ISSN 1330-3651 (Print), ISSN 1848-6339 (Online) https://doi.org/10.17559/TV-20221109032034
Original scientific paper

Error Model and Forecasting Method for Electronic Current Transformers Based on LSTM

Kun WANG*, Hu LI, Huan LI, Jinggeng GAO

Abstract: As an important metering apparatus at the trade settlement gate in intelligent substations, the operating error of electronic current transformers can have an important impact on the electric energy trade settlement, so it is necessary to predict the error state of electronic current transformers. In this paper, a Long Short-Term Memory (LSTM) neural network is used to construct an error prediction model for electronic current transformers, characterizing their errors in the form of multiple input variables and single output variables. In order to reduce the training scale of the LSTM neural network, the partitioning around medoid (PAM) clustering algorithm is used to cluster and analyze the input variables. The analysis results of the algorithm show that the prediction results of the ratio and phase errors meet the requirements of online monitoring and provide information on the change of the error state of electronic current transformers to prevent the risk of electricity trade settlement.

Keywords: electronic current transformer; error state; LSTM; PAM clustering algorithm

1 INTRODUCTION

As one of the main sensing devices of the smart grid, electronic current transformers are widely used in smart substations because of their simple insulation structure, digital interface, wide measurement band, large dynamic range,and other advantages [1-5], which support the development of the digital construction of the smart grid in Li et al. [6].

The primary sensing elements and secondary remote modules of electronic current transformers are installed in situ in the high-voltage equipment area, where electromagnetic interference and environment are more serious compared to the control room. Field application research shows that the stability and reliability of the electronic current transformer error state are poor, and environmental factors such as temperature, humidity, vibration, and magnetic field can have an impact on the electronic current transformer output [7, 8], which in turn affects the fairness of the electric energy trade settlement and the safe operation level of the power grid in Liu et al. [9].

Electronic current transformer error evaluation usually uses offline calibration or online calibration methods to obtain the ratio error and phase error of electronic current transformers by direct comparison method [10-12]. However, these methods have long calibration cycles, complex field wiring, and low work efficiency. In order to improve the electronic current transformer error state evaluation system, it is urgent to establish the prediction method of the electronic current transformer error state, to discover its error overrun problem in time, to reduce the electronic current transformer error overrun operation time, and to guide the electronic transformer testing work, so as to ensure the fairness of electric energy metering.

In order to evaluate the operating errors of electronic transformers, the correlation between primary transmission networks is analyzed to establish a new evaluation standard without a standard transformer. The standard can be calculated as a characteristic statistic in Zhang et al. [13], by using the measured data based on the VN-MWPCA proposed. In order to solve the problem that there is no effective way to predict the error-developing trend of electronic transformers, two kinds of short-term prediction methods for electronic transformer error based on the backpropagation neural network and the prophet model are proposed in Ye et al. [14], respectively. The data fitting and short-term prediction of electronic transformer error are made on the basis of the backpropagation neural network and the Prophet model, and the fitting and prediction results of the two methods are compared and analyzed in combination with four evaluation indexes. Aiming to monitor the online status frequently and conveniently, an attention mechanism-optimized Seq2Seq network is proposed in Xiong et al. [15] to predict the error state of transformers, which combines an attention mechanism, Seq2Seq network, and bidirectional long short-term memory networks to mine the sequential information from online monitoring data of electronic transformers. At present, there are few studies related to electronic transformer error state assessment and the method is not perfect, so it is necessary to conduct an in-depth study on the electronic transformer error state prediction method.

In this paper, the error degradation model of electronic transformers under the influence of a multidimensional environment is analyzed, and further, proposed an error prediction method for electronic current transformers based on Long Short-Term Memory (LSTM) neural network and PAM clustering analysis to realize the error prediction of electronic current transformer based on the historical data of operating error and environmental parameters of electronic current transformers. The error prediction of the electronic current transformer is achieved based on the historical data of the operating error and environmental parameters. Finally, the feasibility of the method is verified by building an online monitoring platform for electronic current transformers and collecting electronic current transformer error data and environmental monitoring data in real-time.

2 ERROR MODEL CONSTRUCTION

The parameters that affect the error of electronic current transformers mainly include electric field, magnetic field, load, frequency, harmonics, temperature, humidity, vibration, and communication links. The correlation analysis method based on random matrix theory [16] is used to analyze the measured data of a 330 kV digital substation in Gansu. By comparing the average spectral radius of the evaluation matrix and the reference

matrix, it is possible to analyze the correlation between the electronic current transformer error state and the influencing quantity. The I_{MSR} of the electronic current transformer ratio difference is used as the evaluation index of correlation [17], and the larger the I_{MSR}, the greater the degree of influence of the influence quantity on the error state of the electronic current transformer. According to the calculation results in Tab. 1, it is 174.23 for temperature, 57.61 for humidity, 236.18 for load, 29.83 for vibration, and 285.33 for magnetic field. It can be seen that the error state of the electronic current transformer is more correlated with the temperature, magnetic field, and load parameters. According to the comparison results, the temperature, magnetic field, and load parameters are mainly considered in this paper.

Table 1 Correlation calculation results of the influence parameters of electronic current transformers

Influence parameters	Sampling period / min	I_{MSR}
Temperature	15	174.23
Humidity	15	57.61
Load	15	236.18
Vibration	15	29.83
Magnetic field	15	285.33

The electronic current transformer error can be represented as a model containing one output variable and several input variables

$$\begin{cases} \alpha = f_1(x_1, x_2, x_3) \\ \beta = f_2(x_1, x_2, x_3) \end{cases} \tag{1}$$

where, α is the ratio error and β is the phase error; x_1, x_2, and x_3 are the magnetic field, load variable, and temperature, respectively. Assuming that there are n sets of data samples for electronic current transformer error and environmental parameters, the error sequence of electronic transformers can be expressed as:

$$\begin{cases} \alpha_1 = f_1(x_{11}, x_{12}, x_{13})\ \beta_1 = f_2(x_{11}, x_{12}, x_{13})\ \alpha_2 = \\ = f_1(x_{21}, x_{22}, x_{23})\ \beta_2 = f_2(x_{21}, x_{22}, x_{23}); \\ \alpha_n = f_1(x_{n1}, x_{n2}, x_{n3})\ \beta_n = f_2(x_{n1}, x_{n2}, x_{n3}) \end{cases} \tag{2}$$

where, α_i is the i-th group of ratio error data; β_i is the i-th group of phase error data; x_{i1}, x_{i2}, x_{i3} are the i-th group of magnetic field, load variable and temperature parametres. Assuming that there are n sets of data samples for electronic current transformer error and environmental parameters, the error sequence of electronic current transformer can be expressed as:

$$\begin{cases} A = XY_1\ B = XY_2 \end{cases} \tag{3}$$

where, $A = [\alpha_1, \alpha_2, ..., \alpha_n]^T$, $B = [\beta_1, \beta_2, ..., \beta_n]^T$, and Y_1 and Y_2 are the matrices to be solved. In the field operation process, the phase error and the ratio error of the electronic current transformer are the results of the cross action of each environmental parameter, and there is no clear functional relationship between the error and the environmental parameter, so it is difficult to get the analytical values of Y_1 and Y_2.

3 ERROR PREDICTION METHOD

A neural network can be understood as a black box that can fit arbitrary functions. The correspondence between input and output quantities can be established by iterative computation of the neural network on the training set, which is a neural network structure built with input, hidden, and output layers as the core. As long as the training data is sufficient, the desired output results can be obtained given a specific amount of input.

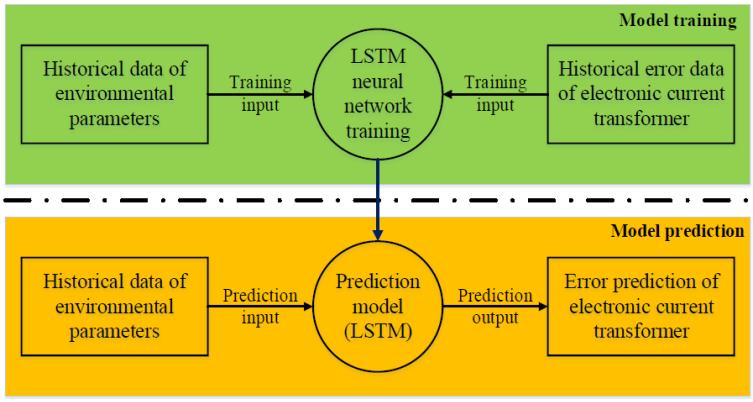

Figure 1 The flow chart of error predicting procedure for electronic current transformer

Since the analytical expression of the electronic current transformer error is difficult to solve, in order to clarify the mapping relationship between the electronic current transformer error state and environmental parameters, a data-driven approach can be considered to establish a prediction model based on the historical data of the electronic current transformer operation error and environmental parameters by introducing LSTM neural network algorithm [18, 19] and PAM cluster analysis algorithm [20, 21], so that the error of electronic current transformer is predicted, as shown in Fig. 1 (the idea of predicting the error state of electronic current transformers).

3.1 LSTM Neural Network

LSTM neural networks are a special form of RNN neural networks, first proposed by Hocheriter et al. [22]

and subsequently improved by Graves [23] to form the current model. Currently, there are two mainstream training methods for recurrent neural network models such as LSTM, the backward error propagation algorithm by time expansion [24, 25] and the real-time recursive learning algorithm [26]. The most important feature of RNN neural networks is that the units of each hidden layer are correlated and also related to the input before the moment of reception of that hidden layer also known as "memorability".

As shown in Fig. 2, an RNN neural network hidden layer unit expansion consists of an input layer, hidden layer, and output layer. U, V, and W are the weights from input to hidden layer, hidden layer to hidden layer, and hidden layer to output, respectively. x is the input of the RNN neural network, o is the output of the RNN neural network, and S is the current moment state of the hidden layer. The training method of iterative computation at each step of RNN greatly reduces the training time of traditional neural networks.

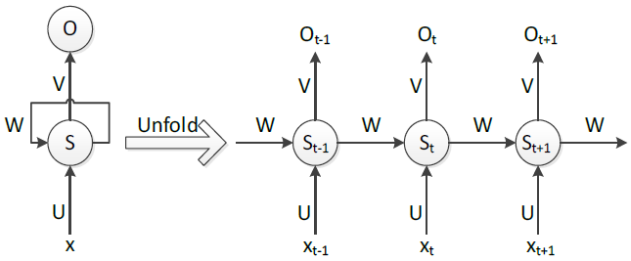

Figure 2 Schematic diagram of single RNN hidden layer unit

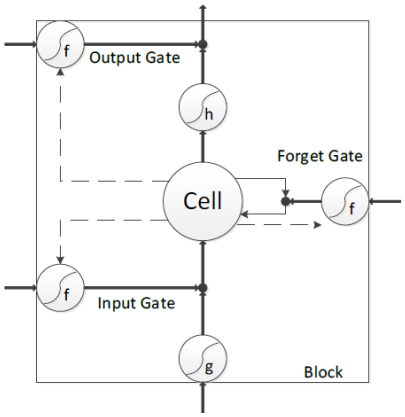

Figure 3 LSTM cell structure

However, RNNs also have shortcomings: for the standard RNN neural network architecture, the limitation of "memorability" has a great impact on the output of the neural network, which is known as the "gradient extinction problem". The reduction of iteration time does not guarantee the improvement of the accuracy of the calculation results. In this regard, LSTM neural network solves this problem. Similar to RNN neural network, LSTM neural network consists of an input layer, hidden layer, and output layer, but the special feature of LSTM is its unique memory unit structure, which is shown in Fig. 3.

Each LSTM neural network cell has a tuple, or memory cell, whose state at time t is c_t. The memory cell of the LSTM neural network can be read and modified by operating on the input gate, forget gate, and output gate, which are generally implemented by the sigmoid or tanh function. The workflow of the LSTM unit is as follows: at each moment, the LSTM unit receives two types of external information input through three gates: the current state x_t and the hidden state h_{t-1} of the LSTM at the previous moment. In addition, each gate receives an internal information input, i.e. the state c_{t-1} of the memory cell. After receiving the input information, each gate performs operations on the inputs from different sources and its logic function decides whether it is activated or not. The input of the input gate is transformed by the nonlinear function and superimposed with the state of the memory cell processed by the forgetting gate to form the new memory cell state c_t. Finally, the memory cell state c_t forms the output h_t of the LSTM cell by the operation of the nonlinear function and the dynamic control of the output gate. The equation between the variables is as follows:

$$i_t = \sigma\left(W_{xi}x_t - W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i\right) \tag{4}$$

$$f_t = \sigma\left(W_{xf}x_t - W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f\right) \tag{5}$$

$$\begin{aligned} c_t &= f_t c_{t-1} + \tan h?\left(W_{xc}x_t + W_{hc}h_{t-1} + b_c\right)c_t = \\ &= f_t c_{t-1} + tanh\tan h\left(W_{xc}x_t + W_{hc}h_{t-1} + b_c\right) \end{aligned} \tag{6}$$

$$o_t = \sigma\left(W_{xo}x_t - W_{ho}h_{t-1} + W_{co}c_{t-1} + b_o\right) \tag{7}$$

$$h_t = o_t\tan h?\left(c_t\right)h_t = o_t\tan h\tan h\left(c_t\right) \tag{8}$$

where, W_{xc}, W_{xi}, W_{xo}, W_{xf} W_{xf} are the weight matrices connecting the input signal x_t, W_{hi}, W_{hf} W_{hf}, $W_{hc}W_{hc}$, $W_{ho}W_{ho}$ are the weight matrices connecting the hidden layer output signal h_t, W_{ci} W_{ci}, $W_{cf}W_{cf}$, $W_{co}W_{co}$ are the diagonal matrices connecting the neuron activation function output vector c_tc_t and the gate function; b_ib_i, b_fb_f, b_cb_c, b_ob_o are the bias vectors; σ is the activation function, usually tanh or sigmoid function.

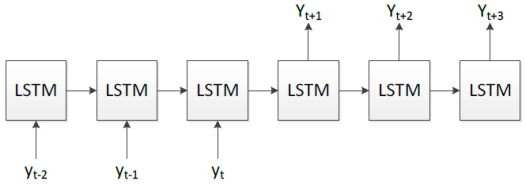

Figure 4 Schematic diagram of BPTT training LSTM

Since the BPTT algorithm is conceptually clear and computationally efficient and has an advantage over RTRL in terms of computation time, it is used to train LSTM networks in this paper. As shown in Fig. 4, the basic idea of the BPTT algorithm is to first unfold the LSTM network into a deep network in temporal order, and then train the unfolded network using the classical error back propagation (BP) algorithm. Like the standard BP algorithm, BPTT also requires iterative application of chain rules. It should be noted that for LSTM neural networks, the loss function (LF) is not only related to the output layer, but also to the hidden layers before and after time points.

The error backpropagation algorithm is basically present in fully connected neural networks. However, RNN

contains not only backpropagation on the spatial axis of layers but also on the temporal axis because it has the transmission of temporal information. Therefore, in RNN, not only the BP but also the BPTT needs to be counted.

Let the total time be T and the loss at moment t be E_t, then for a single sample at all moments the loss E is given as follow:

$$E = \sum_{t=1}^{T} E_t \tag{9}$$

Assuming that the error term is the partial derivative of the loss function with respect to the output value, the output value of RNN at the moment t is h_t, define the partial derivative of the loss E_t with respect to h at the moment t as the error term δ_t:

$$\delta_t = \frac{\partial E_t}{\partial h_t} \tag{10}$$

Decompose the defining formula of RNN into two steps of Eq. (11) and Eq. (12):

$$h_t = g\left(net_t\right) \tag{11}$$

$$net_t = W_{x_t} + Uh_{t-1} + b \tag{12}$$

where, the activation function g is set to choose the tan$h$ function and *net* is the unactivated output value, also known as the logical layer value.

Since the output of the RNN at the previous moment is the input at the current moment, then:

$$\frac{\partial net_t}{\partial net_{t-1}} = \frac{\partial net_t}{\partial h_{t-1}} \frac{\partial h_t}{\partial net_{t-1}} \tag{13}$$

The derivative of the activation function tanh is Eq. (14):

$$\frac{\mathrm{d}\tan h\left(x\right)}{\mathrm{d}x} = 1 - \left[\tan h\left(x\right)\right]^2 \tag{14}$$

Then the partial derivative of the activation function at moment i is:

$$\frac{\partial h_{i-1}}{\partial net_i} = \frac{\mathrm{d}\tan h\left(x\right)}{\mathrm{d}x} = 1 - \left[\tan h\left(x\right)\right]^2 \tag{15}$$

Combining the assumptions and derivations above, the error term δ_k^T for the target moment t is obtained from the chain derivation rule and can be expressed as:

$$\delta_k^T = \frac{\partial E_t}{\partial net_k} = \frac{\partial E_t}{\partial net_t}\frac{\partial net_t}{\partial net_k} = \frac{\partial E_t}{\partial net_k} \cdots \frac{\partial net_{k+1}}{\partial h_k} = \delta_t \prod_{i=k}^{t-1} \frac{\partial h_{t-1}}{\partial net_i} \tag{16}$$

Once the error term δ_k^T is obtained, the gradient at any moment can be calculated:

$$\frac{\partial E_t}{\partial U_{ji}} = \delta_t h_{t-1} \tag{17}$$

After computing the gradient it can be substituted into the gradient descent algorithm to update the gradient, which is similar to the BP algorithm. After repeated iterations to update the gradient values, the final weight parameter model of the neural network can be obtained.

3.2 PAM Clustering Algorithm

In order to simplify the structure of neural networks, a clustering algorithm is needed to categorize the data of environmental covariates in order to standardize the training dataset. The partitioning around medoid (PAM) is a K-medoids algorithm. The basic principle of the PAM algorithm is that k objects in the dataset are randomly selected as initial centroids, and the centroids are continuously replaced in each iteration to find better centroids to improve the clustering [21]. Compared with the K-means algorithm, the PAM algorithm is more robust than the K-means algorithm, insensitive to noise and isolated point data, able to handle different types of data, and more effective in dealing with small data sets. These features are suitable for similar daily clustering of error predictions. The original PAM algorithm requires traversing $n - k$ noncenters for each centroid selection, for a total of k centroids, and each replacement process requires reassigning $n - k$ noncenters. It follows that one iteration yields a time complexity of $O(k^2(n - k)^2)$. When the number of iterations is t, then the total time complexity is $O(tk^2(n - k)^2)$. From the time complexity, it is clear that the PAM algorithm can be very slow on large data sets. Therefore, in order to improve the performance of the PAM algorithm and expand its adaptability to the size of the dataset, several approaches are proposed.

The most tedious stage of the PAM algorithm is that $n - k$ non-centers are traversed in each iteration round, and every time a substitution between centroids and non-centroids occurs, all non-centroids need to be re-traversed to find the nearest centroid. However, in reality, many non-center points nearest centroids do not change and do not need to find the nearest centroid again, and this part of redundancy can be reduced by increasing the spatial complexity to reduce it. Therefore, a method for storing the nearest distance is proposed, which requires storing the distance of each non-centroid from the nearest and second closest centroids noted as D_j and E_j, respectively. In each iteration round, the centroid O_i and the non-centroid O_h are replaced, and the distance between the data point j and the original centroid i, denoted as $dist(j, i)$, needs to be calculated first. The distance between the data point j and the non-center point h which is the new center point after the replacement is then calculated and is denoted as $dist(j, h)$. C_{ijh} is used to denote the cost difference resulting from replacing the centroid O_i. Depending on the values of the cost difference and distance, four cases can be discussed:

(1) $dist(j, i) = D_j$, indicates that data point i is the nearest centroid to data point j:

1) If after replacement, $dist(j, h) < E_j$, i.e., the distance from data point j to data point h is less than the data point j to the distance between the second closest centroid, then for $C_{ijh} = E_j - D_j$, the C_{ijh} uncertainty positive and negative;

2) If after the replacement, $dist(j, h) \geq E_j$, that is, the distance from data point j to data point h is greater than the distance between data point j and the second closest centroid, then for $C_{ijh} = E_j - D_j$, C_{ijh} is positive.

In summary, the expression for C_{ijh} can be written: $C_{ijh} = \{dist(j, h), E_j\} - D_j$.

(2) $dist(j, i) > D_j$, indicates that data point i is not the nearest centroid to data point j:

1) If after the replacement, $dist(j, h) \geq D_j$, that is, the distance from data point j to data point h is greater than the distance between data point j and the nearest centroid, then for $C_{ijh} = 0$;

2) If after the replacement, $dist(j, h) < D_j$, that is, the distance from data point j to data point h is less than the distance between data point j and the nearest centroid, then for $C_{ijh} = dist(j, h) - D_j$, C_{ijh} is negative.

In summary, the expression for C_{jih} can be written: $C_{jih} = \{dist(j, h), D_j\} - D_j$.

The expression of C_{jih} can be rationalized to simplify the original PAM algorithm that repeatedly calculates the distance between the centroid and the non-centroid many times to simplify the process of deciding whether to replace or not. The main process of the PAM clustering algorithm is as follows:

(1) Randomly select k objects $c_1, c_2, ..., c_k$, among n objects of the data set as initial centroids.

(2) The remaining objects in the dataset are assigned to the nearest centroids according to their distances, here the distances are still using the Manhattan distance formula, and the total cost function J_c is calculated and the nearest distance D_j, the second nearest distance E_j and the cost function J_c are stored for each non-centroid.

(3) Select a center point O_i that has not been selected.

(4) Select a non-center point O_h that has not been selected.

(5) Calculate the cost function generated by replacing B with A and record it in TC_{ih}, $TC_{ih} = J_c + \sum_{j=0}^{n-k} C_{jih}$.

For the case of a centroid O_i under replacement, if $dist(j, h) > E_j$, then $TC_{ih} = TC_{ih} + E_j$, i.e., it is grouped into the cluster where the second closest centroid is located. If $dist(j, h) \leq E_j$, then $TC_{ih} = TC_{ih} + dist(i, h)$, that is, to the cluster in which the replaced centroid O_h is located. The new expression for TC_{ih} is $TC_{ih} = TC_{ih} + \{ E_j, dist(j, h)\}$. Choose the smallest TC_{ih}. If $\min\{TC_{ih}\}$ is smaller than the current cost function J_c, perform the corresponding substitution and redistribute the centroids.

(6) Update the closest distance D_j, the second closest distance E_j, and the cost function J_c for each point;

(7) Repeat the substitution until $\min\{TC_{ih}\}$ is greater than the cost function J_c. Stop the iteration and output the classification data.

4 IMPLEMENTATION PROGRAM

The implementation steps of the electronic current transformer error prediction method proposed in the paper are shown in Fig. 5.

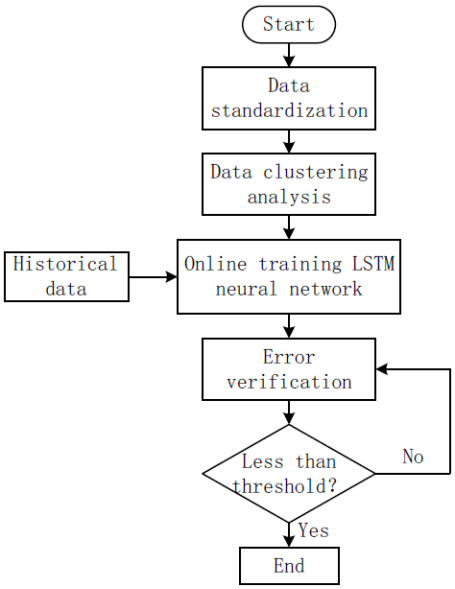

Figure 5 Implementation step for the forecasting method

(1) The variables of the electronic current transformer operation error prediction model are selected. The input variables are temperature, magnetic field, and load parameters, and the output variables are the ratio and phase errors in the operating errors of electronic current transformers.

(2) Variable standardization. Since the units and magnitudes of the input and output variables are different, it is necessary to standardize the data of each variable. The standardization process using Z-score is as follows:

$$x' = (x - \mu)/\sigma \tag{18}$$

where μ is the mean of the data; σ is the standard deviation of the data.

(3) Clustering of input variable data sets. In order to simplify the network size of the prediction model, the input variable data set needs to be clustered to divide the environmental covariate data into different classifications.

(4) LSTM neural network training. Before training the model, the parameters of the prediction model need to be set, which are: the number of input layer time steps, the number of input layer dimensions, the number of hidden layers, the number of dimensions of each hidden layer, and the number of output variable dimensions. Then the LSTM neural network model is trained using the BPTT algorithm. The model training also requires error verification, and when the error is greater than the set value the prediction model needs to be trained iteratively.

(5) Electronic current transformer operation error prediction. According to the neural network model obtained in step (4), the temperature parameter, magnetic field parameter data, and load parameter data are used as input variables for the evaluation of electronic current transformer errors.

In order to quantitatively characterize the error between the prediction results and the actual electronic current transformer operating error values, the root mean square error (RMSE) and mean relative error (MAPE) are used in this paper to evaluate the accuracy of the prediction model:

$$E_{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[\underline{P}(i) - P(i)\right]^2} \tag{19}$$

$$E_{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{\left[\underline{P}(i) - P(i)\right]}{P(i)} \tag{20}$$

where, $\underline{P}(i)$ and $P(i)$ are the actual and predicted values of the electronic current transformer error, respectively, n is the number of predicted validation data, and i is the prediction point sequence number.

5 SIMULATION STUDIES

In order to verify the effectiveness of the electronic current transformer error prediction method, the online monitoring data of an electronic current transformer in a digital substation in Gansu 330 kV were used as training samples to establish an LSTM neural network model for electronic current transformer error prediction.

The electronic current transformer operation error data and environmental monitoring data are collected in the data processing unit and sent to the server uniformly after passing through the switch, and the data model is constructed after data collection, screening, and storage processing, and the error prediction algorithm is retrieved from the algorithm library to perform data mining tasks to realize the analysis of the massive electronic current transformer online monitoring data and thus obtain its error change trend. LSTM The training samples of the neural network come from the collection results of the monitoring platform from February 2022 to June 2022, including temperature parametric data, magnetic field parametric data, and load parametric data. In order to simplify the network size of the prediction model, the input variable data set needs to be clustered to divide the environmental covariate data into different classifications. The PAM algorithm is used to normalize the sample data of temperature, load, and magnetic field and then perform the clustering calculation. According to the clustering results in Tab. 2, four clusters from A to D were calculated for each environmental parameter, and each cluster center was obtained according to the corresponding cluster division.

Table 2 Calculation results of clustering center for environmental parameter

Data type	Clustering center			
	A	B	C	D
Temperature	1.4763	0.7949	−0.0138	−1.4625
Load	−2.0087	−1.7613	−0.2365	0.8159
Magnetic field	−0.4181	−1.8230	2.1633	1.1372

The establishment of an LSTM neural network-based prediction model for electronic current transformer operation error requires the determination of five parameters in the prediction model: the number of input layer time steps, the number of input layer dimensions, the number of hidden layers, the number of dimensions of each hidden layer, and the number of output variable dimensions.

The amount of real measurement data used to predict the operating error of electronic current transformers is the number of input layer time steps, and the setting of this parameter should consider both the comprehensiveness of the prediction model and the effectiveness of the model training. On the one hand, too little measured data will bring about a lack of training knowledge, thus affecting the accuracy of the prediction data; on the other hand, too many measured data will reduce the training efficiency of the model.

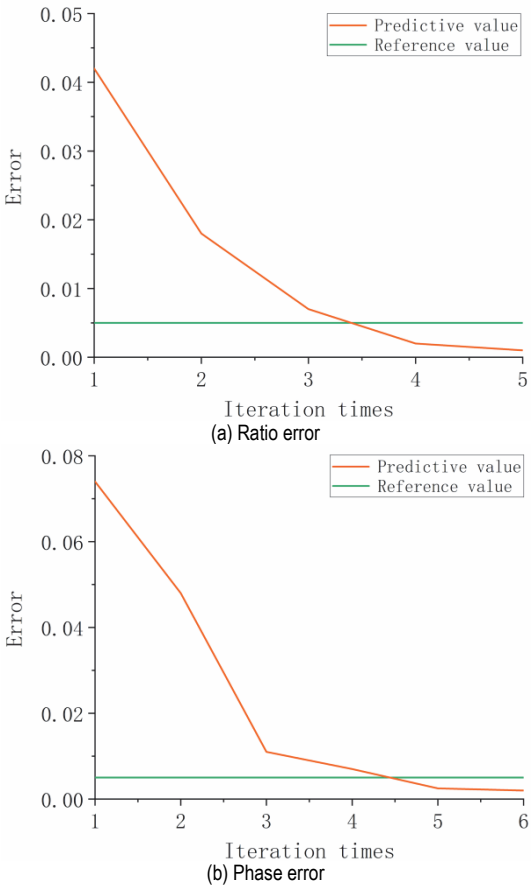

Figure 6 Iteration results of LSTM neural network

The number of input layers is the number of predictor variables, and the number of layers of LSTM neural network is the number of hidden layers. However, the prediction efficiency of the model decreases significantly. In this paper, the value of input layer time steps is finally determined to be 10 by example analysis. The input layer dimension is the number of predictor variables, and the value of this parameter is 2 for bivariate. The number of LSTM neural network layers is the number of hidden layers. With sufficient training samples, the larger the number of hidden layers, the better the nonlinear fitting ability of the model. But the prediction efficiency of the model will be significantly reduced. In this paper, the number of hidden layers is set to 2. Since the prediction of electronic current transformer operation error in this paper is based on actual measurement data, the number of output variable dimensions is set to 2. The iterative results of the LSTM neural network are shown in Fig. 6, which shows that the mean square error is less than 0.3% after four iterations of the electronic current transformer ratio error data, and the mean square error is less than 0.3% after four iterations of the electronic current transformer phase error data. The mean square error is less than 0.3% after 4 iterations of the electronic current transformer ratio error

data and less than 0.3% after 5 iterations of the electronic current transformer phase error data.

The prediction results of the corresponding electronic current transformer errors are obtained based on the historical data training set, as shown in Fig. 7 and Fig. 8.

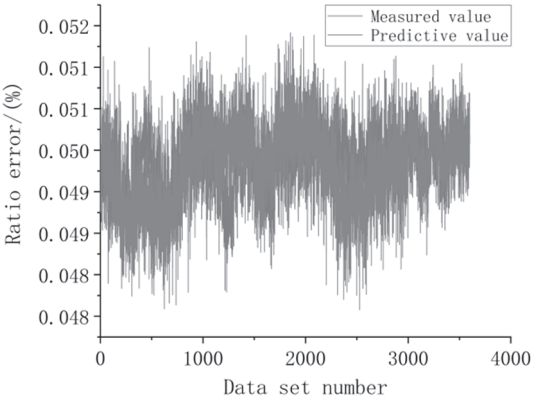

Figure 7 Forecasting results of ratio error

According to the prediction results, the prediction result of the ratio error of the electronic current transformer based on the LSTM neural network model is 2.28% for E_{RMSE} and 4.73% for E_{MAPE}; the prediction result of the angular difference of electronic current transformer is 3.61% for E_{RMSE} and 6.12% for E_{MAPE}. According to the comparison between the predicted and actual values of electronic current transformer error based on LSTM neural network in Fig. 7 and Fig. 8, the absolute value of prediction error of ratio error is less than 0.04%, and the absolute value of prediction error of phase error is less than 8′.

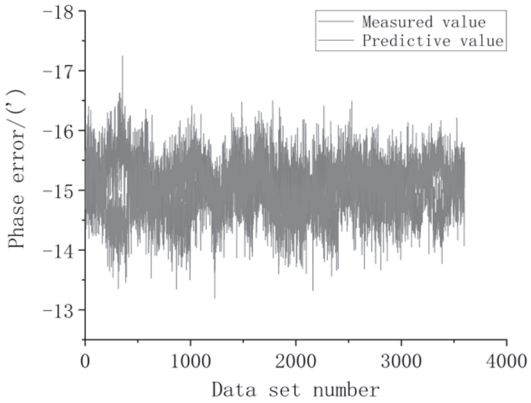

Figure 8 Forecasting results of Phase error

To further verify the effectiveness of the LSTM neural network model, the prediction values generated by two conventional artificial intelligence algorithms, ANN and SVM, were compared. Tab. 3 and Tab. 4 show the prediction error metrics obtained from the three prediction models for predicting the operating error of electronic current transformers.

As can be seen from Tab. 3 and Tab. 4, the two indicators of the prediction error of the LSTM neural network model prediction method are smaller than those obtained by both ANN and SVM, two traditional artificial intelligence methods, precisely because the LSTM neural network circumvents the gradient extinction problem faced by RNNs, thus reflecting well the error variation trend of electronic current transformers.

Table 3 Comparison of the prediction error of the ratio error of 3 prediction methods

Data set range	LSTM Model error / %		ANN Model error / %		SVM Model error / %	
	MAPE	*RMSE*	*MAPE*	*RMSE*	*MAPE*	*RMSE*
1 ~ 600	4.69	2.46	5.44	3.34	5.80	3.99
601 ~ 1200	3.40	1.70	5.14	2.90	5.59	3.71
1201 ~ 1800	4.91	2.28	5.90	3.76	6.90	4.40
1801 ~ 2400	5.09	2.87	6.01	3.47	7.44	4.51
2401 ~ 3000	5.48	3.57	5.58	3.89	8.47	5.44
3001 ~ 3600	5.68	3.79	8.06	4.97	8.55	5.26

Table 4 Comparison of the prediction error of the phase error of 3 prediction methods

Data set range	LSTM Model error / %		ANN Model error / %		SVM Model error / %	
	MAPE	*RMSE*	*MAPE*	*RMSE*	*MAPE*	*RMSE*
1 ~ 600	5.92	3.37	7.22	5.23	7.38	5.88
601 ~ 1200	4.58	2.33	6.9	4.82	7.23	5.62
1201 ~ 1800	5.47	3.21	7.52	5.63	8.4	6.27
1801 ~ 2400	6.34	3.06	7.84	5.32	8.93	6.4
2401 ~ 3000	5.75	3.82	7.36	5.76	9.97	7.41
3001 ~ 3600	5.82	4.06	9.74	6.9	10.11	7.16

6 CONCLUSION

In the paper, an electronic current transformer error prediction model is constructed based on LSTM neural network algorithm and PAM algorithm with environmental parameters as input variables and electronic current transformer error parameters as output variables, and the results of the study show that:

(1) The prediction method can provide the prediction information of the electronic current transformer error state, timely find and deal with the electronic current transformer error state problem, and ensure the normal operation of power trade settlement.

(2) Compared with the prediction models based on two traditional artificial intelligence algorithms, ANN and SVM, the results obtained by the prediction method based on LSTM neural network are more accurate.

(3) Compared with the traditional K-means clustering method, the PAM clustering algorithm has stronger robustness and can better reduce the complexity of the electronic current transformer error prediction model.

(4) The error prediction method proposed in the paper can provide information on the changes in the error states of

electronic current transformers, and provide an effective decision basis for timely detection and treatment of the error state of electronic current transformers to prevent the risk of electricity trade settlement brought by them.

The limitations of the method proposed in the paper are that the selection of the initial clustering centers of the clustering algorithm is random and may fall into the trap of local optimum, while the generalizability of the LSTM prediction model needs to be further verified, and subsequent improvements to the algorithm are needed.

Acknowledgements

This work was supported by the Science and Technology Project of State Grid Gansu Electric Power Company(No. 52273120000A).

7 REFERENCES

[1] Zhao, S., Yang, S., Lu, S., Chen, M., & Xu, M. (2016). Method for online measurement of optical current transformer onsite errors. *Measurement Science and Technology*, *27*(2), 025014. https://doi.org/10.1088/0957-0233/27/2/025014

[2] Li, Z., Yu, C., Abu-Siada, A., Li, H., Li, Z., Zhang, T., & Xu, Y. (2021). An online correction system for electronic voltage transformers. *International Journal of Electrical Power & Energy Systems*, *126*, 106611. https://doi.org/10.1016/j.ijepes.2020.106611

[3] Ghayebloo, A., Ghaleghovand, M., & Jalilvand, A. (2021). A novel identification approach for classic controller design applied on flyback converter. *Journal Européen des Systèmes Automatisés*, *54*, 105-113. https://doi.org/10.18280/jesa.540112

[4] Murali, D. (2021). Steady state behavior of a single-switch non-isolated DC-DC SEPIC converter topology with improved static voltage gain. *Journal Européen des Systèmes Automatisés*, *54*, 445-452. https://doi.org/10.18280/jesa.540307

[5] Eddine, A. T., Ameur, A., & Atallah, B. (2021). RNA identification technique and RST control of a hybrid indirect matrix converter with a flying capacitor three level inverter in power active filtering application. *Journal Européen des Systèmes Automatisés*, *54*, 743-749. https://doi.org/10.18280/jesa.540509

[6] Li, Z., Zheng, Y., Abu-Siada, A., Lu, M., Li, H., & Xu, Y. (2020). Online evaluation for the accuracy of electronic voltage transformer based on recursive principal components analysis. *Energies*, *13*(21), 5576. https://doi.org/10.3390/en13215576

[7] Liu, G., Zhao, P., Zhao, M., Yang, Z., & Chen, H. (2020). Electromagnetic disturbed mechanism of electronic current transformer acquisition card under high frequency electromagnetic interference. *Electronics*, *9*(8), 1293. https://doi.org/10.3390/electronics9081293

[8] Li, B., Zhang, Z., & Ding, L. (2021). Self-detecting the measurement error of electronic voltage transformer based on principal component analysis-wavelet packet decomposition. *IEEE Transactions on Instrumentation and Measurement*, *70*, 3526609. https://doi.org/10.1109/TIM.2021.3124065

[9] Liu, G., Zhao, P., Qin, Y., Zhao, M., Yang, Z., & Chen, H. (2020). Electromagnetic immunity performance of intelligent electronic equipment in smart substation's electromagnetic environment. *Energies*, *13*(5), 1130. https://doi.org/10.3390/en13051130

[10] Hu, C., Chen, M. Z., Li, H. B., Zhang, Z., Jiao, Y., & Shao, H. (2018). An accurate on-site calibration system for electronic voltage transformers using a standard capacitor. *Measurement Science and Technology*, *29*(5), 055901. https://doi.org/10.1088/1361-6501/aaa6a0

[11] Xu, Y., Li, Y., Xiao, X., Xu, Z., & Hu, H. (2017). Monitoring and analysis of electronic current transformer's field operating errors. *Measurement*, *112*, 117-124. https://doi.org/10.1016/j.measurement.2017.08.015

[12] Li, Z. H., Li, H. B., & Zhang, Z. (2013). An accurate online calibration system based on combined clamp-shape coil for high voltage electronic current transformers. *Review of Scientific Instruments*, *84*(7), 075113. https://doi.org/10.1063/1.4815831

[13] Zhang, Z., Chen, Q., Hu, C., Li, H., & Chen, M. (2018). Evaluating the metering error of electronic transformers on-line based on VN-MWPCA. *Measurement*, *130*, 1-7. https://doi.org/10.1016/j.measurement.2018.07.083

[14] Ye, Y., Yang, A., Wu, Y., Hu, C., Li, M., Li, Y., & Deng, X. (2020). Short-term prediction of electronic transformer error based on intelligent algorithms. *Journal of Control Science and Engineering*, *2020*, 9867985. https://doi.org/10.1155/2020/9867985

[15] Xiong, G., Przystupa, K., Teng, Y., Xue, W., Huan, W., Feng, Z., & Beshley, M. (2021). Online measurement error detection for the electronic transformer in a smart grid. *Energies*, *14*(12), 3551. https://doi.org/10.3390/en14123551

[16] Han, F. J., Ashton, P. M., Li, M. Z., Pisica, L., Taylor, G., Rawn, B., & Ding, Y. (2021). A data driven approach to robust event detection in smart grids based on random matrix theory and kalman filtering. *Energies*, *14(8)*, 2166. https://doi.org/10.3390/en14082166

[17] Hu, C., Zhang, Z., Jiao, Y., Li, H. B., & Chen, G. (2018). Error state correlation analysis based on random matrix theory for electronic transformer. *Electric Power Automation Equipment*, *38(9)*, 45-53.

[18] Azzam, R., Alkendi, Y., Taha, T., Huang, S., & Zweiri, Y. (2020). A stacked LSTM-based approach for reducing semantic pose estimation error. *IEEE Transactions on Instrumentation and Measurement*, *70*, 2502614. https://doi.org/10.1109/TIM.2020.3031156

[19] Ciechulski, T. & Osowski, S. (2021). High precision LSTM model for short-time load forecasting in Power Systems. *Energies*, *14*(11), 2983. https://doi.org/10.3390/en14112983

[20] Jin, Y., Zhang, X., & Jia, Y. (2021). Application of optimized partitioning around medoid algorithm in image retrieval. *International Journal of Distributed Systems and Technologies*, *12*(1), 77-94. https://doi.org/10.4018/IJDST.2021010106

[21] Thrun, M. C. & Ultsch, A. (2021). Swarm intelligence for self-organized clustering. *Artificial Intelligence, 290*, 103237. https://doi.org/10.1016/j.artint.2020.103237

[22] Hochreiter, S., Bengio, Y., Frasconi, P., & Schmidhuber, J. (2001). Gradient flow in recurrent nets: the difficulty of learning long-term dependencies. *A field Guide to Dynamical Recurrent Networks*, *Wiley-IEEE Press*, 237-243.

[23] Kawakami, K. (2008). *Supervised sequence labelling with recurrent neural networks.* Technical University of Munich.

[24] Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks*, *5*(2), 157-166. https://doi.org/10.1109/72.279181

[25] Chang, F. J., Chang, L. C., & Huang, H. L. (2002). Real-time recurrent learning neural network for stream-flow forecasting. *Hydrological processes*, *16*(13), 2577-2588. https://doi.org/10.1002/hyp.1015

[26] Zhu, X., Zhang, J., & Feng, J. (2015). Multiobjective particle swarm optimization based on PAM and uniform design. *Mathematical Problems in Engineering*, *2015*, 126404. https://doi.org/10.1155/2015/126404

Contact information:

Kun WANG, intermediate engineer
(Corresponding author)
1. State Grid Gansu Electric Power Company Metrology Center,
2. State Grid Gansu Electric Power Company Marketing Service Centerl,
Lanzhou 730050, China
E-mail: jlzx_wk@163.com

Hu LI, advanced engineer
State Grid Gansu Electric Power Company Marketing Service Center,
Lanzhou 730050, China
E-mail: gs_msc@126.com

Huan LI, intermediate engineer
1. State Grid Gansu Electric Power Company Metrology Center,
2. State Grid Gansu Electric Power Company Marketing Service Center,
Lanzhou 730050, China
E-mail: gs_jlzx@126.com

Jinggeng GAO, advanced engineer
State Grid Gansu Electric Power Company Marketing Service Center,
Lanzhou 730050, China
E-mail: gao2541@163.com