2025 5th International Conference on Artificial Intelligence and Industrial Technology Applications (AIITA)

A Method for Forecasting the DC Support Capacitance of Traction Converters Based on Time-Frequency Analysis and Improved LSTM Networks

Runze Lian
Intelligentization Research Institute
China Academy of Industrial Internet
Beijing, China
lianrunze@china-aii.com

Hongli Shi
Research and Development Department
RDFZ Shijingshan School
Beijing, China
shihongli@rdfzsjsxx.cn

*Weixi Gu
Intelligentization Research Institute
China Academy of Industrial Internet
Beijing, China
lrz_1988@163.com

Lin Li
Data management and application Institute
China Academy of Industrial Internet
Beijing, China
lilin@china-aii.com

***Abstract*—This paper introduces an innovative approach for predicting the DC support capacitor value in train traction converters by synergizing spectrum analysis, data mining, and sophisticated machine learning techniques. The methodology is underpinned by the application of the Fast Fourier Transform (FFT) for spectral decomposition and the Long Short-Term Memory (LSTM) deep learning model for predictive analytics. A pivotal contribution of this research is the employment of Grey Relational Analysis (GRA) algorithm, which transforms the causal analysis of long-time series sequences with single discrete variables into a modeling of low-order correlation. The paper delineates the refinement of an LSTM architecture, marked by an optimized network structure and meticulously calibrated hyperparameters, culminating in a resilient prediction model. The methodology's effectiveness is corroborated through extensive testing against actual aging data of DC support capacitors. The results indicate that the proposed LSTM model excels in the accurate forecasting of capacitor values, thereby endorsing the approach's practicality and reliability in the maintenance and enhancement of railway traction converters.**

Keywords-railway traction converters; DC support capacitors; spectrum analysis; causal analysis of long-time series sequences; LSTM model

I. Introduction

DC support capacitors are key components in the traction converters of rail transit vehicles, as shown in the Figure 1, playing a crucial role in the vehicle's traction rectification and inversion processes [1]:

1) Compensation for the reactive power of the motor

2) Energy regulation during load changes

3) Support the intermediate DC circuit voltage to keep it stable

However, as the usage time increases, support capacitors experience aging [2]. Changes in capacitance can cause severe faults in vehicle operation, leading to prolonged downtime for maintenance and inspection, as well as economic losses due to Vehicle shutdown.

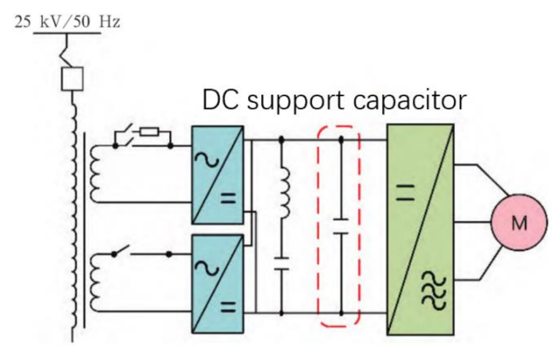

Figure 1. Traction converter topology of a certain high-speed train.

For example, in the operation of domestic urban subways, there have been instances where the aging of support capacitors has caused the traction system of vehicles to shut down, resulting in trains being unable to leave the depot [2]. Therefore, timely analysis of the capacitance of support capacitors is extremely important.

Normally, the capacitance can only be directly measured when the vehicle is in the depot for maintenance and inspection at night, and professional testing equipment is required. Additionally, the testing process is time-consuming, complex, and when the locomotive stops operating, even if the pantograph has been lowered, the support capacitor may still carry a dangerous high voltage for several days [1]. Traditional methods of testing support capacitors cannot promptly reflect changes in their capacitance.

The working state of the traction converter circuit is closely related to the capacitance of the support capacitors. This paper proposes to use measurable and collectible time-series data of the traction converter, such as grid voltage, Filter Capacitor (FC) voltage and the three-phase (u, v, w) alternating current, to conduct data analysis and mining. Through machine learning and deep learning algorithms, we aim to find the relationship between the relevant data and the capacitance of the support capacitors.

Due to the advancement in chip computing power and the development of artificial intelligence algorithms, time-series data analysis and prediction technologies [28] based on machine learning and deep learning have seen rapid improvement [25, 27]. The ARIMA model [3, 4] addresses the challenge of prediction by converting non-stationary processes into stationary ones through the process of differencing. Additionally, a filtering technique is implemented for forecasting purposes [5, 6]. LSTM networks [7] have inherited most characteristics of traditional recurrent neural networks (RNNs), mitigating the issue of vanishing gradients during backpropagation. However, due to their complex structure, LSTMs struggle with efficient parallel data processing. Autoregressive methods and RNNs are combined to construct DeepAR model [8] for the probabilistic distribution of future series. LSTNet [9] incorporates CNNs with recurrent-skip connections to identify both short-term and long-term temporal patterns. Attention-based RNNs [10, 11, 12] utilize temporal attention mechanisms to uncover long-range dependencies crucial for forecasting. Furthermore, numerous studies leveraging temporal convolutional networks (TCN) [13, 14, 15, 16] aim to capture temporal causality through causal convolution. These advanced forecasting models [26] primarily concentrate on modeling temporal relationships through recurrent connections, temporal attention mechanisms, or causal convolution techniques.

However, the traction converter time-series data related to the capacitance of the support capacitors have a high sampling frequency and are extremely voluminous, making it challenging to extract long sequence time-series features and model the relationship with capacitance using the aforementioned methods. Additionally, studying the short-term and long-term patterns of time series changes is also difficult to apply to the establishment of models between long sequence time-series and discrete variables, namely, the capacitance of the support capacitors. Experiments have shown that directly applying these methods to conduct time-domain data analysis and time series forecasting for the capacitance of DC support capacitors yields poor results.

This paper proposes a method for analyzing and estimating the capacitance of DC support capacitors in traction converters based on Fast Fourier Transform (FFT) and an improved LSTM. Applying FFT to the original time-series signals of the traction converter circuit, we obtain spectral data while eliminating high-frequency noise present in the original signals. Grey Relational Analysis (GRA) is used to analyze the correlation between spectral data and the target variable. An improved LSTM algorithm is employed to establish a model for analyzing and estimating the capacitance of support capacitors. Using actual measurement and aging experiment data from vehicle traction converters, the data-driven approach proposed in this paper is used to predict the capacitance of DC support capacitors. This method is compared and evaluated against other methods such as GRU [17, 18], Conformer [19], and Autoformer [20, 25], to verify the effectiveness of the proposed support capacitor estimation method. This provides a safeguard plan for the stable and safe operation of locomotive traction converters.

II. METHODS

This study proposes an analysis and prediction method for the capacitance of support capacitors in traction converters, with LSTM as the backbone network, as shown in Figure 2. Initially, the converter time-series signals after Fast Fourier Transform (FFT) are subjected to GRA feature analysis. A unique streamlined LSTM model is designed, with corresponding normalization layers and fully connected layers adjusted. The selected feature dataset is used to train and optimize the model, achieving analysis and predicting of the DC support capacitor capacitance.

Compared with traditional time-series analysis models, the proposed DC support capacitor capacitance predicting model has the following advantages:

(1) Converting continuous traction converter time-series signals into finite discrete data features through Fast Fourier Transform, the complex long-sequence and discrete numerical causality association analysis is dimensionally reduced, facilitating the establishment of a mapping model between data features and target variables. This also reduces the amount of computation and eliminates the interference of high-frequency noise from the traction converter circuit signals.

(2) The introduction of time-frequency domain data analysis methods for processing raw data achieves long-sequence data enhancement and feature extraction, addressing the complexity of constructing input sequences for the LSTM network in this method.

(3) The design of multiple LSTM network units [29, 30], with optimization of network normalization layers and fully connected layer design, balances the model's high-order nonlinear expression capability and computational efficiency.

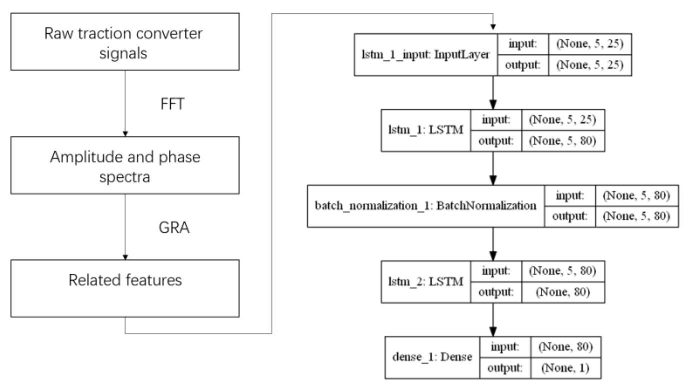

Figure 2. The network structure of proposed model.

A. Fast Fourier Transform

Fourier transform is a form of transformation that converts signals from the time domain to the frequency domain; whereas FFT is an efficient method for implementing the Discrete Fourier Transform (DFT) [21]. For a finite-length digital signal x(n), its frequency domain signal X(k) after DFT can be expressed as:

$$X(k) = \sum_{n=0}^{N-1} x(n) W_N^{nk} \quad (k = 0,1,2,\dots,N-1; n = 0,1,2,\dots,N-1) \tag{1}$$

where W_N is the twiddle factor; N is the length of x(n); n is the discrete time index of x(n), representing the sampling point sequence number of the time-domain signal; k is the frequency index of X(k), representing different frequency components in the frequency domain.

The twiddle factor can be expressed as:

$$W_N = e^{\frac{-j2\pi}{N}} \tag{2}$$

The imaginary unit is denoted by j; e represents the base of the natural logarithm. The FFT leverages the symmetry and periodicity of the twiddle factor to decompose the DFT matrix, breaking down the computation of a long sequence DFT into multiple short sequence DFTs.

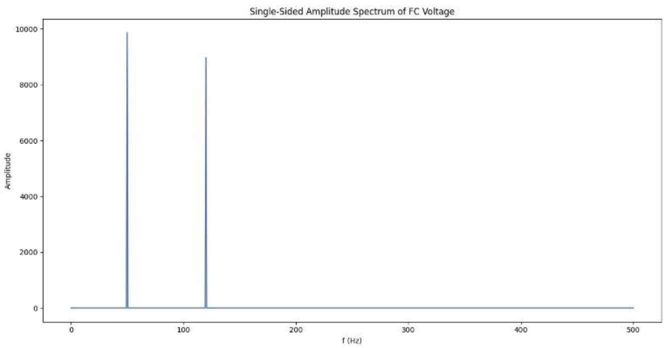

Figure 3. Amplitude spectra of traction converter signal (FC Voltage).

The precision of the FFT computation matches that of the direct DFT computation. This method employs the Fast Fourier Transform to convert the high-frequency signals (10^5Hz）of the traction converter circuit into their corresponding spectral information. That is, through spectral analysis, one can obtain the amplitude or phase spectra of each traction converter signal in the frequency domain, as shown in the Figure 3.

B. Grey Relation Analysis (GRA)

The Grey Relation Analysis (GRA) serves as a sophisticated statistical technique designed for the quantitative evaluation of relationships among multiple variables. Applying the GRA methodology, the analysis quantifies the degree of association between a primary dataset, which represents the objective variable of interest, and a secondary dataset comprising the spectral values of related traction converter signals. This approach allows for the assessment of how closely the various signals from the traction converter are related to the target variable [22].

The process involved scaling the DC capacitor support values and the potential feature sequences through Min-Max normalization. Subsequently, the grey relational degrees were calculated for each pair of traction converter signal sequences. The outcomes of the grey relational analysis for the traction converter signals are visually summarized in a heat map format, as depicted in Figure 4. Within this heat map, the values at each intersection point correspond to the grey relational degree of the converter signals aligned along the x and y axes, while the color saturation of each cell reflects the magnitude of the grey relational degree among the traction converter parameters.

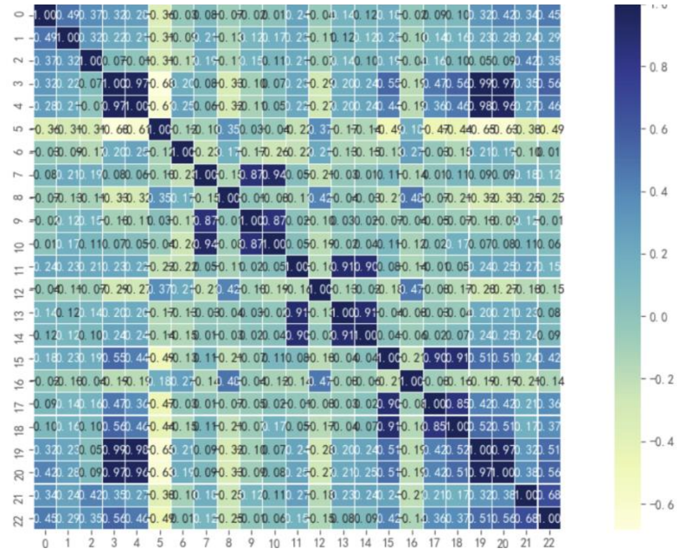

Figure 4. GRA of traction converter signal and DC support capacitance.

In Figure 4, the label 19 is associated with the DC capacitor support value. The top five traction converter parameters (Peak values of FFT spectrum) that exhibit the strongest grey relational degrees with the target variable were identified and are detailed in Table 1.

TABLE 1. GRA OF DC SUPPORT CAPACITANCE

GRA	DC support capacitance
FC voltage	0.91
FC current	0.98
alternating current u	0.91
alternating current v	0.95
alternating current w	0.92

These features include the FC voltage, FC current, and the three-phase (u, v, w) alternating current.

C. Improved LSTM Prediction Model

The predictive model for estimating the DC capacitor support value was developed using a Long Short-Term Memory (LSTM) deep neural network. Unlike conventional machine learning approaches that may falter in recognizing long-term patterns, LSTMs, equipped with their memory cells, excel at capturing extended temporal relationships, which significantly boosts the precision of predictions. This attribute is crucial for managing the intricate dynamics of traction converter performance and stands as a major advantage of LSTMs in this sector [23, 24]. The integration of LSTM with different neural network architectures can also be employed to incorporate spatial dimensions of the parameters. This adaptability allows the model to cater to the unique demands of a variety of traction converters and DC capacitor support systems.

In the development of the DC support capacitance prediction model presented in this paper, a configuration of two sequentially arranged LSTM recurrent structures was implemented, as shown in Table 2. Each of these recurrent structures was designed with a capacity of 80 neuron nodes. The training process involved an adaptive learning rate approach, starting at 0.01, to refine the LSTM network's parameters through iterative backpropagation. The optimization algorithm of choice for the LSTM model's training was the Adam function, which dynamically adjusts the learning rate for each parameter based on the magnitude of the gradient during training. To promote stable gradient flow and expedite the convergence of the model, batch normalization was incorporated prior to the input of the second LSTM recurrent structure.

TABLE 2. CONFIGURATION OF THE IMPROVED LSTM MODEL

Model configuration	Particular set
Number of hidden layer nodes	80
Initial learning rate	0.01
Dropout rate	0.25
Optimization function	Adam
Batch normalization	included

III. EVALUATION OF IMPROVED LSTM MODELS

The efficacy of the enhanced LSTM model in forecasting the DC support capacitance is gauged through two metrics: the root mean square error (RMSE) and the coefficient of determination, often denoted as R^2 Score. The RMSE, which quantifies the average magnitude of the error between the predicted and actual values, is calculated using the formula presented as equation (3):

$$RMSE = \sqrt{\frac{1}{q}\sum_{i=1}^{q}(y_i - \hat{y}_i)^2} \tag{3}$$

This metric represents the average numerical deviation for each of the q predictions made by the LSTM prediction model for DC support capacitance after conducting q predictions. The R^2 Score is defined by the following equation (4):

$$R^2 = 1 - \frac{SSE}{SST}, SSE = \sum_{i=1}^{r}(y_i - \hat{y}_i)^2, SST = \sum_{i=1}^{r}(y_i - \bar{y})^2 \tag{4}$$

When the LSTM model is applied to predict r data samples (feature sequences), where y_i represents the real DC support capacitor value corresponding to the i^{th} data sample, $\hat{y}_i$ is the predicted DC support capacitor value for the i^{th} data sample, and $\bar{y}$ is the average real DC support capacitor value for the r data samples. The R^2 Score reflects the percentage (in percentage form, with a maximum of 100%) to which the changes of DC support capacitor value can be explained by the feature parameters using the LSTM model.

IV. EXPERIMENT AND EVALUATION

The prediction performance of the LSTM model was evaluated using the RMSE and R^2 Score metrics on practical data, i.e. aging experiment data of DC support capacitor.

Conducting aging experiments and setting the variation range of the DC support capacitor of the locomotive traction converter to 0.01mF.

A	B	C	D	E	F	G	H	I
	Timestamp	a	b	c	d	e	f	g
0	0.5	770.4259	110.9956	768.031	110.9495	-10.4299	-269.824	280.2541
1	0.50001	770.4882	110.9908	768.1454	110.9447	-10.4666	-269.598	280.0645
2	0.50002	770.5526	110.9858	768.2598	110.9397	-10.5034	-269.372	279.875
3	0.50003	770.619	110.9805	768.3741	110.9344	-10.5402	-269.146	279.6857
4	0.50004	770.6873	110.9749	768.4885	110.9288	-10.577	-268.92	279.4965
5	0.50005	770.7574	110.9692	768.6028	110.923	-10.6138	-268.694	279.3075
6	0.50006	770.8292	110.9632	768.7172	110.9171	-10.6506	-268.468	279.1185
7	0.50007	770.9027	110.957	768.8315	110.9109	-10.6874	-268.242	278.9297
8	0.50008	770.9778	110.9506	768.9459	110.9045	-10.7243	-268.017	278.7411
9	0.50009	771.0544	110.944	769.0602	110.8979	-10.7611	-267.791	278.5526
10	0.5001	771.1325	110.9373	769.1745	110.8911	-10.798	-267.566	278.3642
11	0.50011	771.212	110.9303	769.2888	110.8842	-10.8349	-267.341	278.176
12	0.50012	771.2927	110.9232	769.4031	110.8771	-10.8718	-267.116	277.9879
13	0.50013	771.3746	110.916	769.5175	110.8698	-10.9088	-266.891	277.7999
14	0.50014	771.4577	110.9085	769.6317	110.8623	-10.9457	-266.666	277.6121
15	0.50015	771.5418	110.9009	769.746	110.8548	-10.9827	-266.442	277.4244

Figure 5. A case of experimental data display.

Taking the capacitance of DC support capacitor as the target variable, the data sampling frequency is 10^5Hz, we accumulate circuit signals during the charging and discharging process of the locomotive traction converter under different capacitances. For example, when the current DC support capacitance is 6.78mF, data such as grid voltage and grid current during the charging and discharging process are collected: a. Grid voltage; b. Grid current; c. FC voltage; d. FC current; e. Alternating current u; f. Alternating current v; g. Alternating current w, and saved in CSV format files, as shown in Figure 5.

The experimental results are presented in Table 3. There are a total of 15 experimental groups, with each group containing 200 test cases, amounting to a total of 3,000 test samples for the traction converter capacitance.

TABLE 3. EVALUATION RESULTS OF 15 EXPERIMENTAL GROUPS

Exp.#	RMSE	R^2(%)
1	0.002	97.1
2	0.005	96.2
3	0.006	95.3
4	0.008	93.1
5	0.010	94.2
6	0.011	92.2
7	0.018	90.4
8	0.019	90.1
9	0.012	91.0
10	0.007	94.0
11	0.009	93.8
12	0.013	91.2
13	0.017	91.9
14	0.021	90.7
15	0.020	90.6

Upon examining Table 3, it becomes clear that the refined LSTM model demonstrates superior performance in the task of predicting the DC support capacitance. The RMSE values are consistently below 0.005 mF, while the R^2 Score remains above 90%, indicating high prediction quality. For a more graphical understanding, Table 4 presents the outcomes of five individual tests from Experiment #1, compared with those of alternative models. Parallel assessments were conducted with GRU [17, 18], Conformer algorithms [19] and Autoformer [20, 26].

TABLE 4. COMPARISON OF 5 INDIVIDUAL TESTS WITH ALTERNATIVE MODELS

Ground Truth(/mF)	9.72	9.74	9.76	9.77	9.80
Novel LSTM	9.713	9.742	9.773	9.772	9.795
GRU	3.678	5.786	7.980	7.897	2.678
Conformer	9.145	8.956	9.243	9.683	9.233
Autoformer	8.978	8.781	8.987	8.654	9.198

V. CONCLUSIONS

This paper elucidates a comprehensive methodology for forecasting the DC support capacitor value of a traction converter, leveraging particularly the Fast Fourier Transform (FFT) and the Long Short-Term Memory (LSTM) deep neural network model. The study employs the Grey Relational Analysis (GRA) method for feature extraction, which quantitatively measures the interrelationships among traction converter signals and the DC support capacitor. The construction of an improved LSTM neural network involves the design of a streamlined network structure, the training of the prediction model, and the meticulous adjustment and optimization of LSTM model parameters. To validate the methodology, tests were conducted using aging experiment data of the DC support capacitor. The evaluation results demonstrate that the proposed LSTM model is capable of accurately predicting the DC support capacitor value of the traction converter, thereby affirming the efficacy and reliability of the proposed approach.

REFERENCES

[1] Xiangling Chen. DC support capacitor in HXD1 locomotive converter [J]. electrical technology, 2011(12): 77-79.DOI: 10.3969/j.issn.1673-3800.2011.12.043.

[2] Xiaoming Xu, Runhao Ren. Life Evaluation Method of DC-Link Capacitor Considering ESR Aging [J]. Railway rolling stock, 2024, 44(04): 44-50.

[3] G. E. P. Box and Gwilym M. Jenkins. Time series analysis, forecasting and control. 1970.

[4] George EP Box and Gwilym M Jenkins. Some recent advances in forecasting and control. J. R. Stat. Soc. (Series-C), 1968.

[5] Emmanuel de Bézenac, Syama Sundar Rangapuram, Konstantinos Benidis, Michael Bohlke-Schneider,Richard Kurle, Lorenzo Stella, Hilaf Hasson, Patrick Gallinari, and Tim Januschowski. Normalizing kalman filters for multivariate time series analysis. In NeurIPS, 2020.

[6] Richard Kurle, Syama Sundar Rangapuram, Emmanuel de Bézenac, Stephan Günnemann, and Jan Gasthaus.Deep rao-blackwellised particle filters for time series forecasting. In NeurIPS, 2020.

[7] Rob J Hyndman and George Athanasopoulos. Forecasting: principles and practice. 2018.

[8] David Salinas, Valentin Flunkert, Jan Gasthaus, and Tim Januschowski. DeepAR: Probabilistic forecasting with autoregressive recurrent networks. Int. J. Forecast., 2020.

[9] Guokun Lai, Wei-Cheng Chang, Yiming Yang, and Hanxiao Liu. Modeling long-and short-term temporal patterns with deep neural networks. In SIGIR, 2018.

[10] Shun-Yao Shih, Fan-Keng Sun, and Hung-yi Lee. Temporal pattern attention for multivariate time series forecasting. Mach. Learn., 2019.

[11] Huan Song, Deepta Rajan, Jayaraman Thiagarajan, and Andreas Spanias. Attend and diagnose: Clinical time series analysis using attention models. In AAAI, 2018.

[12] Q. Yao, D. Song, H. Chen, C. Wei, and G. W. Cottrell. A dual-stage attention-based recurrent neural network for time series prediction. In IJCAI, 2017.

[13] Shaojie Bai, J Zico Kolter, and Vladlen Koltun. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. arXiv preprint arXiv:1803.01271, 2018.

[14] Anastasia Borovykh, Sander Bohte, and Cornelis W Oosterlee. Conditional time series forecasting with convolutional neural networks. arXiv preprint arXiv:1703.04691, 2017.

[15] Rajat Sen, Hsiang-Fu Yu, and Inderjit S. Dhillon. Think globally, act locally: A deep neural network approach to high-dimensional time series forecasting. In NeurIPS, 2019.

[16] Aäron van den Oord, S. Dieleman, H. Zen, K. Simonyan, Oriol Vinyals, A. Graves, Nal Kalchbrenner,A. Senior, and K. Kavukcuoglu. Wavenet: A generative model for raw audio. In SSW, 2016.

[17] CHO K, VAN MERRIËNBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[J]. arXiv:1406.1078, 2014.

[18] FU R, ZHANG Z, LI L. Using LSTM and GRU neural network methods for traffic flow prediction[C]//Proceedings of the 2016 31st Youth Academic Annual Conference of Chinese Association of Automation, Wuhan, Nov 11-13, 2016. Piscataway: IEEE, 2016: 324-328.

[19] LI Y, LU X, XIONG H, et al. Towards long-term time-series forecasting: feature, pattern, and distribution[J]. arXiv:2301.02068, 2023.

[20] Wu, H., Xu, J., Wang, J., & Long, M. (2021). Autoformer: Decomposition Transformers with Auto-Correlation for Long-Term Series Forecasting. Neural Information Processing Systems.

[21] Wu Zhiquan, Yu Haiying, Wang Song. RESEARCH ON HEALTH STATUS MONITORING OF HIGH-SPEED SHAFT GEARS IN WIND TURBINE GEARBOXES BASED ON DSP-LightGBM ALGORITHM. Solar Energy. DOI: 10.19911/j.1003-0417.tyn20231208.01

[22] Tian, X.; Wang, Z.; Taalab, E.; Zhang, B.; Li, X.; Wang, J.; Ong, M.C.; Zhu, Z. Water Quality Predictions Based on Grey Relation Analysis Enhanced LSTM Algorithms. Water 2022, 14, 3851.

[23] Zhu, Hehua, Wang, Xin, Chen, Xueqin, and Zhang, Lianyang. "Similarity Search and Performance Prediction of Shield Tunnels in Operation through Time Series Data Mining." Automation in Construction 114 (2020): 103178.

[24] Elbaz, Khalid, Yan, Tao, Zhou, Annan, and Shen, Shui-Long. "Deep Learning Analysis for Energy Consumption of Shield Tunneling Machine Drive System." Tunnelling and Underground Space Technology 123 (2022): 104405.

[25] Haoyi Zhou, Shanghang Zhang, Jieqi Peng, Shuai Zhang, Jianxin Li, Hui Xiong, and Wancai Zhang. Informer: Beyond efficient transformer for long sequence time-series forecasting. In AAAI, 2021.

[26] Fuzhao Xue, Jianghai Chen, Aixin Sun, Xiaozhe Ren, Zangwei Zheng, Xiaoxin He, Yongming Chen, Xin Jiang, and Yang You. A study on transformer configuration and training objective, 2023.

[27] Antoine Simoulin and Benoit Crabbe. How many layers and why? An analysis of the model depth in transformers. Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: Student Research Workshop, pp. 221–228, Online, August 2021.

[28] Bi J , Cheng H , Zhang W ,et al. Artificial Intelligence in Ship Trajectory Prediction[J]. Journal of Marine Science & Engineering, 2024, 12(5).DOI:10.3390/jmse12050769.

[29] Yi Y , Huang Z , Bao M ,et al. Multi-step Short-term Load Forecasting Based on Attention Mechanism, TCN-BiLSTM Network and Decomposition-based Error Correction[C]. 2024 7th Asia Conference on Energy and Electrical Engineering (ACEEE).0[2024-12-16]. DOI:10.1109/ACEEE62329.2024.10651918.

[30] Qian T, Liu Y, Chao M, et al. Multi-scale network traffic prediction based on attention mechanism and long short-term memory network[J/OL]. The Journal of China Universities of Posts and Telecommunications, 2024