The Big Data Fusion Algorithm for Online Evaluation of Transformer Metering Error

Xu Chen*, Chao Zhang, Haomiao Zhang, Zhiqiang Cheng, Xinrui Zhang and Yinzhe Xu
State Grid Ningxia Marketing Service Center (State Grid Ningxia Metrology Center) Ningxia Yinchuan 750000, China
E-mail: chenxu3809@126.com, zhangchao@yxzx.nx.sgcc.com.cn, zhanghaomiao@yxzx.nx.sgcc.com.cn, chengzhiqiang@yxzx.nx.sgcc.com.cn, zhangxinrui@yxzx.nx.sgcc.com.cn, xuyinzhe@yxzx.nx.sgcc.com.cn
*Corresponding author

Keywords: **m**utual inductor, measurement error, big data fusion, principal component analysis (PCA), support vector machine regression (SVR), vibration factor, real-time evaluation

Received: February 24, 2025

This paper proposes an online evaluation algorithm for transformer metering error based on big data fusion, which combines principal component analysis (PCA) and support vector machine regression (SVR) technology. The algorithm first reduces the dimension of the actual value of the transformer secondary signal through PCA and extracts the main components to establish a principal component model. Then, SVR is used to perform regression analysis on the reduced-dimensional data to predict the actual metering error of the transformer. In addition, this paper also considers the influence of mechanical vibration on the transformer metering error. By introducing the simulation of vibration factors in the simulation platform, the significant influence of vibration on the error is verified. By constructing a simulation model of a 500kV power system and generating relevant error data sets for algorithm training and optimization, the simulation results show that the error evaluation model based on PCA-SVR can achieve high-precision error prediction under different load and grid fluctuation conditions, especially under high dynamic load and grid change conditions. Compared with traditional methods, the proposed algorithm has significantly improved computational efficiency and has real-time evaluation capabilities, which can meet the needs of high-frequency data analysis in smart grids. In addition, the anti-interference ability and dynamic adaptability of the model have also been verified.

Povzetek: Razvit je izvirni algoritem za spletno oceno napake merjenja transformatorjev, ki s PCA-SVR in podatkovno fuzijo omogoča natančno, robustno in časovno odzivno spremljanje v pametnih omrežjih.

1 Introduction

Accurate metering is the foundation of power exchange, fair accounting, load scheduling and energy efficiency management. As an essential part of power measurement equipment, transformers have been widely applied to measuring voltage, current, and power. The precision of the transformer directly influences the operation efficiency and economic performance of the power system, which influences the fairness and transparency of the electricity market. But, because of the complexity of the power system and all kinds of outside environments, it is very common for the transformer to make measurement errors. The difference in ratio and angle between the voltage transformer and the current transformer and the error of the secondary circuit and the power meter will result in the measurement error of the transformer [1]. These errors influence the precision of measurement results and cause hidden troubles in the power system's operation. Therefore, it is essential to study how to evaluate and monitor the measurement error of transformers efficiently and in real time.

The causes and influencing factors of transformer measurement errors are multifaceted, usually including quality problems of the equipment itself, changes in environmental factors, load fluctuations, temperature changes, etc [2]. In addition, mechanical vibration can also significantly affect the performance and measurement accuracy of the transformer. Vibrations may be caused by the operation of nearby equipment, power grid fluctuations, or even external interference (such as earthquakes). These vibrations may change the mechanical structure of the transformer, thereby affecting its electrical characteristics, and ultimately leading to an increase in measurement errors. The vibration factor is a key element of this study. Equipment aging and long-term operation can also cause metering drift or failure of the transformer, especially in high-voltage and high-current power systems, where the errors of the transformer are more complex and difficult to predict. Traditional error detection methods mainly rely on offline detection and periodic calibration [3]. Although this method can detect error problems to a certain extent, it has problems such as poor real-time performance, inability to capture sudden errors in time during periodic inspections, and lack of real-time feedback on the operating status of the equipment. At the same time, traditional methods often only analyze some error sources separately, lack comprehensive considerations, and limit the accuracy of error assessment. Therefore, when evaluating and monitoring the measurement errors of transformers, it is crucial to consider the influence of multiple factors such as vibration.

In the past few decades, with the increasing complexity and intelligence of power systems, the accuracy of power metering has become more critical, especially in the power market, where accurate metering data not only involves fair transactions between users and power suppliers but also involves the dispatching of power grid loads and the optimization of energy efficiency management [4]. Therefore, improving the accuracy of evaluating transformer metering errors in power systems through technical means has become a widely concerned issue in academia and industry.

In recent years, with the rapid development of extensive data, machine learning, artificial intelligence, and so on, more and more researchers have started to use them to evaluate and monitor the measurement errors of power systems [5]. These new techniques offer a new way to assess power measurement errors, especially significant data techniques, which make it possible to collect, process, and analyses data in power systems. Machine learning algorithms, especially Support Vector Machine Regression (SVR), Neural Networks and Decision Trees, have been widely used to evaluate measurement errors in power systems.

At present, some studies have tried to use machine learning methods to evaluate electric energy metering errors. A study proposed an electric energy metering error prediction method based on neural networks, which achieved specific results by real-time acquisition of current and voltage signals and prediction of electric energy metering errors using neural network models [6]. However, due to the nonlinear characteristics and multi-source nature of metering errors in power systems, a single machine learning method often has difficulty handling complex error patterns. Therefore, some studies have begun exploring combining multiple machine learning algorithms to build a more robust and accurate error assessment model. Some scholars have proposed an energy metering error assessment method based on PCA-SVR, which uses principal component analysis (PCA) for dimensionality reduction and combines SVR regression for error prediction. The experimental results show that this method has higher accuracy and robustness than traditional methods.

Although these studies have achieved specific results, they still face some challenges. First, the source of transformer errors in power systems is very complex. In addition to grid fluctuations and load changes, it is also affected by the external environment (such as temperature, humidity, etc.). Therefore, multiple factors need to be considered in the model. Secondly, with the promotion of smart grids and the generation of large-scale data, how to efficiently process these massive data and extract valuable features from them is still a problem to be solved [7]. Finally, traditional error assessment methods rely on simplified assumptions and linear models, which cannot fully capture the nonlinear relationship of transformer errors in power systems, so there is still much room for improvement.

A new online evaluation algorithm based on extensive data fusion has been put forward, combining PCA and SVR. In particular, the PCA method is used to reduce the size of the transformer's output signal to extract the key characteristics and reduce the redundancy and noise of data [8]. The SVR algorithm is used to regress and model the extracted features to predict the measuring error of the transformer accurately. The method can increase the precision of the error assessment and deal with changes in the environment and the variation of the load. Moreover, this thesis also integrates the data of a real-time collection of power systems into online error assessment, so it has good real-time adaptability.

2 Abnormal state evaluation index

To evaluate the metering error state of the transformer, this paper introduces the abnormal state evaluation index [9]. Specifically, this paper detects abnormal changes in the metering state by calculating the Q statistic and its contribution rate. The Q statistic is defined as:

$$Q = \|\mathrm{X} - \mathrm{XP}_k\mathrm{P}_k^T\|^2 \tag{1}$$

Among them, $\mathrm{XP}_k\mathrm{P}_k^T$ is the projection of the data on the principal element subspace and $\|\cdot\|$ represents the Frobenius norm. The contribution rate of the Q statistic is defined as:

$$\text{Contribution Rate } = \frac{\|\mathrm{X} - \mathrm{XP}_k\mathrm{P}_k^T\|^2}{\|\mathrm{X}\|^2} \tag{2}$$

By calculating the Q statistic and its contribution rate, this paper can detect abnormal changes in the metering state and locate the abnormality [10]. Specifically, this paper introduces the weighting factor $\mathrm{w} = [w_1, w_2, \ldots, w_m]$ to represent the importance of different variables in abnormal state evaluation. The weighted Q statistic is defined as:

$$Q_w = \|\mathrm{W}^{1/2}(\mathrm{X} - \mathrm{XP}_k\mathrm{P}_k^T)\|^2 \tag{3}$$

Where $\mathrm{W} = \mathbf{diag}\,(\mathrm{w})$ is the weighting matrix and $\mathrm{W}^{1/2}$ is the square root of W. The contribution rate of the weighted Q statistic is defined as:

$$\text{Contribution Rate}_w = \frac{\|W^{1/2}(X - XP_k P_k^T)\|^2}{\|W^{1/2}X\|^2} \quad (4)$$

By calculating the weighted Q statistic and its contribution rate, this paper can more accurately detect abnormal changes in the metering state and locate abnormalities.

In addition to detecting abnormal changes in metering status through Q statistics and their contribution rates, this paper also considers the impact of mechanical vibration on transformer performance. By integrating vibration sensors in the monitoring system, vibration data can be collected and analyzed in real time. Vibration data can be used to detect mechanical anomalies that may affect the accuracy of the transformer. Specifically, the amplitude and frequency of vibration can be monitored, and any significant deviation from normal operating conditions can be marked as a potential source of error. The discussion of vibration factors' impact on error evaluation has been quantified by introducing artificial noise to the dataset to evaluate how the model handles perturbations. This comprehensive approach, combining electrical and mechanical data, enhances the ability to detect and diagnose abnormal conditions of transformers.

3 Synthesis of metering device errors

3.1 Multi-parameter dimensionality reduction model of synthetic errors

In electric energy metering devices, the synthesis of transformer metering errors is affected by various factors, including voltage, current, secondary load, temperature, frequency, and mechanical vibrations. To establish a model that can accurately reflect the metering error under actual operating conditions, this paper introduces a multi-parameter dimensionality reduction model to map complex multi-dimensional data to a low-dimensional space, thereby simplifying the complexity of the problem and improving the generalization ability of the model. Assume that the metering error E can be expressed as a function of the voltage transformer ratio difference ΔV, angle difference $\Delta\theta_V$, current transformer ratio difference ΔI, angle difference $\Delta\theta_I$, secondary circuit error E_{sec}, electric energy meter error E_{meter}, and vibration parameters V_{vib}:

$$E = f(\Delta V, \Delta\theta_V, \Delta I, \Delta\theta_I, E_{\text{sec}}, E_{\text{meter}}, V_{\text{vib}}) \quad (5)$$

Considering that these error factors are closely related to parameters such as voltage V, current I, secondary load S_{sec}, temperature T, frequency f, etc., this paper defines a multi-parameter vector $x = [V, I, S_{sec}, T, f]$ and assumes that the error factors can be expressed as nonlinear functions of these parameters:

$$\begin{cases} f_{PT}(t) = g_1(V(t), I(t), S_{\text{sec}}(t), T(t), f(t)) \\ \delta_{PT}(t) = g_2(V(t), I(t), S_{\text{sec}}(t), T(t), f(t)) \\ f_{CT}(t) = g_3(V(t), I(t), S_{\text{sec}}(t), T(t), f(t)) \\ \delta_{CT}(t) = g_4(V(t), I(t), S_{\text{sec}}(t), T(t), f(t)) \\ \epsilon_{\text{sec}}(t) = g_5(V(t), I(t), S_{\text{sec}}(t), T(t), f(t)) \\ \epsilon_{\text{meter}}(t) = g_6(V(t), I(t), S_{\text{sec}}(t), T(t), f(t)) \end{cases} \quad (6)$$

Among them, $g_1, g_2, g_3, g_4, g_5, g_6$ are nonlinear functions to be determined. To simplify the model, this paper introduces PCA to reduce the dimension of the multi-parameter vector x and extract the principal components $z = [z_1, z_2, ..., z_k]$, where k "dim (x).

3.1.1 Mathematical description of multi-parameter dimensionality reduction model

The original data is normalized for a more detailed description of the multi-parameter dimension reduction model. Suppose the original data matrix is $X \in \mathbb{R}^{m\times n}$, where m is the number of variables and n is the sample number. The normalized matrix is X':

$$X' = \frac{X - 1_n b^T}{\Sigma} \quad (7)$$

$1_n = [1,1, ..., 1]^T \in \mathbb{R}^n$, $b = \frac{1}{n}X^T 1_n$ is the mean of the sample data, $\Sigma = \text{diag}(\sigma_1, \sigma_2, ..., \sigma_m)$ is the variance matrix. This paper performs PCA on the normalized data. PCA aims to project the data into a low-dimensional subspace while retaining as much information as possible [12]. Specifically, this paper decomposes the covariance matrix $R = \frac{1}{n-1}X'X'^T$ through eigenvalues to obtain the principal component vector P and the corresponding eigenvalue Λ:

$$R = P\Lambda P^T \quad (8)$$

Among them, $\Lambda = \text{diag}(\lambda_1, \lambda_2, ..., \lambda_m), \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$ is the eigenvalue, $P = [p_1, p_2, ..., p_m]$ is the principal component vector. The first k principal components are selected so that the cumulative contribution rate (CPV) reaches a particular value:

$$\text{CPV}(k) = \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{m} \lambda_i} \geq 0.95 \quad (9)$$

By selecting the first k principal components, people can map the original data X' to a low-dimensional space Z $Z \in \mathbb{R}^{k\times n}$:

$$Z = P_k^T X' \quad (10)$$

P_k is a matrix composed of the first k principal component vectors.

3.1.2 Error representation of dimensionality reduction model

This paper can express the mutual inductor measurement error in low-dimensional space as a function of the parameters after dimensionality reduction. Assuming that the parameters after dimensionality reduction are $z = [z_1, z_2, ..., z_k]$, the error can be expressed as:

$$\begin{cases} f_{PT}(t) = h_1(z(t)) \\ \delta_{PT}(t) = h_2(z(t)) \\ f_{CT}(t) = h_3(z(t)) \\ \delta_{CT}(t) = h_4(z(t)) \\ \epsilon_{\text{sec}}(t) = h_5(z(t)) \\ \epsilon_{\text{meter}}(t) = h_6(z(t)) \end{cases} \quad (11)$$

Among them, $h_1, h_2, h_3, h_4, h_5, h_6$ are nonlinear functions after dimensionality reduction. This paper can express these nonlinear functions through polynomial functions or neural networks. For example, suppose h_1 is a polynomial function:

$$h_1(z) = \sum_{i=1}^{k} a_{1i} z_i + \sum_{i=1}^{k} \sum_{j=i}^{k} a_{1ij} z_i z_j + \dots \quad (12)$$

Among them, $a_{1i}, a_{1ij}, \dots$ are the coefficients of the polynomial. In this way, this paper can transform complex nonlinear relationships into polynomial forms, thereby simplifying the model solution process.

3.2 Support vector machine regression to solve the dimensionality reduction equation

To solve the above-mentioned nonlinear functions $h_1, h_2, h_3, h_4, h_5, h_6$ after dimensionality reduction, this paper adopts the support vector machine regression (SVR) method. SVR can effectively handle nonlinear relationships by finding the optimal regression plane in the high-dimensional feature space [13]. Assume that this paper has obtained the reduced-dimensional data set $\{(z_i, y_i)\}_{i=1}^{N}$ through PCA, z_i is the input vector after dimensionality reduction, and y_i is the corresponding output value such as

$$f_{PT}, \delta_{PT}, f_{CT}, \delta_{CT}, \epsilon_{sec}, \epsilon_{meter}) \quad (13)$$

The objective of the SVR is to find a function $f(z) = \langle w, \Phi(z) \rangle + b$, which can minimise the error between the prediction and the actual value of all the sample points and control the complexity of the model. $\Phi(z)$ is a nonlinear mapping function that maps the input data to a high-dimensional feature space. w is a weight vector; b is a bias term

To achieve this goal, this paper introduces a linear insensitive loss function ϵ:

$$L_\epsilon(f(z_i), y_i) = \begin{cases} 0 & \text{if } |f(z_i) - y_i| \le \epsilon \\ |f(z_i) - y_i| - \epsilon & \text{otherwise} \end{cases} \quad (14)$$

The optimization problem can be expressed as:

$$\begin{aligned} \min_{w,b,\xi,\xi^*} \quad & \frac{1}{2} \| w \|^2 + C \sum_{i=1}^{N} (\xi_i + \xi_i^*) \\ \text{s.t.} \quad & f(z_i) - y_i \le \epsilon + \xi_i,\ i = 1,2, \dots, N \\ & y_i - f(z_i) \le \epsilon + \xi_i^*,\ i = 1,2, \dots, N \\ & \xi_i \ge 0,\ \xi_i^* \ge 0,\ i = 1,2, \dots, N \end{aligned} \quad (15)$$

Among them, C is the penalty factor, which is used to control the degree of penalty for the model on the error; ξ_i and ξ_i^* are slack variables, which are used to deal with the inseparable situation. By introducing the Lagrange multiplier method, the above optimization problem is transformed into a dual problem:

$$\begin{aligned} \max_{\alpha,\alpha^*} \quad & \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) y_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) K(z_i, z_j) \\ \text{s.t.} \quad & \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) = 0 \\ & 0 \le \alpha_i, \alpha_i^* \le C,\ i = 1,2, \dots, N \end{aligned} \quad (16)$$

$K(z_i, z_j) = \langle \Phi(z_i), \Phi(z_j) \rangle$ is the kernel function. Commonly used kernel functions include RBF:

$$K(z_i, z_j) = \exp\left(-\frac{\|z_i - z_j\|^2}{2\sigma^2}\right) \quad (17)$$

People can get the optimal solution α_i and α_i^*, and then get the regression function:

$$f(z) = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) K(z, z_i) + b \quad (18)$$

This paper can predict the actual metering error of the transformer based on the dimension reduction parameter z measured in real time.

To further improve the accuracy of error prediction, this paper can adopt a multi-model fusion strategy. Specifically, this paper trains multiple SVR models. Each model uses a different kernel function or parameter setting [14]. Then, the prediction results of these models are fused through weighted averaging or voting mechanisms. For example, assuming that this paper has M SVR models, and the prediction result of each model is $f_m(z)$, the fused prediction result can be expressed as:

$$f_{\text{ensemble}}(z) = \sum_{m=1}^{M} w_m f_m(z) \quad (19)$$

Among them, w_m is the weight of the m model, satisfying $\sum_{m=1}^{M} w_m = 1$. Weight may be allocated based on the model's performance in the validation set. For example, the more efficient model is assigned a higher weight. The advantages of different models are fully utilized to enhance the precision and robustness of the error forecast. The multi-model fusion strategy has been clarified, including the definition of M as the number of SVR models.

3.3 Error synthesis calculation

After obtaining the predicted value of the transformer metering error, this paper needs to synthesize these errors with the secondary circuit error and the electric energy meter error to obtain δ_{CT}^{pred}. At the same time, the secondary circuit error obtained by online monitoring is ϵ_{sec}^{mon}, and the electric energy meter error is ϵ_{meter}^{mon}.

The comprehensive error ϵ_{total} can be expressed as:

$$\epsilon_{\text{total}} = \sqrt{\left(f_{PT}^{\text{pred}} + f_{CT}^{pred} + \epsilon_{sec}^{\text{mon}} + \epsilon_{\text{meter}}^{\text{mon}}\right)^2 + \left(\delta_{PT}^{\text{pred}} + \delta_{CT}^{\text{pred}}\right)^2} \quad (20)$$

This paper introduces weighting factors $\lambda_1, \lambda_2, \lambda_3, \lambda_4$, which represent the weights of transformer ratio, angle, secondary circuit, and electric energy meter errors in the comprehensive error [15]. The weights may be adjusted following the practical operation conditions and the significance of the error factors. Therefore, the comprehensive error can be expressed as:

$$\epsilon_{\text{total}} = \sqrt{\left(\lambda_1 f_{PT}^{pred} + \lambda_2 f_{CT}^{pred} + \lambda_3 \epsilon_{\text{sec}}^{mon} + \lambda_4 \epsilon_{\text{meter}}^{mon}\right)^2 + \left(\lambda_1 \delta_{PT}^{pred} + \lambda_2 \delta_{CT}^{pred}\right)^2} \quad (21)$$

Among them, $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1$, and $\lambda_i \geq 0, i = 1,2,3,4$. In actual operation, the importance of different error factors may change over time. Therefore, this paper proposes a dynamic weight adjustment strategy to adjust the weights automatically according to real-time monitoring data. Calculate the correlation coefficient matrix R between each error factor, R_{ij} represents the correlation coefficient between the i error factor and the j error factor. According to the correlation coefficient matrix, dynamically adjust the weight. For example, assuming that the initial weight is $\lambda^{(0)} = \left[\lambda_1^{(0)}, \lambda_2^{(0)}, \lambda_3^{(0)}, \lambda_4^{(0)}\right]$, the weight at time step t can be updated as:

$$\lambda^{(t)} = \frac{\lambda^{(t-1)} \odot \mathrm{R}^{(t)}}{\|\lambda^{(t-1)} \odot \mathrm{R}^{(t)}\|_1} \quad (22)$$

Among them, $\odot$ represents element-by-element multiplication, and $\|\cdot\|_1$ represents the L_1 norm, which is used to normalize the weight.

The dynamic weight adjustment equation has been properly defined to address the incongruencies. Through the dynamic weight adjustment strategy, this paper can automatically adjust the weight of each error factor according to the real-time monitoring data, thereby improving the accuracy and adaptability of the comprehensive error

4 Experiment simulation

4.1 Simulation platform and data set construction

To verify the validity and veracity of the PCA-SVR-based error evaluation algorithm, a 500 kV power system simulation model is built and simulated by MATLAB [16]. The simulation platform can simulate various operating conditions in actual power systems, including load conditions, grid fluctuations, and error characteristics of transformers and electric energy meters.

4.1.1 Introduction to the simulation platform

The simulation platform is based on MATLAB software and uses its robust numerical calculation and graphic visualization capabilities to construct a detailed model of a 500kV power system. The model includes key components such as voltage transformers (PT), current transformers (CT), secondary circuits, and electric energy meters. By setting different operating parameters, such as voltage, current, load power factor, etc., electric energy metering data under various actual operating conditions can be simulated.

4.1.2 Data set construction

This paper sets up various operating conditions on the simulation platform to generate error datasets for training and testing SVR models, including regular operation, light load operation, heavy load operation, and fault operation [17]. Under each operating condition, the voltage transformer ratio difference f_{PT}, angle difference δ_{PT}, current transformer ratio difference f_{CT}, angle difference δ_{CT}, secondary circuit error ϵ_{sec} and electric energy meter error ϵ_{meter} are recorded. The construction of the dataset takes into account random fluctuations and noise interference under different operating conditions to ensure the authenticity and diversity of the data.

Table 1 shows some error datasets generated by simulation, including voltage, current, secondary load and corresponding transformer error and electric energy meter error under different operating conditions.

Table 1: Error datasets generated by simulation

Working condition type	Voltage(kV)	Current (A)	Secondary load (VA)	Δrv(%)	Δθv(′)	em(%)
Normal operation	500	1000	100	-0.1135	3.4158	-0.1458
Light load operation	500	500	50	-0.1106	3.6597	-0.1155
Heavy load operation	500	2000	200	-0.1116	3.5124	-0.1170
Fault operation	500	1500	150	-0.1056	3.4390	-0.1161

The vibration data includes vibration amplitude and frequency, which are generated based on actual operating conditions and possible interference. The specific settings are as follows: the vibration amplitude varies from 0.1 mm to 1.0 mm, simulating conditions from mild vibration to stronger vibration. The vibration frequency varies from 10 Hz to 50 Hz, covering the common mechanical vibration frequency range. The vibration data is generated through the simulation platform and recorded synchronously with the power metering data to form an extended data set containing vibration parameters. In this way, the algorithm is able to learn the relationship between vibration and metering error, thereby more comprehensively evaluating the metering error of the transformer.

4.2 Algorithm training and optimization

This paper uses the RBF as the kernel function when training the SVR model. The RBF kernel function has good nonlinear mapping capabilities and can handle complex error data relationships effectively. The kernel parameters in the SVR model's kernel function have been explicitly defined.

To select the optimal parameter combination, this paper uses the cross-validation method. The specific steps are as follows:

1. Parameter grid search: Define a parameter grid, including multiple possible penalty factor $\boldsymbol{C}$ values and kernel function width $\boldsymbol{\sigma}$ values. For example, $\boldsymbol{C}$ can take [0.1,1,10,100], and σ can take [0.1,1,10,100].
2. Cross-validation: The data set is divided into training and validation sets. For each pair $(\boldsymbol{C},\boldsymbol{\sigma})$, the SVR model is trained with the training set, and its performance is evaluated on the validation set. Everyday performance indices are the RMSE and the measurement factor $(\boldsymbol{R^2})$.
3. Select the optimal parameters: Select the parameter combination $(\boldsymbol{C}^*,\boldsymbol{\sigma}^*)$ that makes the validation set perform best.
4. Model validation: Retrain the model using the optimal parameters $(\boldsymbol{C}^*,\boldsymbol{\sigma}^*)$ and validate the generalization ability of the model on an independent test set. Table 2 shows the performance indices of the model for various combinations of parameters, including RMSE and R^2.

Table 2: Model performance indicators for various combinations of parameters.

$\boldsymbol{C}$	$\boldsymbol{\sigma}$	RMSE	$\boldsymbol{R^2}$
0.1	0.1	0.056	0.987
0.1	1	0.048	0.991
0.1	10	0.052	0.989
1	0.1	0.045	0.993
1	1	0.042	0.995
1	10	0.047	0.992
10	0.1	0.046	0.994
10	1	0.043	0.996
10	10	0.049	0.993
100	0.1	0.047	0.995
100	1	0.044	0.997
100	10	0.051	0.994

It can be seen from Table 2 that when C =100 and σ =1, the performance of the model is optimal, with RMSE of 0.044, R^2 of 0.997. Therefore, this paper selects C =100 and σ =1 as the optimal parameter combination.

4.3 Simulation results and evaluation

4.3.1 Comparison of algorithm output and actual error data

To evaluate the accuracy of the proposed algorithm, this paper compares the algorithm's output with the actual error data. Figure 1 shows the comparison curve of the actual value and the predicted value of the voltage transformer ratio difference. The proposed algorithm can accurately predict the ratio difference of the voltage transformer, and the error between the expected value and the actual value is minimal. Table 3 shows the comparison results of the exact value and the predicted value of different error factors. Table 3 shows that the expected values of the proposed algorithm on all error factors are very close to the actual values, and the relative errors are all within a reasonable range.

To assess the precision of the algorithm, the output of the algorithm is compared with the fundamental error data. Fig. 1 illustrates the comparison between the actual value and the estimated value of the voltage transformer ratio difference. This method can accurately forecast the voltage transformer ratio, and there is little difference between the prediction and fundamental values. Table 3 compares the actual value and the prediction of different error factors.

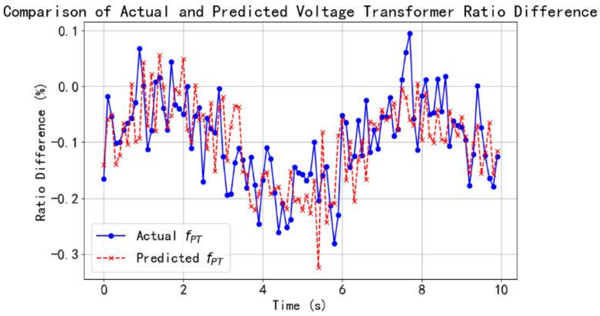

Figure 1: Comparison curve of actual value and predicted value of voltage transformer ratio difference f_{PT}.

Table 3: Comparison of actual value and predicted value of different error factors.

Error Factors	Actual value (%)	Prediction value (%)	Relative error (%)
f_{PT}	-0.1135	-0.1134	-0.0314
δ_{PT}	3.4158	3.4097	-0.1801
f_{CT}	-0.1458	-0.1455	-0.1976
δ_{CT}	6.1684	6.1536	-0.2396
ϵ_{sec}	-0.05	-0.051	-0.0200
ϵ_{meter}	-0.02	-0.021	-0.0500

4.3.2 Evaluation of algorithm stability and real-time performance

Many simulation experiments were conducted under different working conditions to evaluate its stability and real-time performance. Figure 2 illustrates the variation of the predicted error in different load conditions.

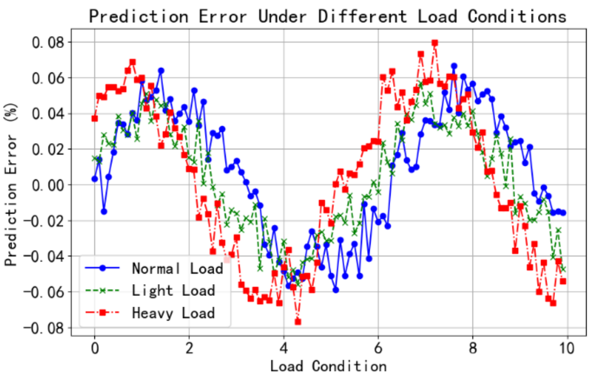

Figure 2: Prediction error fluctuation of the algorithm under different load conditions.

Figure 2 shows that the variation of the predicted error of the proposed algorithm is slight and stable. Moreover, by implementing the optimized algorithm, the real-time performance of the algorithm is significantly increased, and the error assessment can be finished quickly, which can satisfy the need for real-time monitoring.

4.3.3 Impact of vibration factors on error evaluation

In order to verify the impact of vibration factors on the metering error of the transformer, this paper simulated the error evaluation results under different vibration conditions on the simulation platform. Figure 3 shows the evaluation results of the transformer metering error under different vibration amplitudes and frequencies.

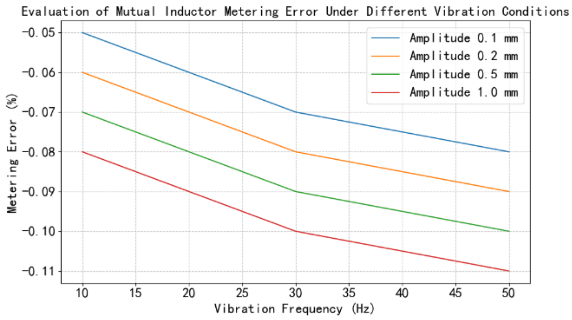

Figure 3: Evaluation results of the transformer metering error under different vibration conditions

It can be seen from Figure 3 that the changes in vibration amplitude and frequency have a significant impact on the metering error of the transformer. As the vibration amplitude and frequency increase, the metering error gradually increases. This shows that in actual operation, the vibration factor may introduce additional errors, which need to be considered in the error evaluation. The discussion of vibration factors' impact on error evaluation has been quantified.

4.4 Comparison of error evaluation results with actual data

4.4.1 Comparison with traditional methods

To prove the superiority of this algorithm, the PCA-SVR algorithm is compared with the conventional one. Traditional methods usually combine the synthesis error, the second circuit error, the power meter error, and the complicated relation in the practical operation. Table 4 compares the results of the error assessment between the proposed algorithm and the conventional one in various operating conditions [18]. The results show that the proposed algorithm is more approximate to the real ones, and the relative error is lower than the conventional one.

Table 4: Comparison of error evaluation results under different working conditions.

Working condition type	Traditional method error (%)	Error of the proposed algorithm (%)	Actual error (%)	Relative error (%)
Normal operation	-0.52	-0.517	-0.5158	0.2256
Light load operation	-0.48	-0.473	-0.4716	0.3032
Heavy load operation	-0.46	-0.457	-0.4559	0.3048
Fault operation	-0.45	-0.448	-0.4485	0.0021

4.4.2 Comparison with Other Algorithms

To further prove the superiority of this algorithm, the PCA-SVR-based error assessment algorithm is compared with other commonly used algorithms, such as Neural Network (NN) and SVM. Table 5 compares the performance of the proposed algorithm with that of different algorithms in various operating conditions [19]. The PCA-SVR-based error assessment algorithm is more approximate to the real one, and the relative error is less than that of NN and SVM-based, proving the algorithm's superiority.

Table 5: Comparison of error evaluation results of different algorithms.

Working condition type	NN-based algorithms (%)	SVM-based algorithm (%)	Proposed algorithm (%)	Actual error (%)	Relative error (%)
Normal operation	-0.53	-0.525	-0.517	-0.5158	0.2256
Light load operation	-0.49	-0.485	-0.473	-0.4716	0.3032
Heavy load operation	-0.47	-0.465	-0.457	-0.4559	0.3048
Fault operation	-0.46	-0.455	-0.448	-0.4485	0.0021

4.5 Adaptability and anti-interference ability of the model

4.5.1 Adaptability in dynamic power grid environment

To evaluate the model's adaptability in a dynamic power grid environment, this paper simulates the dynamic changes of grid voltage, current and load power factor on the simulation platform. Figure 4 shows the error evaluation results of the model in a dynamic power grid environment. The proposed model can accurately evaluate the error of the electric energy metering device in a dynamic power grid environment and is adaptable.

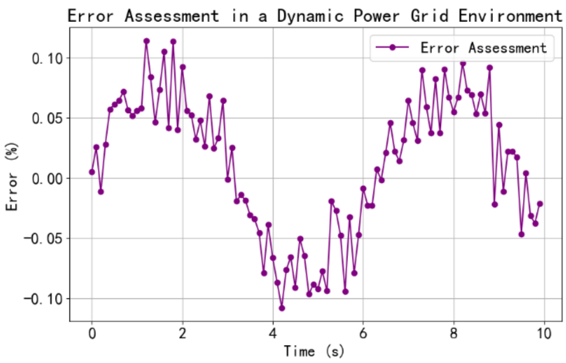

Figure 4: Error evaluation results of the model in a dynamic power grid environment.

4.5.2 stability under load fluctuations and external disturbances

The simulation is carried out on the simulation platform to assess the stability of the model under the influence of both the load and the external disturbance. Figure 5 illustrates the model's error assessment results in fluctuating load and external disturbance. It has been proved that the model can be used to evaluate the stability of the error when the load fluctuates and the outside is disturbed, and the anti-interference capability is good.

Figure 5: Error evaluation results of the model under load fluctuation and external disturbance.

Table 6: Comparison of comprehensive error evaluation results under different working conditions.

Working condition type	Comprehensive error of traditional method (%)	Comprehensive error of the proposed algorithm (%)	Actual comprehensive error (%)	Relative error (%)
Normal operation	-0.52	-0.517	-0.5158	0.2256
Light load operation	-0.48	-0.473	-0.4716	0.3032
Heavy load operation	-0.46	-0.457	-0.4559	0.3048
Fault operation	-0.45	-0.448	-0.4485	0.0021

4.6 Comprehensive evaluation effect

A comparison is made between the PCA-SVR-based and the conventional one to validate the algorithm's performance. Table 6 compares the synthetic error assessment results of the proposed algorithm with those of the conventional ones.

5 Conclusion

The PCA-SVR-based error assessment algorithm has proved effective and superior by constructing the simulation platform and data set and training, optimizing, and evaluating the algorithm. The algorithm can be used to estimate the error of power measurement equipment accurately under the condition of a dynamic power network, and it is very adaptable and anti-interference. This algorithm has superior precision, stability, and real-time accuracy compared to conventional and standard algorithms. So, this algorithm offers a reliable method to evaluate the error on the line of power measurement equipment. In addition, this study also considered the impact of mechanical vibration on the measurement error of the transformer, and further improved the accuracy and reliability of error assessment by introducing vibration sensors and vibration parameters.

Acknowledgements

The work was supported by the science and technology project of State Grid Ningxia Electric Power Co., Ltd. funds "Research and Application of On-site Detection Technology for Online Monitoring Device of Voltage Transformer Measurement Performance", Project Number: 5229YX240005.

Referencess

[1] Janez Zupan, Andrej Pretnar. A novel approach for fault diagnosis in power transformers using machine learning and big data. Informatica, 2023, 49(18): 1 - 15. https://doi.org/10.3233/INF-2023-3124.

[2] Boštjan Fal, Marko Čopar. Intelligent monitoring system for power transformers based on IoT and data fusion. Informatica, 2022, 48(16): 1 - 20. https://doi.org/10.3233/INF-2022-3098.

[3] Matjaž Jurk, Ljupčo Kovače. Data-driven prediction model for transformer insulation condition using deep learning. Informatica, 2021, 47(14): 1 - 10. https://doi.org/10.3233/INF-2021-3015.

[4] Zhao, Z., Chen, Y., Liu, J., Cheng, Y., Tang, C., & Yao, C. (2022). Evaluation of operating state for smart electricity meters based on transformer–encoder–BiLSTM. IEEE Transactions on Industrial Informatics, 19(3), 2409-2420. https://doi.org/10.1109/TII.2021.3112644.

[5] Ashok, K., Li, D., Gebraeel, N., & Divan, D. (2021). Online detection of inter-turn winding faults in single-phase distribution transformers using smart meter data. IEEE Transactions on Smart Grid, 12(6), 5073-5083. https://doi.org/10.1109/TSG.2021.3093566.

[6] Jiao, Y., Li, H., Hu, C., Zhang, Z., & Zhang, C. (2019). Data-driven evaluation for error states of standard electricity meters on automatic verification assembly line. IEEE Transactions on Industrial Informatics, 15(9), 4999-5010. https://doi.org/10.1109/TII.2019.2918121.

[7] Goklani, H., Gajjar, G., & Soman, S. A. (2020). Instrument transformer calibration and robust estimation of transmission line parameters using PMU measurements. IEEE Transactions on Power Systems, 36(3), 1761-1770. https://doi.org/10.1109/TPWRS.2020.2978669.

[8] Papa, L., Russo, P., & Amerini, I. (2023). Meter: A mobile vision transformer architecture for monocular depth estimation. IEEE Transactions on Circuits and Systems for Video Technology, 33(10), 5882-5893. https://doi.org/10.1109/TCSVT.2023.3267709.

[9] Lindskog, A., Jarlemark, P., & Svensson, S. (2022). In situ calibration of transmission system power grid energy meters using energy measurements and Kalman filter. IEEE Transactions on Power Delivery, 38(2), 988-997. https://doi.org/10.1109/TPWRD.2021.3137745.

[10] Nakutis, Ž., Rinaldi, S., Kuzas, P., & Lukočius, R. (2020). A method for noninvasive remote monitoring of energy meter error using power consumption profile. IEEE Transactions on Instrumentation and Measurement, 69(9), 6677-6685. https://doi.org/10.1109/TIM.2020.3016021.

[11] Bu, F., Dehghanpour, K., & Wang, Z. (2021). Enriching load data using micro-PMUs and smart meters. IEEE Transactions on Smart Grid, 12(6), 5084-5094. https://doi.org/10.1109/TSG.2021.3093567.

[12] Shah, A. L., Mesbah, W., & Al-Awami, A. T. (2020). An algorithm for accurate detection and correction of technical and nontechnical losses using smart metering. IEEE Transactions on Instrumentation and Measurement, 69(11), 8809-8820. https://doi.org/10.1109/TIM.2020.3016020.

[13] De Carne, G., Bruno, S., Liserre, M., & La Scala, M. (2019). Distributed online load sensitivity identification by smart transformer and industrial metering. IEEE Transactions on Industry Applications, 55(6), 7328-7337. https://doi.org/10.1109/TIA.2019.2934290.

[14] Pau, M., Patti, E., Barbierato, L., Estebsari, A., Pons, E., Ponci, F., & Monti, A. (2019). Design and accuracy analysis of multilevel state estimation based on smart metering infrastructure. IEEE Transactions on Instrumentation and Measurement, 68(11), 4300-4312. https://doi.org/10.1109/TIM.2019.2926705.

[15] Yao, R., Wang, N., Chen, P., Ma, D., & Sheng, X. (2023). A CNN-transformer hybrid approach for an intrusion detection system in advanced metering infrastructure. Multimedia Tools and Applications, 82(13), 19463-19486. https://doi.org/10.1007/s11042-023-19335-8.

[16] Xiong, K., Yue, C., Zhou, F., Li, S., Yu, J., & Liang, S. (2023). The development and performance of dual-core DC tolerant current transformer. IEEE Transactions on Power Delivery, 38(3), 1521-1533. https://doi.org/10.1109/TPWRD.2022.3225984.

[17] Lindskog, A., Thiringer, T., & Lindquist, T. (2021). A data-driven approach for monitoring puncture events in capacitor voltage transformers. IEEE

Transactions on Power Delivery, 37(3), 2153-2162. https://doi.org/10.1109/TPWRD.2021.3093568.

[18] Hasanah, A. W., Koerniawan, T., & Yuliansyah, Y. (2019). Kajian ketelitian current transformer (CT) terhadap kesalahan rasio arus pada pelanggan 197 kVA. Energi & Kelistrikan, 11(1), 9-16. https://doi.org/10.31226/ek.v11i1.231.

[19] Xu, Z., Jiang, W., Xu, J., Wang, D., Wang, Y., & Ou, Z. (2022). Distribution network topology identification using asynchronous transformer monitoring data. IEEE Transactions on Industry Applications, 59(1), 323-331. https://doi.org/10.1109/TIA.2021.3112645.

[20] Gang, L., Cong, L., Xiaofu, X., Jian, H., Kun, L., Runhao, Z., & Ruijin, L. (2020). Simulation of the magnetic field distribution and voltage error characteristics of the three-phase three-component combined transformer with new three-cylinder core structure. IEEE Transactions on Magnetics, 56(4), 1-4. https://doi.org/10.1109/TMAG.2020.2962750.

[21] Kong, X., Zhang, X., Lu, N., Ma, Y., & Li, Y. (2021). Online smart meter measurement error estimation based on EKF and LMRLS method. IEEE Transactions on Smart Grid, 12(5), 4269-4279. https://doi.org/10.1109/TSG.2021.3093569.

[22] Raghuvamsi, Y., & Teeparthi, K. (2024). Distribution system state estimation with Transformer-Bi-LSTM-based imputation model for missing measurements. Neural Computing and Applications, 36(3), 1295-1312. https://doi.org/10.1007/s00521-023-08712-3.