www.rericjournal.ait.ac.th

Electronic Voltage Transformer Metering Error Prediction Model Based on VMD and Correction Model

Zhenhua Li*, Jiuxi Cui*,[1], Zhenxing Li*, Yedong Mao[#,^], Yanfeng Xiao[#,^], and Yijun Xiao[#,^]

ARTICLE INFO

Article history:
Received 08 February 2024
Received in revised form
27 July 2024
Accepted 30 November 2024

Keywords:
Electronic voltage transformer
Metering error prediction
Random forest
Two-stage deep model
VMD

ABSTRACT

The deployment of Electronic Voltage Transformers (EVTs) within ultra-high voltage smart grids exceeding 110kV is progressively expanding, rendering their operational stability paramount for precisely acquiring voltage signals across power grids. This paper proposes a dual-stage metering error prediction model for EVTs, which integrates Variational Mode Decomposition (VMD) with a Convolutional Neural Network-Bidirectional Gated Recurrent Unit-Attention Mechanism (CNN-BiGRU-AM). Initially, an optimization algorithm is applied to determine the optimal decomposition parameters in the variational mode. Non-stationary and stochastic raw ratio error sequences are then decomposed into different modal components through VMD. Subsequently, each component is modeled with a predictive framework incorporating BiGRU and an Attention Mechanism to obtain error deviation data based on the first stage's predicted ratio error results. Next, highly correlated features with EVT metering errors are identified using a random forest algorithm to construct a feature set. These features, combined with the error deviation obtained from the first stage, are used as joint inputs to establish an error correction process (second stage), resulting in the final prediction value. Empirical analysis indicates that the proposed method improves prediction accuracy and stability.

1. INTRODUCTION

Electronic voltage transformers (EVTs), crucial for measuring equipment in power systems, are extensively utilized in voltage measurements across transmission lines and substations, ensuring real-time monitoring and control of these systems. Their stability is critical to the grid's safe, reliable, and efficient functioning [1]-[4].

However, influenced by environmental temperature and humidity, electromagnetic fields, and equipment aging [5], performance deteriorates over time, affecting long-term stability and causing a decline in measurement accuracy. According to relevant standards, periodic calibration of the measuring instruments' accuracy is required [6]. Traditional verification is primarily conducted through offline checks [7], a relatively mature method. However, since offline testing requires planned power outages, it leads to low efficiency, high cost, and poor timeliness and is not conducive to batch testing. Consequently, some scholars have proposed connecting standard transformers to live circuits for online error evaluation of transformers without power shutdowns, as referenced in [8]. An online verification system for transformers was developed, using an SF6 cylinder capacitor as a standard capacitor voltage divider connected to operating lines. An electronic transformer verifier compares the output signals of the standard transformer and the tested device, obtaining the measurement ratio error of the EVT under operating conditions. Due to reduced verification costs, the calibration cycle has been significantly shortened. However, due to the high-voltage environment, it can only be performed in the short term.

To address the challenges above, scholars have studied evaluation methods based on transformers without standards, focusing on error detection and risk warnings for measurement devices. For instance, references [9,10] employ equivalent circuit modeling, using parameters from transmission and distribution lines to model and solve the operational errors of transformers. While this method is theoretically feasible, accurately obtaining the specific parameters of operating equipment poses a significant challenge. The accuracy of these parameters directly impacts the evaluation results of the modeling, leading to low accuracy rates in engineering applications.

In recent years, with the rapid development of information technology, evaluation methods centered on big data technology have gained widespread attention. Collecting operational data and environmental parameters of transformers and using data statistics or

*College of Electrical Engineering & New Energy, Three Gorges University, Hubei 443002, China; Provincial Engineering Research Center of Intelligent Energy Technology, Three Gorges University, Hubei 443002, China.

[#]Hubei Technology Innovation Center for Smart Hydropower, Hubei, 430000, China.

^ China Yangtze Power Co., Ltd., Hubei 443002, China.

[1]Corresponding author:
Email: cuiup@ctgu.edu.cn

modeling to extract associated information from the data, a preliminary evaluation of the transformers' measurement status has been achieved. For example, references [11], [12] use the short-term invariant characteristics of three-phase imbalance as constraints and apply principal component analysis to analyze the secondary output information of transformers. These studies achieve an online evaluation of error states by calculating statistical measures and their control thresholds. However, this method assumes that the output data of transformers follow a normal distribution, and it may fail when the fluctuations in the three-phase imbalance greatly exceed the operational errors of the transformers. Reference [13] improves upon this approach by employing "in-phase measurement consistency" among groups of voltage transformers. Using principal component analysis on the secondary output information of the same phase voltage transformer groups, it eliminates the impact of random fluctuations in the power system, thereby preventing the need to consider the distribution of the output signals. However, this method requires at least three groups of transformers on the same line, whereas many substations along the same line are equipped with only two groups.

Although the principal component analysis method is effective, its stringent prerequisites limit its application scenarios. Additionally, since it evaluates statistical thresholds to determine the possibility of degradation, the evaluation results are binary ("yes" or "no"). They cannot reflect changes in transformers' most important measurement index—measurement accuracy (ratio error).

Similarly, based on big data technology, references [14], [15] employ neural network modeling to build prediction models using historical data of transformers' operational ratio errors and environmental parameters. While neural network-based prediction models can directly forecast changes in the measurement accuracy of transformers at future times, the models' simplicity and tuning limitations lead to slow convergence and issues such as local optima, thereby affecting the prediction results.

The rapid development of deep learning technologies in recent years has facilitated the application of advanced neural network models in engineering fields. These models, with adjustable topologies, fast parallel processing speeds, and strong adaptability, address the shortcomings of traditional neural network modeling. For instance, convolutional neural networks (CNNs) [16], recurrent neural networks (RNNs) [17], and attention mechanisms [18] are now widely used in natural language processing and time series prediction. They have also extended to the power industry and power equipment fields, such as wind power prediction [19], [20], photovoltaic output [21], [22], and electric load [23], demonstrating excellent performance. Some deep learning-based approaches also predict transformer measurement errors. For example, reference [24] uses the BiLSTM neural network to directly predict the measurement errors of voltage transformers, and reference [25] employs the gated recurrent unit (GRU) neural network to predict the ratio errors of voltage transformers. However, these approaches apply models directly and without considering the impact of actual operating environment characteristics on transformers.

To address the shortcomings of the methods above, this paper proposes a prediction model based on variational mode decomposition (VMD) and a bidirectional GRU (BiGRU) structure with an attention mechanism. The northern goshawk optimization (NGO) algorithm is used to assist in selecting parameters for the VMD and neural network models to predict the ratio error of electronic voltage transformers. Additionally, an error correction model that considers environmental characteristic factors is proposed to enhance the stability of the model further. The main contributions are as follows:

1) The complexity of input data affects the model's prediction results. Using VMD to decompose complex, non-stationary input data into simple components improves the model's generalization ability.
2) Considering that the model's accuracy loss arises from the decomposition residual loss of VMD, the impact of measurement sequence fluctuations on the model, and the association between the transformers' operating environment and the equipment's measurement accuracy, this paper adopts a hybrid model that adds an error correction phase.

In the first stage, the NGO finds the decomposition parameters that minimize the VMD decomposition loss, and the original data is decomposed. Neural network models based on BiGRU and attention mechanisms (BiGRU-AM) are constructed separately to obtain predicted ratio error data. In the second stage, the error sequence obtained from the first stage and the variational mode components and feature vectors affecting the transformer's operation are selected as input vectors to construct a transformer error prediction model based on the BiGRU-AM network. The model's prediction errors are obtained and finally converted into ratio error prediction results. The proposed method's effectiveness was validated through case analysis, demonstrating its ability to predict future measurement error changes for 0.2-level EVTs.

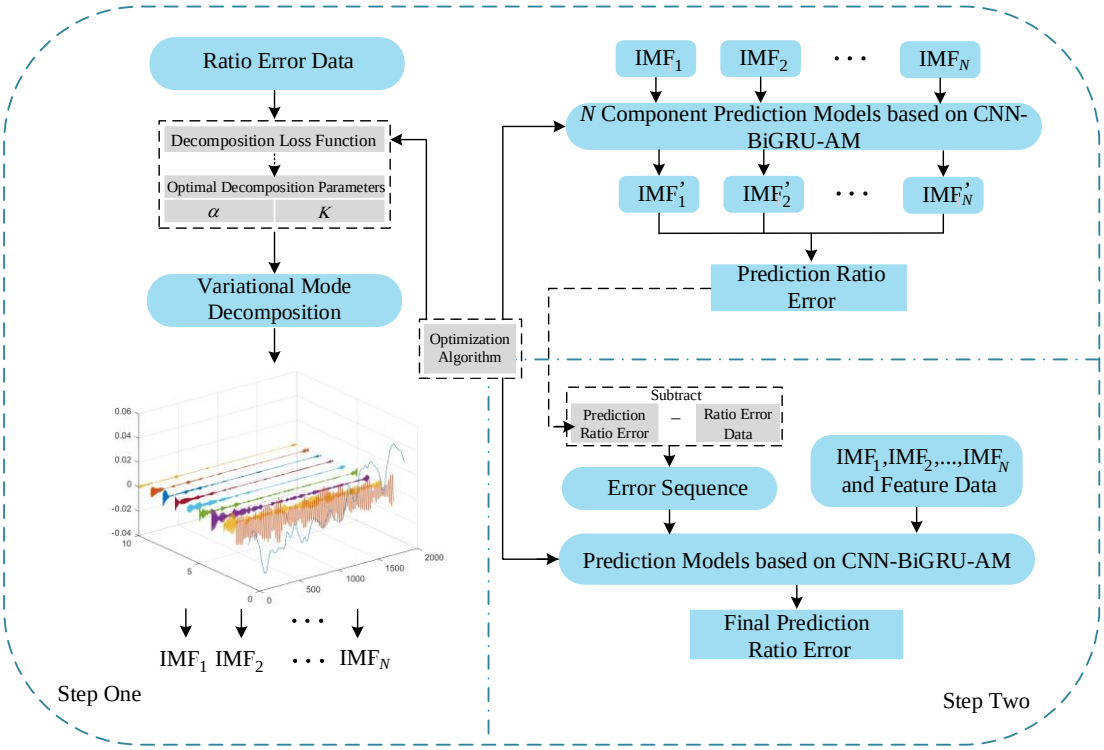


Fig. 1. EVT metering error prediction model.

2. ANALYSIS OF TRANSFORMER RATIO ERROR

The EVT transforms primary voltage signals from the power grid to digital signals for secondary systems. Typically, there's a discrepancy between the actual voltage and EVT's secondary measured data, with this metering error usually expressed as the ratio error [26], [27]:

$$f = \frac{K_r U_2 - U_1}{U_1} \times 100\% \tag{1}$$

In Equation (1), f denotes the ratio error; K_r is the EVT's transformation ratio; U_1 is the actual value of the primary voltage signal, and U_2 is the secondary side's measured voltage. Since EVTs are calibrated pre-installation, the ratio error f is nearly zero. EVTs, composed of highly integrated electronic units, face inevitable challenges in industrial environments, including environmental contamination, temperature and humidity fluctuations, electric and magnetic field effects, and vibration changes. These elements contribute to a gradual drift in measurement error. The physical network model depicting the increase in EVT ratio error due to these factors is as follows:

$$U_2 = mU_1 / K_r + v_s + f_s \tag{2}$$

Here, m denotes multiplicative noise; v_s additive noise, and f_s represents EVT's inherent measurement error under abnormal conditions.

Considering the complex temporal signals in transformer measurement data and assessing factors affecting the transformer's gradual error, the historical ratio error is represented as $\boldsymbol{Y}$, $\boldsymbol{Y} = (Y_1, Y_2, \cdots, Y_T) \in \mathbf{R}^T$. Environmental factors impacting measurement stability at any time are input variables; the matrix at T historical moments is thus denoted as.

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 \\ \vdots \\ \boldsymbol{X}_i \\ \vdots \\ \boldsymbol{X}_N \end{bmatrix}_{N\times T} = \begin{bmatrix} X_{1,1} & \cdots & X_{1,j} & \cdots & X_{1,T} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ X_{i,1} & \cdots & X_{i,j} & \cdots & X_{i,T} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ X_{N,1} & \cdots & X_{N,j} & \cdots & X_{N,T} \end{bmatrix}_{N\times T} \tag{3}$$

Here, $X_{i,j}$ is the value of the ith influencing factor at time j. Using historical ratio error $\boldsymbol{Y}$, the model predicts the EVT ratio error at the next moment with multiple variables. Defined by the function mapping F, the next moment's predicted ratio error, Y_{T+1}, is expressed as:

$$Y_{T+1} = F(Y_1, Y_2, Y_3 \cdots Y_T, \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3, \cdots, \boldsymbol{X}_T) \tag{4}$$

3. PROCESSING OF DATA

3.1 Variational Mode Decomposition

VMD is an adaptive, non-recursive signal processing method [28], utilizing Gaussian smoothing and Hilbert transformation for decomposing and extracting sequence components. VMD optimizes each mode's central frequency and bandwidth by calculating the variational model's optimal solution, simplifying complex time series signal decomposition. The core steps include:

The Hilbert transform calculates the unilateral spectrum of modes, aligning the spectrum to the base frequency through exponent estimation. A constraint equation is formed to minimize the total bandwidths of modal components:

$$\min_{\{u_k\},\{\omega_k\}}\left\{\sum_k \partial_t \left\| \left[(\delta(t)+\frac{j}{\pi t})*u_k(t)\right] e^{-j\omega_k t}\right\|_2^2\right\} \quad (5)$$

$$\text{s.t.}\sum_k u_k = f(t) \quad (6)$$

Here, $\{u_k\}=\{u_1,u_2,\ldots,u_N\}$ represents the N decomposed modal components, $\{\omega_k\}=\{\omega_1,\omega_2,\ldots,\omega_N\}$ are the central frequencies of each modal component, $f(t)$ is the original signal function, $\delta(t)$ is the Dirac delta function.

To address the constraint problem, introducing a penalty factor and Lagrange multipliers transforms the constraint equation into an unconstrained one. This ensures the variational model's nonlinearity and thus enables precise signal decomposition, as shown in Equation (7).

$$L[u_k,\omega_k,\lambda]=\alpha\sum_k \left\| \partial_t\left[\delta(t)+\frac{j}{\pi t}*u_k(t)\right]e^{-j\omega_k t}\right\|_2^2 + \left\| f(t)-\sum_k u_k(t)\right\|_2^2 + <\lambda(t), f(t)-\sum_k u_k(t)> \quad (7)$$

Here, λ is the Lagrange multiplier, and α is the penalty factor.

Finally, the modal components and central frequencies are iteratively updated using the Alternating Direction Method of the Multipliers algorithm:

$$u_k^{n+1}(\omega)=\frac{f(\omega)-\sum_{i\neq k}u_i(\omega)+\frac{\lambda(\omega)}{2}}{1+2\alpha(\omega-\omega_k)^2} \quad (8)$$

$$\omega_k^{n+1}=\frac{\int_0^\infty \omega\,|u_k^{n+1}(\omega)|^2\,d\omega}{\int_0^\infty |u_k(\omega)|^2\,d\omega} \quad (9)$$

3.2 Optimization and Reconstruction

The effectiveness of the VMD depends on the choice of decomposition parameters, primarily related to K and α. The number of decompositions, K, determines the number of modal separations, while the penalty factor α affects the decomposition process. Thus, the VMD algorithm can be regarded as a function $f_v(\ )$ related to (K, α), transforming it into a two-dimensional extremum optimization problem. The decomposition and reconstruction process is then represented as $\{u_k(t)\,|\,k=1,2,\ldots,K;t=1,2,\ldots,T\}=f_v(f(t),K,\alpha)$. The reconstructed signal $f'(t)$ can be expressed as $f'(t)=\sum_{k=1}^{K}u_k(t)$. The standard for evaluating the decomposition loss between the reconstructed signal and the original signal is usually the mean absolute error L, which is formulated as:

$$L=\frac{1}{T}\sum_{t=1}^{T}|f(t)-f'(t)| \quad (10)$$

To mitigate modal mixing and ensure signal reconstruction accuracy, an optimization algorithm is used to tune VMD parameters iteratively. The optimization is set to two dimensions. By inputting the measured ratio error sequence and setting parameters such as the maximum number of iterations and the initial population, the fitness function is calculated, retaining the globally optimal individuals and their fitness. To achieve effective decomposition and improve the predictive accuracy of subsequent neural network models, the minimum envelope entropy is selected as the fitness function for evaluation to identify the optimal solution.

$$\text{s.t.}\begin{cases} K\in\{2,3,\cdots,30\} \\ \alpha\in[100,10^5] \end{cases} \quad (11)$$

To validate the effectiveness of the decomposition via the optimization algorithm, the system's default VMD decomposition values are compared to the NGO [29] meta-heuristic algorithm. The number of decompositions and penalty factors are rounded to the nearest integer, and the population and iteration count are uniformly set to 25 and 100, respectively. The results are presented in Table 1.

Table 1. Comparison of VMD decomposition results.

Model	Decomposition		Fitness
	K	α	
VMD	5	1000	0.08466
NGO-VMD	8	4128	0.06448

Experimental results show that NGO-optimized VMD effectively minimizes signal loss in decomposition, adaptively determines optimal parameters, and guarantees accuracy, overcoming the randomness and ambiguity of empirical settings. Figure 2 illustrates the initial ratio error sequence segmentation into intrinsic mode functions (IMFs), each showing distinct change perns. The primary component, IMF_1, displays a lower frequency and higher amplitude, indicating the overall trend of the ratio error sequence. This research separates non-stationary stochastic time series signals into stable, independent component sequences, enhancing model accuracy by finely modeling the signal's trends.

Figure 3 illustrates the primary component IMF1, representing the overall trend, and compares it with the reconstructed signal and decomposition residuals. We can see that the original sequence is divided into two main parts. The grey columnar fill represents the "Reconstructed Signal," which results from all intrinsic mode functions (IMFs) overlaid, with its magnitude detailed on the left Y-axis. The red columnar fill denotes the "Residual," which is the remaining irregular part that cannot be decomposed, with its magnitude shown on the right Y-axis.

Figures 2 and 3 reveal that IMF1 aligns with the overall trend. The emergence of residuals post-reconstruction is deemed inevitable; the magnitude of residuals is related to the precision of the decomposition.

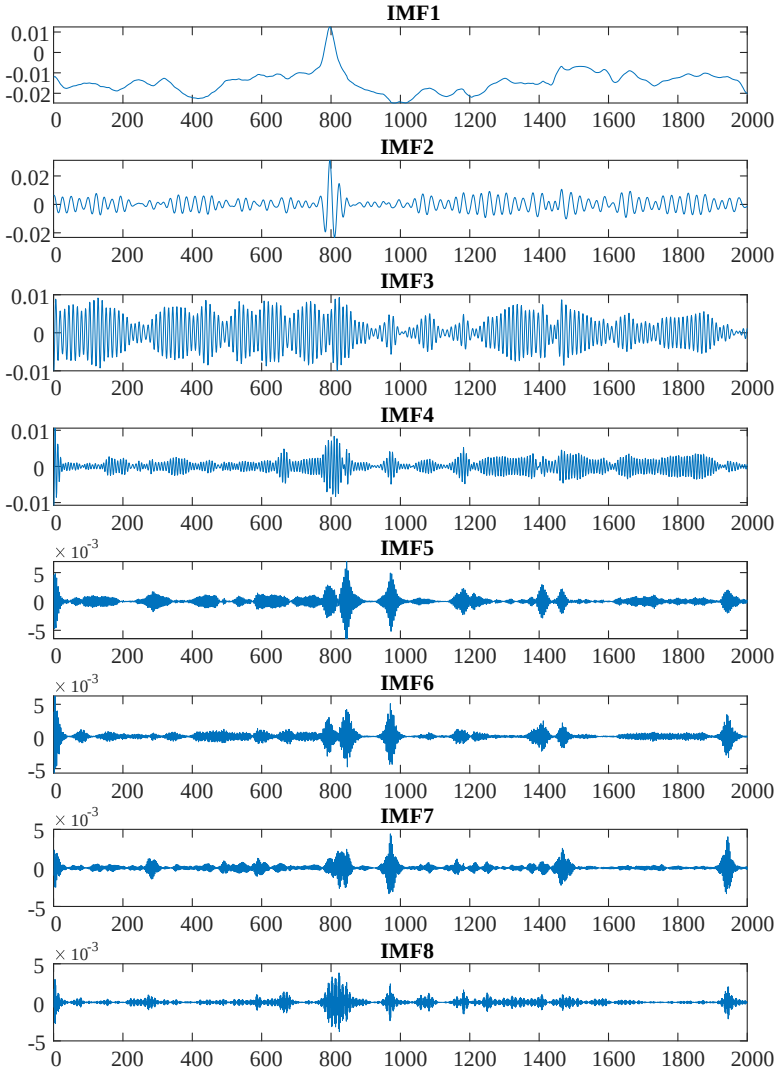


Fig. 2. VMD decomposition.

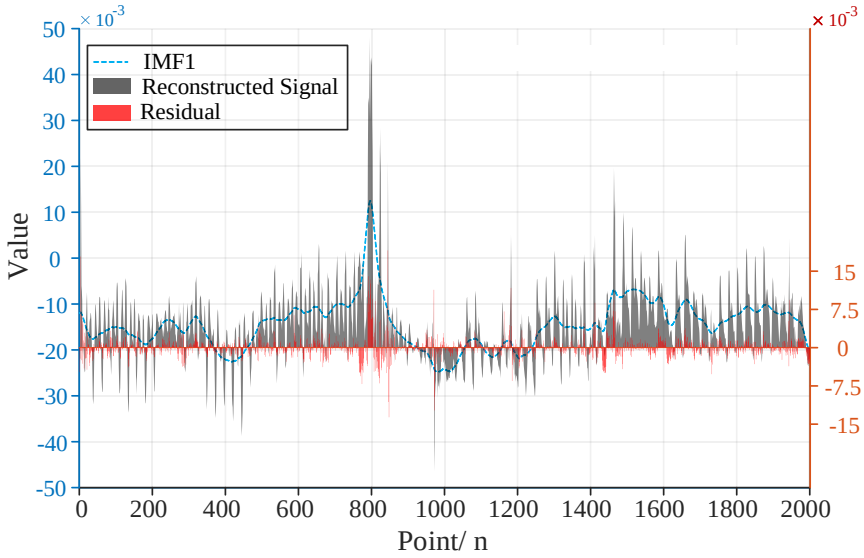


Fig. 3. Decomposition, reconstruction, and residuals post-reconstruction.

3.3 NGO-BiGRU-AM Model

The GRU demonstrates notable advancements in sequence modeling through the amalgamation of the long short-term memory (LSTM) neural network's forget and input gates into a unified update gate. This integration facilitates efficient information flow

management, parameter reduction, overfitting mitigation, and expedited convergence. Furthermore, the BiGRU surpasses traditional GRU limitations by assimilating information bidirectionally, thereby capturing both antecedent and subsequent contexts within each iteration. This bidirectional processing amalgamates past and future insights, ensuring the maximal retention of pivotal features. The operational dynamics of this mechanism are depicted below, with the update and reset gates symbolized by z_t and r_t, respectively. Here, x_t signifies the input, h_t denotes the hidden state propagated in reverse, and $\tilde{h}_t$ encapsulates the candidate's hidden state. The corresponding mathematical formulations are delineated as follows:

$$\left.\begin{aligned} z_t &= \sigma(W_z h_{t-1} + U_z x_t + b_z) \\ r_t &= \sigma(W_r h_{t-1} + U_r x_t + b_r) \\ \tilde{h}_t &= \tan h\left[W_h \cdot x_t + U_h (r_t \odot h_{t-1}) + b_h\right] \\ h_t &= (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \end{aligned}\right\} \quad (12)$$

Here, σ is the sigmoid function, W_z, W_r, W_h, U_z, U_r, *and* U_h are the weight matrices respectively; b_z, b_r, b_h are the corresponding bias vectors.

By dynamically concentrating on varying segments within the sequence, the attention mechanism enhances the model's capability to comprehend mid-to-long-term dependencies, consequently bolstering predictive stability.

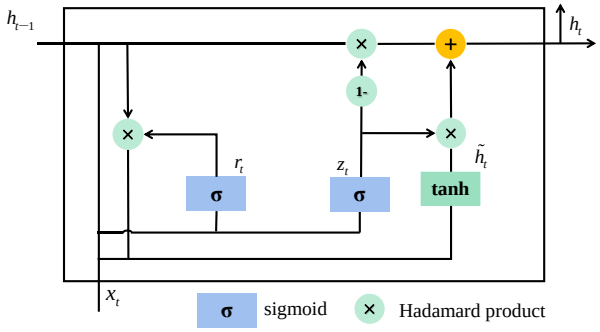


Fig. 4. Basic unit of GRU.

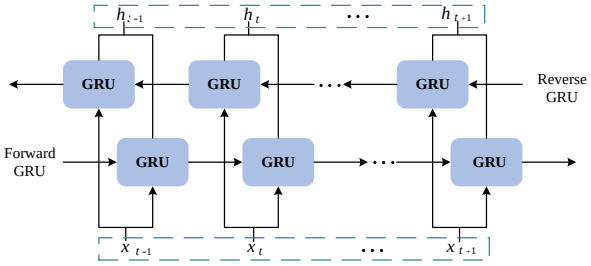


Fig. 5. Bi-directional GRU propagation structure.

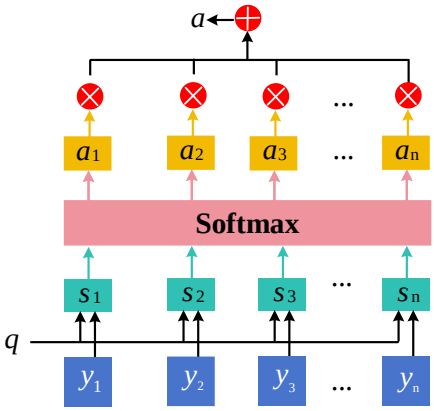


Fig. 6. Attention mechanism.

The methodology entails initially encoding the input sequence to procure the query vector q, followed by the allocation of weights based on similarity:

$$s(y_i, q) = V^T \tanh(W y_i + U q) \quad (13)$$

Here, $\boldsymbol{V}$, $\boldsymbol{W}$, and $\boldsymbol{U}$ are designated as matrices pending parameter learning. Utilization of the Softmax function for normalization facilitates the derivation of the weight $\boldsymbol{a_i}$ for each group of input vectors:

$$a_i = \text{softmax}(s(y_i, q)) = \frac{\exp(s(y_i, q))}{\sum_{i=1}^{n} \exp(s(y_i, q))} \quad (14)$$

Subsequently, the computed weighted sum is based on the weights and their corresponding value vectors:

$$a = \sum_{i=1}^{n} a_i \quad (15)$$

To improve the model's predictive accuracy, this paper introduces a VMD-CNN-BiGRU-AM voltage transformer metering prediction model, utilizing error correction as depicted in Figure 1. Initially, different components of a VMD-CNN-BiGRU-AM prediction model are reconstructed. The algorithm calculates the optimal (K,α) solution, partitioning the transformer's non-stationary random ratio error sequence into stable modal components IMF_1, IMF_2, IMF_3, …, IMF_N.

Separate CNN-BiGRU-AM algorithm models are constructed for each intrinsic modal component, yielding predictions IMF_1, IMF_2, IMF_3,…, IMF_N. These component model predictions are then aggregated to forecast the ratio error sequence.

Considering the accuracy losses due to VMD decomposition, non-stationary random measurement sequence fluctuations, and the transformer's operating environment, the second stage involves error correction. A highly correlated set of features, intrinsic modal components, and the first-stage neural network error residuals are used as inputs. The CNN-BiGRU-AM model establishes nonlinear relationships among these inputs. Finally, the final predicted ratio error is obtained.

4. NUMERICAL EXAMPLE ANALYSIS

All experiments in this paper are based on Matlab 2023a and focus on EVTs at a 110kV substation in Henan. Time series data on ratio error and characteristics like magnetic field, temperature, humidity, and electric field phase were collected hourly. A total of 2000 datasets (Figure 8) were compiled, and divided into training, testing, and validation sets in proportions of 0.8, 0.1, and 0.1, respectively.

Evaluation criteria such as Root Mean Square Error (X_{RMSE}), Mean Absolute Error (X_{MAE}), and the coefficient of determination R assess model accuracy.

$$X_{MAE} = \frac{1}{m}\sum_{i=1}^{m}|\hat{y}_i - y_i| \tag{16}$$

$$X_{RMSE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(\hat{y}_i - y_i)^2} \tag{17}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{m}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{m}(y_i - \frac{1}{m}\sum_{i=1}^{m}y_i)^2} \tag{18}$$

Lower values of X_{RMSE} and X_{MAE}, and R^2 closer to 1 indicate higher model prediction accuracy.

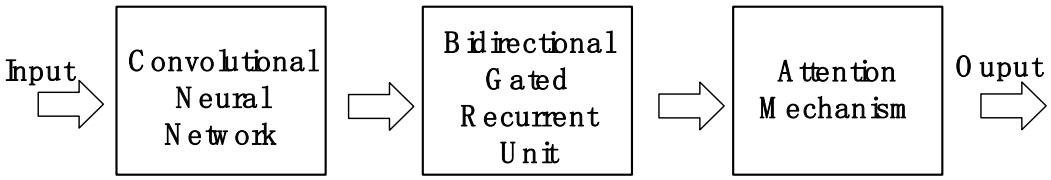


Fig. 7. Main prediction model.

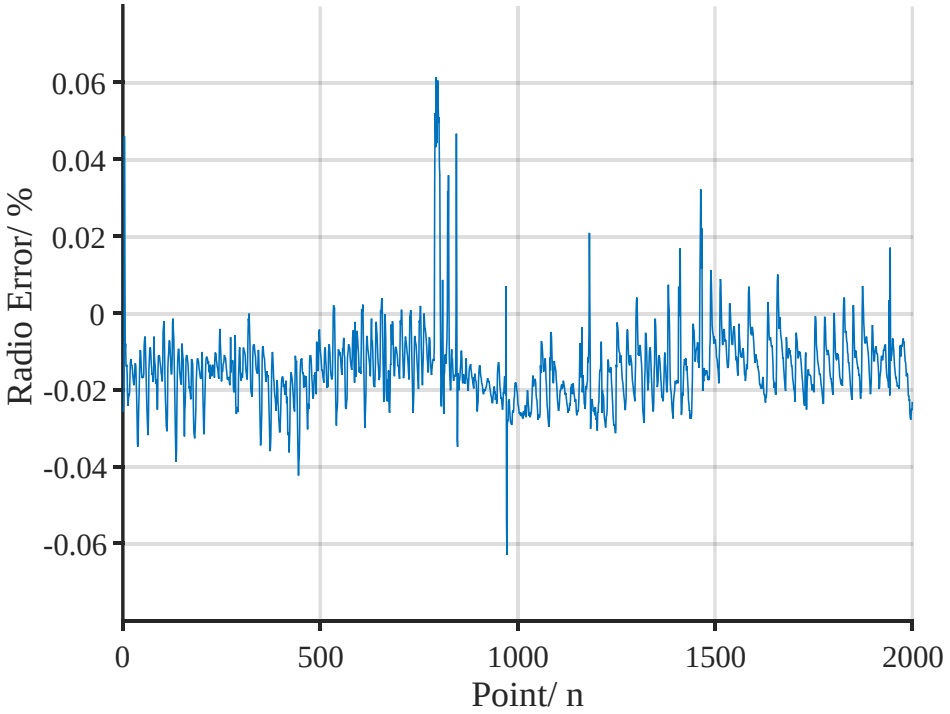


Fig. 8. Ratio error.

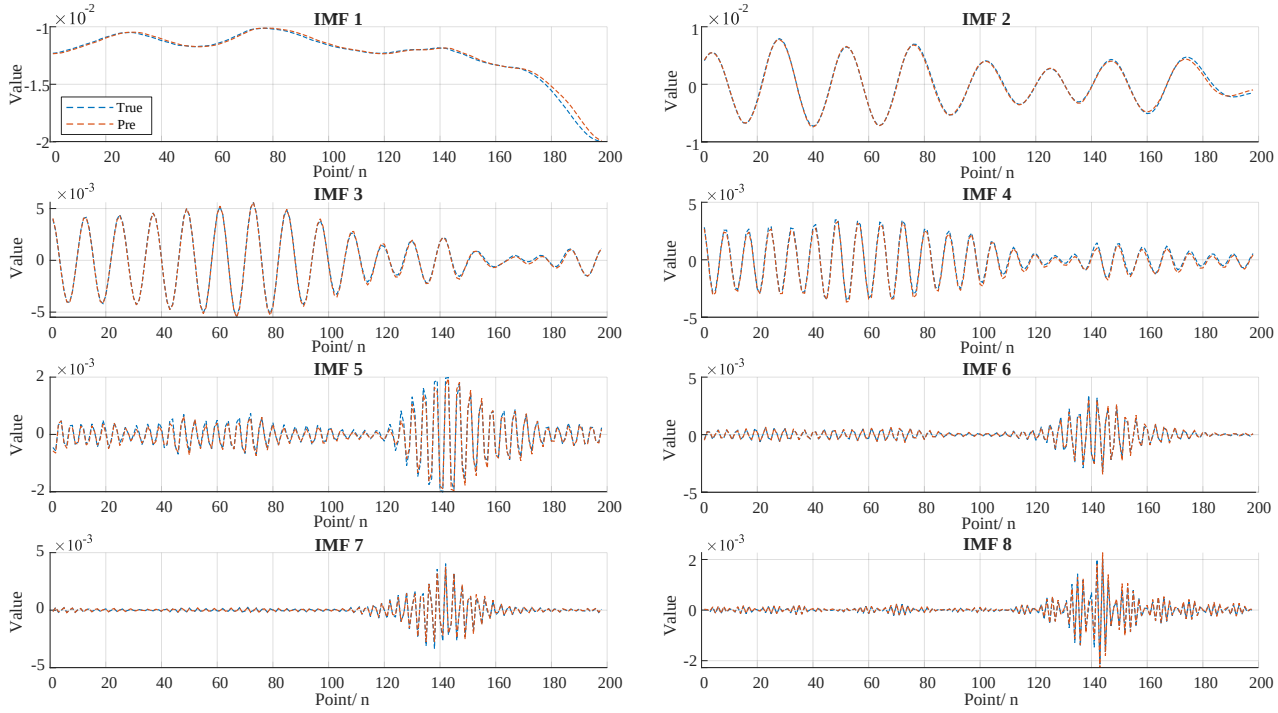


Fig. 9. Predicted and true values after VMD decomposition.

4.1 VMD Sequence Decomposition

Figure 9 shows the VMD decomposition of the test set, with true values indicated by a blue line and predictions by a red line. The figure demonstrates that, following algorithm-based decomposition, the predictions closely follow the exact values. Models for each decomposed sequence were established to validate the VMD decomposition's efficacy further. A comparison was made by randomly selecting 100 consecutive points, and contrasting it with a model directly using the original ratio error sequence.

Table 2 and Figure 9 clearly show that the prediction results after VMD decomposition are significantly better than the direct prediction results without decomposition. The predicted values closely follow the true values. Specifically, the RMSE decreased by 47.51%, the MAE decreased by 20.1%, and the coefficient of determination (R^2) increased by 62.9%. This indicates that using VMD to decompose the ratio error sequence into different simple components significantly reduces the learning difficulty of the model, thereby improving the accuracy of the prediction. Moreover, the decomposition method increases the sample size and reduces the risk of model overfitting.

Table 2. Comparison of VMD Experiments

Model	Evaluation		
	X_{RMSE}	X_{MAE}	R^2
BiGRU	4.61×10^{-3}	2.09×10^{-3}	0.5351
VMD- BiGRU	2.42×10^{-3}	1.67×10^{-3}	0.8719

4.2 Comparative Experiment

To verify the performance of the hybrid model proposed in this paper, 100 randomly selected sequential data points were used to compare the VMD-CNN-BiGRU and VMD-CNN-BiGRU-AM models in predicting the ratio error data after VMD decomposition. Results in Figure 10(a) and Table 3 show that compared to VMD-BiGRU, CNN-BiGRU and CNN-BiGRU-AM models reduced RMSE by 14.05% and 17.77%, and MAE by 9.58% and 14.37%, while improving R^2 by 3.83% and 4.75%, respectively. This improvement is primarily attributed to the convolution layer's transformation of lengthy sequence data from a low to high-dimensional space, optimizing the sequential data processing. Additionally, the attention mechanism dynamically weights important sequences, reducing model forgetting by enhancing medium and long-term dependencies, which improves prediction stability.

4.3 Error Correction Experiment and Analysis

Considering the primary causes of error, which include the residuals from VMD decomposition and the inherent model influences, it is evident from the analysis above that the VMD-CNN-BiGRU-AM neural network exhibits the most minor error. Based on this model, an error correction experiment was conducted. To quantify the effectiveness of error correction and high-correlation

feature sets, the following model construction and validation methods have been employed in comparative experiments:

1) The model without error correction (VMD-CNN-BiGRU-AM). 2) Error correction using only the modal components derived from the decomposition as additional inputs, denoted as S1. 3) Error correction using both modal components and high-correlation feature sets as inputs to the model, denoted as S2. Experiments and analyses were conducted using the above three models.

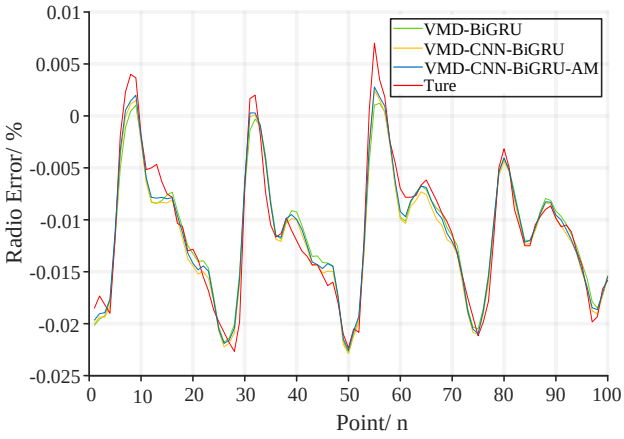
4.3.1 High-correlation feature

The operational stability of transformers is influenced by the long-term effects of their operating environment. To identify the strongly correlated and redundant features in the collected substation environment data, this paper employs a Random Forest [30] algorithm to recognize the critical feature set that is highly correlated with the dependent variable. As shown in Figure 10(b), temperature and humidity are identified as strongly correlated feature parameters for ratio error prediction. Based on the results of the Random Forest, we excluded weakly correlated features and selected temperature (3.587) and humidity (2.004) to form the feature set used as inputs for the correction model.

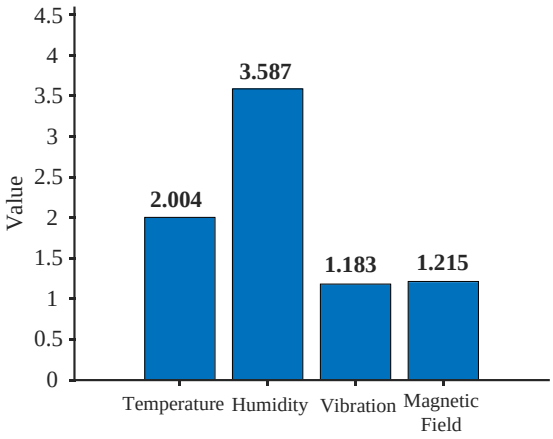
4.3.2 Experiment and analysis

In Figure 11(a), the MAE comparison of three models is presented, where the blue, red, and orange lines represent the absolute values of error amplitudes between model predictions and true values, displayed on the y-axis. Models S1 and S2 exhibit lower single-point relative errors over different time intervals, indicating that, after two stages of correction, they can more accurately capture contextual correlations and show superior performance. Figure 11(b) compares the predictions of the three models with the true values, showing that the predictions of S1 and S2 are closer to the true values. When faced with abnormal fluctuations in measurement errors, these models can quickly identify and respond to such changes. The figure shows that the S2 model, which incorporates highly correlated feature sets, demonstrates the highest correlation with true values. This improvement is attributed to the S2 model's in-depth exploration of the nonlinear relationships among features, ratios, and losses, and the hidden information they contain, thereby accurately reflecting EVT measurement errors and enhancing prediction precision.

The deployment of the S1 and S2 models has markedly enhanced predictive accuracy. This is evidenced by a 15.08% decrease in RMSE, a corresponding 15% reduction in MAE, and a 2.63% enhancement in R^2 for the S1 model. Similarly, the S2 model demonstrated a 19.1% reduction in RMSE, a 17% decrease in MAE, and a 3.31% increase in R^2. These advancements not only elucidate the models' capability to minimize error magnitudes but also to bolster correlation significantly, thereby affirming their enhanced predictive prowess through the assimilation of sophisticated error correction methodologies and intricate feature sets.



(a). Ablation Experiments.

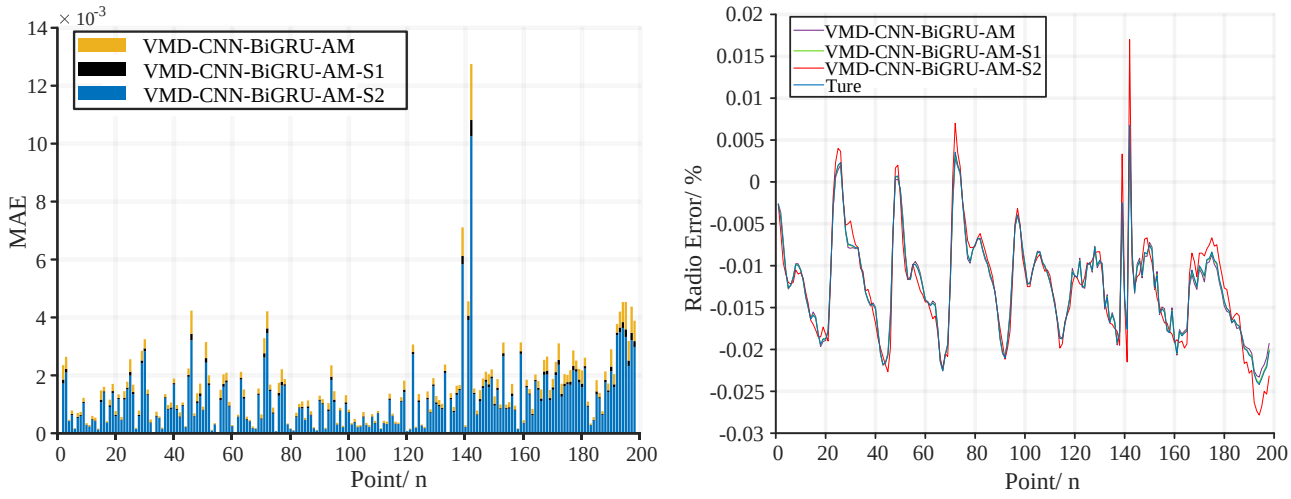


(b). Feature Selection.

Fig. 10. Ablation experiments and feature selection.

Table 3. Comparison of ablation experiments.

Model	Evaluation		
	X_{RMSE}	X_{MAE}	R^2
VMD-CNN-BiGRU	2.08×10^{-3}	1.51×10^{-3}	0.9053
VMD-CNN-BiGRU-AM	1.99×10^{-3}	1.43×10^{-3}	0.9133



(a). Comparison of prediction MAE.

(b). Comparison of prediction error.

Fig. 11. Ablation experiments and feature selection

Table 4. Comparison of prediction experiments.

Model	Evaluation		
	X_{RMSE}	X_{MAE}	R^2
VMD-CNN-BiGRU-AM-S1	1.69×10^{-3}	1.22×10^{-3}	0.9372
VMD-CNN-BiGRU-AM-S2	1.61×10^{-3}	1.16×10^{-3}	0.9434

5. CONCLUSION

This paper establishes a prediction model utilizing VMD-CNN-BiGRU-AM with a two-stage correction method. Key conclusions include:

1) VMD usage enhances data predictability. Decomposing sequences into varied sub-sequences via optimized VMD simplifies learning and ensures reconstruction accuracy.
2) Integrating bidirectional and attention mechanisms, expanding the data volume, and focusing on key data significantly improves the accuracy of the model.
3) By employing a random forest to construct the feature set and incorporating an error correction stage, the model's inherent errors are significantly minimized, thereby further enhancing the model's predictive stability.

ACKNOWLEDGEMENT

This work is supported by the National Key R&D Program of China (2022YFB2403800), and the Open Research Fund of the Hubei Technology Innovation Center for Smart Hydropower (HBCXZX-JJ-202407).

DECLARATION OF COMPETING INTEREST

The authors declare that there are no conflicts of interest.

REFERENCES

[1] Li Z., Cui J., and Chen H., 2025. Research progress of all-fiber optic current transformers in novel power systems: A review. *Microwave and Optical Technology Letters* 67: e70061.

[2] Long Z., Li W., and Zhou F., 2022. Device of 1200 kV wideband capacitive divider based on high voltage standard capacitor. *High Voltage Engineering* 48: 1826–1835.

[3] Li Z., Jiang W., Abu-Siada A., Li Z., Xu Y., and Liu S., 2020. Research on a composite voltage and current measurement device for HVDC networks. *IEEE Transactions on Industrial Electronics* 68(9): 8930–8941.

[4] Li Z., Zhang Y., Abu-Siada A., Chen X., Li Z., Xu Y., Zhang L., and Tong Y., 2021. Fault diagnosis of transformer windings based on decision tree and fully connected neural network. *Energies* 14(6): 1531.

[5] Ruan S., Wang D., Xu K., Xu H., Li F., and Lv P., 2018. ECT fault statistical analysis of dc transmission system. *Power System Technology* 42: 3170–3175.

[6] Ye D., 2012. *General Terms in Metrology and Their Definitions: JJF 1001-2011*. Beijing: China Quality Inspection Press.

[7] Brandolini A., Faifer M., and Ottoboni R., 2009. A simple method for the calibration of traditional and electronic measurement current and voltage transformers. *IEEE Transactions on Instrumentation and Measurement* 58(5): 1345–1353.

[8] Li Z., Yu C., Abu-Siada A., Li H., Li Z., Zhang T., and Xu Y., 2021. An online correction system for electronic voltage transformers. *International Journal of Electrical Power & Energy Systems* 126: 106611.

[9] Pegoraro P.A., Brady K., Castello P., Muscas C., and von Meier A., 2019. Compensation of systematic measurement errors in a PMU-based monitoring system for electric distribution grids. *IEEE Transactions on Instrumentation and Measurement* 68(10): 3871–3882.

[10] Wang H., Tang K., Xu R., Li H., Zhu X., and Li X., 2012. Diagnosis of soft fault of electronic transformer in digital substation. *Power System Protection and Control* 40(24): 53-58.

[11] Zhang Z., Li H., Tang D., Hu C., and Jiao Y., 2017. Monitoring the metering performance of an electronic voltage transformer on-line based on cyber-physics correlation analysis. *Measurement Science and Technology* 28(10):105015.

[12] Zhang Z., Chen Q., Hu C., Li H., and Chen M., 2018. Evaluating the metering error of electronic transformers on-line based on vn-mwpca. *Measurement* 130: 1–7.

[13] Zhang C., Li H., Yang J., Chen M., and Jiao Y., 2019. Detecting measurement error drifts of a capacitor voltage transformer on-line and its field application. *Measurement Science and Technology* 30(10):105109.

[14] Lin X., Xing J., Chen D., Li Z., Rong Z., Liu Y., Liu S., Su S., Tong N., and Jing R., 2018. A novel synchronous fault identification strategy of electronic transformer based on synergy of historical data. *International Journal of Electrical Power & Energy Systems* 103: 247–256.

[15] Hu C., Zhang Z., Yang A., Li M., Jiao Y., and Li D., 2020. Error model and error state prediction method of electronic current transformers. *Power Engineering Technology* 39(4): 187–193.

[16] Medsker L.R. and L.C., Jain. 2001. Recurrent neural networks. *Design and Applications* 5(64–67): 2.

[17] Medsker L. and L.C. Jain. 1999. Recurrent Neural Networks: Design and Applications. CRC Press.

[18] Han K., Xiao A., Wu E., Guo J., Xu C., and Wang Y., 2021. Transformer in transformer. *Advances in Neural Information Processing Systems* 34: 15908–15919.

[19] Tuerxun W., Xu C., Guo H., Guo L., Zeng N., and Gao Y., 2022. A wind power forecasting model using LSTM optimized by the modified bald eagle search algorithm. *Energies* 15(6): 2031.

[20] Chen Y., Yu S.S., Lim C.P., and Shi P., 2024. Multi-objective estimation of optimal prediction intervals for wind power forecasting. *IEEE Transactions on Sustainable Energy* 15(2): 974–985.

[21] Song Z., Huang Y., Xie H., and Li X., 2023. Generation method of multi-regional photovoltaic output scenarios-set using conditional generative adversarial networks. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* 13(3): 861–870.

[22] Sun Y., Zhou Y., Wang S., Mahfoud R.J., Alhelou H.H., and Sideratos G., 2023. Nonparametric probabilistic prediction of regional PV outputs based on granule-based clustering and direct optimization programming. *Journal of Modern Power Systems and Clean Energy* 11(5): 1450–1461.

[23] Hafeez G., Alimgeer K.S., and Khan I., 2020. Electric load forecasting based on deep learning and optimized by heuristic algorithm in smart grid. *Applied Energy* 269: 114915.

[24] Zhou F., Zhao P., Lei M., Yue C., Yu J., and Liang S., 2022. Capacitive voltage transformer measurement error prediction by improved long short-term memory neural network. *Energy Reports* 8: 1011–1021.

[25] Zhang W., Shi Y., Yu J., Yang B., and Lin C., 2023. Online measurement of capacitor voltage transformer metering errors based on GRU and MTL. *Electric Power Systems Research* 221: 109473.

[26] IEC, 2011. *IEC 61869-5: Instrument Transformers - Part 5: Additional Requirements for Capacitive Voltage Transformers*. Geneva: International Electrotechnical Commission.

[27] IEC, 2011. *IEC 61869-3: Transformers - Part 3: Additional Requirements for Inductive Voltage Transformers.* Geneva: International Electrotechnical Commission.

[28] Dragomiretskiy K. and D. Zosso, 2013. Variational mode decomposition. *IEEE Transactions on Signal Processing* 62(3): 531–544.

[29] Dehghani M., Hubalovsky S., and Trojovsky P., 2021. Northern goshawk optimization: a new swarm-based algorithm for solving optimization problems. *IEEE Access* 9: 162059–162080.

[30] Aria M., Cuccurullo C., and Gnasso A., 2021. A comparison among interpretative proposals for Random Forests. *Machine Learning with Applications* 6:100094.