Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Measurement Error Evaluation Method for Distribution Network Voltage Transformers Based on an Attention-RBF Physics-Informed Neural Network

Dongsheng Li
State Grid Wuhan Power Supply Company
Xiujuan Zeng
zengxiujuan2024@163.com

State Grid Wuhan Power Supply Company
Tong Liu
State Grid Wuhan Power Supply Company
Qilin Zhang
State Grid Wuhan Power Supply Company
Wei Wu
State Grid Wuhan Power Supply Company

Research Article

Keywords: Distribution network, Voltage transformer, Error evaluation, Physics-informed neural network, Radial basis function, Attention mechanism

Posted Date: October 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11016584/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Measurement Error Evaluation Method for Distribution Network Voltage Transformers Based on an Attention-RBF Physics-Informed Neural Network

Dongsheng Li, Xiujuan Zeng*, Tong Liu, Qilin Zhang, Wei Wu

State Grid Wuhan Power Supply Company, Wuhan, 430014, China

*** Correspondence:** Email: zengxiujuan2024@163.com.

Abstract: To address the problems of power interruption and low efficiency in the conventional verification of distribution network voltage transformers (VTs), this paper proposes a physics-informed neural network integrating an attention mechanism and radial basis functions (Attention-RBF-PINN) method for online error evaluation of distribution network VTs. In this method, Kirchhoff's voltage law (KVL) topological constraints of switching stations and the cable thermal equation are embedded into the neural network as soft loss terms. Meanwhile, radial basis functions (RBFs) are adopted to extract multi-scale temporal features, and a residual attention mechanism is employed to adaptively weight multi-branch inputs, thereby enabling VT error estimation. Experiments conducted on a laboratory physical simulation platform demonstrate that the proposed method outperforms comparable methods in all evaluation metrics. Compared with the second-ranked MRSA-DGCN, the RMSE is reduced by 35.2%, the MAE is reduced by 36.4%, and the R^2 is improved by 2.63 percentage points. In field application, one out-of-tolerance VT was successfully identified, and the evaluation result was consistent with the de-energized verification result. This study indicates that the fusion of physical guidance and data-driven modeling can effectively improve the accuracy and physical interpretability of transformer error evaluation, providing an effective approach for online error evaluation of distribution network voltage transformers.

Keywords: Distribution network; Voltage transformer; Error evaluation; Physics-informed neural network; Radial basis function; Attention mechanism

1. Introduction

VTs are core equipment for electric energy metering, condition monitoring, and operation control in distribution network power systems. Their performance directly affects the accuracy of electric energy metering, the security of power systems, and the reliability of power supply for consumers [1-3]. Subject to the complex operating environment of distribution networks, the metering performance of VTs gradually deteriorates during long-term service, eventually leading to out-of-tolerance errors [4-6]. According to the regulations [7], VTs are subject to periodic mandatory verification [8], which requires de-energization, disassembly, and on-site or laboratory comparison verification using standard transformers. This periodic offline verification mode consumes considerable manpower and material resources, suffers from low operational efficiency, and the resulting frequent power interruptions severely compromise power supply reliability. Consequently, it is difficult to accommodate the operational characteristics of distribution network

VTs, which are numerous, widely distributed, and impose substantial operation and maintenance pressure. Therefore, research on online error evaluation technology for distribution network VTs is of significant engineering value.

Regarding the online error evaluation of VTs, existing studies can be classified into three categories [9]. Early studies mainly focused on physics-based methods, in which mathematical expressions between VT outputs and electrical parameters were established, and measurement data such as PMU data were used to inversely infer the error state or correct systematic errors [10-11]. However, such methods are highly dependent on precise physical model parameters. The on-site electromagnetic environment in the distribution rooms where distribution network VTs are located is complex, and the line parameters are strongly time-varying, which makes it difficult to obtain accurate model prior information. As a result, the accuracy of the physical model is insufficient, thereby restricting the application of such methods in distribution network scenarios. With the development of deep learning techniques, data-driven evaluation methods have gradually become a research focus. Reference [12] combined self-attention and graph convolutional networks to construct a VT error evaluation model based on the transformation topology of distribution transformers. References [13-15] employed a fused convolutional module, a graph convolutional-temporal convolutional network, and a multi-dimensional attribute RBF model, respectively, to achieve CVT error prediction. The advantages of such methods lie in that they do not require an accurate physical model and can adaptively fit complex nonlinear mapping relationships. However, purely data-driven models are highly dependent on the quality and coverage of training data, have limited generalization ability to untrained operating conditions such as severe load fluctuations and abrupt environmental changes in distribution networks, and find it difficult to distinguish whether measurement variations originate from the deterioration of VT intrinsic errors or from external environmental disturbances. This mechanistic ambiguity fundamentally restricts further improvement in evaluation accuracy. To achieve both physical interpretability and data adaptability, many researchers have explored fusion methods that embed physical models into data-driven models. Reference [16] constructed an error evaluation framework for electronic transformers that integrates mechanism models and data correction. References [17-18] utilized the physical constraint relationships among measurements of multiple VTs at the same voltage level to construct mutual verification constraints, effectively suppressing the effects of individual error drift and measurement noise. However, distribution network VTs are dispersedly installed in various distribution rooms. Although each distribution room is fed from the bus of the same upstream switching station, the feeder length can reach several kilometers. During operation, load current fluctuations induce Joule heating in cables, which, superimposed with time-varying ambient temperature disturbances, causes the cable core temperature to drift continuously and dynamically. As a result, the cable impedance exhibits real-time dynamic variations, and the resulting dynamic line voltage drop cannot be neglected. If the line voltage drop is directly ignored in the group solution, the evaluation results will deviate significantly from the true values. Therefore, the key challenge in achieving multi-unit joint error evaluation of distribution network VTs lies in how to effectively address the interference of dynamic line voltage drop on the measurement reference.

To address the aforementioned challenge of dynamic line voltage drop interference in distribution network scenarios, the PINN offers an effective solution framework. By embedding physical differential equation constraints into the loss function, PINNs enable the network

learning process to balance data fitting with consistency to physical laws, and have demonstrated advantages in solving inverse problems such as parameter identification [19]. When applied to distribution network VT error evaluation, the cable thermal balance differential equation is incorporated into the network training as a physical constraint, such that the network simultaneously satisfies the thermal balance law while fitting the voltage transfer relationship. Consequently, the interference component caused by line voltage drop can be effectively identified and separated from the measurement data, thereby enabling accurate estimation of VT errors. However, the accuracy of solving differential constraints depends on the higher-order differentiability of the activation function. Moreover, if numerical differentiation is employed to approximate the differential constraints during network training, the truncation error further amplifies measurement noise and degrades the solution accuracy. Furthermore, when multiple constraints such as the thermal balance equation and the voltage transfer equation are jointly trained, the significant magnitude differences among the residual terms can readily induce gradient imbalance, thereby compromising convergence stability [20]. To overcome these limitations, reference [21] exploits the infinitely continuously differentiable property of radial basis functions (RBFs), thereby enhancing the accuracy of differential constraint solving. Reference [22] introduces an attention mechanism and a temporal modeling module to suppress interference during high-noise periods through adaptive weighting. Reference [23] adopts parameter balancing transformation and Fourier feature augmentation to alleviate gradient imbalance in multi-constraint joint training. Although these improvements provide feasible strategies from the perspectives of differential fitting accuracy, temporal noise suppression, and training stability, the influence of dynamic line voltage drop has not yet been considered in distribution network VT error evaluation.

To address this, an Attention-RBF-PINN method integrating radial basis functions and an attention mechanism is proposed to enhance its applicability in distribution network VT error evaluation. The main contributions of this paper are summarized as follows:

1.An error evaluation model for distribution network VTs that accounts for line voltage drop is established, and error evaluation conditions based on the voltage transfer relationship between adjacent distribution rooms are derived, providing a theoretical basis for error evaluation.

2.An Attention-RBF-PINN method is proposed, in which RBF-based multi-scale temporal encoding is utilized to enhance the temporal feature representation capability, and a residual attention mechanism is employed to adaptively weight the input features, thereby achieving VT error evaluation in the presence of dynamic voltage drop interference.

3.A multi-physics constrained weighted loss function is designed, in which the KVL topological constraint, the cable thermal balance constraint, and the data fitting loss are unified into a single formulation. This enables the network to be simultaneously constrained by the distribution network topology mechanism and the cable thermal dynamic characteristics during data-driven fitting, thereby achieving synergistic fusion of physical priors and data-driven modeling.

4.Comparative experiments are conducted based on simulation and field data to verify that the proposed method outperforms other methods in terms of evaluation accuracy and robustness.

The remainder of this paper is organized as follows. Section II elaborates on the theoretical foundations of the proposed error evaluation method. Section III details the architecture and training strategy of the proposed Attention-RBF-PINN model. Section IV presents the experiments and result analysis. Section V summarizes the paper.

2. Related work

2.1 Voltage transfer characteristic error evaluation theory

In the distribution network power system, each distribution room is connected to the busbar of the same switching station via an independent feeder and is supplied by the 10kV feeder of the upstream substation. Therefore, by calculating the line voltage drop from the switching station to each distribution room using line parameters and combining it with the VT voltage of the distribution room, the busbar voltage at the upstream switching station can be inversely derived. According to KVL, the switching station busbar voltages inversely derived from the VTs of all distribution rooms should be exactly equal. However, when a metering deviation exists in the VT of a certain distribution room, an obvious offset will appear in the inversely derived switching station busbar voltage, thereby enabling error evaluation of distribution network VTs. The schematic diagram of the distribution network VT connection topology is shown in Fig. 1.

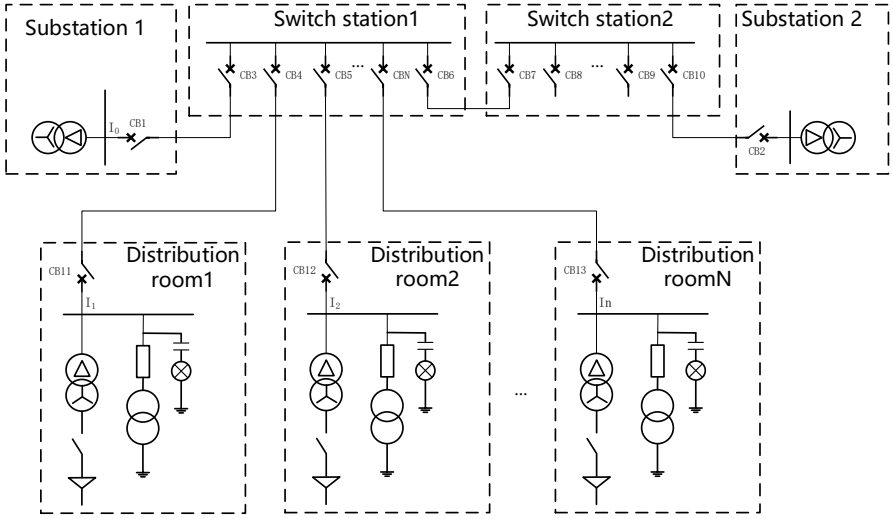


Fig.1 Electrical topology diagram of distribution network

According to the distribution network topology shown in Fig.1, multiple distribution rooms are supplied from the 10kV feeder of the substation through the switching station. Thus, the line voltage drop ΔU_i from the switching station to each distribution room is expressed as follows:

$$\Delta U_i = I_i Z_i, \qquad i = 1,2\ldots,n \tag{1}$$

In this formula, I_i represents the current flowing into distribution room i, and $Z_i = R_i + jX_i$ is the complex impedance of the lines from each distribution room to the switching station.

According to KVL, the busbar voltages at the switching station can be regarded as equal, which yields:

$$U_i + \Delta U_i = U_j + \Delta U_j, \qquad i = 1,2\ldots,n-1; j = 2,3\ldots,n \tag{2}$$

In the formula, U_i and U_j represent the primary voltage values of the distribution rooms i and j respectively.

For the substation side, the following holds:

$$U_0 - \Delta U_0 = U_i + \Delta U_i \tag{3}$$

where U_0 is the primary voltage of the substation busbar, $\Delta U_0 = I_0 Z_0$ is the line voltage drop from the substation feeder to the switching station, I_0 is the current flowing into the switching station, and Z_0 is the complex impedance of the incoming line from the substation side to the switching station.

The measured secondary-side voltage values of the VTs in the distribution rooms and the substation are expressed as follows:

$$u_i = \frac{1}{k}(1+f_i)U_i, \qquad i = 0,1,2\ldots,n \tag{4}$$

where k is the rated transformation ratio of the voltage transformer, and $f_i = \varepsilon_i + j\delta_i$ is the complex error of each VT.

Then, the ratio of the measured secondary-side output voltages of the VTs in adjacent distribution rooms is given by:

$$\frac{u_i}{u_j} = \frac{\frac{1}{k}(1+f_i)U_i}{\frac{1}{k}(1+f_j)U_j}, \qquad i = 1,2\ldots,n-1; j = 2,3\ldots,n \tag{5}$$

Substituting (1) – (4) into (5) and simplifying yields:

$$u_i = [\frac{1+f_i}{1+f_j} + \frac{I_j Z_j - I_i Z_i}{k u_j}(1+f_i)]u_j \tag{6}$$

From (6), it can be observed that the measured secondary-side voltage of the distribution room VT is related to the VT error, line current, and impedance parameters. When the line parameters are known, the value of $\frac{I_j Z_j - I_i Z_i}{k u_j}$ can be calculated. Therefore, it can be treated as a coefficient λ, and the simplified result of (6) is expressed as (7):

$$u_i = [\frac{1+f_i}{1+f_j} + \lambda(1+f_i)]u_j \tag{7}$$

According to (7), $[\frac{1+f_i}{1+f_j} + \lambda(1+f_i)]$ can be regarded as a dynamic parameter β that varies solely with the errors, and a certain equilibrium between u_i and u_j is achieved through this dynamic parameter. Even if the errors of adjacent VTs drift in the same direction, the term $\frac{1+f_i}{1+f_j}$ in β does not change significantly, whereas $\lambda(1+f_i)$ varies with the drift of f_i. Therefore, the value of β can be adopted as an indicator of VT error variation. In an actual distribution network, VTs operating normally constitute the majority, while abnormal VTs are the minority. The characteristics among multiple adjacent distribution room VTs can serve as mutual references, and abnormal VTs can be identified by detecting deviations from the group characteristics. If the specific error of a certain VT is known, it can be used as a reference to identify the error states of the remaining VTs in the group.

2.2 Line parameter identification

Feeder lines are an important component of distribution network systems, and cable parameters directly affect the voltage drop. Under actual distribution network operating conditions, loads fluctuate randomly in real time, and the cable conductor resistance is not constant but rather a time-varying parameter affected by the cable core temperature. Conventional constant-impedance modeling approaches are difficult to satisfy the requirements of high-precision voltage transformer error evaluation. Therefore, an electro-thermal coupling model of the cable is established in this paper to describe the relationship among load, temperature, and line impedance.

The cable conductor resistance is strongly correlated with temperature. The temperature-corrected cable resistance is expressed as (8):

$$R_i(T)=\rho_0\frac{L}{A}[1+\alpha(T-T_0)] \tag{8}$$

where ρ_0 is the conductor resistivity at the reference temperature, L is the line length, A is the conductor cross-sectional area, α is the temperature coefficient of resistance, $T=T_{env}+kI^2$ (T_{amb} is the ambient temperature, k is the equivalent heat dissipation coefficient accounting for the thermal resistance of the insulation layer, sheath, and surrounding environment, which can be calibrated through a limited number of field measurements, and I is the load current), and T_0 is the reference temperature.

Considering the variation of cable core temperature under load fluctuations and the thermal inertia lag effect, the transient temperature rise equation of the cable core is established as (9):

$$\frac{dT}{dt}=\frac{1}{mc}\left[I^2R_{i,0}(1+\alpha(T-T_0))-k(T-T_{env})\right] \tag{9}$$

where $R_{i,0}$ is the conductor resistance at the reference temperature, m is the mass of the cable conductor, and c is the specific heat capacity.

2.3 PINN

The PINN [19] is a class of deep learning frameworks that integrates data-driven modeling with physical priors. Its core idea is to embed the algebraic and differential equations governing system evolution into the loss function in residual form as soft constraints to guide network training. Compared with purely data-driven models, PINNs directly solve the differential terms of physical equations by relying on automatic differentiation, without requiring a large number of labeled samples, and thus exhibit significant advantages in inverse problem parameter identification under low signal-to-noise ratio scenarios.

A standard PINN is based on a fully connected network and adopts the Tanh activation function. The input is the time-series measurement variable $x=[t,I,U]$, and the ambient temperature T_{env} is used as an auxiliary boundary parameter input. The model outputs the set of parameters to be solved:

$$\Theta=[f_0,f_1,\cdots,f_n,R_1(T),X_1,\cdots,R_n(T),X_n] \tag{10}$$

where $f_i=\varepsilon_i+j\delta_i$ is the complex error of the i-th voltage transformer, and $R_i(T)$ and X_i are the time-varying impedance of the i-th feeder.

The total loss function of the standard PINN consists of a data fitting loss and a physical constraint loss:

$$L_{total}=L_{data}+\lambda L_{phys} \tag{11}$$

where L_{data} is the mean squared error between the network output and the measured values, L_{phys} is the residual loss of the physical equation, and λ is the constraint weight coefficient.

2.4 Attention mechanism

The attention mechanism [24] dynamically assigns differentiated weights to input feature sequences, thereby enhancing key information and suppressing redundant information, and can adaptively capture local importance differences in time-series data. Its core lies in computing the correlation between features and generating weight coefficients through three sets of linear mappings, namely Query, Key, and Value, which enables adaptive capture of local importance differences in time-series sequences.

For an input time-series feature sequence $H=[h_1,h_2,\cdots h_T]\in R^{T\times d}$ (where T is the time-series

length and d is the feature dimension), the self-attention is computed as follows:

$$Attention(Q,K,V) = softmax(\frac{QK^T}{\sqrt{d_k}})V \quad (12)$$

where Q, K, and V are the Query, Key, and Value matrices, respectively, obtained by linear transformation of the input features; d_k is the dimension of the key vector, which is used to scale the dot-product result to avoid gradient anomalies.

3. Error evaluation method based on Attention-RBF-PINN

To address the difficulty of decoupling the coupling between dynamic cable voltage drop and intrinsic metering errors of VTs in multi-feeder distribution network scenarios, an Attention-RBF-PINN evaluation method with embedded multi-head self-attention is proposed in this section. The overall architecture is illustrated in Fig. 2.

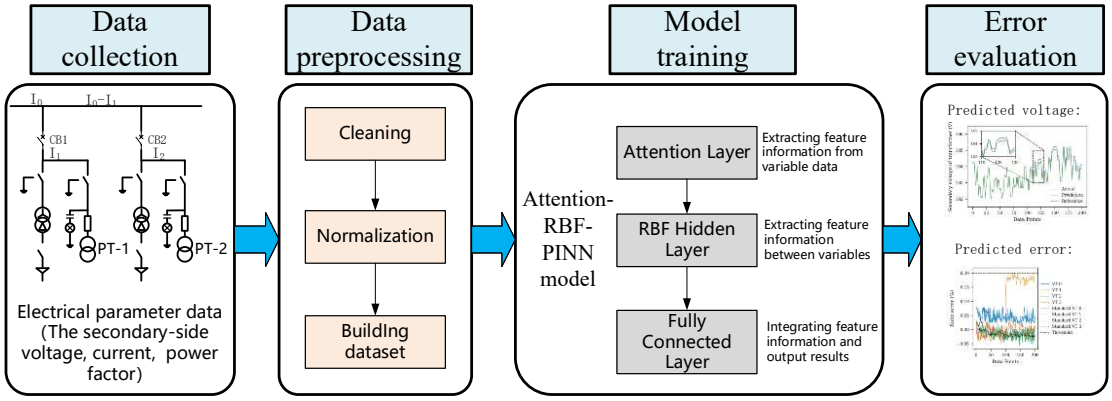


Fig.2 Schematic diagram of the Attention-RBF-PINN

The proposed method takes the voltage topology constraints and the cable thermal balance differential constraints derived as its physical basis, and adopts a three-layer network architecture of "feature weighting – nonlinear mapping – physical constraint." The inputs consist of minute-level feeder current and VT secondary time-series data, while ambient temperature is embedded into the constraint terms as a slowly varying boundary parameter. The outputs take the complex error of each distribution room VT as the core identification target, and simultaneously output the cable core temperature and time-varying impedance as intermediate variables for decoupling. This section focuses on the design of three aspects: the RBF backbone network, multi-head self-attention embedding, and the multi-physics constraint loss function. A joint training strategy is adopted to ultimately achieve synergistic decoupled identification of VT ratio error, phase error, and line time-varying impedance.

3.1 RBF backbone network and automatic differentiation mechanism

PINNs typically employ global activation functions such as Tanh, which apply a uniform nonlinear mapping across all input regions and thus struggle to adaptively capture the multi-timescale fluctuation characteristics ranging from minutes to hours in distribution network load sequences. In addition, if numerical differentiation is adopted to approximately solve the differential constraints, the truncation error of the difference scheme tends to amplify measurement noise, leading to degraded accuracy in solving the time-varying impedance. To this end, Gaussian RBFs are employed in this paper to construct a multi-scale temporal encoding layer. By means of multiple RBF basis functions uniformly distributed along the time axis, the scalar time input is mapped into a high-dimensional multi-scale feature vector, thereby providing rich

temporal representations for the subsequent network. The expression of the Gaussian RBF is given as:

$$\phi(||x - c_i||) = exp(-\frac{||x - c_i||^2}{2\sigma_i^2}) \tag{13}$$

where c_i is the center vector of the i-th hidden layer neuron, and σ_i is the width parameter.

Gaussian RBFs possess the mathematical property of being infinitely continuously differentiable. Combined with automatic differentiation, they can accurately solve the first-order differential term dT/dt of the cable core temperature rise in (9), without relying on numerical differentiation. This fundamentally avoids the interference of difference noise on identification results and provides support for the accurate calculation of cable time-varying impedance and dynamic voltage drop.

3.2 Residual feature attention embedding mechanism

Distribution network loads exhibit strongly time-varying characteristics. During heavy-load periods, the line voltage drop features are pronounced with a high signal-to-noise ratio (SNR), containing richer identification information; during light-load periods, the voltage drop is weak and measurement noise accounts for a high proportion, making identification bias readily introduced. The fully connected structure of a standard PINN treats all input feature channels equally, and the standard self-attention mechanism requires computing a complete attention score matrix, resulting in a high parameter count and computational complexity, which can easily lead to overfitting for small-scale time-series data.

To this end, a more lightweight residual feature attention module is introduced between the input feature layer and the RBF hidden layer in this paper, achieving key feature enhancement through channel-level weight learning. In this module, the importance weights of each input channel are learned through a two-layer fully connected network, and feature enhancement is achieved by means of a residual connection:

$$\tilde{x} = x \odot (1 + \alpha \cdot tanh(W_2 tanh(W_1 x + b_1) + b_2)) \tag{14}$$

where x is the input feature, α is the attention scaling coefficient, generally in the range of [0.3, 0.7], and is taken as $\alpha = 0.5$ in this paper, $\tilde{x}$ is the weighted feature, W_1,W_2,b_1, and b_2 are trainable parameters, and $\odot$ denotes element-wise multiplication. The last layer of the attention branch adopts zero initialization, so that the weights are zero at the initial stage of training and the module degenerates into an identity mapping, thereby ensuring training stability. As training proceeds, the network adaptively learns the importance of each feature channel, strengthening the contribution of high-SNR features and suppressing noise interference. The weighted feature sequence is fed into the RBF hidden layer for nonlinear mapping, and the set of parameters to be identified is output.

This embedding approach enhances the SNR of effective features while preserving the integrity of physical constraints over the entire time period, thereby accommodating the strongly time-varying operating characteristics of distribution network loads.

3.3 Design of the Multi-Physics constrained loss function

Combining the distribution network voltage topology constraint and the cable thermal balance differential constraint derived in Section II, a multi-scale physics-constrained loss function is constructed. The total loss function is defined as:

$$L_{total} = \lambda_1 L_{data} + \lambda_2 L_{topo} + \lambda_3 L_{thermal} \quad (15)$$

where L_{data} is the data fitting loss, L_{topo} is the busbar topology constraint loss, and $L_{thermal}$ is the cable thermal balance differential constraint loss; λ_1, λ_2, and λ_3 are the weight coefficients of the respective constraint terms.

The data fitting loss is constructed based on the mean squared error between the measured and calculated secondary voltage values of the VTs, constraining the model output to fit the measured data:

$$L_{data} = \frac{1}{nT} \sum_{t=1}^{T} \sum_{i=1}^{n} ||\hat{u}_{i,t} - u_{i,t}||^2 \quad (16)$$

where $\hat{u}_{i,t}$ and $u_{i,t}$ are the calculated and measured secondary voltage values of the i-th VT at time t, respectively, n is the number of distribution rooms, and T is the time-series length.

Based on the busbar voltage consistency constraint at the switching station side in (2), the topology constraint loss is constructed from the deviation of the busbar voltages inversely derived from each feeder:

$$L_{topo} = \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} ||(\hat{u}_{i,t} + \Delta\hat{U}_{i,t}) - (\hat{u}_{j,t} + \Delta\hat{U}_{j,t})||^2 \quad (17)$$

where $\Delta\hat{U}_{i,t}$ is the calculated line voltage drop of the i-th feeder at time t.

Based on the transient temperature rise equation of the cable core in (9), the first-order derivative of temperature with respect to time is directly solved by exploiting the automatic differentiation capability of the RBF network, and the differential constraint loss is constructed to constrain the cable core temperature rise process to conform to the thermal balance mechanism:

$$L_{thermal} = \frac{1}{nT} \sum_{t=1}^{T} \sum_{i=1}^{n} \left(\frac{d\hat{T}_{i,t}}{dt} - \frac{1}{mc}\left(I_{i,t}^2 \hat{R}_{i,t} - k(\hat{T}_{i,t} - T_{env})\right)\right)^2 \quad (18)$$

where $\hat{T}_{i,t}$ is the identified cable core temperature of the i-th feeder at time t, $\hat{R}_{i,t}$ is the corresponding time-varying resistance, and $\frac{d\hat{T}_{i,t}}{dt}$ is directly solved by the network through automatic differentiation without numerical differentiation.

The model adopts an end-to-end joint training approach, taking the multi-physics constrained joint loss shown in (15) as the optimization objective, and employs the Adam optimizer for gradient descent, combined with learning rate decay and an early stopping mechanism, until the model converges. Gradient clipping is set during training to prevent gradient explosion, and the model parameters corresponding to the lowest validation set loss are retained as the final result. The key hyperparameter configurations of the proposed model are listed in Table 1.

Table 1. Key hyperparameters of the proposed model

Hyperparameter	**Value**
Optimizer	Adam
Initial Learning Rate	0.00016
Max Epochs	8000
Number of RBF Centers	30
RBF Width σ	0.05
Hidden Dimension	256
Dropout Rate	0.1

4. Experiment

4.1 Experiment data and environment

4.1.1 Dataset

To verify the effectiveness of the proposed algorithm, a typical 10kV distribution network topology in a distribution area of Wuhan is adopted as the engineering prototype in this paper, and a full-link physical simulation platform with adjustable parameters and reproducible operating conditions is designed and constructed in the laboratory, covering the substation outgoing line, switching station, distribution lines, and terminal dedicated transformers, as shown in Fig. 3.

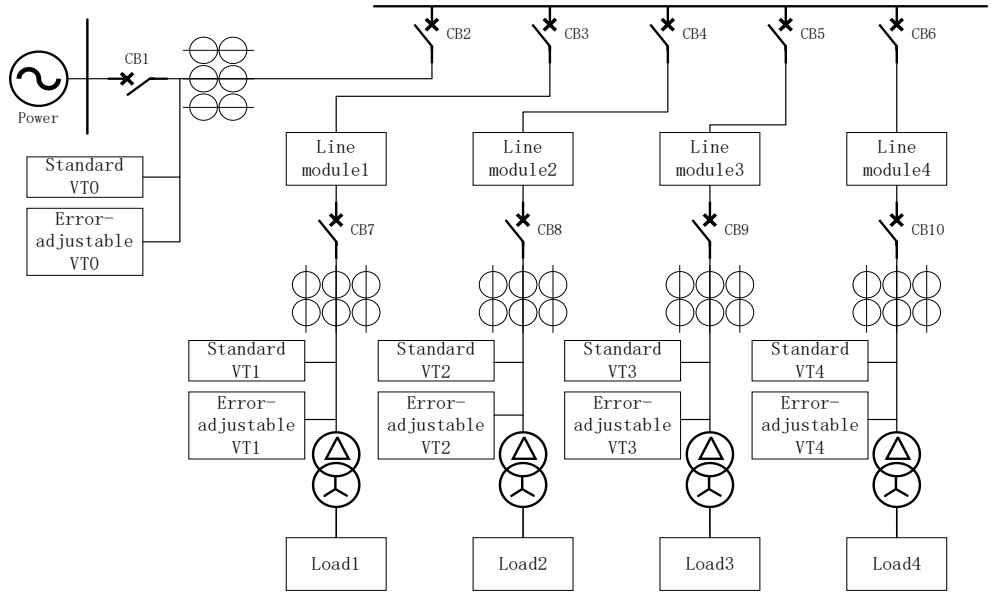


Fig.3 Framework diagram of the physical simulation platform.

The platform consists of a power source, programmable loads, adjustable line modules, standard voltage transformers, voltage transformers with adjustable errors, standard current transformers, and transformers. The injection magnitude and mode of transformer errors can be precisely controlled, providing controllable experimental conditions for quantitative verification of the algorithm. The core parameters of the platform are listed in Table 2.

Table 2. Core module parameters of the simulation platform.

Number	Module	Parameters
1	Three-phase power source	Three-phase four-wire, input 380V, output 0–400V at 37.5 A
2	Programmable load	Three-phase four-wire, input 380V, 0–6A, stability 2%, power factor adjustable in three levels: 0.6/0.85/1
3	Transformer	Rated transformation ratio: 380V/380V; Capacity: 4000VA; Rated frequency: 50Hz
4	Standard VT	Rated transformation ratio: 380V/100V; Accuracy: class 0.02; Rated frequency: 50Hz
5	Standard current transformer(CT)	Rated transformation ratio: 5A/5A, 20A/5A; Accuracy: class 0.02; Rated frequency: 50Hz
6	VT with adjustable error	Rated input voltage: 100V; Rated frequency: 50Hz; Ratio error adjustment range: ±0.01% to ±10%; Ratio error adjustment resolution: 0.01%;

		Phase error adjustment range: ±1′ to ±100′; Phase error adjustment resolution: 1′
7	Line module	Rated current: 6A; Rated frequency: 50Hz; Voltage drop adjustment range: 0–12%

The platform simulates the topology of one substation feeder passing through a switching station and its three subordinate distribution lines. Known ratio errors and phase errors are injected into any branch through the voltage transformers with adjustable errors. Programmable loads are used to simulate the time-series load fluctuations of different types of users, and the acquisition period covers a typical load variation range from valley to peak. The voltage drop characteristics of the transmission cables are simulated by adjusting the equivalent line length parameter. The platform is equipped with 0.05-class high-precision synchronous acquisition equipment to ensure the temporal consistency of multi-node measurement data. Before the experiment begins, the system is operated at full load for no less than 30 minutes until the cable core temperature reaches thermal steady state. Subsequently, the secondary voltage and current signals of each transformer are synchronously acquired at a sampling frequency of 1 Hz, yielding a total of 10,800 valid records. In the experiment, the first 90% of the time-series data are randomly divided into a training set and a validation set at a ratio of 8:1, and the remaining last 10% of the time-series data are used as the test set.

4.1.2 Experimental environment

The GPU used in the experiments of this paper is an NVIDIA GeForce RTX 4060Ti with 8 GB of video memory, combined with CUDA 12.1 for accelerated computing, which satisfies the experimental environment requirements for constructing high-precision composite models. The deep learning models used in the experiments are implemented based on Python 3.11 and PyTorch 2.1.2.

4.1.3 Evaluation indicators

To objectively evaluate the proposed method, mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and the coefficient of determination R^2 are selected as evaluation metrics, with their formulas given as follows:

$$MSE = \frac{1}{n}\sum_{k=1}^{n}(\hat{y}_k - y_k)^2 \tag{18}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(\hat{y}_k - y_k)^2} \tag{19}$$

$$MAE = \frac{1}{n}\sum_{k=1}^{n}|\hat{y}_k - y_k| \tag{20}$$

$$R^2 = 1 - \frac{\sum_{k=1}^{n}(\hat{y}_k - y_k)^2}{\sum_{k=1}^{n}(\bar{y} - y_k)^2} \tag{21}$$

In the above formulas, $\hat{y}_k$ and y_k are the predicted and actual values of the ratio error at the k-th sampling point, respectively, and n is the total number of intercepted samples. MSE, RMSE, and MAE are inversely proportional to the prediction performance of the algorithm; the smaller the metric, the more accurate the prediction result. In contrast, the coefficient of determination R^2 is directly proportional to the prediction performance of the algorithm; the larger the metric, the

better the prediction result matches the actual values, and the more accurate the result.

4.2 Algorithm ablation experiment

To verify the effectiveness of each module in the proposed method, eight groups of ablation experiments are designed. All methods use the same data partitioning and evaluation metrics. The specific experimental comparison results are listed in Table 3:

Table 3. Comparison of ablation experiment results.

Model name	RBF	Attention	PhysicsInformed	Evaluation index			
				MSE	RMSE	MAE	R^2
MLP				0.000019	0.004307	0.003354	0.8728
RBF	√			0.000016	0.004016	0.002935	0.8901
Attention		√		0.000017	0.004169	0.003073	0.8804
PINN			√	0.000014	0.003689	0.002672	0.9064
Attention-RBF	√	√		0.000014	0.003779	0.002604	0.9034
RBF-PINN	√		√	0.000009	0.002952	0.001741	0.9406
Attention-PINN		√	√	0.000006	0.002454	0.001223	0.9587
Attention-RBF-PINN	√	√	√	**0.000003**	**0.001662**	**0.000757**	**0.9810**

As shown in Table 3, the proposed Attention-RBF-PINN method achieves the best performance in terms of MSE, RMSE, and R^2. Taking the optimal data-side combination Attention-RBF as the baseline, after introducing physical constraints, the ratio error RMSE decreases from 0.003779 to 0.001662 (a reduction of 56.0%), and R^2 increases by 7.76 percentage points. The purely physics-constrained method PINN also reduces the RMSE by 14.3% compared with MLP, indicating that the KVL topological constraint and the thermal balance differential constraint themselves provide positive gains. Both RBF and Attention exhibit stable gains with and without physical constraints. With physical constraints, their individual introduction reduces the RMSE by 20.0% and 33.5%, respectively, while their synergistic effect reduces it by 54.9%, outperforming their individual effects. Without physical constraints, they also reduce the RMSE by 6.8% and 3.2%, respectively, indicating that the gains of the data-side modules are independent and that the two are complementary.

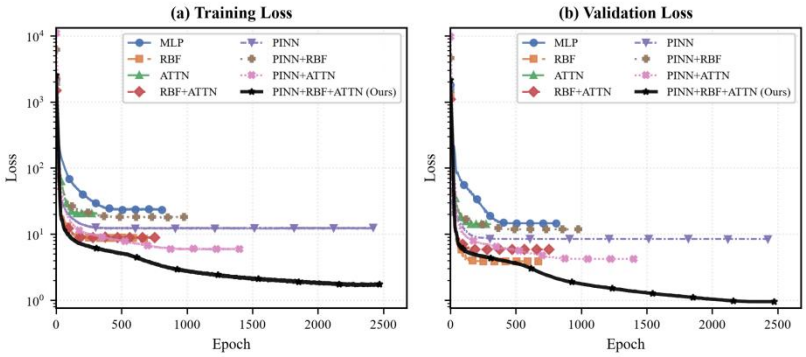


Fig.4 Training and validation loss curves of different models.

To further analyze the training convergence behavior of different modules, Fig. 4 presents the loss variation curves of the eight methods on the training and validation sets. As shown in Fig. 4(a), the four methods without physical constraints, namely MLP, RBF, ATTN, and RBF+ATTN, decrease rapidly in the early stage of training but typically enter a plateau within several hundred epochs and trigger early stopping. The PINN method exhibits the slowest loss decline, remaining

stable at a relatively high level even after more than 2000 epochs of training, reflecting the limited fitting capability of the physics-constrained model without enhancement. In contrast, the proposed method shows a continuous downward trend during training without an obvious plateau, and its final validation loss is significantly lower than those of the other seven methods. In Fig. 4(b), the validation losses of all methods decrease synchronously with the training losses without rebound, indicating that dropout regularization and the early stopping mechanism effectively suppress overfitting, further verifying the enhancement effect of RBF multi-scale features and the attention mechanism on the physics-constrained model.

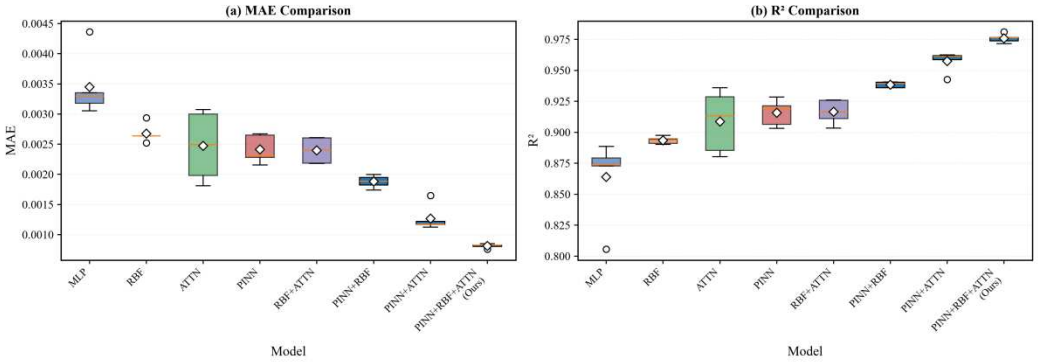


Fig.5 Evaluation metrics of different models

To verify the stability of the experimental results, five independent repeated experiments are conducted for each of the eight methods, and boxplots of MAE and R^2 are plotted, as shown in Fig. 5. Overall, the group without physical constraints exhibits relatively higher box positions and greater dispersion; after introducing physical constraints, the boxes shift downward noticeably and become narrower. The proposed method achieves the optimal box position and small dispersion in both MAE and R^2, with R^2 remaining stable around 0.98 across the five experiments, verifying the robustness and reproducibility of the method.

4.3 Algorithm comparison experiment

The above ablation experiments have verified the effectiveness of the proposed method. To further validate its superiority, six mainstream method models are selected for experimental comparison with the proposed Attention-RBF-PINN algorithm model. All methods use the same input features and data partitioning and are independently tuned. The comparison results of the experimental evaluation metrics are listed in Table 4:

Table 4. Comparison of experimental results of different methods

Model name	Evaluation index			
	MSE	**RMSE**	**MAE**	**R^2**
GradientBoosting	0.000015	0.003811	0.001659	0.9002
PINN	0.000014	0.003687	0.002670	0.9065
TCN	0.000025	0.004950	0.003492	0.8331
MLP	0.000019	0.004307	0.003354	0.8728
Transformer	0.000012	0.003509	0.002605	0.9149
MRSA-DGCN	0.000007	0.002559	0.001191	0.9547
Attention-RBF-PINN	**0.000003**	**0.001662**	**0.000757**	**0.9810**

As shown in Table 4, the proposed Attention-RBF-PINN method achieves the best performance across all evaluation metrics. The compared methods cover three categories: traditional machine

learning, deep learning, and physics-constrained fusion. Among them, MRSA-DGCN achieves the best result among the compared methods through the multi-module fusion of dynamic graph convolution and multi-head self-attention (R^2 = 0.9547), indicating that multi-module collaborative modeling can effectively enhance feature representation capability. Single-structure networks such as MLP and TCN have R^2 values below 0.88 due to limited feature expression capability. The traditional machine learning method GradientBoosting achieves an R^2 of 0.9002, which is comparable to neural network performance. The pure PINN model outperforms MLP and GradientBoosting in terms of R^2, indicating that physical constraints themselves provide positive gains; however, its accuracy remains inferior to that of the best data-driven method, MRSA-DGCN, confirming that physical constraints must be combined with strong feature extraction capability to fully exert their effectiveness. In this paper, RBF multi-scale feature extraction and a residual attention mechanism are further introduced within the physics-constrained framework, raising R^2 to 0.9810. Compared with the second-ranked MRSA-DGCN, the proposed method reduces RMSE by 35.2% and MAE by 36.4%, and improves R^2 by 2.63 percentage points. These results indicate that physical priors provide an interpretable constraint foundation for the model, while RBF multi-scale features and the attention mechanism further enhance representation capability. The organic integration of the two enables the proposed method to comprehensively outperform comparable algorithms in the task of distribution network transformer error evaluation.

4.4 Effect verification

To quantitatively verify the capability of the proposed algorithm in identifying voltage transformer errors, errors are injected through voltage transformers with adjustable errors on the laboratory physical simulation platform, and the secondary-side data of standard voltage transformers, voltage transformers with adjustable errors, and current transformers are synchronously acquired. Among them, the synchronous measurement data of the standard voltage transformers and the voltage transformers with adjustable errors are used to calculate the reference error. The proposed model takes current data and voltage data containing errors as inputs, outputs predicted voltages, and calculates the output evaluation error based on the sampled values, which is then compared with the reference error. Considering that normal transformers constitute the majority and abnormal transformers the minority in actual field conditions, one transformer is selected to superimpose errors on the simulation platform in this paper.

To simulate the abrupt error condition of the transformer, the target transformer is adjusted to a fixed error of 0.2% during the acquisition process. The measurement data containing the abrupt error are input into the proposed model to obtain the restored predicted voltage values, which are then compared with the measured values and reference values to output the evaluation error. The results before and after the abrupt change are shown in Fig. 6:

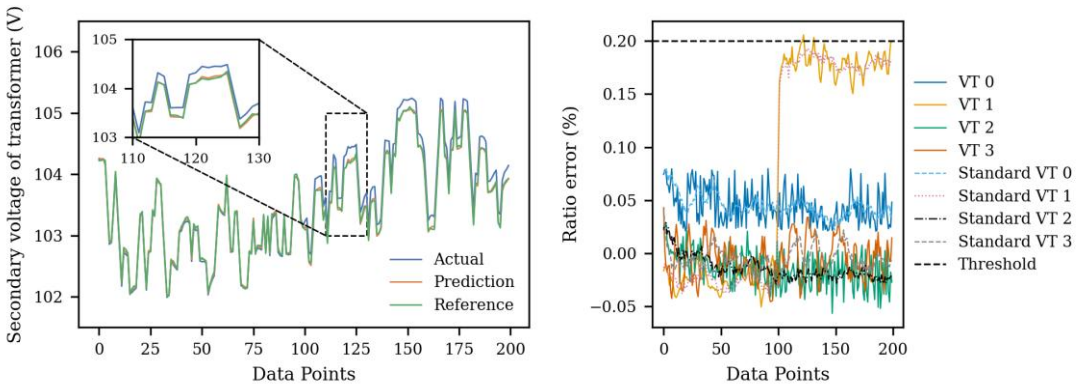


Fig.6 Comparison curves of abrupt error

Furthermore, the deviation between the adjusted evaluation results and the measured errors is calculated to obtain the evaluation result deviation, and a boxplot is drawn as shown in Fig. 7.

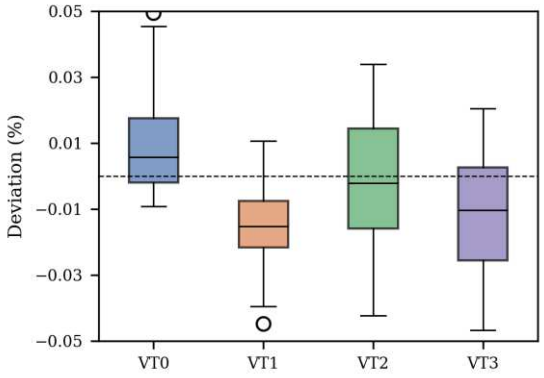


Fig.7 Evaluation deviation diagram of abrupt error

As shown in Fig. 6, under normal operating conditions before the abrupt change, the predicted voltage values of the model agree well with the measured values. After error injection, an obvious step deviation appears in the measured values, whereas the model predicted values can still effectively track the true voltage, identifying the abrupt error to around 0.2%. Fig. 7 shows that the maximum absolute value of the evaluation ratio error deviation does not exceed 0.05%, satisfying the evaluation requirements of class 0.2 VTs.

The gradual degradation condition of transformer performance is further simulated, in which the error of the target transformer is linearly adjusted from 0 to 0.4% during the acquisition process. The comparison between the model output and the measured results is shown in Fig. 8:

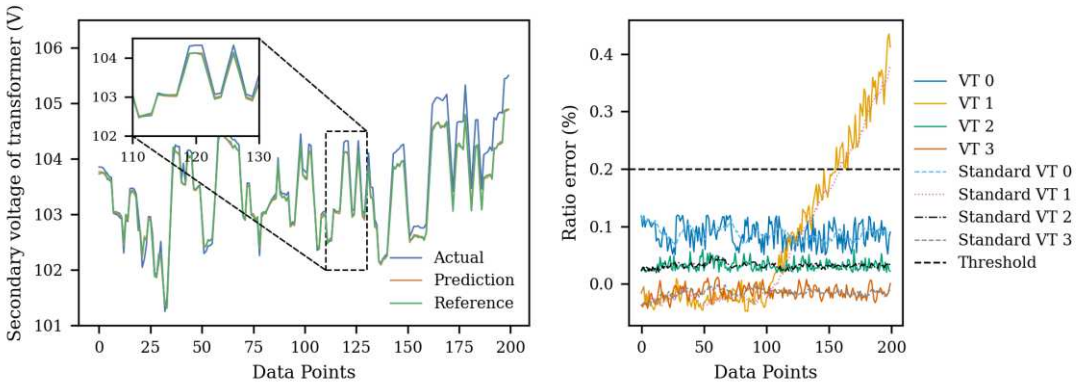


Fig.8 Comparison curves of gradual error

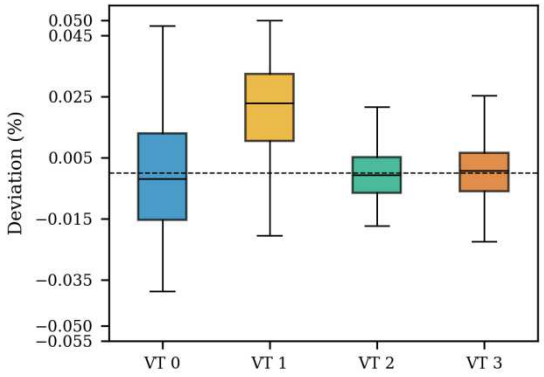


Fig.9 Evaluation deviation diagram of gradual error

As shown in Fig. 8, as the gradual error continues to increase, the measured values gradually deviate from the true values, and the deviation accumulates increasingly over time. The proposed model can continuously identify and correct the gradual drift, and the prediction curve always remains close to the true values, with the error accumulation effectively suppressed. Fig. 9 shows that the maximum absolute value of the gradual error deviation does not exceed 0.05%. The verification results of the above two conditions indicate that the proposed method exhibits good identification capability for both abrupt errors and gradual drift of transformers, and can satisfy the requirements of online evaluation of distribution network transformers.

4.5 Field application

A substation feeder and its three subordinate distribution rooms under a typical 10 kV distribution network topology in a distribution area of Wuhan were selected as the pilot, and a data synchronous acquisition system and an online evaluation system for distribution network transformers with the proposed Attention-RBF-PINN algorithm embedded were deployed. The schematic diagram of the transformer connections in the pilot distribution area is shown in Fig. 10.

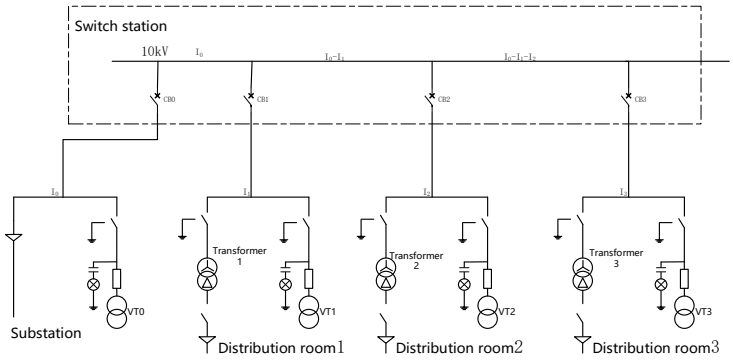


Fig.10 Schematic diagram of TV in the pilot distribution area

Since the proposed algorithm is modeled only for a single phase, phase A of the 10kV TVs at four sites (one substation feeder and three distribution room incoming lines) was selected as the evaluation object. The evaluation workflow is as follows:

(1)The secondary output signals of the four VTs and the corresponding CTs are acquired by the data synchronous acquisition system and synchronously transmitted to the database.

(2)The stored raw data are cleaned, abnormal data are removed, and data are then extracted from the raw database at 5-min intervals, stored in the evaluation database, and updated in real time.

(3)The first 3000 synchronous sample data records are extracted from the evaluation database as the training set to establish the initial evaluation model.

(4)The newly added data are input into the evaluation model in real time to complete the online evaluation of VT operating errors.

Based on the above workflow, online VT error evaluation was carried out when a total of 3375 synchronous sample data records had been accumulated in the evaluation database. The first 3000 synchronous sample data records were used to train the evaluation model, and the remaining 375 records were used for error evaluation. The ratio error evaluation results of the four VTs are shown in Fig. 11:

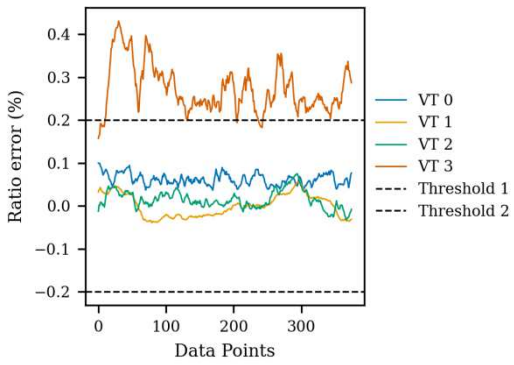


Fig.11 Ratio error evaluation curves of the pilot VTs

As shown in Fig. 11, the ratio error fluctuation of VT3 among the four VTs exceeds the allowable error limit of 0.2% for class 0.2 VTs, indicating that its operating state may be abnormal, while the errors of the other three VTs are within ±0.2%, indicating normal operating states. Based on the above evaluation results, on-site de-energized verification was carried out on the four VTs, and the verification results are listed in Table 5.

Table 5. Comparison of verification and evaluation results

Number	Offline evaluation results			Online evalution results			Compare results
	Ratio error (%)	Phase error (')	condition	Ratio error (%)	Phase error (')	condition	
VT0	0.13	1.3	Normal	0.06	0.5	Normal	Consistent
VT1	0.05	3.2	Normal	0.01	2.2	Normal	Consistent
VT2	0.11	2.8	Normal	0.02	1.5	Normal	Consistent
VT3	0.39	0.8	Abnormal	0.28	0.3	Abnormal	Consistent

The verification results show that the ratio errors and phase errors of VT0, VT1, and VT2 are all within the normal allowable error range of class 0.2 VTs, indicating that these three VTs are in normal states, whereas the ratio error of VT3 exceeds the allowable range of class 0.2 VTs, indicating that the VT is abnormal. Overall, the maximum absolute deviation of the ratio error between the online evaluation and the de-energized verification is 0.11%, and the maximum absolute deviation of the phase error is 1.3′ . The above offline verification results are all consistent with the online evaluation results, demonstrating that the proposed online evaluation method achieves favorable performance in practical application scenarios.

5. Conclusion

This paper proposes an Attention-RBF-PINN method that integrates physical constraints and

data-driven modeling. In this method, the KVL topological constraints of the switching station and the cable thermal equation are first adopted as physical priors, and the circuit laws are embedded into the neural network loss function. Subsequently, radial basis functions are introduced for multi-scale temporal feature extraction to capture the temporal dynamics of load fluctuations. Finally, a residual attention mechanism is employed to adaptively weight the multi-branch input features, thereby achieving online evaluation of VT error states.

Experiments based on the laboratory physical simulation platform demonstrate that the proposed method outperforms the compared methods in terms of MSE, RMSE, MAE, and R^2, and exhibits favorable identification capability for 0.2% abrupt errors and 0.4% gradual errors. In the field application in a distribution area of Wuhan, the proposed method successfully identified one out-of-tolerance VT, and the evaluation results were consistent with the de-energized verification results. However, certain limitations remain in the proposed method: the current model is established only for single-phase modeling, and its performance is constrained by the synchronously acquired voltage and current data. In practice, sensor faults, missing data, and environmental noise may affect the evaluation accuracy. Future research can be extended to three-phase unbalanced scenarios, incorporate the joint evaluation of CT errors, and explore the fusion of data augmentation and physical constraints to reduce dependence on data quality.

In summary, the proposed method provides a new idea for the online evaluation of distribution network VT error states, and offers a theoretical basis and technical support for condition assessment and anomaly warning of distribution network metering equipment.

Declarations

Availability of data and materials

Not applicable.

Competing interests

Not applicable.

Funding

The study was supported by “Research and Application of Online Identification Technology for Metering Inaccuracy of Medium-Voltage Distribution Network Transformers” (5700-202422270A-1-1-ZN).

Authors' contributions

In this study, the contributions of the authors are as follows:

Dongsheng Li: Contributed to the conception and design of the study, performed the experiments/data collection, analyzed and interpreted the data, and drafted the manuscript.

Xiujuan Zeng: Contributed to the design of the study, provided critical intellectual input during the experimental phase, performed statistical analyses, and critically revised the manuscript for important intellectual content.

Tong Liu、Qilin Zhang: Assisted in data collection and analysis, reviewed relevant literature, provided valuable feedback on the manuscript, and approved the final version for submission.

Wei Wu: Supervised the study, ensured compliance with ethical standards, provided administrative and logistical support, and reviewed and approved the manuscript.

Acknowledgements

Not applicable.

Authors' information

Dongsheng Li (1975-), male, Senior Engineer, engaged in power technology development, power marketing and customer service, and integrated energy management.

Xiujuan Zeng (1976-), female, bachelor, main research interests include power system automatic control and measurement technology.

Tong Liu (1972-), male, bachelor, Engineer, has been engaged in electric energy metering for more than 30 years.

Qilin Zhang (1976-), female, bachelor, main research interest is electric energy metering.

Wei Wu (1985-), female, bachelor, main research interest is electric energy metering.

References

1. Battal F, Balci S, Sefa I (2021) Power electronic Transformers: A review. Measurement 171. https://doi.org/10.1016/j.measurement.2020.108848

2. Kong T-S et al (2020) Study on reliability improvement of voltage Transformers by increasing voltage factor. J Electr Eng Technol 15:1463-1469. https://doi.org/10.1007/s42835-020-00374-0

3. De Reis A, Leite R, Washington Luiz Araújo Neves, and, Lopes FV (2019) Coupling capacitor voltage transformers models and impacts on electric power systems: A review. IEEE Transactions on Power Delivery 34: 1874 – 1884. https://doi.org/10.1109/tpwrd.2019.2908390

4. Crotti G et al (2021) Instrument transformers for power quality measure-ments: A review of literature and standards. IEEE 11th International Workshop on Applied Measurements for Power Systems (AMPS). IEEE, 2021. https://doi.org/10.1109/amps50177.2021.9586028

5. Chen Y, Liu K, Feng Y et al (2022) Review on harmonic monitoring technology of capacitor voltage transformer. Power Syst Prot Control 43:70–75. https://doi.org/10.14044/j.1674-1757.pcrpc.2022.05.011

6. Liu S, Zhang Y, Yang Y et al (2023) Review on common faults of electromagnetic potential transformers and the countermeasures. J Xihua Univ (Nat Sci Ed) 42:47–58. https://doi.org/10.12198/j.issn.1673-159X.4779

7. National Technical Committee for Electromagnetic Metrology of China, Subcommittee on High Voltage Metrology (2022) Instrument transformers — part 4: voltage transformers in power system. JJG 1189.4-2022. China Standards Press.

8. Xu Y, Li H, Zhang C (2024) Error quantitative evaluation method of potential transformer based on optimal modeling and deep belief network. Electr Meas Instrum 61:55–62. https://doi.org/10.19753/j.issn1001-1390.2024.08.007.

9. Zhang B (2024) Research on the impact of new distribution network harmonics on metering and online assessment method of transformers. Dissertation, Wuhan University of Technology. https://doi.org/10.27381/d.cnki.gwlgu.2024.002366.

10. Khandeparkar KV, Soman SA, Gajjar G (2017) Detection and correction of systematic errors in instrument transformers along with line parameter estimation using PMU data. IEEE Trans Power Syst 32:3089–3098. https://doi.org/10.1109/TPWRS.2016.2620990

11. Goklani H, Gajjar G, Soman SA (2018) Phase segregated soft calibration of instrument transformers using synchronised phasor measurements. In: 2018 Power Systems Computation Conference (PSCC), Dublin, Ireland. IEEE, 2018:1-7. https://doi.org/10.23919/PSCC.2018.8442518
12. Zeng X, Liu T, Xie H et al (2025) Measurement error evaluation method for voltage transformers in distribution networks based on self-attention and graph convolutional networks. Energy Inform 8:69. https://doi.org/10.1186/S42162-025-00525-5.
13. Cao Z, Yang J, Zhang W et al (2024) Capacitor voltage transformer error prediction method based on fusion convolution module. In: 2024 6th International Conference on Energy, Power and Grid (ICEPG). IEEE, 2024:1–5. https://doi.org/10.1109/ICEPG63230.2024.10775656
14. Gao L, Zheng K, Liao T (2024) Error prediction for capacitor voltage transformer based on graph convolution and temporal convolution. In: 2024 4th International Conference on Energy, Power and Electrical Engineering (EPEE). IEEE, 2024:695–698. https://doi.org/10.1109/EPEE63731.2024.10875391
15. Wang P, Shi C, Yu B et al (2025) Voltage transformer error state prediction method considering multi-dimensional attributes. In: 2025 7th International Conference on Energy Systems and Electrical Power (ICESEP). IEEE, 2025:1-5. https://doi.org/10.1109/ICESEP66633.2025.11155479
16. Zhang Z, Li H, Tang D et al (2017) Monitoring the metering performance of an electronic voltage transformer on-line based on cyber-physics correlation analysis. Meas Sci Technol 28:105015. https://doi.org/10.1088/1361-6501/aa8024
17. Zhang C, Li H, Yang J et al (2019) Detecting measurement error drifts of a capacitor voltage transformer on-line and its field application. Meas Sci Technol 30:105109. https://doi.org/10.1088/1361-6501/ab199a
18. Bao Z, Jiao Y, Chen M et al (2021) On-line anomaly detection for the measurement error of HVPT in the multi-bus structure. Measurement 170:108722. https://doi.org/10.1016/j.measurement.2020.108722
19. Ryu I, Park GB, Lee Y et al (2024) Physics-informed neural network for engineers: a review from an implementation aspect. J Mech Sci Technol 38:3499–3519. https://doi.org/10.1007/s12206-024-0624-9
20. Liu M et al (2021) Activated gradients for deep neural networks. IEEE Trans Neural Netw Learn Syst 34:2156–2168. https://doi.org/10.1109/tnnls.2021.3106044
21. Chen X, Li Y, Bai Y et al (2025) Identification algorithm for positive- and zero-sequence parameters of double-ended transmission lines based on RBF-PINN. J Tianjin Univ (Sci Technol) 58:1242–1254. https://doi.org/10.11784/tdxbz202412025
22. Ren Z, Li C, Qi P et al (2026) A two-stage hybrid prognostic method for power MOSFET based on PINN and attention-Mamba. Power Electron. https://doi.org/10.20222/j.cnki.cn61-1124/tm.20260610.004.
23. Zou Y, Wang Q, Yan Q et al (2024) Structural dynamic response analysis based on parameter-balanced fourier-feature PINN. Eng Mech. https://doi.org/10.6052/j.issn.1000-4750.2024.12.0935
24. Niu Z, Zhong G, Yu H (2021) A review on the attention mechanism of deep learning. Neurocomputing 452:48 – 62. https://doi.org/10.1016/j.neucom.2021.03.091