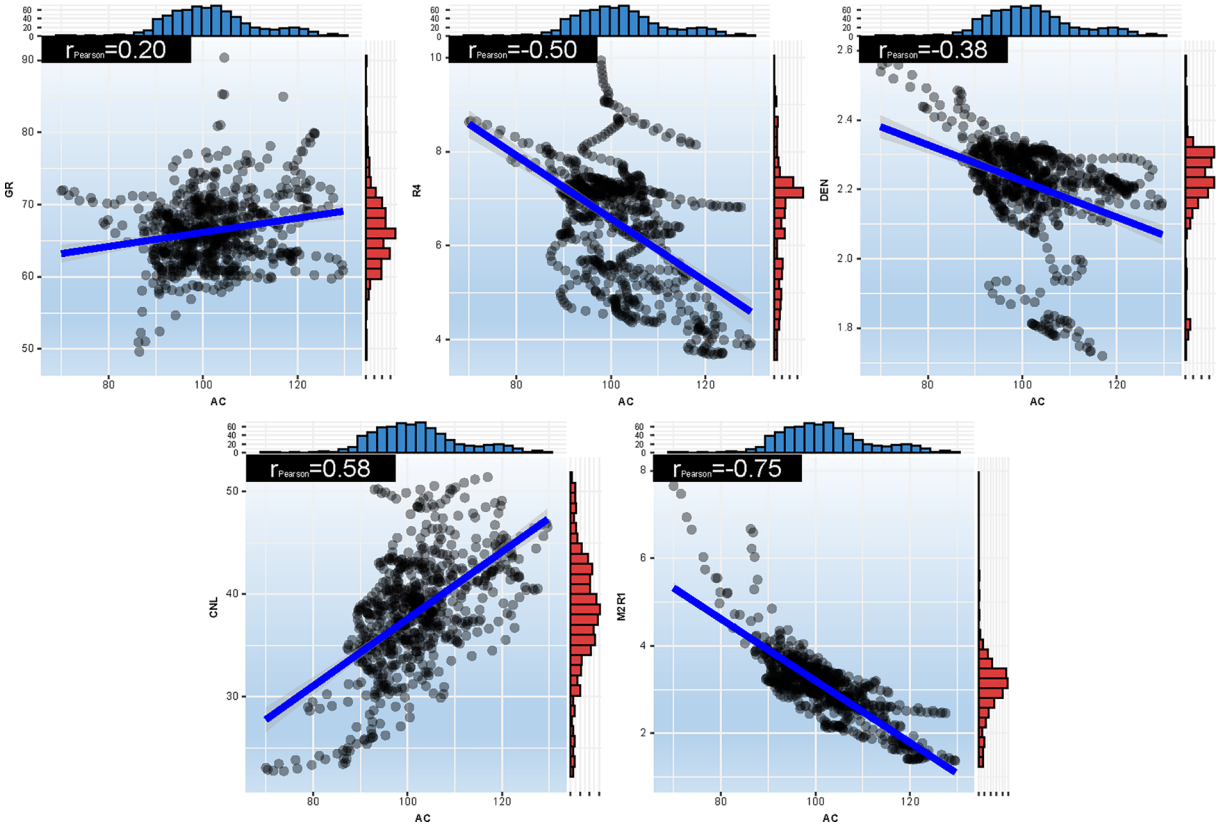


Figure 6. Well 2 logging parameters and AC scatter plot distribution chart.

where μ_B is the average of the training batches, σ_B^2 is the variance of the training batch, m is the amount of data in a single channel of the training batch, and x_i, $\hat{x_i}$, and y_i are the i-th data on the feature map before normalization, after standardization, and

https://doi.org/10.1021/acsomega.3c03677
ACS Omega 2023, 8, 27710−27724

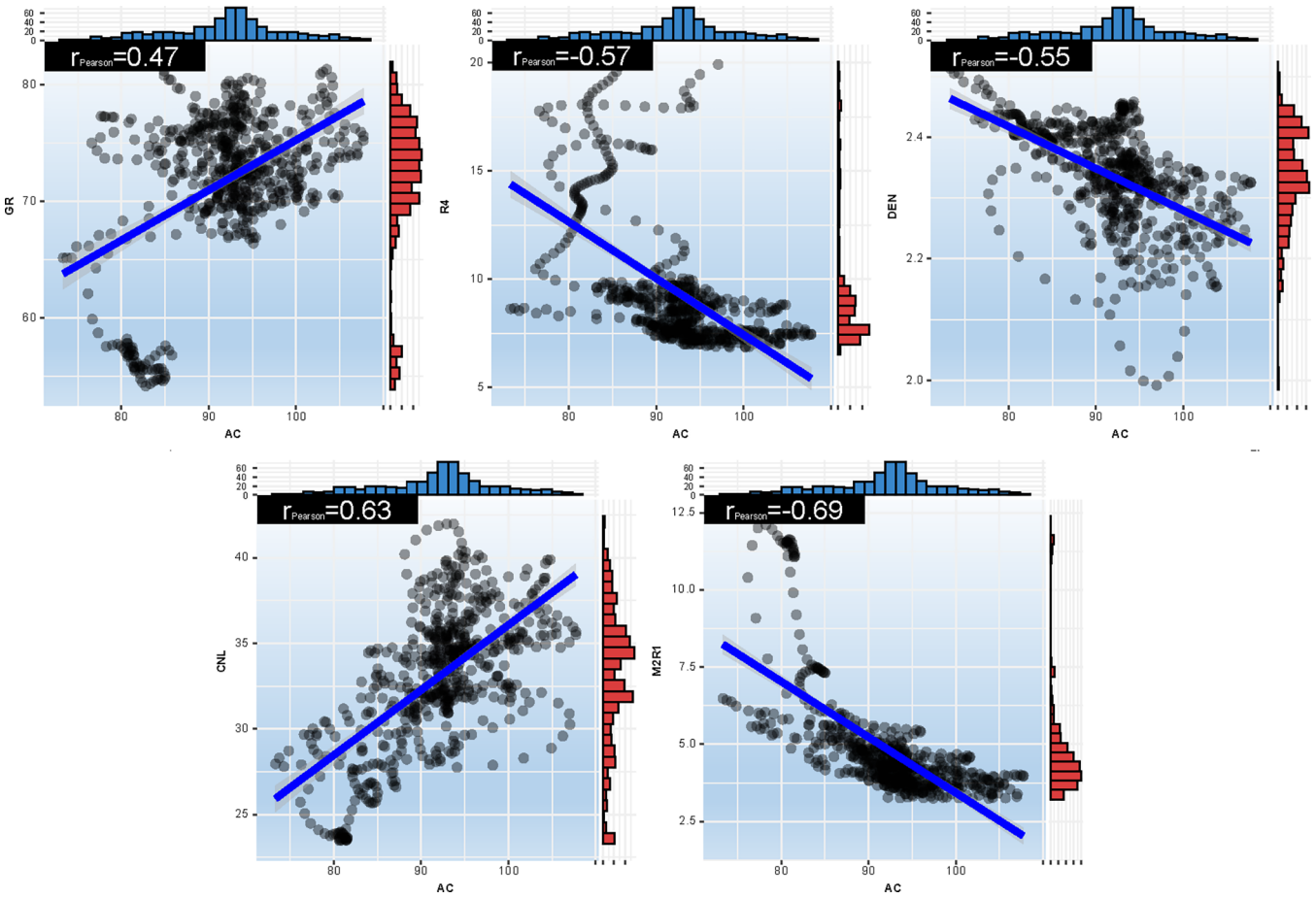


Figure 7. Well 3 logging parameters and AC scatter plot distribution diagram.

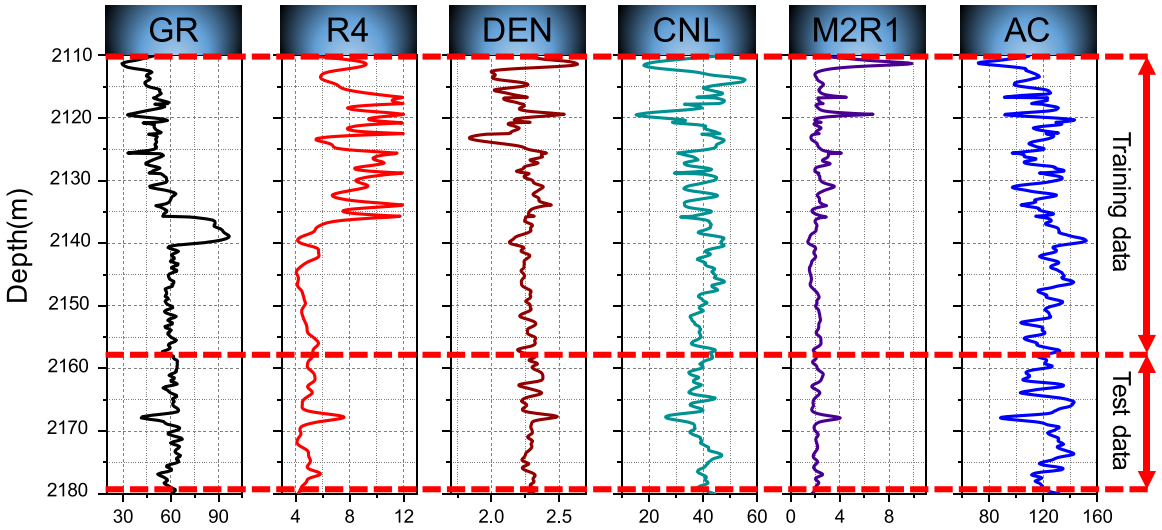


Figure 8. Line chart corresponding to the depth and value of each logging parameter of well 1.

after zooming and panning, respectively. γ and β are the control factors for spatial scaling and spatial translation, respectively; ε is the constant that can be calculated when the variance is 0.

2.2. BiGRU. GRU[23] and LSTM[24] are both developed by RNN.[25] LSTM can capture long-term dependencies and is suitable for analyzing series data. The LSTM internal unit structure is mainly composed of forget gate, input gate, output gate, and memory cell composition. The GRU simplifies on the basis of LSTM to obtain a network structure that only includes reset gate and update gate, which reduces network complexity and improves computational efficiency.[26] The internal structure of the GRU neural network is shown in Figure 2.

In Figure 2, z_t and r_t are update gate and reset gate, respectively, h_{t-1} is the input from the previous moment, and h_t is the output. The calculation process of the output value h_t of GRU is as follows.

https://doi.org/10.1021/acsomega.3c03677
ACS Omega 2023, 8, 27710−27724

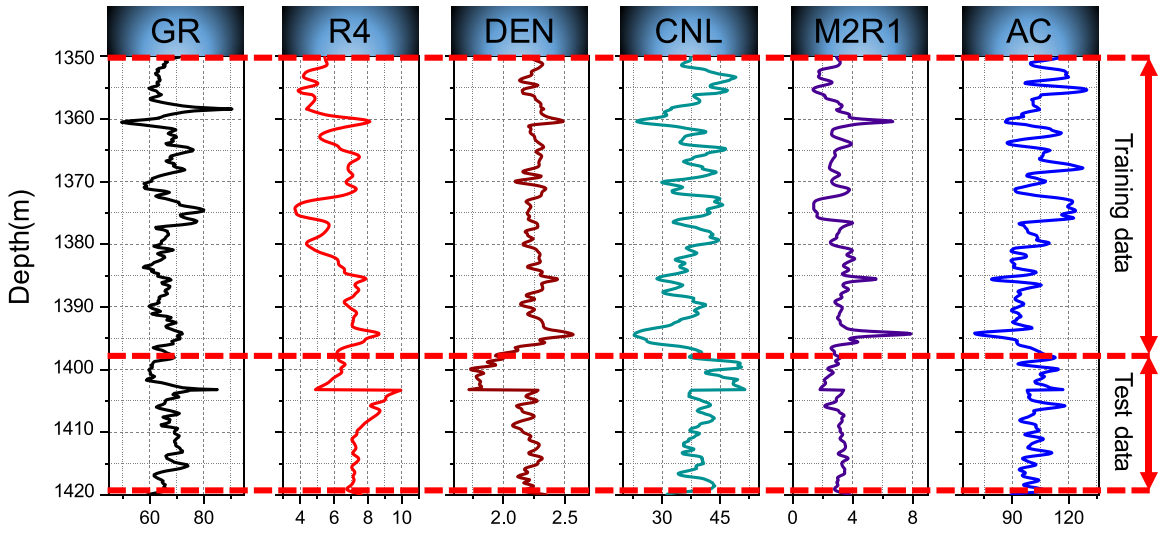


Figure 9. Line chart corresponding to the depth and value of each logging parameter in well 2.

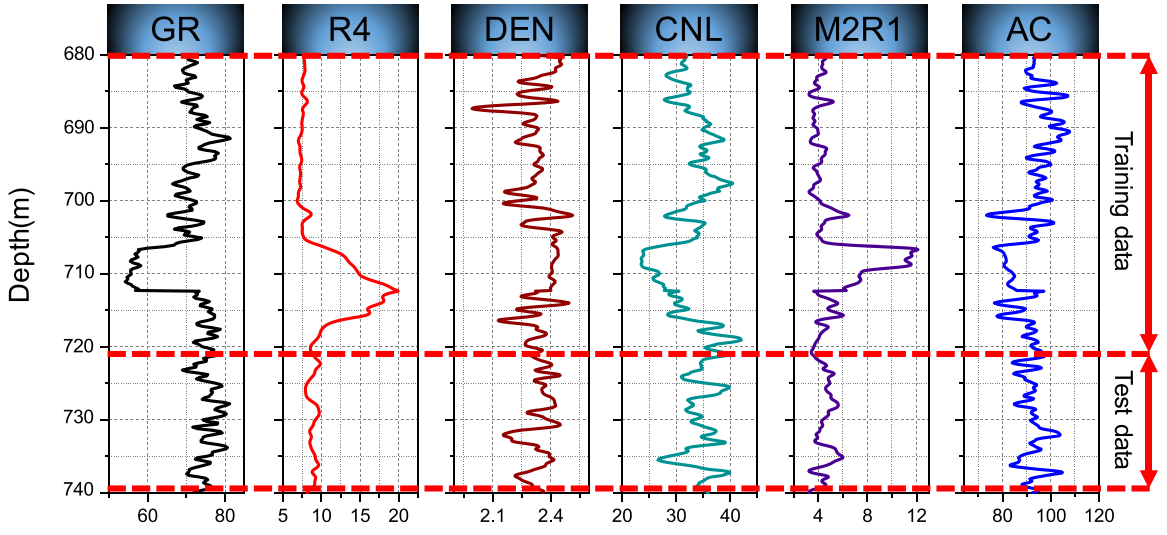


Figure 10. Line chart corresponding to the depth and value of each logging parameter in well 3.

$$z_t = \sigma(W_{xz}x_t + W_{hz}h_{t-1} + b_z) \quad (5)$$

$$r_t = \sigma(W_{xr}x_t + W_{hr}h_{t-1} + b_r) \quad (6)$$

$$\widetilde{h_t} = \tanh(W_{xh}x_t + r_t \cdot h_{t-1}W_{hh} + b_h) \quad (7)$$

$$h_t = (1 - z_t)\cdot\widetilde{h_t} + z_t \cdot h_{t-1} \quad (8)$$

where W is the weight matrix between the update gate and the reset gate, σ is the sigmoid activation function, and r_t is the reset gate and its value is [0,1]. When $r_t = 0$, it indicates that all information passed at the previous moment has been forgotten; b is the bias vector.

In the series, to fully consider the information law of the positive and reverse of the data, BiGRU[27] is composed of a bidirectional recurrent neural network with forward propagation and reverse propagation (Figure 3). Compared with the unidirectional GRU, BiGRU considers the change law of the data, and there is no connection between the forward hidden layer and the backward hidden layer, so BiGRU can better mine the series characteristics of the data. The network structure of BiGRU is expressed as follows.

$$\vec{h}_t = GRU(x_t, \vec{h}_{t-1}) \quad (9)$$

$$\overleftarrow{h}_t = GRU(x_t, \overleftarrow{h}_{t-1}) \quad (10)$$

$$h_t = W_{\vec{h}_t}\vec{h}_t + W_{\overleftarrow{h}_t}\overleftarrow{h}_t + b_t \quad (11)$$

where GRU is a traditional GRU network computing process, $W_{\vec{h}_t}$ and $\vec{h}_t$ are the states and weights of the forward hidden layer at the moment, respectively, $\overleftarrow{h}_t$ and $W_{\overleftarrow{h}_t}$ are the backward hidden layer states and weights at the moment, respectively, and b_t is the offset of the hidden layer state at time t.

2.3. Inception-BiGRU. In this paper, the Inception-BiGRU[28] module is built, as shown in Figure 4. First, the logging data is reconstructed, and the two-dimensional data obtained by the reconstruction is used as the input of the network model. In the figure, $D_{1,1}$ in the input part represents the first logging data of the first logging parameter, and $D_{N,N}$ represent the N-th logging data of the N-th logging data. First, the Inception module is used to extract multi-scale features from the input samples, and the Inception module can adaptively extract deep features from the input samples because it contains

https://doi.org/10.1021/acsomega.3c03677
ACS Omega 2023, 8, 27710−27724

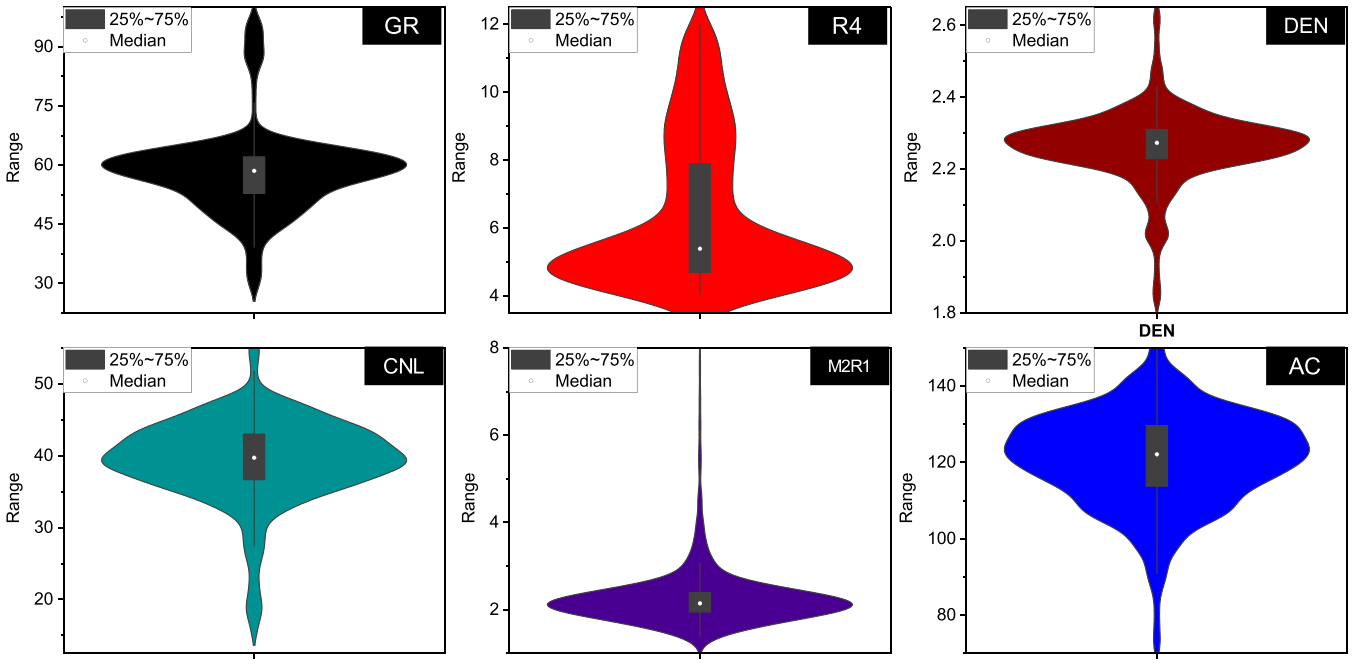


Figure 11. Violin diagram of each logging parameter in well 1.

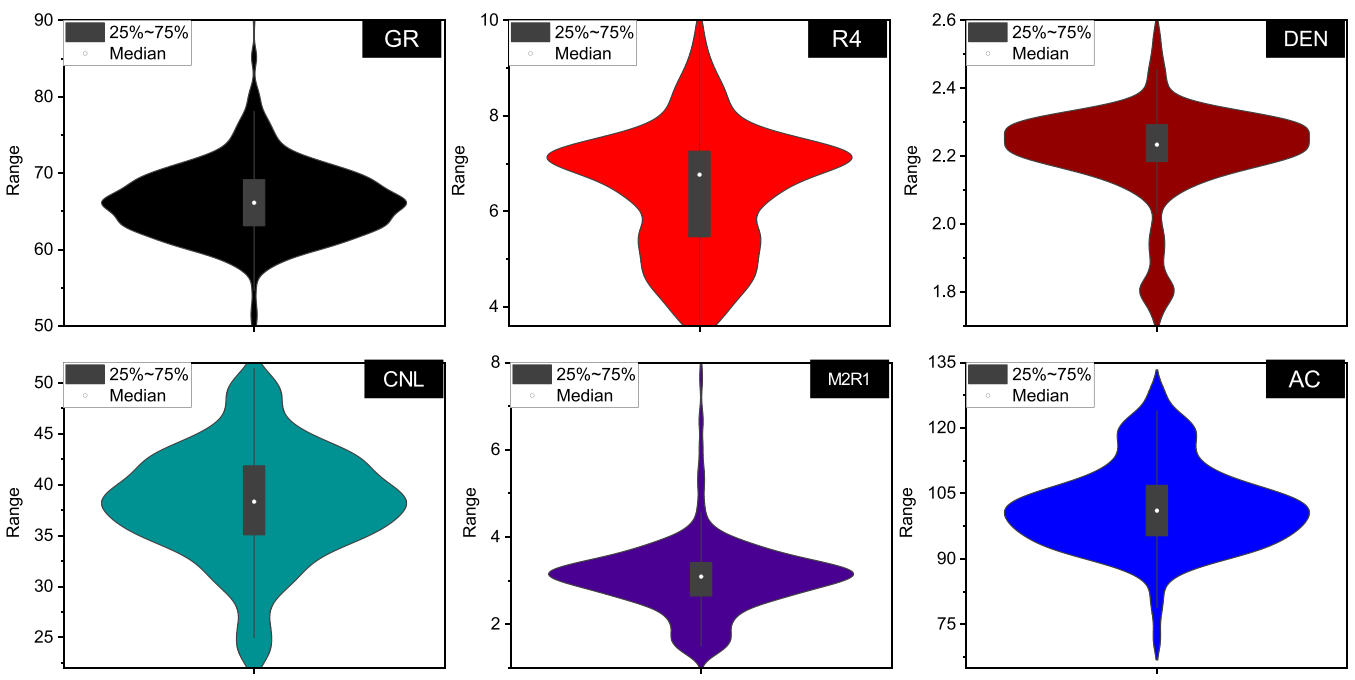


Figure 12. Violin diagram of each logging parameter of well 2.

multiple convolution kernels of different scales. On the one hand, there is no need to extract artificial features, avoiding the dependence on expert knowledge and experience. On the other hand, the convolution check of different scales is used to extract features from the input, which can maximize the extraction of features in the logging data. Then, the features extracted by the Inception module are input into BiGRU to further learn the sequence dependencies in the logging data.

In the Inception-BiGRU model proposed in this paper, the input dimension is 10 × 5, where 10 represents the number of sample points and 5 represents the number of logging parameters. There are a total of 4 parallel channels in the Inception module, which perform feature extraction separately, and the structural parameters of the convolutional layer (1,1,32)/(3,3,32) mean the following: Data first passes through the convolution kernel of 1 × 1 (the number of channels is 32) and then passes serial connection 3 × 3 convolution kernels (the number of channels is 32). The number of neurons in BiGRU is set to 50. The specific model parameters are shown in Table 2. In the table, the output dimensions (10,5,1) refer to the height, width, and depth of the data, respectively. The Adam[29] optimizer is used to optimize the cross-entropy loss function with a learning rate of 0.01, and Dropout[30] is used to suppress overfitting.

https://doi.org/10.1021/acsomega.3c03677
ACS Omega 2023, 8, 27710−27724

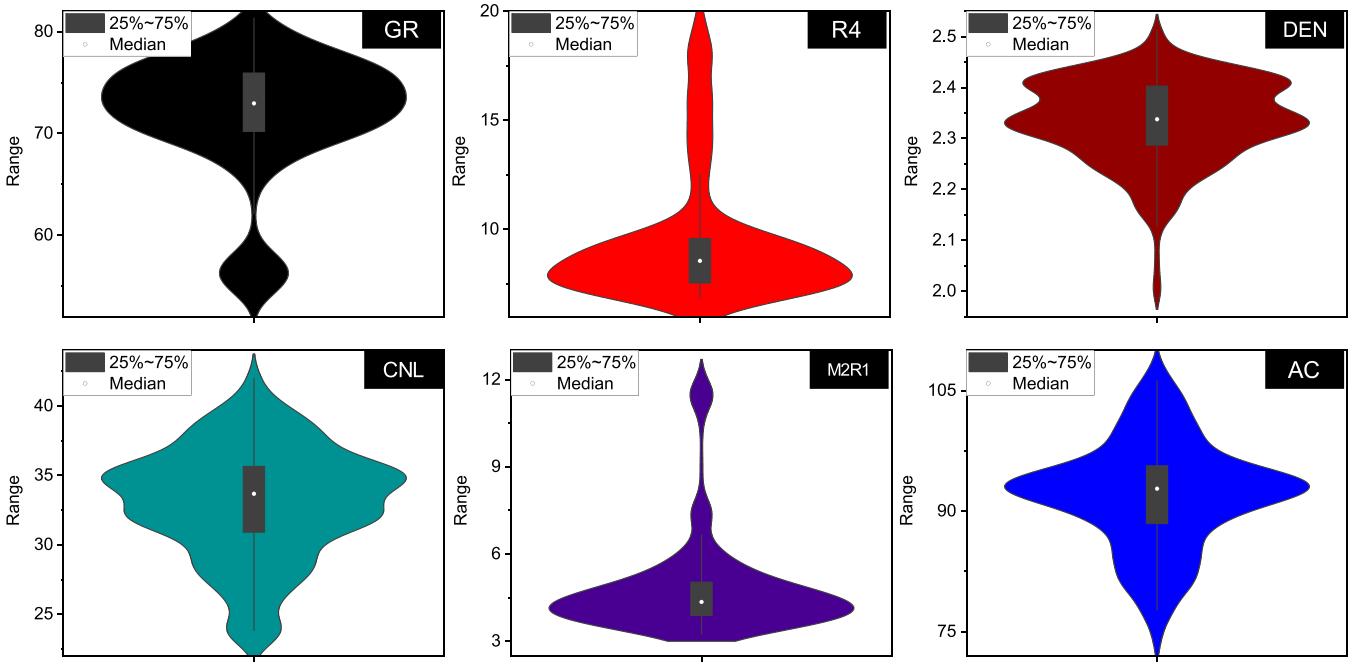


Figure 13. Violin diagram of each logging parameter of well 3.

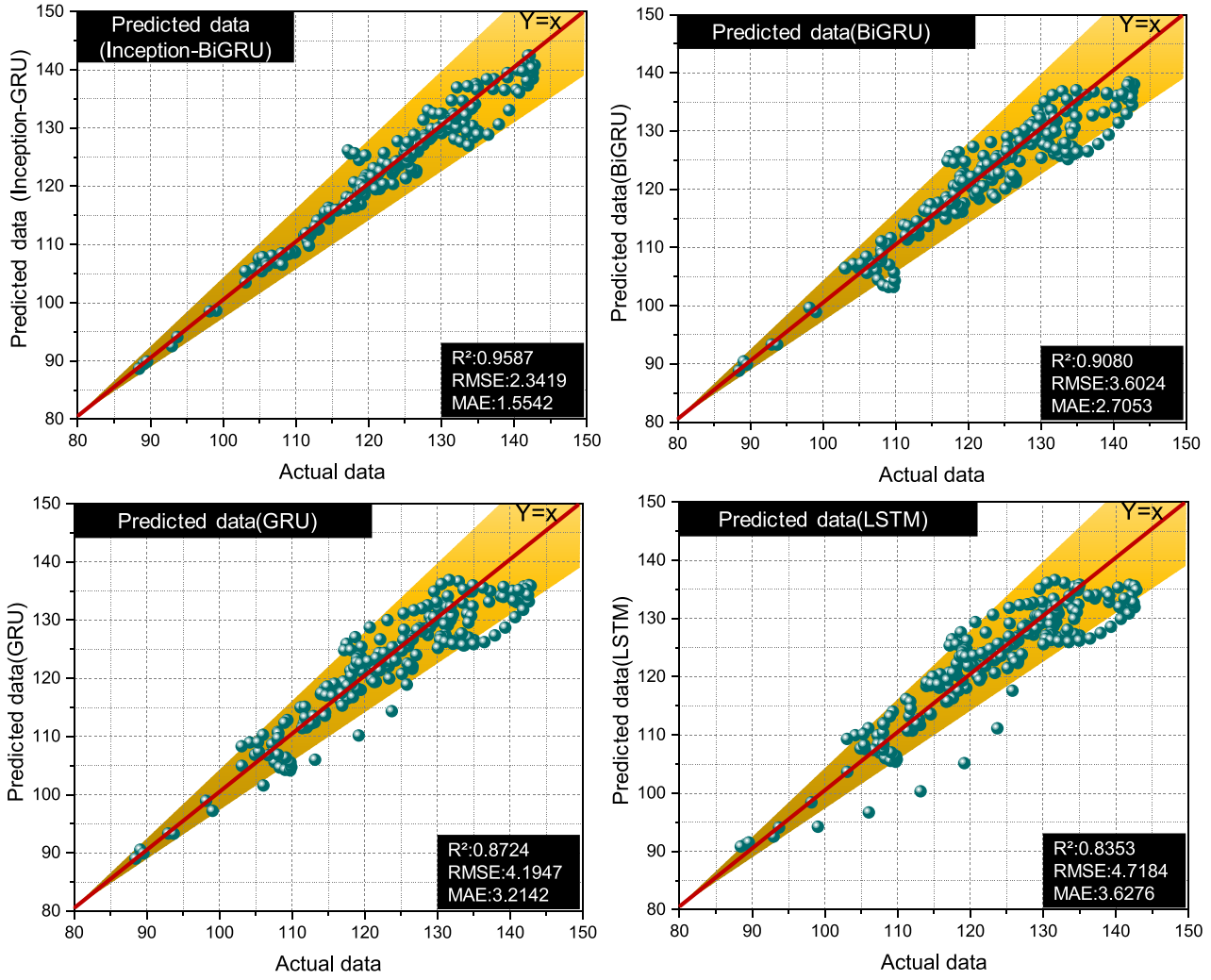


Figure 14. Scatter distribution of forecast data and actual data in well 1.

https://doi.org/10.1021/acsomega.3c03677
ACS Omega 2023, 8, 27710−27724

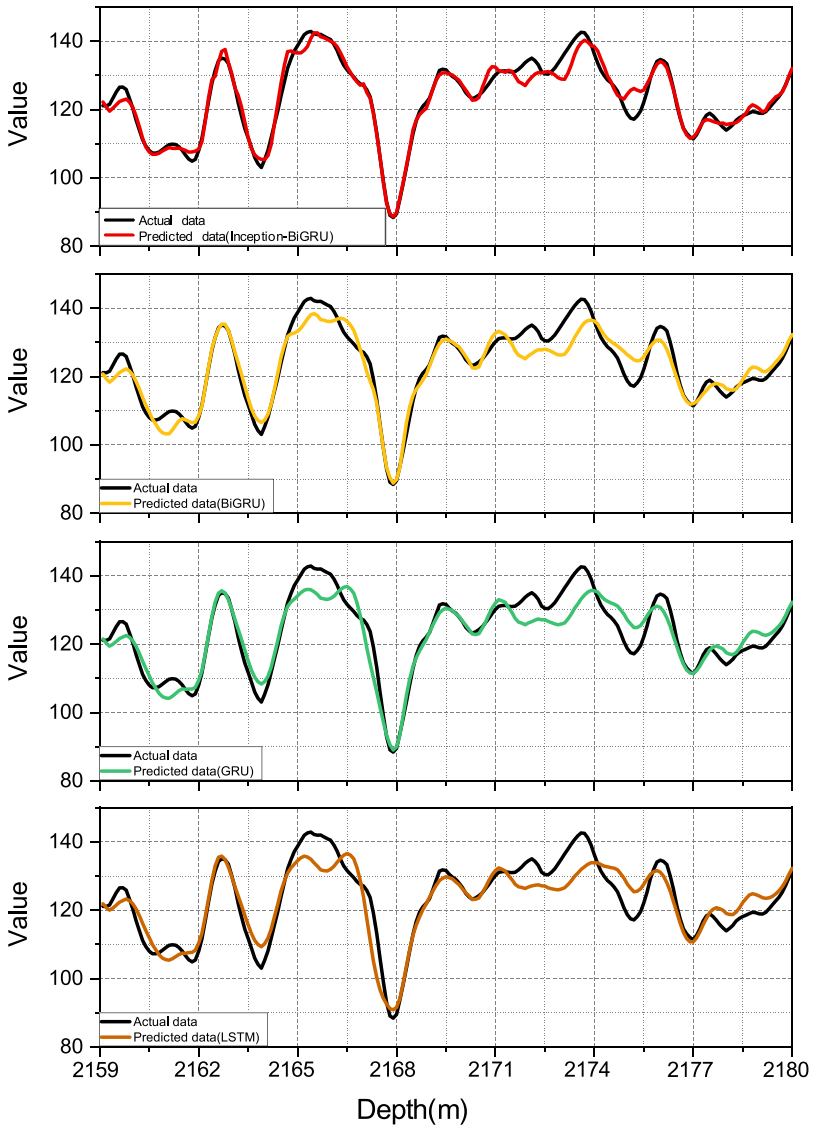


Figure 15. Line chart of forecast data and actual data for well 1.

The prediction process of missing AC logging data proposed in this paper is mainly divided into the following steps:

Table 3. Well 1 Model Test Results

	R^2	RMSE	MAE	rank
Inception-BiGRU	0.9587	2.3419	1.5542	1
BiGRU	0.908	3.6024	2.7053	2
GRU	0.8724	4.1947	3.2142	3
LSTM	0.8353	4.7184	3.6276	4

(1) The Z-score standardizes the collected logging data and divides the training and test sets according to 7:3 ratio.

(2) Initialize the network parameters, train the network model using the training set, extract feature information, and update the network parameters through back-propagation. Use the cross-entropy function as the loss function for model training. Its expression is shown in the following equation, and the Adam optimization algorithm is used to update the parameters.

$$L = -\frac{1}{M}\sum_{i=1}^{M}[y_i \ln \tilde{y}_t + (1-y_i)\ln(1-\tilde{y}_t)] \quad (12)$$

where y_i and $\tilde{y}_i$ are the actual probability distribution and predicted probability distribution of the i-th sample, respectively, M is the sample size, and L is the loss value.

(3) Determine whether the network training number k reaches the number of iterations N. If so, stop the training and save the model; otherwise, continue training. This article sets the number of iterations to 100.

(4) Input the test set data into the trained network model to predict AC data and output the prediction results.

2.4. Evaluation Indicators. In this paper, the R^2,[31] RMSE,[32] and MAE[33] evaluation indicators are used to evaluate the prediction results. R^2 shows how well the data fit the regression model, and a higher R^2 indicates that the model explains more variability. The root mean square error (RMSE) is the square root of the square of the deviation of the predicted value from the true value to the ratio of n number of observations. The smaller the RMSE of the model is, the higher the accuracy of the model is. The mean absolute error difference

https://doi.org/10.1021/acsomega.3c03677
ACS Omega 2023, 8, 27710−27724

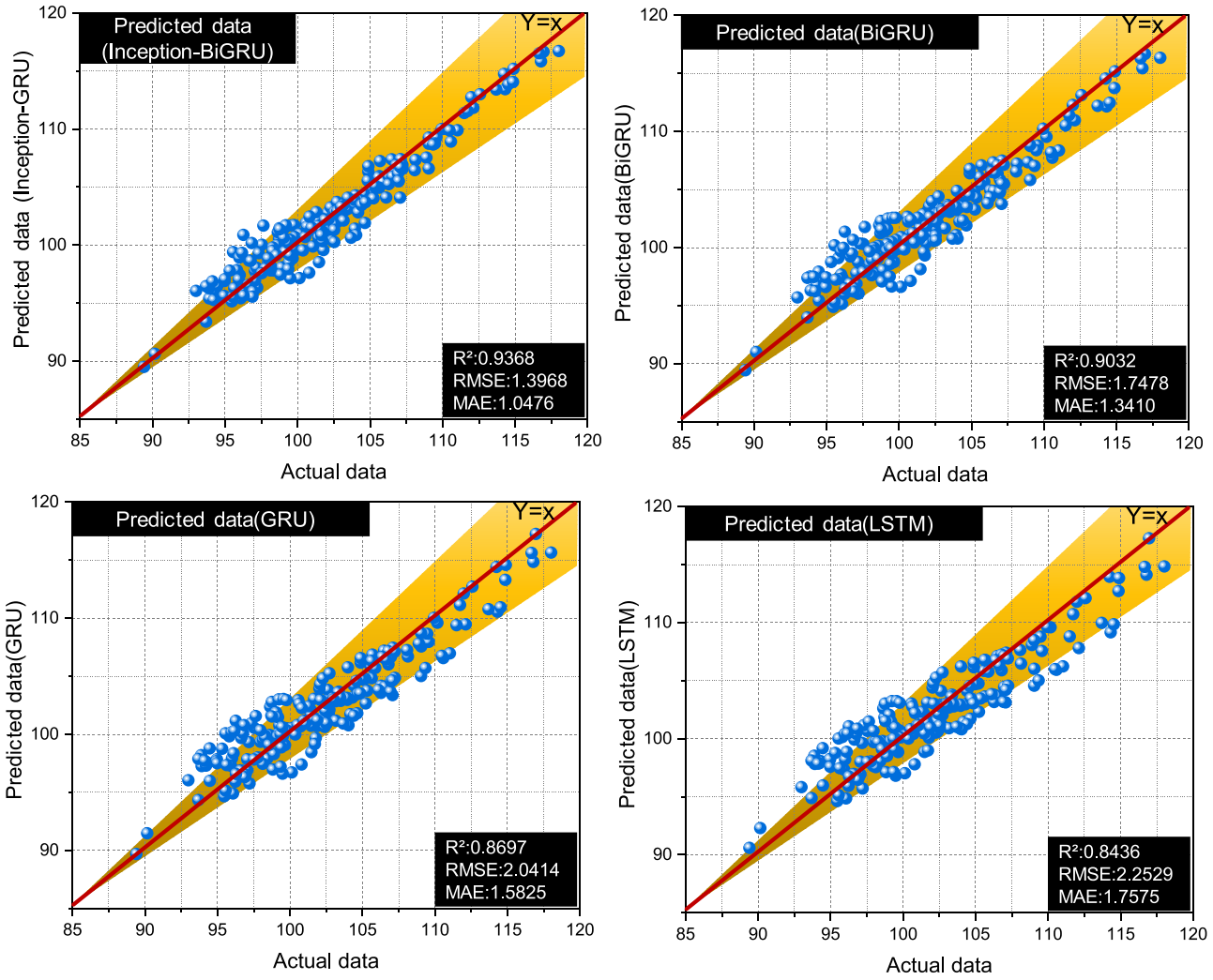


Figure 16. Scatter distribution of forecast data and actual data in well 2.

(MAE) is the average of the absolute error, which can better reflect the actual situation of the predicted value error. The smaller the MAE is, the smaller the error is.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2} \tag{13}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_t - y_i| \tag{14}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_t)^2} \tag{15}$$

where n is the number of samples, y_i is the true value corresponding to the i-th sample, $\hat{y}_i$ is the predicted value corresponding to the i-th sample, and $\overline{y}$ is the average of the sample data.

3. DATA

The data used in this article is from the Western China exploration area. The well logging data collected are AC (acoustic), GR (gamma), R4 (resistivity), DEN (density), CNL (neutron), and M2R1 (high-resolution array sensing). Three wells' data is selected for model testing. Well 1 data is selected at depths ranging from 2110 to 2180 m. The sampling interval is 0.1 m, and it contains 700 sample sites. Well 2 data is selected at a depth of 1350−1420 m, and the sampling interval is 0.1 m and it contains 700 sample sites. Well 3 data is selected at a depth of 680−740 m, and the sampling interval is 0.1 m and it contains 600 sample sites.

The training and test sets are selected by 7:3 ratio.[34] Well 1 data is trained at depths ranging from 2110 to 2159 m, with a total of 490 sample points, and the test depth is 2159 to 2180 m, with a total of 210 sample points. Well 2 data is trained at depths ranging from 1350 to 1399 m, with a total of 490 sample points. The test depth ranges from 1399 to 1420 m with a total of 210 sample points. Well 3 data is trained at depths ranging from 680 to 722 m, with a total of 420 sample points. The test depth ranges from 722 to 740 m, with a total of 180 sample points.

The purpose of this paper is to test the effect of each model, so it is assumed that the AC curve test segment is unknown (because in actual oilfield production, AC data is easy to be missing, this paper proposes machine learning models to complete AC data) and needs to be reconstructed. The input data of machine learning is GR, R4, DEN, CNL, and M2R1, and the output data is AC data.

Figures 5−7 show the scattered distribution of the three well logging parameters AC, where r_{Pearson}[35] is the correlation coefficient of the two parameters, and the magnitude of the

https://doi.org/10.1021/acsomega.3c03677
ACS Omega 2023, 8, 27710−27724

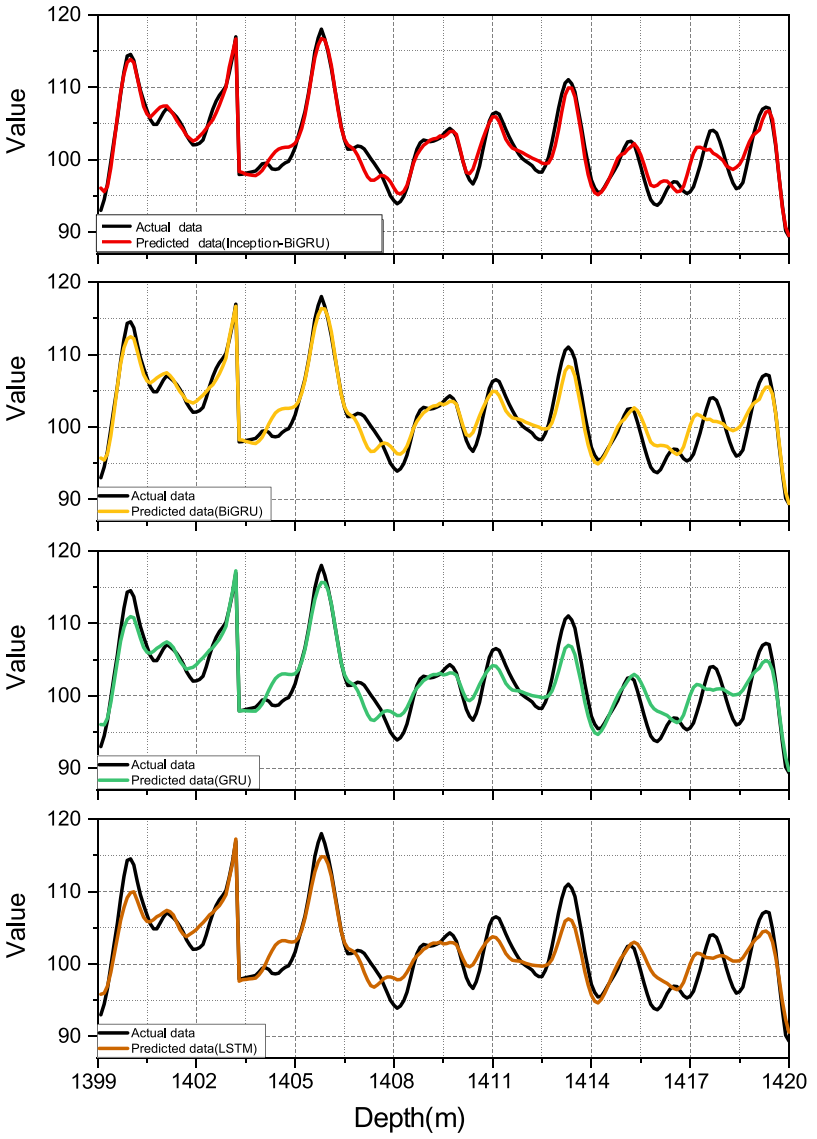


Figure 17. Line chart of forecast data and actual data for well 2.

Table 4. Well 2 Model Test Results

	R^2	RMSE	MAE	rank
Inception-BiGRU	0.9368	1.3968	1.0476	1
BiGRU	0.9032	1.7478	1.341	2
GRU	0.8697	2.0414	1.5825	3
LSTM	0.8436	2.2529	1.7575	4

absolute value reflects the magnitude of the correlation, where the larger the absolute value is, the greater the correlation is.

Figure 5 shows the scatter distribution of well 1, from which it can be seen that the AC curve was positively correlated with GR and CNL, with correlations of 0.59 and 0.49, respectively. AC has a negative correlation with R4, DEN, and M2R1, with AC having the best correlation with M2R1 at −0.73. However, if only M2R1 is used to linearly fit the AC curve, a high accuracy cannot be achieved. Therefore, this text introduces machine learning algorithms to fit the nonlinear relationship between individual logging parameters. Figure 6 shows the scatter distribution of well 2, and it can be seen from the figure that the AC curve is also the best correlation with M2R1 and the absolute value reaches 0.75. Figure 7 shows the scatter distribution of well 3, from which it can be seen that the AC curve correlates best with M2R1, with an absolute value of 0.69. The correlation between the AC curve and GR was the worst at 0.47.

Figure 8 is a line chart of each logging parameter of well 1, in which the data used for training is 2110 to 2159 m, because the log response of 2110 to 2140 m in this section is violent, but the response from 2140−2159 m logging is very stable, which can well reflect the situation of reservoirs or non-reservoirs. The test well section was selected from 2159 to 2180 m, and there was an AC curve for the stationary section and an AC curve reflecting the vigorous section. Well 2 was selected for a section for a similar reason to well 1 (Figure 9). Well 3 selects well sections from 680 to 722 m, of which 690−712 m are reservoir sections, so the logging response is violent, which is used as training data and training noise data, because the test well section is relatively stable, which is used to evaluate the noise immunity of BiGRU proposed in this paper (Figure 10). Figures 11−13 are violin diagrams of well 1, well 2, and well 3, respectively, and the violin diagram shows the distribution of data, from which it can be seen that the logging parameters basically conform to the normal distribution.

https://doi.org/10.1021/acsomega.3c03677
ACS Omega 2023, 8, 27710−27724

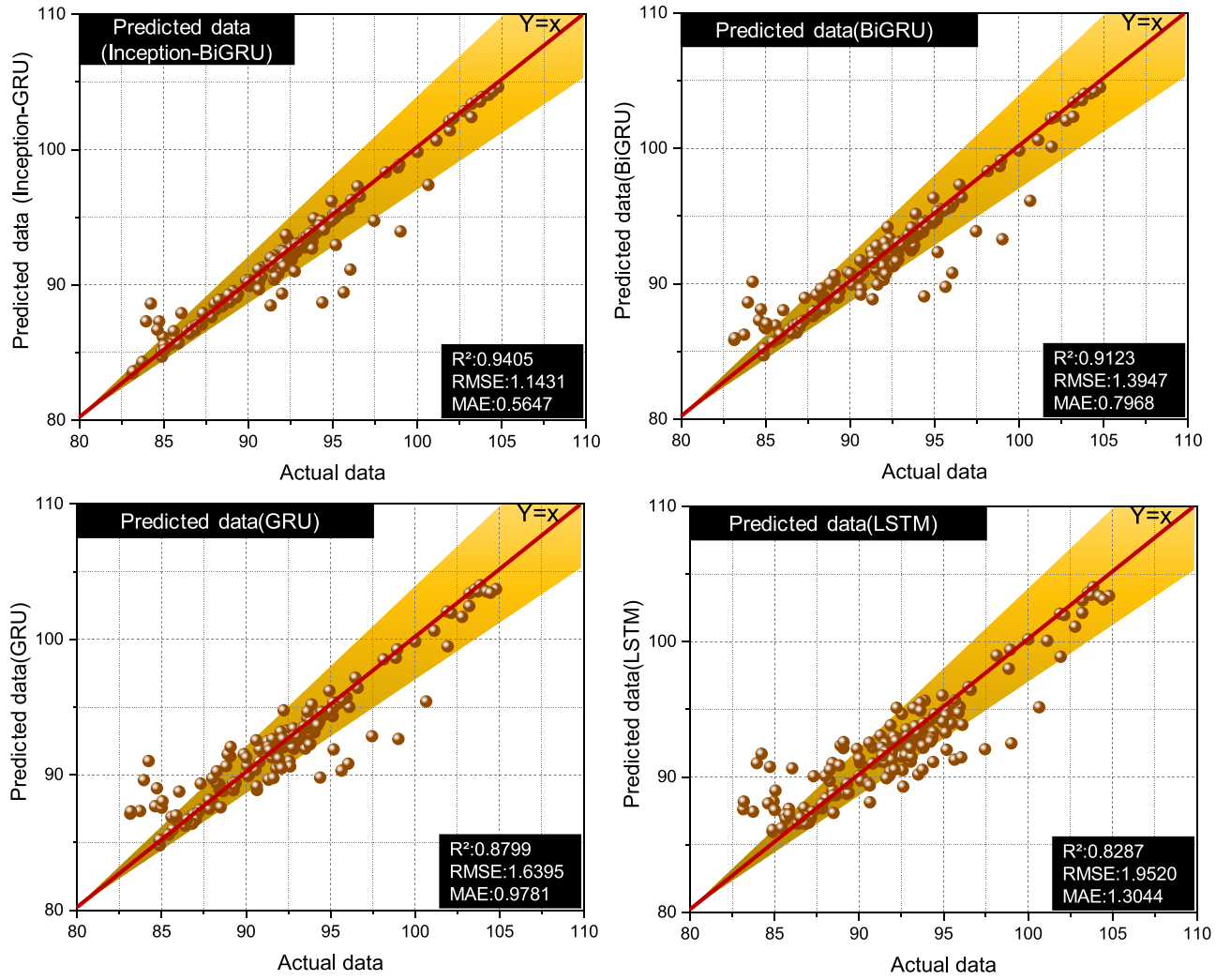


Figure 18. Scatter distribution of forecast data and actual data in well 3.

4. RESULTS AND DISCUSSION

Figures 14 and 15 and Table 3 show the predicted results in well 1. Figure 14 is a scatterplot of predicted values and actual values, and it can be seen from the plot that the best effect is Inception-BiGRU, with an RMSE of only 2.3419 and R^2 reaching 0.9587. Compared with the BiGRU model (RMSE = 3.6024), the accuracy has been improved. The results show that the data features extracted by the Inception module proposed in this paper make the error of the model lower and the generalization better. From the figure, we can also clearly see that the bidirectional gated recurrent unit network (BiGRU) proposed in this paper has higher accuracy and better model effect than the traditional gated recurrent unit network. This result shows that the BiGRU layer can learn forward data and backward data compared to GRU. In the formation, the logging parameters have good continuity, especially in the field of curve reconstruction. Through the adjustment of the bidirectional layer, the forward and backward formation logging data is learned to make the model more accurate. The well sections used for testing range from 2159 to 2180 m. It can be seen from the figure that there is a sharp jitter in the logging curve at about 2166 m, in which LSTM and the traditional GRU network fail to identify the jitter in this area. However, through the bidirectionally regulated GRU network, by learning the forward data and backward data, the jitter situation of about 2166 m can be basically identified.

Figures 16 and 17 and Table 4 show the predicted results in well 2. The model quality and accuracy are sorted from the highest to lowest: Inception-BiGRU (R^2 = 0.9368, RMSE = 1.3968, and MAE = 1.0476), BiGRU (R^2 = 0.9032, RMSE = 1.7478, and MAE = 1.3410), GRU (R^2 = 0.8697, RMSE = 2.0414, and MAE = 1.5825), and LSTM (R^2 = 0.8436, RMSE = 2.2529, and MAE = 1.7575). Figure 17 shows the line chart of the predicted data and actual data of the test section of well 2. As can be seen from the figure, the Inception-BiGRU prediction results are the most consistent with the real results.

Figures 18 and 19 and Table 5 show the predicted results in well 3. The best model is Inception-BiGRU. From the tests of these three wells, it can be seen that the accuracy of the model and optimization method proposed in this paper has been improved compared with the traditional model. In Figure 19, the LSTM network could not identify mutations in the logging curve of small stratigraphic sections from 725 to 731 m. The BiGRU network makes predictions more accurate by learning from forward and backward data. In addition, the data features extracted by the Inception module also make the BiGRU network have higher prediction accuracy and smaller model errors.

https://doi.org/10.1021/acsomega.3c03677
ACS Omega 2023, 8, 27710−27724

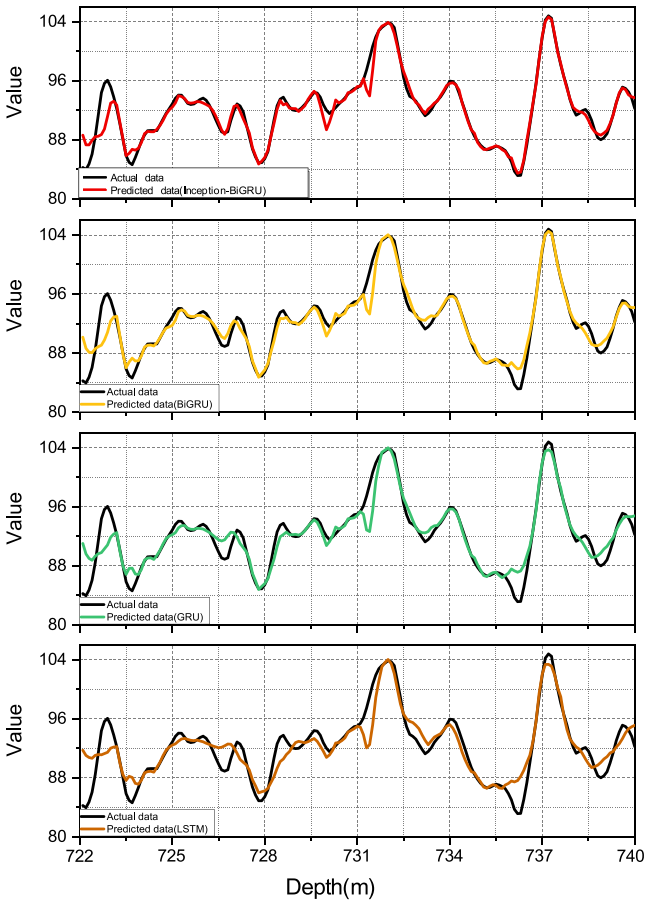


Figure 19. Line chart of forecast data and actual data in well 3.

Table 5. Well 3 Model Test Results

	R^2	RMSE	MAE	Rank
Inception-BiGRU	0.9405	1.1431	0.5647	1
BiGRU	0.9123	1.3947	0.7968	2
GRU	0.8799	1.6395	0.9781	3
LSTM	0.8287	1.952	1.3044	4

Figure 20 shows the evaluation index of the model prediction results, where panel (a) is the model evaluation index of well 1, panel (b) is the model evaluation index of well 2, and panel (c) is the model evaluation index of well 3. The larger the R^2 is, the better the model is. The smaller the RMSE and MAE are, the smaller the model error is and the higher the accuracy is. From the figure, it can be seen that the Inception-BiGRU model has the smallest error and the highest accuracy among the four sequence prediction models.

5. CONCLUSIONS

Logging curve interpretation can obtain stratigraphic parameters such as lithology, porosity, and permeability to locate the oil layer, but in practical applications, AC logging data is often missing due to logging instrument failure, and re-logging is not only expensive but difficult to implement. In this paper, the logging data (GR, R4, DEN, CNL, and M2R1) is used to predict the missing AC data. Based on the previous research on AC logging data prediction using sequence models, this paper innovatively proposes BiGRU networks based on the Inception module to predict AC logging data. This paper comes to the following conclusions:

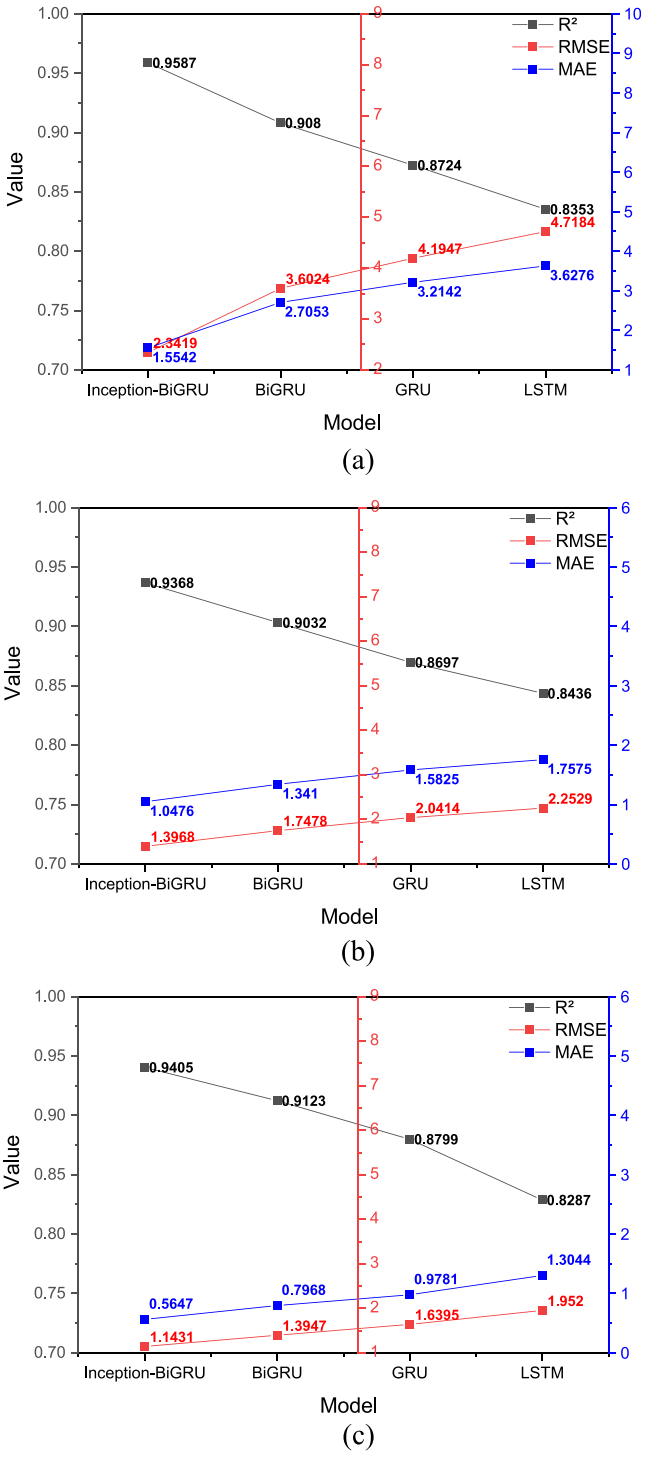


Figure 20. Model prediction result evaluation index, where panel (a) is the model evaluation index of well 1, panel (b) is the model evaluation index of well 2, and panel (c) is the model evaluation index of well 3.

This paper compares the existing sequence prediction models (LSTM, GRU, and BiGRU), and in the prediction test experiment, the BiGRU model has higher accuracy, followed by GRU and LSTM. This also shows that the bidirectional mechanism of the BiGRU model can effectively improve the prediction accuracy of the GRU model.

In the AC data prediction experiment on three wells, the accuracy of the Inception-BiGRU model is higher than

https://doi.org/10.1021/acsomega.3c03677
ACS Omega 2023, 8, 27710−27724

that of the BiGRU model. This shows that the Inception module, as a module in deep learning, can effectively extract data features, and the prediction accuracy of the model is improved after passing the data features to the BiGRU network. This also provides a new idea for future research, that is, combining deep learning modules with sequence models to improve the accuracy of prediction.

In future research, this paper has two recommendations. The first suggestion is about the study of missing log curve prediction. In actual production, there may also be missing logging data such as density, neutron, and gamma, so it is recommended that the absence of other logging parameters can be studied in the future. The second suggestion is about model studies. The model proposed in this paper innovatively combines the deep learning module and the sequence prediction models to predict AC data and has a relatively good prediction effect. In the future, deep learning modules and sequence models are constantly updated, and this paper proposes to replace the proposed modules with efficient deep learning modules and sequence models that appear in the future.

AUTHOR INFORMATION

Corresponding Author

Junhua Zhang − *College of Earth Science and Technology, China University of Petroleum, Qingdao 266555, China*; Email: z21010088@s.upc.edu.cn

Authors

Youzhuang Sun − *College of Earth Science and Technology, China University of Petroleum, Qingdao 266555, China*; orcid.org/0000-0001-5622-997X

Zhengjun Yu − *Shengli Oilfield Geophysical Research Institute, Dongying 257000, China*

Yongan Zhang − *College of Computer Science, China University of Petroleum, Qingdao 266555, China*

Zhen Liu − *College of Earth Science and Technology, China University of Petroleum, Qingdao 266555, China*; orcid.org/0000-0003-3528-5264

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.3c03677

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors are very grateful to China University of Petroleum (East China) for the experimental equipment. The authors thank the Shengli Oilfield Geophysical Institute for providing experimental data. The authors also thank the State Key Laboratory for providing the platform.

ABBREVIATIONS

AC acoustic
GRU gate recurrent unit network
BiGRU bidirectional gate recurrent unit network
LSTM long short-term memory network
GR gamma
R4 resistivity
DEN density
CNL neutron
M2R1 high-resolution array sensing
MAE mean absolute error
RMSE root mean square error

REFERENCES

(1) Chen, Z. H.; Zha, M.; Jin, Q. Application of natural gamma ray logging and natural gamma spectrometry logging to recovering paleoenvironment of sedimentary basins. *Chin. J. Geophys.* **2004**, *47*, 1286−1290.

(2) Soltani, S.; Kordestani, M.; Aghaee, P. K. Improved estimation for well-logging problems based on fusion of four types of Kalman filters. *IEEE Trans. Geosci. Electron.* **2018**, *56*, 647−654.

(3) Maiti, S.; Krishna Tiwari, R.; Kümpel, H. J. Neural network modelling and classification of lithofacies using well log data: a case study from KTB borehole site. *Geophys. J. Int.* **2007**, *169*, 733−746.

(4) Bader, S; Wu, X; Fomel, S Missing well log estimation by multiple well-log correlation.*80th EAGE Conference and Exhibition*. 2018.

(5) Maji, S; Berg, A C; Malik, J *Classification using intersection kernel support vector machines is efficient*. IEEE conference on computer vision and pattern recognition. 2008.

(6) Banchs, R.; Jiménez, J. R.; Del Pino, E. Nonlinear estimation of missing logs from existing well log data. *SEG Annual Meeting*. **2001**.

(7) Jordan, M. I.; Mitchell, T. M. Machine learning: Trends, perspectives, and prospects. *Science* **2015**, *349*, 255−260.

(8) Srisutthiyakorn, N. Redefining the standard of missing log prediction: Neural network with bayesian regularization (nnbr) with stratigraphic constraint−a case study from laminated sand-shale reservoir. *SEG Technical Program Expanded Abstracts*. **2012**.

(9) Mo, X.; Zhang, Q.; Li, X. Well logging curve reconstruction based on genetic neural networks. *International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*. **2015**.

(10) Ayoub, M.; Mohamed, A. A. Estimating the lengthy missing log interval using group method of data handling (GMDH) technique. *Appl. Mech. Mater.* **2014**, *695*, 850.

(11) Jian, H.; Chenghui, L.; Zhimin, C.; Haiwei, M. Integration of deep neural networks and ensemble learning machines for missing well logs estimation. *Flow Meas. Instrum.* **2020**, *73*, 101−112.

(12) Kwon, M.; Yoo, J.; Kang, C. Accurate and Convenient Missing Well Log Synthesis using Generative Model. *Abu Dhabi International Petroleum Exhibition & Conference*. **2020**.

(13) Fan, P.; Deng, R.; Qiu, J.; Zhao, Z.; Wu, S. Well logging curve reconstruction based on kernel ridge regression. *Arabian J. Geosci.* **2021**, *14*, 1−10.

(14) Zhang, D.; Yuntian, C.; Jin, M. Synthetic well logs generation via Recurrent Neural Networks. *Pet. Explor. Dev.* **2018**, *45*, 629−639.

(15) Pham, N.; Wu, X. Missing sonic log prediction using convolutional long short-term memory. *SEG International Exposition and Annual Meeting*. **2019**.

(16) Pham, N.; Wu, X.; Zabihi, N. E. Missing well log prediction using convolutional long short-term memory network. *Geophysics* **2020**, *85*, 159−171.

(17) Cheng, C.; Gao, Y.; Chen, Y.; Jiao, S.; Jiang, Y.; Yi, J.; Zhang, L. Reconstruction Method of Old Well Logging Curves Based on BI-LSTM Model—Taking Feixianguan Formation in East Sichuan as an Example. *Coatings* **2022**, *12*, 113−123.

(18) Dong, N.; Zhao, L.; Wu, C. H.; Chang, J. F. Inception v3 based cervical cell classification combined with artificially extracted features. *Appl. Soft Comput.* **2020**, *93*, 106−113.

(19) Zhou, X.; Cao, J. X.; Wang, X. J. Acoustic log reconstruction based on bidirectional Gated Recurrent Unit (GRU) neural network. *Progress in Geophysics* **2022**, *37*, 357−366.

(20) Zhao, R.; Wang, D.; Yan, R. Machine health monitoring using local feature-based gated recurrent unit networks. *IEEE Trans. Ind. Electron.* **2018**, *65*, 1539−1548.

(21) Lu, B.; Luktarhan, N.; Ding, C.; Zhang, W. Iclstm: Encrypted traffic service identification based on inception-lstm neural network. *Symmetry* **2021**, *13*, 1080−1092.

(22) Szegedy, C; Vanhoucke, V; Ioffe, S Rethinking the inception architecture for computer vision. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

(23) Dey, R; Salem, F M Gate-variants of gated recurrent unit (GRU) neural networks. *IEEE 60th international midwest symposium on circuits and systems*. 2017.
(24) Yu, Y.; Si, X.; Hu, C.; Zhang, J. A review of recurrent neural networks: LSTM cells and network architectures. *Neural, Comput.* **2019**, *31*, 1235−1270.
(25) Fang, W.; Chen, Y.; Xue, Q. Survey on research of RNN-based spatio-temporal sequence prediction algorithms. *J. Big Data* **2021**, *3*, 97−110.
(26) Fu, R; Zhang, Z; Li, L. Using LSTM and GRU neural network methods for traffic flow prediction. *Youth Academic Annual Conference of Chinese Association of Automation*. 2016.
(27) Lin, X.; Quan, Z.; Wang, Z. J.; Huang, H.; Zeng, X. A novel molecular representation with BiGRU neural networks for learning atom. *Briefings Bioinf.* **2020**, *21*, 2099−2111.
(28) Han, X.; Chen, F.; Ban, J. Music Emotion Recognition Based on a Neural Network with an Inception-GRU Residual Structure. *Electronics* **2023**, *12*, 978−985.
(29) Jais, I. K. M.; Ismail, A. R.; Nisa, S. Q. Adam optimization algorithm for wide and deep neural network. *Knowl. Eng. Data Sci.* **2019**, *2*, 41−46.
(30) Srivastava, N.; Hinton, G.; Krizhevsky, A. Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929−1958.
(31) Nakagawa, S.; Schielzeth, H. A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in ecology and evolution* **2013**, *4*, 133−142.
(32) Chai, T.; Draxler, R. R. Root mean square error (RMSE) or mean absolute error (MAE)? −Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.* **2014**, *7*, 1247−1250.
(33) Willmott, C. J.; Matsuura, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* **2005**, *30*, 79−82.
(34) Guan, X.; Zhang, B.; Fu, M.; Li, M.; Yuan, X.; Zhu, Y.; Peng, J.; Guo, H.; Lu, Y. Clinical and inflammatory features-based machine learning model for fatal risk prediction of hospitalized COVID-19 patients: results from a retrospective cohort study. *Ann. Med.* **2021**, *53*, 257−266.
(35) Zhou, H.; Deng, Z.; Xia, Y.; Fu, M. A new sampling method in particle filter based on Pearson correlation coefficient. *Neurocomputing* **2016**, *216*, 208−215.

https://doi.org/10.1021/acsomega.3c03677
ACS Omega 2023, 8, 27710−27724