RESEARCH ARTICLE | APRIL 01 2025

An anti-noise block and frequency-aware framework in deep learning for formation resistivity prediction by transient electromagnetic data

Yongan Zhang (张永安); Jian Li (李健) ; Junfeng Zhao (赵俊峰); Xuanran Wang (王轩然); Youzhuang Sun (孙有壮) ; Yizheng Li (李奕政); Yuntian Chen (陈云天) ; Dongxiao Zhang (张东晓)

Physics of Fluids 37, 047107 (2025)
https://doi.org/10.1063/5.0256397

Physics of Fluids ARTICLE pubs.aip.org/aip/pof

An anti-noise block and frequency-aware framework in deep learning for formation resistivity prediction by transient electromagnetic data

Cite as: Phys. Fluids **37**, 047107 (2025); doi: 10.1063/5.0256397
Submitted: 4 January 2025 · Accepted: 8 February 2025 ·
Published Online: 1 April 2025

View Online Export Citation CrossMark

Yongan Zhang (张永安),[1,2] Jian Li (李健),[1] Junfeng Zhao (赵俊峰),[3] Xuanran Wang (王轩然),[3] Youzhuang Sun (孙有壮),[4] Yizheng Li (李奕政),[1,2] Yuntian Chen (陈云天),[1,5,a)] and Dongxiao Zhang (张东晓)[1,5,a)]

AFFILIATIONS

[1]Ningbo Institute of Digital Twin, Eastern Institute of Technology, Ningbo 315000, China
[2]The Hong Kong Polytechnic University, Hong Kong 999077, China
[3]Sinopec Matrix Corporation, Qingdao 266001, China
[4]China University of Petroleum (East China), Qingdao 266580, China
[5]Zhejiang Key Laboratory of Industrial Intelligence and Digital Twin, Eastern Institute of Technology, Ningbo 315200, China

[a)]**Authors to whom correspondence should be addressed:** ychen@eitech.edu.cn and dzhang@eitech.edu.cn

ABSTRACT

Formation resistivity prediction is critical for understanding subsurface fluid behavior. Traditional methods struggle to accurately measure subsurface fluid parameters from cased wells, where a transient electromagnetic method can be effectively applied. However, resistivity prediction using transient electromagnetic method data faces two significant challenges: high-frequency disaster and environmental noise. These factors collectively hinder the ability of neural networks to capture high-frequency features and to handle noise interference, diminishing prediction accuracy. To address these challenges, this study proposes a frequency-aware framework and a temporal anti-noise block. The frequency-aware framework addresses high-frequency disaster by employing a dual-stream structure with wavelet transformation to isolate and learn high-frequency components. Meanwhile, the temporal anti-noise block mitigates environmental noise by denoising temporal features through a soft threshold attention mechanism. Using a long short-term memory model as a baseline, these enhancements are integrated to conduct two experiments. The ablation experiment demonstrates that the proposed block and framework significantly improve prediction accuracy, achieving an R^2 of 0.91 (0.22 higher than the baseline) with significant gains in handling high-frequency features (an improvement of 0.30 in R^2 at high-frequency part). The robustness experiment shows that the temporal anti-noise block reduces the impact of Gaussian and impulse noise by 1/8 compared to the baseline, confirming its strong noise resistance. This study achieves accurate formation resistivity prediction using transient electromagnetic method data, paving the way for advanced subsurface fluid behavior analysis in complex geological settings.

I. INTRODUCTION

The resistivity of geological formations is a critical parameter for characterizing the properties and behavior of subsurface fluids. Specifically, formation resistivity offers valuable insight into fluid distribution, migration, and interaction within porous rock formations. In the context of subsurface fluids, formation resistivity serves as a proxy for estimating lithology, porosity, fluid saturation, and other key characteristics. Moreover, dynamic changes in formation resistivity provide critical information regarding the quantity, quality, and movement of fluids in groundwater monitoring, helping track resource levels and detect contamination or depletion trends.[1,2] Additionally, formation resistivity is essential for identifying fluid flow patterns within geological formations[3] and plays a key role in understanding how fluids interact with rocks.[4] Furthermore, formation resistivity can also be used to assess the sensitivity of fluids in subsurface pores to changes in temperature and pressure, offering a more nuanced understanding of subsurface fluid behavior.[5]

In open-hole logging data prediction, artificial intelligence (AI) technologies have been successfully applied to various aspects of series data problems.[6,7] Numerous studies have demonstrated the applicability of AI in handling diverse data sources in the field of well logging[8–11] and petrophysics.[12–14] In 2010, Karimpouli *et al.*[15] developed a supervised committee machine neural network voting model for reconstructing petrophysical parameter curves. In 2016, Korjani *et al.*[16] introduced a reservoir modeling method based on deep neural networks to predict rock physics features. In 2017, Parapuram *et al.*[17] proposed a multi-stage curve generation scheme where each stage's log curve serves as a constraint in predicting the next stage's curve, enhancing the reliability of the final interpretation. In 2018, Akinnikawe *et al.*[18] proposed an approach using variable clustering analysis and data transformation for differentiating input features in well log data. This process enables the training and testing of various machine learning methods. In 2018, Zhang *et al.*[19] utilized a cascaded long short-term memory neural network for the reconstruction of logging curves, introducing multiple layer mapping relationships to enhance the accuracy of logging curve reconstruction. This study was the first to employ temporal neural networks to learn the temporal characteristics of logging curves. Wu *et al.*[20] introduced a workflow combining cross-entropy clustering, Gaussian mixture model, and hidden markov model to propagate information from training wells to other wells by identifying locally stationary clustering clusters corresponding to formation zones. In 2018, Akkurt *et al.*[21] built an automated process for rapid learning and enhanced analysis of rock physics to detect outliers in well logging data, evaluate inter-well similarity post-outlier removal using clustering information, and subsequently reconstruct well log curves. In 2020, Chen and Zhang[22] proposed an ensemble long short-term memory network capable of processing sequential data on a small dataset. By combining the ensemble neural network and cascaded long short-term memory network, an ensemble long short-term memory network (EnLSTM) reduces costs and saves time. In the same year, Chen and Zhang[23] introduced a physics-constrained long short-term memory network, incorporating the physical mechanism behind geomechanically parameters as prior information. This model directly estimates geomechanically logs based on easily available data, demonstrating improved prediction accuracy by leveraging domain knowledge. In 2023, Sun *et al.*[24] developed the inception bidirectional gated recurrent unit (Inception-BiGRU) model, a gated recurrent unit network based on the inception module, to improve the accuracy of missing log value prediction by extracting spatial relationships within well log curves. However, these methods typically rely on open-hole well data. In many subsurface fluid studies, most of already-drilled wells are cased wells, where traditional techniques cannot provide the necessary data for resistivity prediction. This makes the task of predicting formation resistivity more inaccurate. For cased wells, a current method to obtain accurate formation resistivity signals is through the transient electromagnetic method (TEM). TEM through-casing logging technology is a new technique that overcomes the limitations of traditional logging methods.

The main theory of TME is a transmission line equation (TLE), and the traditional way to obtain resistivity through TLE is mainly through theoretical models to calculate it. In 1994, Schenkel and Morrison[25] introduced the transmission line equation, which calculates numerical solutions for potentials and their derivatives. The method incorporates finite-length casing, variable-radius annular zones, and vertical discontinuities to calculate formation resistivity, achieving an accuracy of up to 60%. In 2005, Yeten *et al.*[26] proposed estimating leakage current by measuring the differential voltage on the casing, introducing the concept of cased transient resistivity logging to calculate formation conductivity. However, due to incomplete theories and methodologies at the time, this technology did not undergo extensive development and widespread application. In 2006, Pardo *et al.*[27] simulated measurement results under different frequencies using a multi-frequency cased resistivity tool through a goal-oriented hp-adaptive finite-element method to evaluate rock properties. Experimental results demonstrated the capability of this method for logging simulations in the presence of highly conductive casings, but numerical calculations from empirical formulas proved inadequate in certain scenarios, leading to insufficient accuracy. In 2018, Liu *et al.*,[28] addressing the limitation of the conventional TLE in calculating resistivity logging responses in fractured formations through casing, proposed a new method utilizing both TLE and the current flux tube model. This approach considered fracture conditions and extended to the calculation of lateral formation resistivity. However, it still falls within the realm of empirical formulas, resulting in limitations in accuracy and applicability. In 2023, Liu and Chen,[29] addressing the inability of the conventional TLE to investigate current and potential distributions in borehole fluids, proposed an improved TLE. The enhanced TLE incorporates formation information outside the casing and fluid inside the well, allowing for the analytical solution of casing wall potential distribution. The advantage of this type of theoretical method is that it can be applied to casing wells. Despite these improvements, the above theoretical methods still have significant shortcomings in accuracy compared to neural network methods. For TEM data, two major challenges arise in predicting formation resistivity by neural networks: (1) environmental noise:[30,31] factors such as casing corrosion, blowouts, and current leakage can introduce fluctuations in the current; (2) high-frequency disaster:[32,33] a phenomenon leads to a reduction in Ai model accuracy. The reason for this phenomenon is that the prediction of resistivity by TEM data is a task that is highly sensitive to high-frequency features, but traditional neural networks are difficult to fit high-frequency features. High-frequency disaster has been observed in various scientific machine learning applications.[34,35] In addressing the challenge of obtaining formation resistivity from cased wells, there has been relevant research.

To address the issues mentioned above, this study proposes a frequency-aware framework and a temporal anti-noise block. For environmental noise, temporal anti-noise block achieves denoising of temporal features through a soft threshold attention mechanism. For the high-frequency disaster, the frequency-aware framework uses a dual-stream structure with wavelet transformation to isolate the high-frequency part and to learn it. Through the proposed block and framework, accurate prediction of subsurface formation resistivity by using TEM data can be achieved, providing a robust data foundation and deeper research basis for subsequent studies on subsurface fluids.

II. METHODOLOGIES

Overall, the work presented in this paper involves proposing a new anti-noise block and frequency-aware framework to predict formation resistivity, thereby providing a data foundation for other formation fluid studies. Formation resistivity is an important parameter that characterizes the physical properties of subsurface fluids and also

serves as a crucial data basis for important studies such as groundwater monitoring and groundwater migration patterns. The innovative aspect of this paper is the introduction of a new neural network block and framework, enabling the common neural network to predict formation resistivity through TEM logging data, which could not be accurately predicted by conventional open-hole logging data in the past. This provides a data foundation for subsequent fluid research. The flow chart for predicting formation resistivity in this study is shown in Fig. 1.

This section mainly introduces the frequency-aware framework, the temporal anti-noise block, and the assembled temporal neural network model. Specifically, the assembled temporal neural network model uses the long short-term memory model (LSTM) as an example to integrate the proposed framework and block with subsequent comparisons conducted using LSTM as the baseline. It is worth noting that the framework and block proposed in this study can also be integrated into any other temporal models.

A. Frequency-aware framework

To address high-frequency disaster, the frequency-aware framework (FAF, as illustrated in the Fig. 2) based on the wavelet transform is proposed. The FAF primarily consists of a wavelet encoder that utilizes wavelet encoder to divide the temporal data into two streams: high-frequency features flow and low-frequency flow features. Both high-frequency and low-frequency flow features undergo separate temporal neural networks before being output to their respective wavelet decoders, also based on the wavelet transform. After transformation into high-frequency and low-frequency flows, the results are combined by the wavelet decoder to form the prediction outcome. The FAF segregates the high-frequency and low-frequency flows of temporal data. Through the dual-stream structure, it simultaneously provides feedback and adjusts both high-frequency and low-frequency flows, combating the challenge of high-frequency disaster and ultimately improving prediction accuracy.

The FAF based on wavelet encoder separates high-frequency and low-frequency features through wavelet transform. It simultaneously engages in gradient descent and feature learning for both high-frequency and low-frequency data, thereby enabling the model to a certain extent, avoid high-frequency disaster.

The encoding formula for the wavelet transform is expressed as follows:

$$Ws(a,b) = |a|^{-1/2}\int_R s(t)\bar{\Psi}\left(\frac{t-b}{a}\right)dt, \tag{1}$$

$$\Psi_M(t) = e^{i\omega_0 t}e^{-\frac{t^2}{2}}, \tag{2}$$

where $W_s(a,b)$ represents the result of continuous wavelet transform and $s(t)$ is a variable representing the time t of the input series data. The parameter a is the scaling parameter, related to the frequency, and b is the position parameter. $\Psi(t)$ is the complex conjugate function of $\Psi(t)$. $\Psi_M(t)$ is the wavelet function. In dealing with series data, a non-orthogonal Morlet wavelet is chosen. ω_0 is a constant representing the center frequency, and i is the imaginary unit determining whether the information in the transformation result is in the time domain or spatial domain.

The Mallat algorithm[36] is employed for wavelet transform to decompose low-frequency data and high-frequency data. The encoder and decoder for wavelet transform are illustrated in Fig. 3.

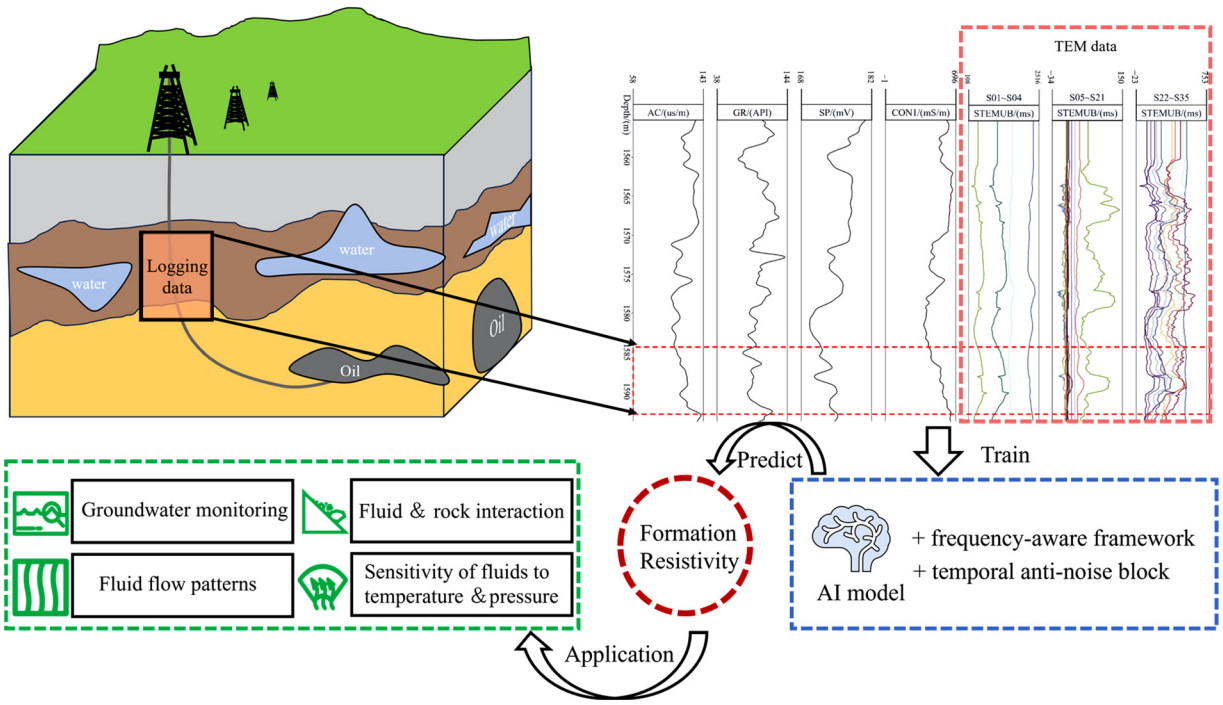


FIG. 1. Flowchart for predicting formation resistivity using deep learning with frequency aware framework and temporal anti-noise block.

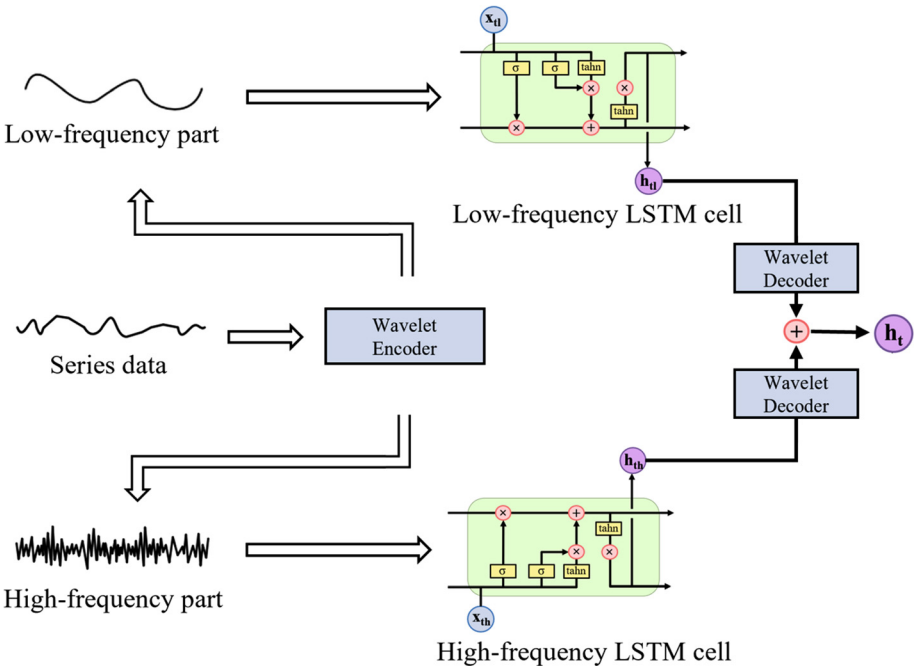


FIG. 2. Frequency-aware framework (FAF, the upper part is for low-frequency flow processing, and the lower part is for high-frequency flow processing).

The calculation formulas for the approximation coefficient $a_{j-1,k}$ and the detail coefficient $d_{j-1,k}$ are expressed as follows:

$$\begin{cases} a_{j-1,k}(n) = \sum_{k=-\infty}^{+\infty} a_{j,k} h(k-2n), n = 0, 1, \ldots 2^{N-j} - 1, \\ d_{j-1,k}(n) = \sum_{k=-\infty}^{+\infty} a_{j,k} g(k-2n), n = 0, 1, \ldots M - 1, \end{cases} \tag{3}$$

where $h(t)$ represents the impulse response of the low-pass filter H, $g(t)$ is the impulse response of the high-pass filter G, M denotes the total number of wavelet decomposition levels (set to 5 in the model), and j is the current wavelet decomposition level.

The reconstruction expression is given by

$$\begin{cases} a_j = \sum_{k=-\infty}^{+\infty} \{h^+(n-2k) a_{j-1,k} + g^+(n-2k) d_{j-1,k}\} \\ j = M-1, M-2, \ldots, 0, \end{cases} \tag{4}$$

where $h^{+(k)}$ represents the impulse response of the reconstruction low-pass filter and $g^{+(k)}$ is the impulse response of the reconstruction high-pass filter.

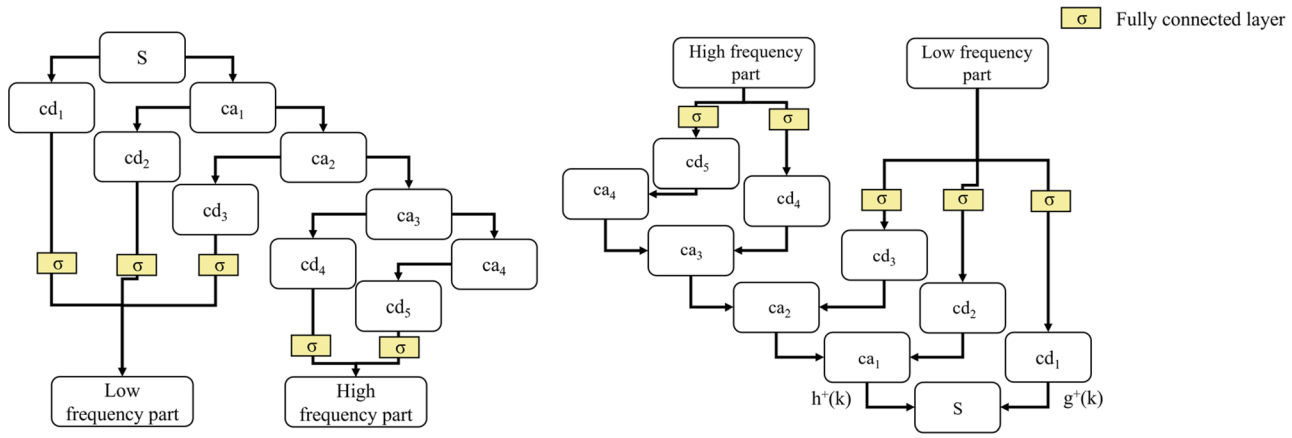


FIG. 3. Wavelet transforms (S, ca, and cd represent sequence data, approximation coefficients, and detail coefficients, respectively).

B. Temporal anti-noise Block

Temporal anti-noise block serves as a denoising block, specifically designed for data with strong noise. Structurally, temporal anti-noise block (shown in Fig. 4) consists of three parts: residual unit, threshold learning unit, and soft-threshold processing unit. The residual unit (the green part of Fig. 4) primarily reduces training difficulties through identity mapping. The threshold learning unit (the blue part of Fig. 4), guided by attention mechanisms, learns appropriate thresholds through feedback adjustment and feeds it into the soft-threshold processing unit as the input of the soft threshold formula. The soft-threshold processing unit (the red part of Fig. 4) is a crucial part for noise data removal, implementing the soft-threshold method[37] to set features within a specific range to zero, thereby eliminating specific features (either noise or redundant features). The soft thresholding method is a denoising method that can automatically learn and adapt to various denoising rules under different noise distributions based on the formula to adjust the threshold value. Essentially, temporal anti-noise block combines residual networks, soft-thresholding functions, and attention mechanisms. It introduces soft-thresholding into the residual module as a shrinkage layer, employs multi-head attention mechanisms to set adaptive thresholds, and ensures that the model, through feedback adjustment, focuses on repeatedly occurring features in multidimensional redundant information.

The residual unit, soft-threshold processing unit, and deep residual shrinkage network share similarities with the deep residual shrinkage network.[38] The expression for the soft-threshold function and its derivative are provided in the following formulas:

$$\mathrm{soft}(x,T)=\begin{cases} x+\tau & x<-\tau, \\ 0 & |x\le\tau|, \\ x-\tau & x>\tau, \end{cases} \tag{5}$$

$$\frac{\partial y}{\partial x}=\begin{cases} 1 & x>-\tau, \\ 0 & |x\le\tau|, \\ 1 & x>\tau, \end{cases} \tag{6}$$

where τ represents the threshold size in the soft-threshold function, x is the input, and y is the output.

The threshold learning unit incorporates the idea of multi-head attention mechanism,[39] and its specific structure is illustrated in Fig. 4:

In Fig. 4, B, S, and F denote Batch size, sequence length, and feature quantity, respectively. Similar to the soft-thresholding under the deep attention mechanism of the deep residual shrinkage network, temporal anti-noise block primarily employs soft-thresholding to set irrelevant features to zero in the feature vector, thereby filtering out irrelevant features. Initially, the input multidimensional sequence data undergo absolute value transformation and global mean pooling to extract global features. These processed features are then input into the multi-head attention mechanism for temporal feature extraction, learning a set of weights to suppress noise, which are subsequently fed into the soft-thresholding processing unit. In the soft-thresholding processing unit, the obtained weights and the temporal data undergo soft-thresholding calculations using formulas (5) and (6), suppressing irrelevant strong noisy features to counteract the high-noise characteristics of the data. Finally, the results of the soft-thresholding calculations are added to the residual results and fed into the fully connected layer. This process ensures that the entire neural network, during the learning of noise filtering, mitigates gradient vanishing and feature elimination. In summary, the temporal anti-noise block has two main characteristics:

(1) Noise and feature filtering in time-series data: Addressing the issue of excessive redundant information and the interference caused by weak and noisy features in well logging input data, the block introduces a multi-head attention mechanism for noise filtering. The multi-head attention mechanism in the threshold learning unit performs feature selection and noise filtering, allowing the model to focus on removing noise and selecting relevant features after feedback adjustment.
(2) Multi-attention mechanism: Combating the problem of insufficient feature learning, the model combines both the multi-head attention mechanism and soft-thresholding attention mechanism. The multi-head attention mechanism is integrated into the soft-thresholding recognition process and assists in setting adaptive thresholds. By identifying repeated features in multidimensional time-series data, it learns a weight matrix capable of removing noise and redundant information, which is then submitted to the soft-thresholding processing unit. Therefore, the

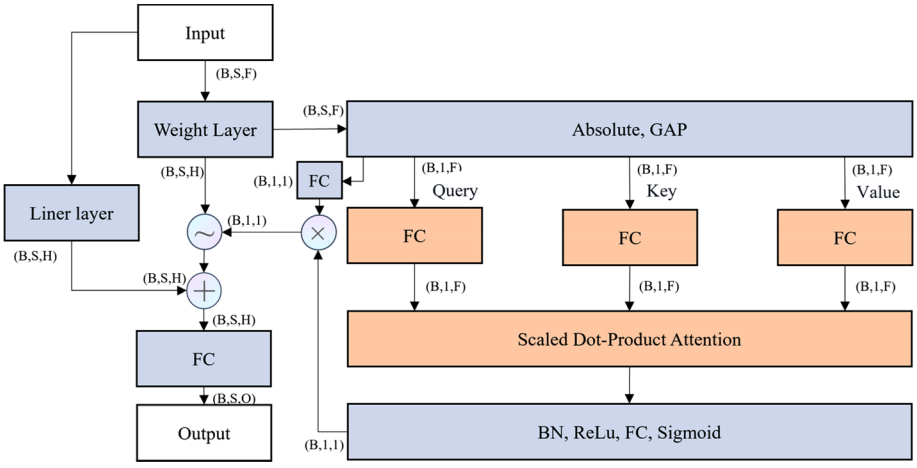


FIG. 4. Temporal anti-noise block (red part: the threshold learning unit, blue part: the soft-threshold processing unit).

model's dual attention mechanisms, coupled with residual and contraction operations, reinforce both the deep learning and sparse representation capabilities of the network.

C. Case model structure

To verify the effectiveness of the anti-noise block and frequency-aware framework, the temporal anti-noise LSTM (TAL, showed in Fig. 5) and frequency-aware LSTM (FAL, showed in Fig. 6) were constructed by using the LSTM model as an example, facilitating subsequent experimental comparisons. Figure 5 illustrates the complete structure of TAL, and Fig. 6 presents the structure of frequency aware LSTM (FAL).

In Fig. 5, input data are initially processed through a conventional LSTM cell (Fig. 5 in green). Subsequently, it enters the temporal anti-noise block (Fig. 5 in blue), where the multidimensional sequence data undergo absolute value transformation and global average pooling to extract global features. This transformed data are then input into the multi-head attention mechanism for temporal feature extraction. The soft threshold calculation is performed using Eqs. (5) and (6), and the resulting soft threshold calculation is added to the residual output before entering the fully connected layer. Finally, the loss is calculated with the label data, and adjustments are made through feedback. In comparison to LSTM, the additional temporal anti-noise block possesses a soft thresholding attention mechanism with multi-head attention, handling the selection of multidimensional features and suppression of temporal data noise, respectively. The residual structure enables TAL to counteract gradient vanishing and feature elimination when filtering strong noise.

In FAL, analogous to the structure of FAF, temporal data are transformed from the time domain to the frequency domain, specifically into high-frequency and low-frequency components, through a wavelet transform encoder. The dual-stream architecture of FAL is employed to handle the high-frequency and low-frequency components, where the neural networks in the FAL dual-stream structure are based on TAL. The network is trained to learn features from the high-frequency data stream and low-frequency data stream separately. The output of the dual-stream structure undergoes reverse transformation using a wavelet transform decoder to convert back to the time domain. Subsequently, a loss function is computed by comparing the output with the label data of normal temporal sequences, and feedback adjustments are made. In comparison to TAL, the additional wavelet transforms encoder and decoder in FAL enable the extraction of high-frequency and low-frequency features. In this way, FAL can avoid high-frequency disaster and impact of noise.

III. EXPERIMENTS AND RESULTS

This section comprises two experiments. Section III A presents the missing well logs prediction experiment, including an ablation study to evaluate each component of our models. Section III B introduces robustness analysis, testing the models on datasets with more noise.

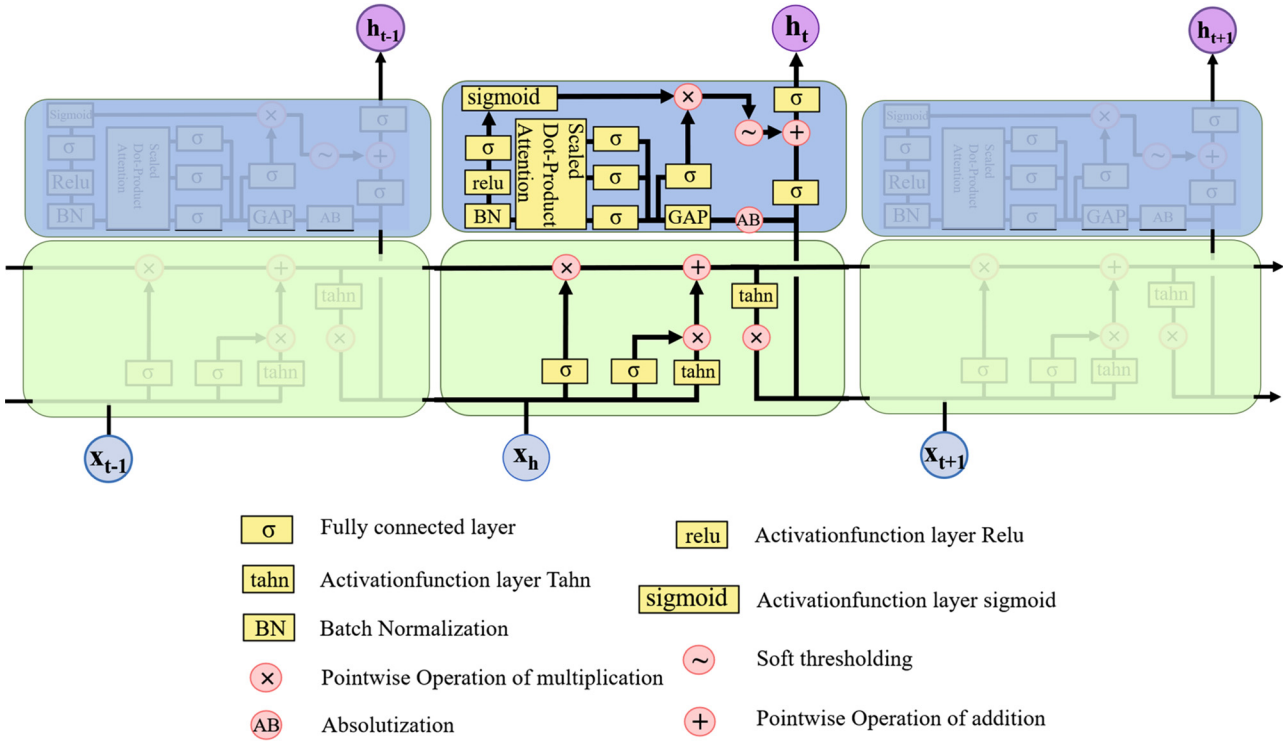


FIG. 5. Temporal anti-noise LSTM (TAL, with the blue part representing the temporal anti-noise block and the green part representing the conventional LSTM cell).

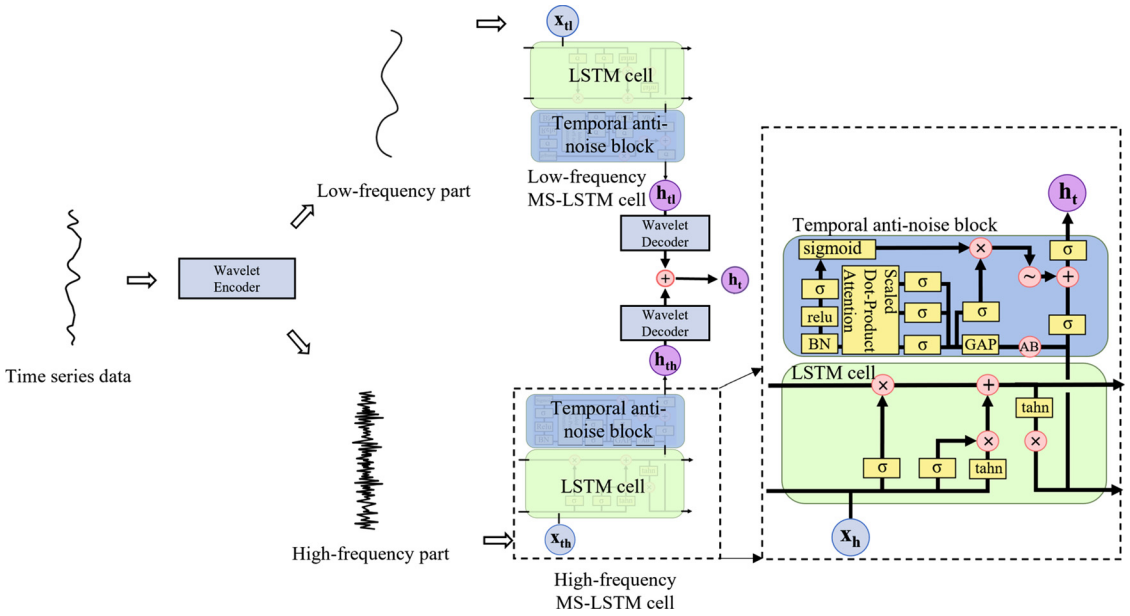


FIG. 6. Frequency aware LSTM (FAL, with TAL cells replacing LSTM cells).

A. Experiment on predicting missing well logs

The dataset is derived from the alluvial plains in eastern China. The input dataset comprises through-casing electromagnetic logging curve data and conventional well logging data, while the output consists of formation resistivity data obtained from barefoot well logging curves. The primary focus is on predicting formation resistivity based on through-casing electromagnetic logging curves. Due to the multi-point distributed measurement instruments of through-casing electromagnetic logging, simultaneous measurements are essential to ensure data accuracy. Therefore, the number of through-casing electromagnetic logging curves is relatively greater compared to conventional well logging curves. In this study, the dataset consists of three wells, encompassing a total of 35 through-casing electromagnetic logging curves and four conventional well logging curves. The prediction target is one set of formation resistivity data. The conventional well logging curves include AC (acoustic logging curve), GR (natural gamma curve), CON (induction logging curve), and SP (spontaneous potential curve). The through-casing electromagnetic logging curves include STEMUB01–STEMUB35 (readings from instrument current meters 01–35), as well as RT (formation resistivity). The past open-hole resistivity log curves of these wells were used as labels (ground truth), while the TEM data and conventional logging curves from the cased-hole state of the same wells served as input. The data preprocessing involves two main components: granularity alignment and standardization (details can be find in Subsection 1 of Appendix). For the test section, three different depth intervals were selected as the test set, while the remaining intervals were used as the training and validation sets. Overall, the ratio of training, validation, and test sets was 7:2:1 (the details of the data and the experiment setting will be elucidated in Subsections 1 and 2 of Appendix). The pictures of three wells data used in experience of missing well logs prediction can be in the supplementary material.

To analyze the contributions of each component, the experiment on predicting missing well logs including the ablation experiment. The purpose of ablation experiments is to selectively remove certain components or features from the model to evaluate their impact on model performance. By gradually eliminating parts from the model, we can delve into the model's performance, thereby gaining a comprehensive understanding of each part of model. Additionally, attention-LSTM with conventional attention mechanism and Res-LSTM with a residual structure are included as baseline for a complete ablation experiment. The prediction results of FAF, TAL, FAL, and baseline on this dataset are presented in Fig. 7 and Table I. In Fig. 7, each subgraph represents a missing portion of the formation resistivity for a certain well. Due to the limited paper space, the detailed test set setup process can be found in Subsection 2 of Appendix.

Comparing to the baseline, it can be observed from the predictions of the third column figures in Fig. 7 that FAF provides a more accurate fit to the actual values in the high-frequency curve section. However, its performance in the low-frequency domain shows little difference from that of conventional models, indicating that the FAF primarily focuses on the extraction and learning of high-frequency features. In contrast to the baseline, TAL enhances the accuracy across the board, with the most significant improvements seen in the predictions for low-frequency and low-amplitude segments (as depicted in the second column of Fig. 7). This demonstrates that the temporal anti-noise block is capable of reducing the impact of noise in both high- and low-frequency bands, although its effectiveness in the high-frequency section is far less than that in the low-frequency section, due to the high-frequency disaster effect. Compared to the baseline, FAL exhibits the best prediction performance in all figures of Fig. 7. It not only maintains stability in the low-frequency band without being affected by noise interference but also demonstrates the same

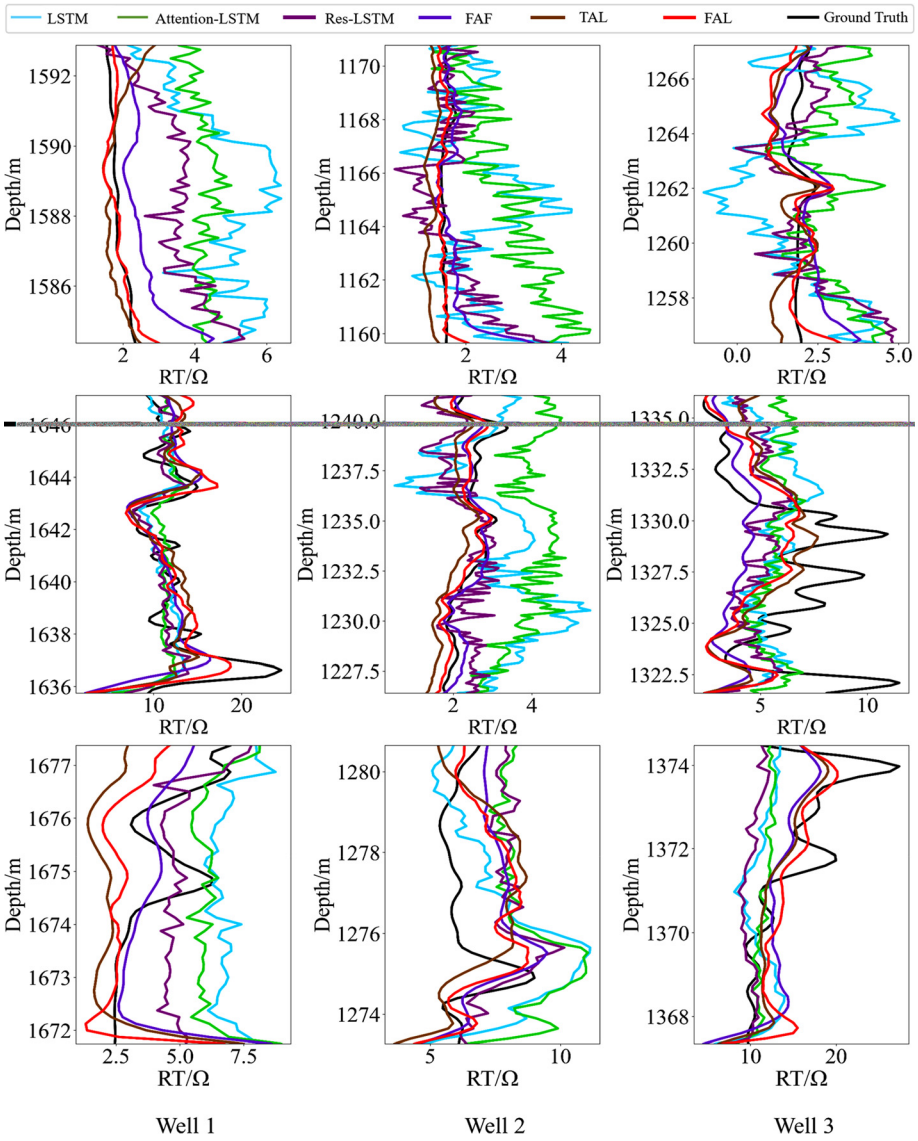


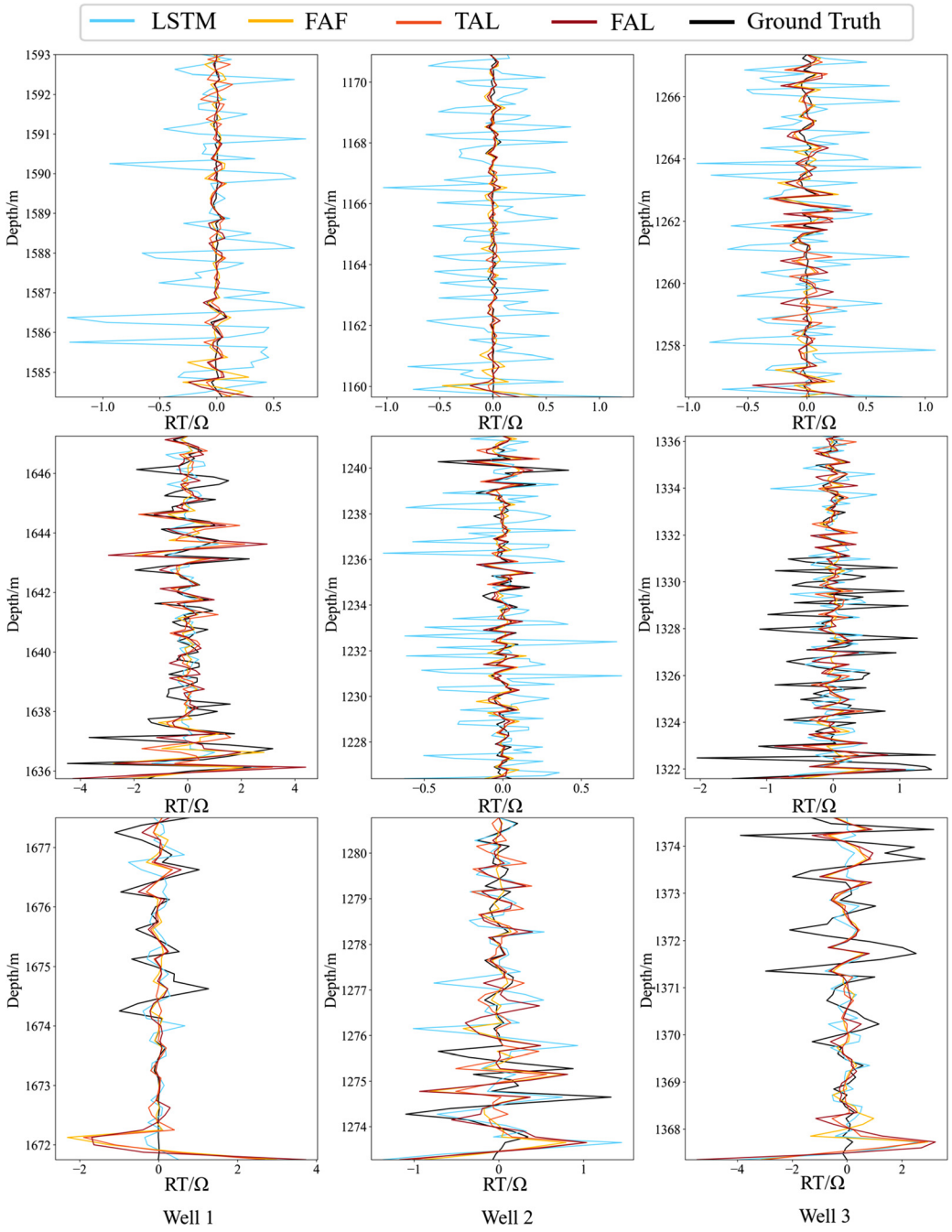
FIG. 7. Prediction results and comparison results.

high-frequency extraction and learning capabilities as FAF in the high-frequency section, highlighting the model's resistance to interference and its ability to combat high-frequency disaster in the presence of strong noise and high redundant time-series data.

In Fig. 7, compared to other models, the model presented in this study shows a distinct accuracy advantage in the first and second rows of Well 1, Well 2, and Well 3. It is worth noting that in the prediction section of the first row, both the conventional model

TABLE I. Results of ablation study. Bold values indicate the results of base line model that integrates the framework and block proposed in this study.

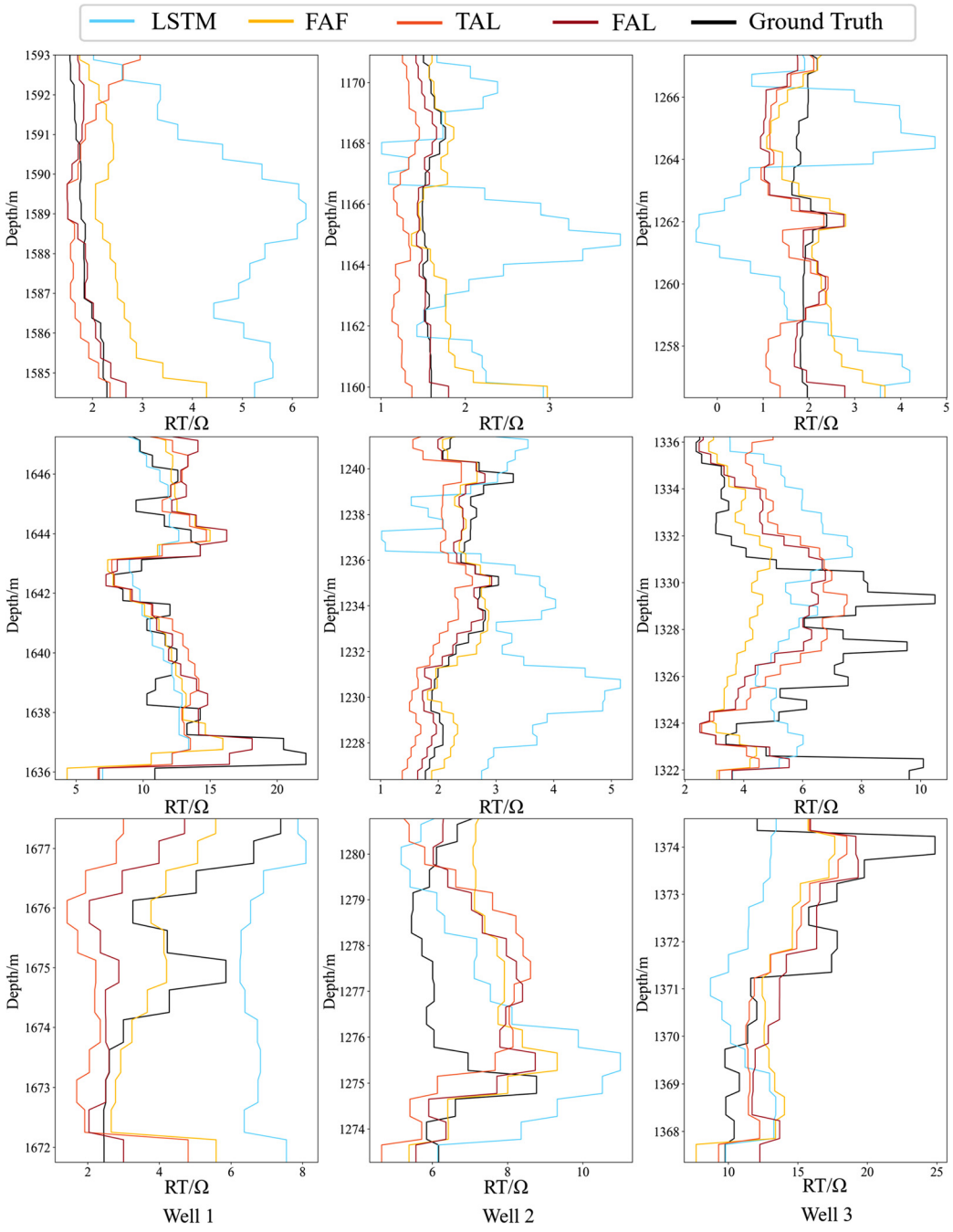
Model	Training loss	$R^2\uparrow$	MAE↓	RMSE↓	MSE↓	Training time (s)	Prediction time (s)
LSTM(baseline)	0.0253	0.69	1.90	2.62	6.86	1.7012	0.0435
Attention-LSTM	0.0135	0.74	1.64	2.35	5.54	2.9954	0.0495
Res-LSTM	0.0102	0.75	1.44	2.34	5.496	1.7698	0.0421
FAF	0.0037	0.80	1.21	2.07	4.29	15.2112	0.0644
TAL	0.0023	0.81	1.40	2.01	4.04	46.7555	0.0563
FAL	**0.0027**	**0.91**	**0.95**	**1.40**	**1.96**	**31.4587**	**0.0536**



(a) Results in high frequency domain 337

FIG. 8. Result of missing well logs prediction in different frequency domains.

and the baseline model exhibit high-frequency prediction fluctuations, resulting in high prediction errors and failure to predict the correct trend of the curve. In contrast, the model presented in this study maintains low-frequency prediction fluctuations and accurately predicts the trend of the curve, indicating its superior ability to resist noise in actual prediction results. However, due to insufficient following depth data in the end test sections (1672–1677 m section of Well 1, the 1274–1280 m section of Well 2, and the 1368–1374 m section of Well 3), both the model proposed in this study and the comparative models exhibited suboptimal performance in these end test sections.



(b) Results in low frequency domain

FIG. 8. (*Continued.*)

Table I presents the results of the ablation experiments (the bold part is our model), with each model group displaying Training Loss, training accuracy, training time, and prediction time. The training accuracy includes four metrics: R^2, MAE, RMSE, and MSE, with both overall R^2 and individual well R^2 values presented in the R^2 metric. For FAF, TAL, and FAL, there is a relative improvement of 0.1120, 0.1272, and 0.2212 in the overall R^2 compared to LSTM. In terms of R^2, FAL achieves the highest R2 of 0.9113 among all models, demonstrating the model's enhanced prediction accuracy. Moreover, FAL improves the original R^2 of LSTM by 31%. Regarding the contributions of individual

TABLE II. Results of missing well logs segments prediction in different frequency domains. Bold values indicate the results of base line model that integrates the framework and block proposed in this study.

Model	Frequency domain	R^2	MAE	RMSE	MSE
LSTM	Low	0.5109	1.8553	2.5395	6.4490
	High	0.4245	0.4197	1.0175	1.0352
FAF	Low	0.5518	1.7842	2.4013	6.0051
	High	0.7248	0.3391	0.5329	0.5668
TAL	Low	0.8136	1.3787	1.9533	3.8153
	High	0.4871	0.3229	1.02	1.0404
FAL	**Low**	**0.8939**	**0.9696**	**1.464**	**2.1433**
	High	**0.7937**	**0.3536**	**0.5223**	**0.5596**

components, the standalone anti-noise block enhances the baseline's original R^2 by 17%, while the standalone frequency-aware framework improves the baseline's original R^2 by 15%. These results indicate that the anti-noise block and the frequency-aware framework enhance the baseline from different perspectives, validating the rationale behind their design motivations. Additionally, significant improvements are observed when compared to Attention-LSTM and Res-LSTM, confirming that the primary contributions of the anti-noise block and the frequency-aware framework do not merely stem from the residual structure or the self-attention mechanism. To compare the proposed model's ability to address the high-frequency disaster, the prediction results are divided into high-frequency and low-frequency components, as shown in Fig. 8 and Table II.

Figure 8 illustrates the performance of the proposed model and baseline in predicting the high-frequency and low-frequency components of the results, while Table II provides detailed error metrics for the models. From the analysis of Fig. 8, the first-row subgraph of Fig. 8(a) highlights FAF's superiority in predicting the high-frequency components. According to the error metrics in Table II, TAL and FAL outperform the baseline in predicting the high-frequency components, with R^2 improvements of approximately 0.30 and 0.37, respectively (corresponding to increases of about 71% and 88% over the original R^2). This validates the effectiveness of the frequency-aware framework in handling high-frequency disaster. Moreover, TAL demonstrates enhanced performance in mitigating noise interference in the low-frequency components, leading to an overall improvement of 0.30 of R^2 in the low-frequency prediction. This indicates that the anti-noise block's noise resistance in the low-frequency range is its primary contribution to improving model accuracy. In Fig. 8, the predicted results of all models exhibit some prediction fluctuations in certain sections, such as the 1672–1677 m section of Well 1, the 1274–1280 m section of Well 2, and the 1368–1374 m section of Well 3. This issue primarily arises because these sections are situated at the end of the training dataset. As a result, the model has limited data available to learn noisy of deeper formation, which leads to inadequate learning of their frequency characteristics. However, our model still has a significant accuracy advantage overall and can reliably predict the basic trends and directions of formation resistivity. In summary, FAL leverages the advantages of FAF for high-frequency predictions while also benefiting from the robust noise resistance provided by the anti-noise block in the low-frequency components. Experiments have found that independently learning the high and low frequency in time-series data can avoid high-frequency disaster, thereby improving the learning ability and prediction accuracy of neural networks.

TABLE III. Results of missing well logs segments prediction with noise. Bold values indicate the results of base line model that integrates the framework and block proposed in this study.

Model	Noise	Training Loss	$R^2\uparrow$	MAE↓	RMSE↓	MSE↓
LSTM	No noise	0.0253	0.69	1.90	2.62	6.86
	Gaussian noise (α of 0.1)	0.0144	0.64	2.14	2.80	7.86
	Gaussian noise (α of 0.2)	0.0186	0.57	2.34	3.07	9.46
	Impulse noise (α of 0.1)	0.0111	0.65	2.11	2.78	7.73
	Impulse noise (α of 0.2)	0.0082	0.61	2.24	2.91	8.46
FAF	No noise	0.0037	0.80	1.21	2.07	4.29
	Gaussian noise (α of 0.1)	0.0043	0.76	1.33	2.26	5.13
	Gaussian noise (α of 0.2)	0.0043	0.71	1.64	2.39	5.72
	Impulse noise (α of 0.1)	0.0043	0.76	1.45	2.15	4.64
	Impulse noise (α of 0.2)	0.0049	0.72	1.59	2.276	5.18
TAL	No noise	0.0023	0.81	1.40	2.01	4.04
	Gaussian noise (α of 0.1)	0.0047	0.79	1.34	2.12	4.50
	Gaussian noise (α of 0.2)	0.0031	0.79	1.45	2.15	4.65
	Impulse noise (α of 0.1)	0.0022	0.80	1.44	2.07	4.32
	Impulse noise (α of 0.2)	0.0018	0.79	1.34	2.11	4.46
FAL	**No noise**	**0.0027**	**0.91**	**0.95**	**1.40**	**1.96**
	Gaussian noise (α of 0.1)	**0.0032**	**0.90**	**1.01**	**1.48**	**2.20**
	Gaussian noise (α of 0.2)	**0.0031**	**0.89**	**1.05**	**1.55**	**2.42**
	Impulse noise (α of 0.1)	**0.0029**	**0.90**	**1.02**	**1.48**	**2.19**
	Impulse noise (α of 0.2)	**0.0030**	**0.89**	**1.03**	**1.51**	**2.29**

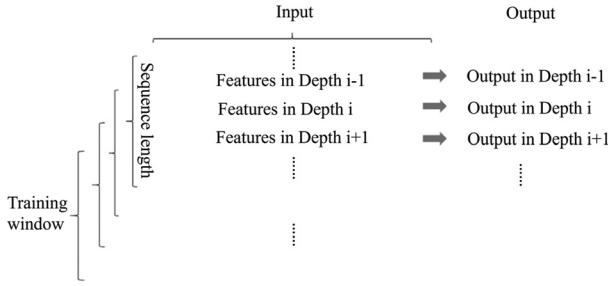


FIG. 9. Sequence prediction methods for depth sequences data.

B. Robustness analysis with noisy data

To compare the predictive capabilities of our proposed model with conventional models under complex well conditions, we simulated two common types of well noise scenarios: (1) Gaussian noise, simulating white noise and electronic noise inherent to logging instruments and (2) impulse noise, simulating transient high-amplitude pulse noise generated during the logging process due to environmental factors.

The formula for added Gaussian noise is expressed as follows:

$$p_i(z_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(z_i-\mu)}{2\sigma^2}}, \tag{7}$$

$$L_i' = L_i + \alpha p_i(z_i), \tag{8}$$

where $p_i(z_i)$ represents Gaussian noise following a normal distribution at depth i. μ, σ denote the mean and variance of the well log curve to which noise is added, and z_i is a uniformly random value between 0 and 1 at depth i. α is the scaling factor used to amplify the noise impact. L_i' represents the well log curve value after adding noise at depth i, and L_i represents the well log curve value before adding noise at depth i. The added impulse noise is expressed by the following equation:

$$n_i(z_i) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(z_i-\mu)^2}{2\sigma^2}} & z_i \le \frac{1}{T}, \\ 0 & z_i > \frac{1}{T}, \end{cases} \tag{9}$$

$$L_i' = L_i + \alpha n_i(z_i), \tag{10}$$

where $n_i(z_i)$ represents impulse noise at depth i. μ and σ denote the mean and variance of the well log curve to which noise is added, and z_i is a uniformly random value between 0 and 1 at depth i. α is the scaling factor used to amplify the noise impact. L_i' represents the well log

TABLE IV. Hyperparameters of our models in experiments.

Parameter	Setting
Optimizer	Adam
Loss function	Mean absolute error (MAE)
Learning rate	{0.001}
Batch size	64
Sequence length	40

TABLE V. Architecture of proposed models in experiments.

FAF	TAL	FAL
Input (64 × 40 × 50)	Input (64 × 40 × 50)	Input (64 × 40 × 50)
Wavelet encoder (50 × 64)	LSTM cell (64 × 32)	Wavelet encoder (50 × 64)
LSTM (64 × 32)	GAP	LSTM cell (64 × 32)
Sigmoid	FC (32 × 32)	GAP
	Multiple Heads attention (32 × 32)	Multiple heads attention (32 × 32)
	Sigmoid	FC (32 × 32)
		BN, relu
		FC (32 × 32)
		sigmoid

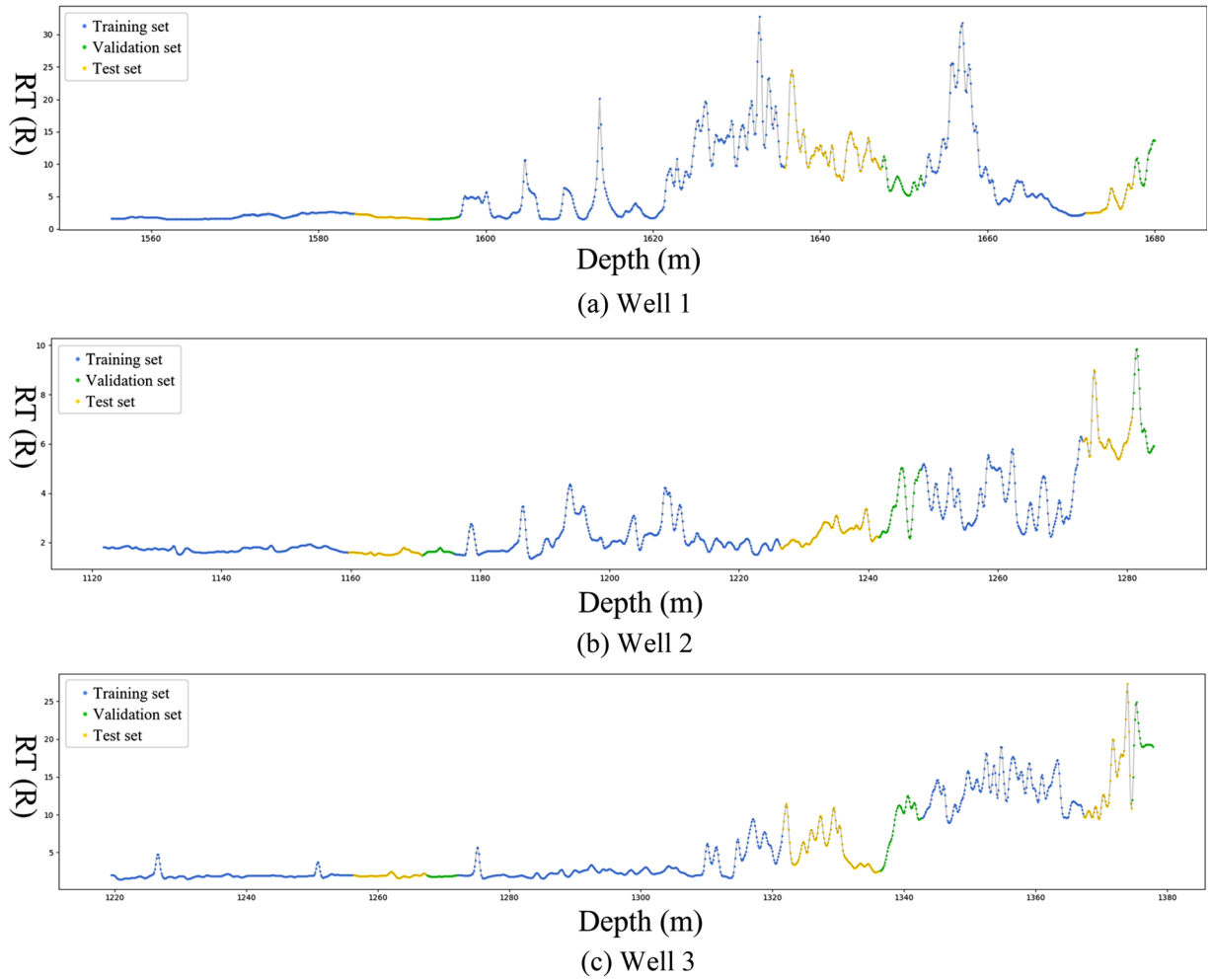


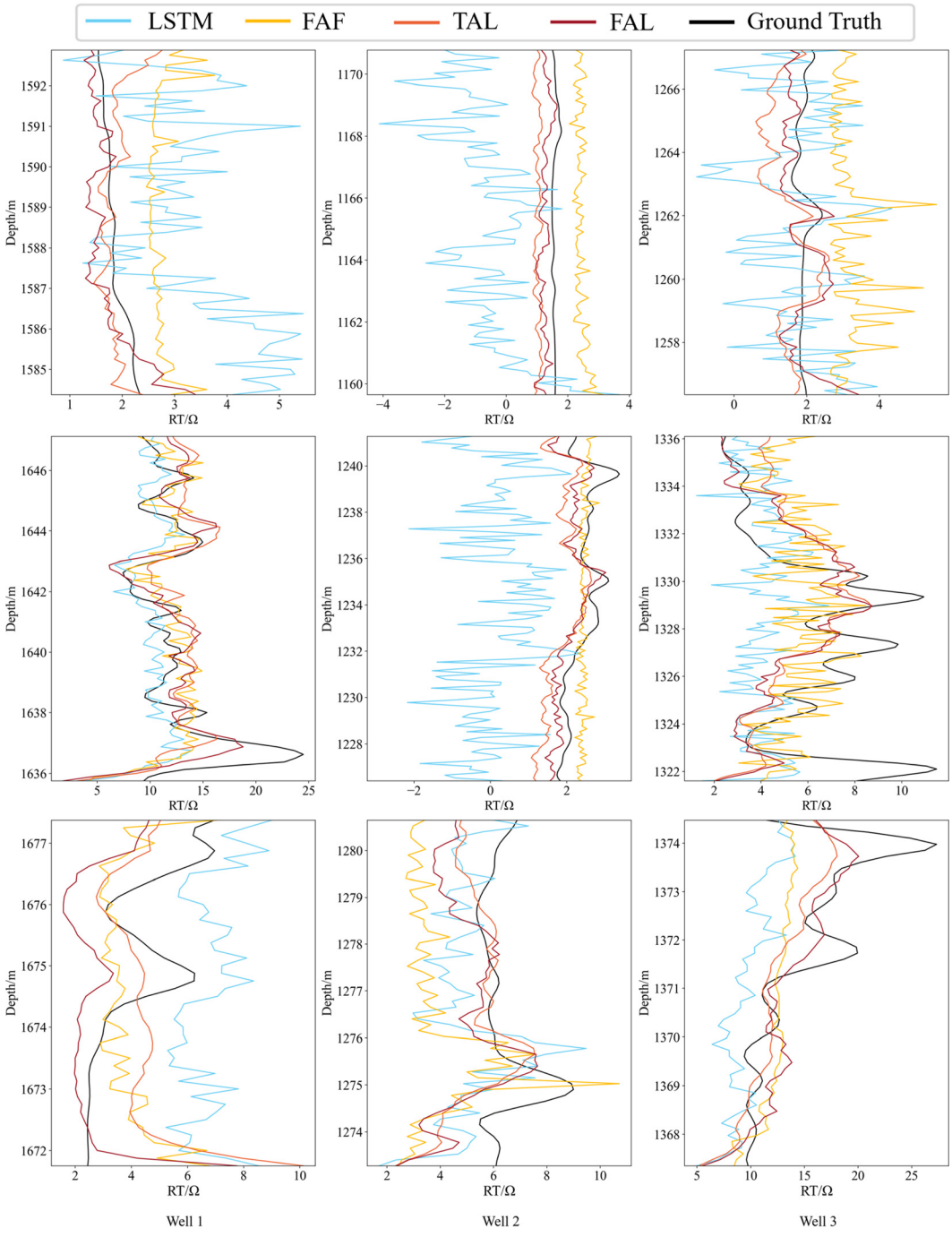
FIG. 10. Data partitioning results (using RT as an example).

curve value after adding noise at depth i, and L_i represents the well log curve value before adding noise at depth i, while T represents the pulse period (set to 0.625 m in this study).

Four types of added noise were considered: Gaussian noise (with α of 0.1), Gaussian noise (with α of 0.2), impulse noise (with α of 0.1), and impulse noise (with α of 0.2). The results are presented in Subsection 3 of Appendix (shown in Figs. 11 and 12). In addition, the images of the input features after adding noise can be found in the supplementary material. We incorporated the addition of two types of noise into the well log curves of the input data. Our model was then compared with the baseline model to analyze their respective capabilities and sensitivities to noise interference. Four types of added noise were considered: Gaussian noise (with α of 0.1), Gaussian noise (with α of 0.2), impulse noise (with α of 0.1), and impulse noise (with α of 0.2).

Table III presents the performance of models in data with noise (the bold part is baseline model with proposed farmwork and block). Analyzing Table III reveals that the R^2 of the LSTM model decreases the most under the influence of Gaussian noise and impulse noise, by 16.9% and 10.4%, respectively. For the other models, they show the following maximum reductions in R^2: FAF (Gaussian noise: 8.0%, impulse noise: 5.0%), TAL (Gaussian noise: 3.3%, impulse noise: 2.2%), and FAL (Gaussian noise: 2.2%, impulse noise: 1.6%). By comparing the decrease in accuracy of models, we can analyze the contribution of the proposed anti-noise block in noise resistance. Under noise of α =0.1, the anti-noise block reduces the impact of noise on the baseline, decreasing the R^2 reduction from 5%–7% to 1%–2%. Similarly, under noise of $\alpha = 0.2$, the anti-noise block reduces the impact of noise on the baseline model, lowering the R^2 reduction from 11%–17% to 2%–3%. Therefore, in terms of noise resistance, the anti-noise block significantly mitigates the effects of both Gaussian noise and impulse noise on the model.

Regarding the noise sensitivity of the anti-noise block, under the same noise level, the influence of Gaussian noise is nearly identical to impulse noise, indicating that the primary sensitivity of the module is attributed to the intensity of the noise. From the results under different noise levels, it is evident that the R^2 of the baseline model decreases by 6%–10% when the noise level increases from $\alpha = 0.1$ to $\alpha = 0.2$. In

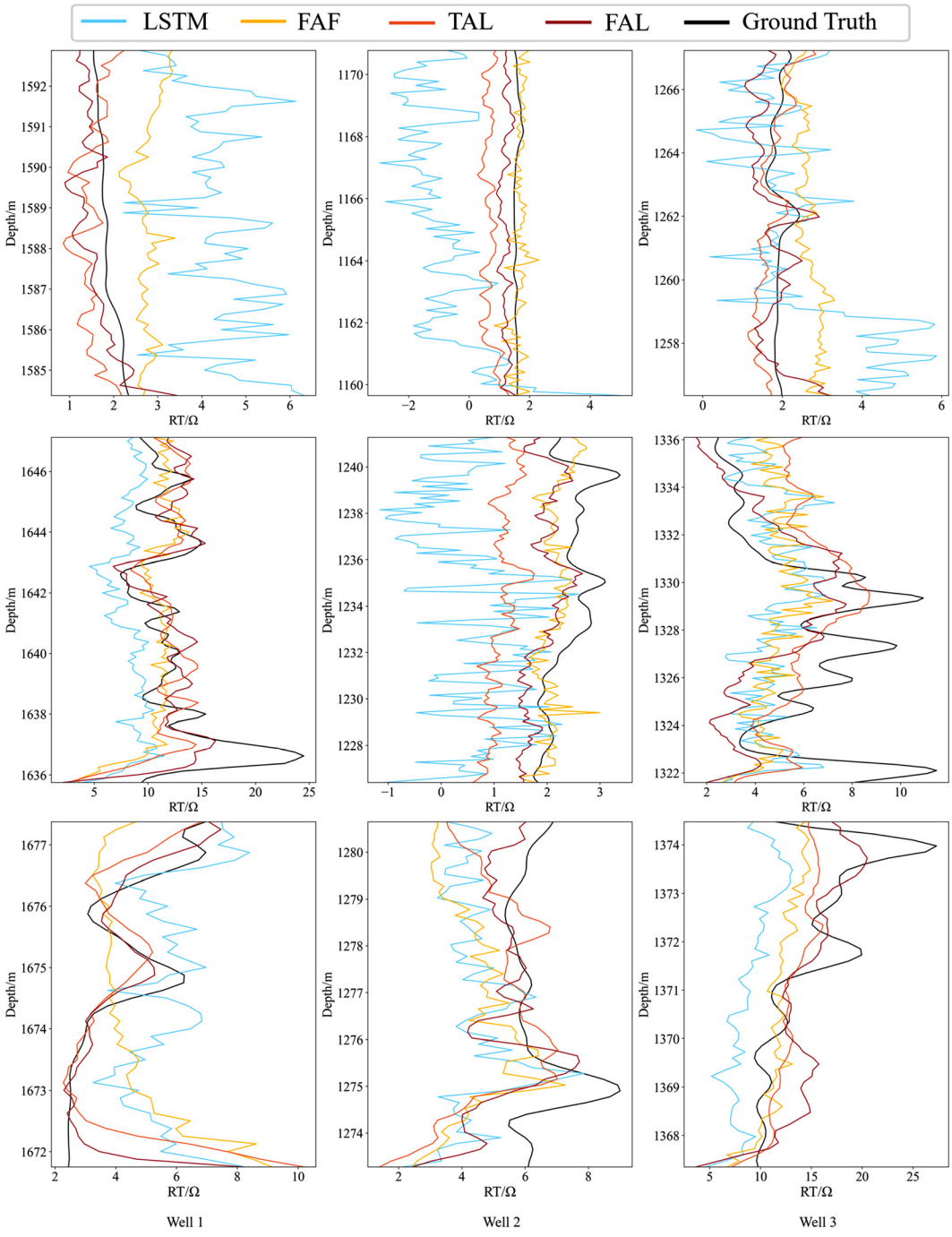


(a) results of data with Gaussian noise of 0.1 in α

FIG. 11. Result of missing well logs prediction with Gaussian noise.

contrast, with the anti-noise block, the decrease in R^2 under the same noise increase is only 1%–2% from $\alpha = 0.1$ to $\alpha = 0.2$. Consequently, in terms of sensitivity to noise intensity, the anti-noise block significantly reduces the model's sensitivity to the intensity.

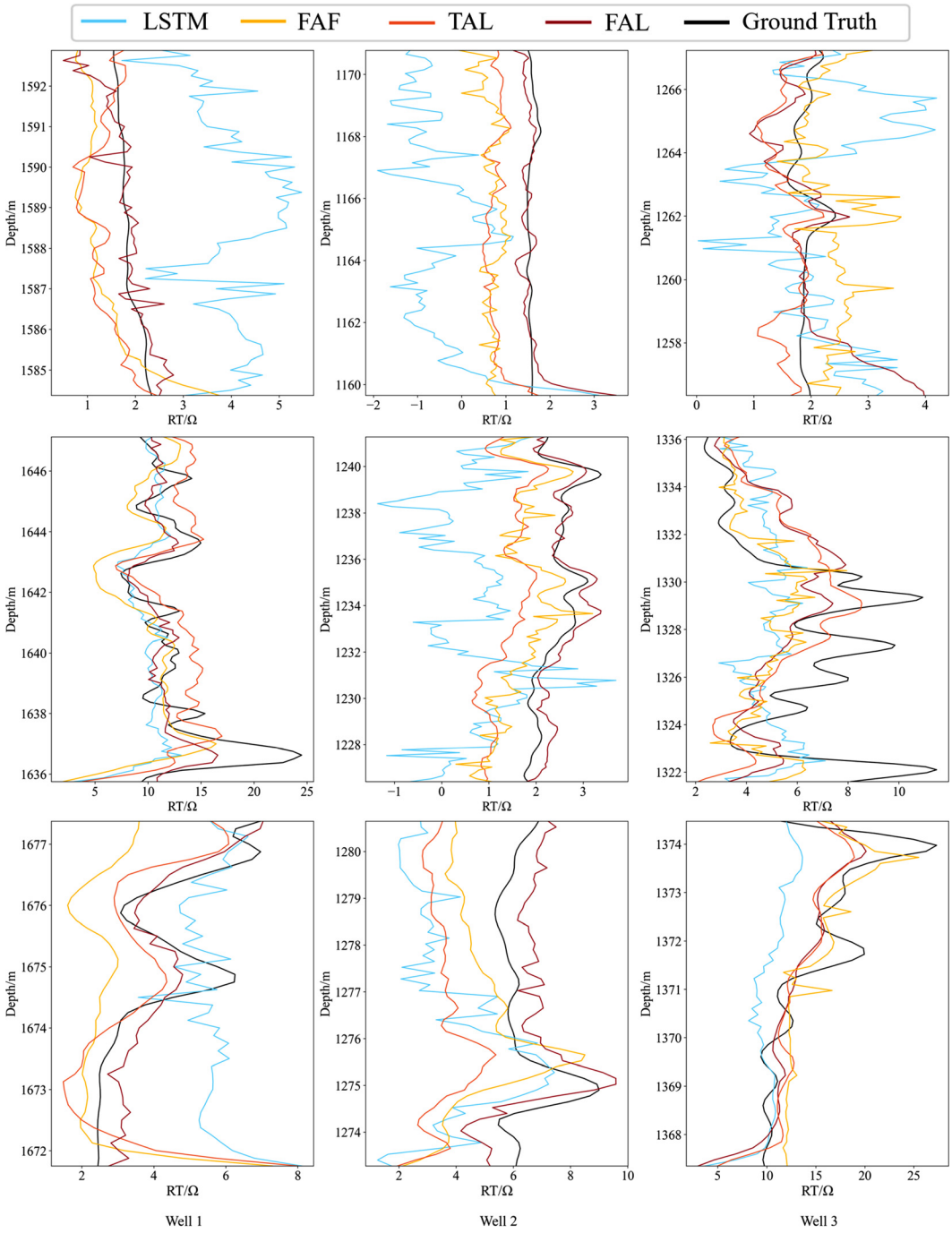
From the engineering perspective, the proposed model effectively addresses the challenges of high-frequency component isolation and noise interference in formation resistivity prediction by introducing a dual-stream structure with wavelet transformation.



(b) results of data with Gaussian noise of 0.2 in α

FIG. 11. (*Continued.*)

This dual-stream structure not only enables precise separation of high-frequency components but also enhances the model's ability to capture complex resistivity features, thus improving the accuracy of subsurface fluid behavior prediction. Additionally, the integration of temporal anti-noise blocks significantly strengthens the model's noise resistance, ensuring stability even in noisy environments. As a result, the model achieves an R^2 value of 0.91 in resistivity prediction, with minimal R^2 reductions of only 2.2% and

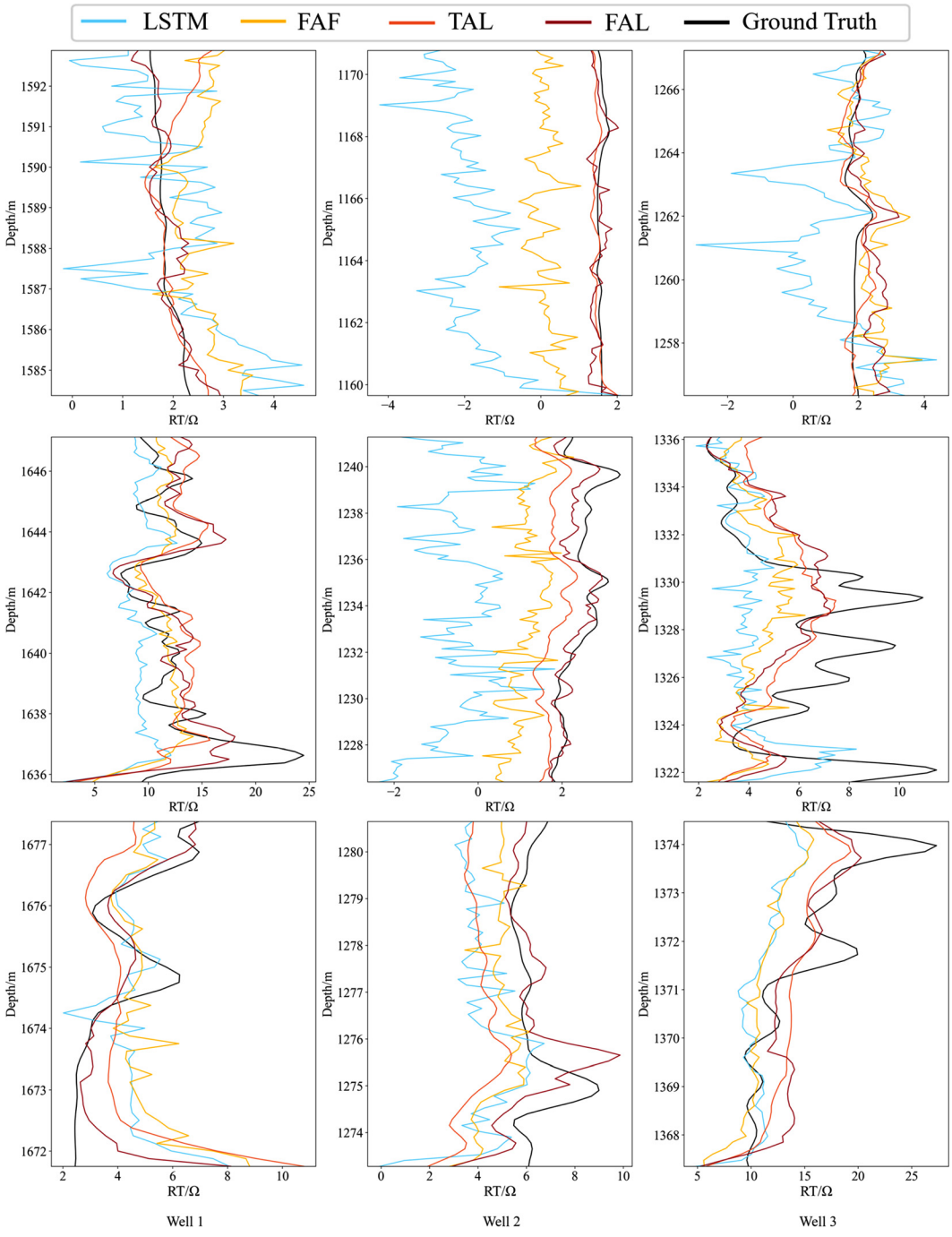


(a) results of data with Impulse noise of 0.1 in α

FIG. 12. Result of missing well logs prediction with impulse noise.

1.6% under Gaussian and impulse noise, respectively (roughly 1/8 of the baseline's reduction), demonstrating its potential for high-accuracy prediction in complex geological settings, especially in high-noise and challenging subsurface layers.

The value of this approach lies in its robustness and adaptability. In practical applications, the ability to accurately predict resistivity from noisy TEM data will greatly enhance the reliability of subsurface fluid dynamics assessments. For instance, in



(b) results of data with Impulse noise of 0.2 in α

FIG. 12. (*Continued.*)

hydrocarbon exploration and groundwater resource management, accurate evaluation of formation resistivity is critical for identifying potential oil and gas reservoirs and monitoring fluid migration. The model's ability to combine high-frequency feature learning with strong noise filtering, as demonstrated in this study, suggests that it can be deployed in noisy field environments. This offers a novel solution for resistivity prediction in high-noise subsurface environments, providing new insight and approaches for

subsurface fluid behavior analysis in such conditions, thus advancing the field of subsurface fluid prediction.

IV. CONCLUSIONS

Traditional methods struggle to accurately measure subsurface fluid parameters from cased wells. Although TEM data can characterize the physical properties of subsurface fluids, noise issues and high-frequency disaster led to inadequate prediction accuracy for these properties. To address these challenges, this paper proposes a frequency-aware framework and a temporal anti-noise block, designed to leverage TEM data for more accurate prediction of a crucial subsurface fluid property, formation resistivity.

The model's ability to independently learn both high and low-frequency components enhances the accuracy of formation resistivity prediction. The integration of multi-head and soft-threshold attention mechanisms effectively mitigates noise interference, ensuring more reliable predictions even under noisy conditions. Validation results further show that frequency-aware framework outperforms baseline model in formation resistivity prediction using TEM data, achieving an R^2 of 0.91, with an improvement of 0.3 in R^2 at high-frequency part prediction. The frequency-aware framework also demonstrates strong noise resistance in formation resistivity prediction using TEM data, with minimal R^2 reductions of 2.2% and 1.6% under Gaussian and impulse noise, respectively, compared to an approximately eightfold greater reduction in the baseline model.

However, the model does face challenges in predicting large-amplitude signals, which may be attributed to the insufficient availability of fluid data. Future work will focus on incorporating physical knowledge embedding and expanding fluid parameter datasets to improve prediction accuracy. Overall, this work establishes a data foundation for advancing the study of subsurface fluids, such as groundwater monitoring, reservoir fluid characterization, and other subsurface fluid-related investigations, while also provides valuable insights for subsurface investigations under noisy environments.

SUPPLEMENTARY MATERIAL

See the supplementary material for additional details on the datasets and results used in the study. The dataset described in Sec. III A includes three wells, as shown in the supplementary materials. In supplementary materials, Fig. 1 illustrates three wells' data used in experience of missing well logs prediction. It includes a total of 35 transient electromagnetic logging curves through the casing and 4 conventional logging curves. The dataset in Sec. III B includes data from the same three wells as in Sec. III A. In supplementary materials, Fig. 2 presents the logging curve diagrams of four conventional logging curves under two types of noise conditions. Similarly, Fig. 3 shows the logging curve diagrams of 35 logging curves through the casing under the same two types of noise conditions. These figures are essential for understanding the robustness of the model under noisy data conditions.

ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Program (Grant No. 2024YFF1500600), the Natural Science Foundation of Ningbo of China (No. 2023J027), and the High Performance Computing Centers at Eastern Institute of Technology, Ningbo, and Ningbo Institute of Digital Twin.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Yongan Zhang and Jian Li contributed equally to this work.

Yongan Zhang: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Validation (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Jian Li:** Conceptualization (equal); Data curation (equal); Investigation (equal); Methodology (equal); Validation (equal); Writing – original draft (equal); Writing – review & editing (equal). **Junfeng Zhao:** Data curation (equal); Formal analysis (equal); Funding acquisition (equal); Resources (equal); Writing – review & editing (equal). **Xuanran Wang:** Data curation (equal); Resources (equal); Validation (equal); Writing – review & editing (equal). **Youzhuang Sun:** Data curation (equal); Methodology (equal); Validation (equal); Writing – review & editing (equal). **Yizheng Li:** Formal analysis (equal); Methodology (equal); Validation (equal); Writing – original draft (equal); Writing – review & editing (equal). **Yuntian Chen:** Conceptualization (equal); Data curation (equal); Funding acquisition (equal); Methodology (equal); Project administration (equal); Supervision (equal); Writing – review & editing (equal). **Dongxiao Zhang:** Conceptualization (equal); Data curation (equal); Funding acquisition (equal); Resources (equal); Supervision (equal); Validation (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data used to support the findings of this study are available from the corresponding authors upon request.

APPENDIX: DATA PREPARATION AND EXPERIMENTAL SETUP

1. Data and pretreatment

The data preprocessing involves two main components: granularity alignment and standardization.

(1) Granularity Alignment: Primarily focusing on the granularity of through-casing transient electromagnetic logging curves, the conventional well logging curves (granularity of 0.01 m) are sparsely aligned with and adjusted to match the granularity of the through-casing transient electromagnetic logging curves (granularity of 0.125 m). The entire process ensures depth consistency, as detailed follows:

$$s_i = t_{sj}, \tag{A1}$$

where s represents the value corresponding to the index i of the sparsely processed data, where the granularity is 0.125m. t_{sj} denotes the value corresponding to the index $8j$ of the original data, with a granularity of 0.01m.

(2) Standardization: Standardization is performed using conventional normalization methods, as specified as

$$X_{std} = \frac{X - u}{\sigma}, \tag{A2}$$

where X_{std} represents the standardized data, X is the original data, u is the mean of the data, and σ is the standard deviation of the data.

2. Experiment setting

Before model training, we employed a time window method with depth as the sequence dimension for data training and prediction, as illustrated in the figure below. The sequence length is determined by the reservoir homogeneity of the logging data source, and for this experiment, the sequence length is set to 5 m (equivalent to the length of 40 samples). Figure 9 depicts the correspondence between the training data samples and labels in the experiment. The "train window" represents the sliding window in the time window method, sliding at intervals of one sample point each time. The data within each window serves as an independent sample in each training iteration, and the corresponding target data for the depth sequence acts as an independent label. Since the experimental sequence output results involve repetitive predictions at various depths, the mean of the output results at each depth in this experiment is considered as the true output result.

In terms of model hyperparameters, the sequence length is set to 40 (corresponding to a depth interval of 5 m), the batch size is set to 64, and the number of features is 50 (indicating the number of inputs well log curves). The parameter settings are summarized in Table IV. The sequence length is set to 40, the batch size is configured as 64, and the number of features is specified as 50. The model training is terminated when the training loss stabilizes (change less than 0.001). The parameters of the network structure are outlined in Table V (BN: Batch Normalization layer, ReLU: Rectified Linear Unit layer, FC: Fully Connected layer, and GAP: Global Average Pooling layer).

During the model training, the input parameters consist of the previously mentioned through-casing transient electromagnetic logging curves and conventional well logging data. The output results comprise formation resistivity well log curves. The three wells are uniformly segmented into training, validation, and testing sets. Specifically, the data are segmented into three sections based on depth order, with the first 70%, middle 10%, and last 20% forming the training, validation, and testing sets in each section. It is noteworthy that the validation set is employed to evaluate the model after each epoch, and the model parameters from the epoch with the highest validation accuracy are selected as the final model parameters. The final dataset partition ratio is as follows: training set: validation set: testing set = 0.7:0.1:0.2, as illustrated in Fig. 10.

3. Detail of robustness analysis with noisy data

Figure 11. Result of missing well logs prediction with Gaussian noise.

Figure 12. Result of missing well logs prediction with impulse noise.

REFERENCES

[1]A. I. Riwayat, M. A. A. Nazri, and M. H. Z. Abidin, "Application of electrical resistivity method (ERM) in groundwater exploration," J. Phys. **995**, 012094 (2018).

[2]P. K. Srivastava and A. K. Bhattacharya, "Groundwater assessment through an integrated approach using remote sensing, GIS and resistivity techniques: A case study from a hard rock terrain," Int. J. Remote Sens. **27**, 4599–4620 (2006).

[3]E. Brothelande *et al.*, "Fluid circulation pattern inside La Soufrière volcano (Guadeloupe) inferred from combined electrical resistivity tomography, self-potential, soil temperature and diffuse degassing measurements," J. Volcanol. Geotherm. Res. **288**, 105–122 (2014).

[4]J. B. Spacapan *et al.*, "Low resistivity zones at contacts of igneous intrusions emplaced in organic-rich formations and their implications on fluid flow and petroleum systems: A case study in the northern Neuquén Basin, Argentina," Basin Res. **32**, 3–24 (2020).

[5]K. Xiao, Z. Duan, Y. Yang, H. Li, Z. Qin, and Q. Luo, "Experimental study of relationship among acoustic wave, resistivity and fluid saturation in coalbed methane reservoir," Acta Geophys. **71**, 1241–1249 (2022).

[6]Y. Chen and D. Zhang, "Theory-guided deep-learning for electrical load forecasting (TgDLF) via ensemble long short-term memory," Adv. Appl. Energy **1**, 100004 (2021).

[7]J. Gao, Y. Chen, W. Hu, and D. Zhang, "An adaptive deep-learning load forecasting framework by integrating transformer and domain knowledge," Adv. Appl. Energy **10**, 100142 (2023).

[8]Z. Yu, Y. Sun, J. Zhang, Y. Zhang, and Z. Liu, "Gated recurrent unit neural network (GRU) based on quantile regression (QR) predicts reservoir parameters through well logging data," Front. Earth Sci. **11**, 1087385 (2023).

[9]Y. Sun, J. Zhang, Z. Yu, Y. Zhang, and Z. Liu, "Bidirectional long short-term neural network based on the attention mechanism of the residual neural network (ResNet–BiLSTM–attention) predicts porosity through well logging parameters," ACS Omega **8**, 24083–24092 (2023).

[10]S. Youzhuang, Z. Junhua, and Z. Yongan, "Echo state neural network based on an improved gray wolf algorithm predicts porosity through logging data," ACS Omega **8**, 21182–21194 (2023).

[11]M. Xiang, P. Qin, and F. Zhang, "Research and application of logging lithology identification for igneous reservoirs based on deep learning," J. Appl. Geophys. **173**, 103929 (2020).

[12]Y. Zhang, X. Zhang, Y. Sun, A. Gong, and M. Li, "An adaptive ensemble learning by opposite multiverse optimizer and its application in fluid identification for unconventional oil reservoirs," Front. Earth Sci. **11**, 1116664 (2023).

[13]C. Xi and R. E. N. Zhanli, "Artificial intelligence logging: Fundamental, principle, technique, and application," Coal Geol. Explor. **52**(8), 15 (2024).

[14]Y. Sun, J. Zhang, and Y. Zhang, "Adaboost algorithm combined multiple random forest models (Adaboost-RF) is employed for fluid prediction using well logging data," Phys. Fluids **36**(1), 016602 (2024).

[15]S. Karimpouli, N. Fathianpour, and J. Roohi, "A new approach to improve neural networks' algorithm in permeability prediction of petroleum reservoirs using supervised committee machine neural network (SCMNN)," J. Pet. Sci. Eng. **73**(3–4), 227–232 (2010).

[16]M. Korjani, A. Popa, E. Grijalva, S. Cassidy, and I. Ershaghi, "A new approach to reservoir characterization using deep learning neural networks," paper presented at the SPE Western Regional Meeting, Anchorage, Alaska, 2016.

[17]G. K. Parapuram, M. Mokhtari, and J. B. Hmida, "Prediction and analysis of geomechanical properties of the upper bakken shale utilizing artificial intelligence and data mining," paper presented at the SPE/AAPG/SEG Unconventional Resources Technology Conference, Austin, Texas, 2017.

[18]O. Akinnikawe, S. Lyne, and J. Roberts, "Synthetic well log generation using machine learning techniques," paper presented at the SPE/AAPG/SEG Unconventional Resources Technology Conference, Houston, Texas (2018).

[19]D. Zhang, C. Yuntian, and M. Jin, "Synthetic well logs generation via recurrent neural networks," Pet. Explor. Dev. **45**(4), 629–639 (2018).

[20]P. Wu, V. Jain, M. S. Kulkarni, and A. Abubakar, "Machine learning-based method for automated well-log processing and interpretation," in *SEG International Exposition and Annual Meeting* (SEG) (2018).

[21]R. Akkurt, T. T. Conroy, D. Psaila, A. Paxton, J. Low, and P. Spaans, "Accelerating and enhancing petrophysical analysis with machine learning: A case study of an automated system for well log outlier detection and reconstruction," paper presented at the SPWLA 59th Annual Logging Symposium, London (2018).

[22]Y. Chen and D. Zhang, "Well log generation via ensemble long short-term memory (EnLSTM) network," Geophys. Res. Lett. **47**(23), e2020GL087685, https://doi.org/10.1029/2020GL087685 (2020).

[23]Y. Chen and D. Zhang, "Physics-constrained deep learning of geomechanical logs," IEEE Trans. Geosci. Remote Sens. **58**(8), 5932–5943 (2020).
[24]Y. Sun, J. Zhang, Z. Yu, Y. Zhang, and Z. Liu, "The bidirectional gated recurrent unit network based on the inception module (Inception-BiGRU) predicts the missing data by well logging data," ACS Omega **8**(30), 27710–27724 (2023).
[25]C. J. Schenkel and H. F. Morrison, "Electrical resistivity measurement through metal casing," Geophysics **59**(7), 1072–1082 (1994).
[26]B. Yeten, A. Castellini, B. Guyaguler, and W. H. Chen, "A comparison study on experimental design and response surface methodologies," paper presented at the SPE Reservoir Simulation Symposium, The Woodlands, Texas (2005).
[27]D. Pardo, C. Torres-VerdinTorres-Verdin, and L. F. Demkowicz, "Simulation of multifrequency borehole resistivity measurements through metal casing using a goal-oriented hp finite-element method," IEEE Trans. Geosci. Remote Sens. **44**(8), 2125–2134 (2006).
[28]X. Liu, F. Liu, J. Chen, Z. Zhao, A. Wang, and Z. Lu, "Resistivity logging through casing response of inclined fractured formation," IEEE Trans. Geosci. Remote Sens. **56**(8), 4919–4929 (2018).
[29]X. Liu and J. Chen, "Electric field and potential distribution in layered homogeneous borehole fluids and its logging responses of resistivity logging through casing," Geoenergy Sci. Eng. **224**, 211532 (2023).
[30]J. Cai and Q. Chen, "De-noising for NMR oil well logging signals based on empirical mode decomposition and independent component analysis," Arabian J. Geosci. **9**, 1–11 (2016).
[31]J. A. Nieto, D. P. Schmitt, and R. G. Keys, "Removal of borehole induced noise from well logs," paper presented at the SPWLA 36th Annual Logging Symposium, Paris, France (1995).
[32]H. Xu, Y. Chen, and D. Zhang, "Worth of prior knowledge for enhancing deep learning," Nexus **1**(1), 100003 (2024).
[33]Y. Chen and D. Zhang, "Integration of knowledge and data in machine learning," arXiv:2202.10337 (2022).
[34]X. Huang *et al.*, "Solving partial differential equations with point source based on physics-informed neural networks," arXiv:2111.01394 (2021).
[35]J. Sirignano and K. Spiliopoulos, "DGM: A deep learning algorithm for solving partial differential equations," J. Comput. Phys. **375**, 1339–1364 (2018).
[36]M. J. Shensa, "The discrete wavelet transform: Wedding the a trous and Mallat algorithms," IEEE Trans. Signal Process. **40**(10), 2464–2482 (1992).
[37]D. L. Donoho, "De-noising by soft-thresholding," IEEE Trans. Inf. Theory. **41**, 3, 613–627 (2002).
[38]M. Zhao, S. Zhong, X. Fu, B. Tang, and M. Pecht, "Deep residual shrinkage networks for fault diagnosis," IEEE Trans. Ind. Inf. **16**(7), 4681–4690 (2020).
[39]A. Vaswani *et al.*, "Attention is all you need," *Advances in Neural Information Processing Systems* (NeurIPS, 2017), Vol. 30.