Short-Term Prediction of Electronic Transformer Error Based on Intelligent Algorithms

Yuanyu Ye,1 Aichao Yang,1 Yu Wu,1 Chen Hu,1 Min Li,1 Yan Li,2 and Xiaosong Deng2

1State Grid Jiangxi Electric Power Company, Nanchang 330077, China
2The State Key Laboratory of Power Transmission Equipment & System Security and New Technology, Chongqing University, Chongqing 400044, China

Correspondence should be addressed to Yan Li; liyanly@cqu.edu.cn

Received 19 November 2019; Revised 18 May 2020; Accepted 18 July 2020; Published 28 August 2020

1. Introduction

As a key equipment for signal acquisition in the power system, the transformer provides a reliable basis for electric energy measurement, condition monitoring, and relay protection. The electronic transformer (ET), as the important equipment for digitizing electric parameter information in the smart grid, has been highly valued. However, compared with the traditional electromagnetic transformer, ET has a more complicated structure and lower stability, which is easy to exceed error tolerance in the actual operation.

During the actual operation process of ET, it is not only necessary to accurately and quickly diagnose the abnormality of the measurement error, but also to make timely predictions on the deterioration trend of measurement error, so that the related personnel can carry out the inspection and maintenance and thus reduces the loss of electric energy measurement and ensures the normal operation of the monitoring protection device, which is of great significance for ensuring the safe, stable, and economical operation of the power system.

So far, many researchers have studied the fault analysis of power transformers and have proposed some typical fault diagnosis methods. Gutten et al. [1] adopt sweep frequency response analysis through time impact test to analyse power transformer winding faults. Rigatos and Siano [2] propose a neural fuzzy network considering hotspot temperature for the detection of incipient faults in power transformers. Yang et al. [3] propose a fault diagnosis method for power transformer based on BA-PNN. Based on the dissolved gas analysis (DGA), the power transformer fault type can be obtained by inputting three key gas ratios through BA-PNN. Mejia et al. [4] applied empirical mode decomposition (EMD) into inrush current signals of power transformer, and thus, three indexes can be obtained providing the proper quantification of the fault severity. However, compared with the power transformer, ET has higher precision and smaller size, and the structures
of ET and power transformer are quite different. Therefore, the fault analysis methods mentioned above are not suitable for ET. Meanwhile, the error trend prediction method driven by data is an important way to detect equipment failure in advance, which can be utilized in the prediction of ET’s metering error.

Intelligent algorithms are often used for data prediction, among which the more common methods include the neural network and the Prophet model. The neural network has been applied in many fields, such as the prediction of power system load [5], oil breakdown voltage [6], and railway freight volume [7]. And the Prophet model has obtained good effect as well in the prediction of bank reserve payment [8], product sales volume [9], daily air quality index [10], and other aspects. However, the research on the prediction of the ET measurement state is almost blank. Because the prediction of secondary output information of ET will be affected by the random fluctuation of the grid state, there is no effective way to predict the trend of the measurement error of ET yet. Error comparison with traditional transformers is still widely adopted in the state analysis of ET for now, which is difficult to meet the requirements of error state analysis under the scenario of large-scale application of ET.

Therefore, under the situation of no standard transformer, making full use of the measurement data of ET in the smart grid, this paper analyses and simulates the development trend of ET error through the backpropagation neural network (BPNN) and the Prophet model, respectively, so as to ensure the safe and stable operation of ET.

2. Data Analysis

2.1. Data Source. The ET error data used in this paper is obtained through a 220kV ET metering error condition monitoring platform shown in Figures 1 and 2.

In the monitoring platform, the electromagnetic current transformer and the air-core coil current transformer are installed at the line interval. The accuracy of the two transformers are both 0.2, and their rated current is 600 A. The electromagnetic transformer’s rated output is 5 A, the rated secondary capacity is 25 V·A, and ratio error and phase error under the rated current are 0.08% and 6′, respectively; the air-core coil current transformer’s ratio error and phase error under the rated current are 0.12% and 6′, respectively, and its output complies with the IEC61850-9-2 protocol [11].

The signal acquisition unit converts the analog signal of the electromagnetic current transformer into a digital signal. The data processing unit, which meets the accuracy requirement of 0.05 level, receives the output signal and sampled value message data of the signal acquisition unit and then obtains the error comparison result by taking the output of the electromagnetic current transformer as the standard.

In this paper, ET metering error prediction methods based on BPNN and the Prophet model are proposed. In the case study, the phase error data are taken as an example to illustrate the effectiveness of the proposed prediction methods. The original transformer phase error data are the actual running data of the ET during the period from July 1 to 31, 2018, and the sampling interval is 1 s, as shown in Figure 3.

2.2. Data Preprocessing. Since the error data of ET are generally irregular and contain noise, the following preprocessing operations need to be performed on the phase error dataset before using the intelligent algorithms to predict the phase error of ET [12]:

(1) Set the upper and lower boundary of phase error data and delete the abnormal value.

(2) Determine the ratio of missing values in the dataset. When the ratio is large, Lagrange interpolation or Newton interpolation will be adopted to supplement data; when the ratio is small, extra processing is not necessary.

(3) Due to the high acquisition frequency of the signal acquisition unit, the data samples of one month contain more than 2.6 million sets of data, and thus, the solving of the prediction model becomes very difficult. Therefore, the average values of the hourly phase error can be obtained by (1), which are used as the training samples to reduce the prediction complexity and improve the prediction efficiency of the model:

\[
p_{i}^{\text{th}} = \frac{1}{3600} \sum_{j=1}^{3600} P_{i}^{j},
\]

where \(p_{i}^{\text{th}}\) represents the average values of hourly phase error and \(P_{i}^{j}\) represents the per-second sample value of phase error.

The ET phase error data after preprocessing is shown in Figure 4.

2.3. Data Visualization. Although Figure 4 can show the hourly variation curve of the ET phase error, the symmetry, the dispersion level, and the general development trend of the phase error data cannot be visually seen through Figure 4. This paper introduces the boxplot to visualize the phase error data. The boxplot [13] is a common method for data visualization, which mainly uses five kinds of statistics to describe the data: the minimum, the first quartile, the median, the third quartile, and the maximum. The preprocessed dataset is divided into 31 groups of datasets at intervals of 24 h, corresponding to the time period from July 1 to 31. The boxplot is drawn as shown in Figure 4.

It can be seen from Figure 5 that in July, the overall distribution of the ET phase error is relatively uniform and the dispersion level is relatively stable. Only a few days (3rd, 5th, 6th, 7th, 8th, and 22nd) have relatively more abnormal values, and thus, the distribution can be seen as concentrated to some degree. On 13th to 21st, the median line is in the middle, and the phase error presents normal distribution; at the beginning and the end of the month, the median line is shifted upwards, and the phase error presents skew
**Figure 1:** Block diagram of ET metering error condition monitoring platform.

**Figure 2:** ET metering error condition monitoring platform (a) and ET in the transformer substation (b).

**Figure 3:** Original ET phase error data.

**Figure 4:** ET phase error data after preprocessing.
distribution. It can be seen from the change of the median line that the phase error of ET presents a fluctuating and decreasing trend on the whole in this month.

3. Prediction Model Based on Intelligent Algorithms

The BPNN algorithm widely used in various fields and the Prophet model which has a huge advantage in processing periodic data are chosen in this paper to make short-term prediction of the ET metering error data.

3.1. Backpropagation Neural Network. The BPNN [14] is a concept proposed by scientists led by Rumelhart and McClelland in 1986. BPNN is a multilayer feedforward neural network trained by the backpropagation algorithm, which is one of the most widely used neural networks. BPNN has a simple structure and fast convergence.

There are many factors influencing the error developing trend of ET, and thus, it is difficult to establish an analytical model. However, BPNN can realize a mapping function from input to output without knowing the analytical expression, which is especially suitable for the simulation of the system with complex internal mechanism. According to the Kolmogorov theorem, as long as the number of hidden layer nodes of a 3-layer BPNN is sufficient, any nonlinear continuous function can be approximated on the closed set with arbitrary precision [15]. Therefore, BPNN is applicable for the prediction of ET measurement error.

A typical BPNN topology [16] is shown in Figure 6, which contains input layer, hidden layer, and output layer. Each layer consists of a certain number of neurons. All the neurons in the upper and lower layers are interconnected, and the connection strength is represented by the weight. But neurons in the same layer are not interconnected.

In the figure, \( x_0, x_1, \ldots, x_{l-1} \) stand for the inputs. It is assumed that there are \( m \) neurons in the hidden layer, the inputs of which are \( x_{0}', x_{1}', \ldots, x_{m-1}' \). There are \( n \) neurons in the output layer, the outputs of which are \( y_0, y_1, \ldots, y_{n-1} \). The weight and threshold of input layer to hidden layer are \( w_{ij} \) and \( \theta_j \), respectively; the weight and threshold of hidden layer to output layer are \( w_{jk}' \) and \( \theta_k' \), respectively. The outputs of neurons in each layer are as follows:

\[
\begin{align*}
x_j &= f \left( \sum_{i=1}^{l-1} w_{ij} x_i - \theta_j \right), & j &= 0, 1, \ldots, m - 1, \\
y_k &= f \left( \sum_{i=1}^{m-1} w_{jk}' x_j' - \theta_k' \right), & k &= 0, 1, \ldots, n - 1,
\end{align*}
\]

where \( f (\bullet) \) is the activation function, and the S-type activation function is shown in the following equation:

\[
f (x) = \frac{1}{1 + e^{-x}}
\]
corrects each weight according to the reducing direction of error of the output value and actual value, so as to approximate the output response value to the actual value with set accuracy level through constant iterations.

3.2. The Prophet Model. The Prophet model is a model developed by Facebook for large-scale time series analysis, which has an open access software package to facilitate the implementation of model algorithms [17, 18].

Through the ET phase error data and its analysis, it can be seen that the fluctuation of ET phase error in July is obviously periodic. The Prophet model is designed to handle complex features in time series, which has very powerful predictive capabilities and is especially good at handling daily periodic data with shifts in trend [19], and thus is suitable for the prediction based on ET measurement error. The Prophet model is robust to missing data and handles large outliers well, which can have good performance when the actual operation data of ET is not complete or has outliers. In addition, this model is also designed to have intuitive parameters that can be adjusted without knowing the details of the underlying model [20].

The Prophet model uses a generalized additive model to fit and predict functions, as shown in (4), where the model regards the time series as a superposition of three parts: the trend $g(t)$, the period $s(t)$, and the mutation $h(t)$:

$$ y(t) = g(t) + s(t) + h(t) + \epsilon_t. $$

(a) If the growth trend of time series is nonlinear, $g(t)$ is a logistic regression function; if the growth trend is linear, $g(t)$ is a linear function. It can be seen from the boxplot of ET phase error that the growth trend is linear, and thus, the linear growth model is adopted in this paper.

(b) $s(t)$ represents the periodic change, which consists of Fourier series:

$$ s(t) = \sum_{n=-N}^{N} c_n \exp(i(2\pi nt / p)), $$

where $p$ indicates the period of the target time series, $c_n$ is the parameters to be estimated. $N$ is the set number of approximation terms. The larger $N$ is, the more complex periodicity can be fitted, but the filtering effect may be not good. The setting of $N$ needs to be considered in combination with $p$. For the periodicity of each week, $p$ is set to 7, and $N$ is set to 3 [8].

(c) $h(t)$ indicates the changes caused by particular reasons:

$$ h(t) = Z(t)k, $$

$$ z(t) = [I(t \in D_1), \ldots, I(t \in D_n)], $$

where $D_i$ indicates the dates with special changes in the time series; $I$ is an indicator function representing whether time $t$ belongs to $D_i$; $k$ follows normal distribution, i.e., $k \sim N(0, \nu)$, and the larger $\nu$ is, the larger fluctuations the model is allowed to adapt to, while the smaller value of $\nu$ will suppress the influence of the model.

(d) $\epsilon_t$ is a noise term that follows normal distribution and is used to represent random and unpredictable fluctuations.

4. Case Study

4.1. Parameter Settings of BPNN. Before using the BPNN to predict the phase error of ET, the preprocessed data need to be standardized. The $Z$-score standardization method is adopted in this paper, as shown in the following equation:

$$ x' = \frac{x - \mu}{\sigma}, $$

where $\mu$ and $\sigma$ are the mean and the standard deviation of ET phase error dataset, respectively.

The parameter settings of BPNN include the number of hidden layers, the number of neurons in each layer, the neuron training function of each layer, and the learning efficiency. These parameters have a great influence on the final fitting and prediction accuracy. Unreasonable parameter settings will easily cause the error convergence of neural network fall into local optimum and thus cannot meet the accuracy requirement. Moreover, it is noteworthy that too high target fitting accuracy may lead to the overfitting phenomenon of the neural network, which will reduce the prediction accuracy.

In the BPNN training, the neural network parameters can be continually modified through looping statements, and multiple sets of experimental results can be obtained, so that the optimal parameter settings can be obtained through comparison. The parameter selection and training process of BPNN are shown in Figure 7. The loop in Figure 7 represents the number of loops. After multiple cycles of the comparison test, the parameter settings with the best prediction effect are shown in Table 1.

4.2. Parameter Settings of the Prophet Model. Before using the Prophet model to predict the ET phase error, the preprocessed data need to be adjusted according to the data requirements of Prophet. The data should be changed to two columns of data frame with fixed name: $d$ and $y$, where the columns of $ds$ are the date and the columns of $y$ are the series of phase error of ET. After the format of input data is processed, parameters of the Prophet model will be set according to Table 2. The process of prediction phase error of ET with the Prophet model is presented in Figure 8.

4.3. Prediction Evaluation Indexes. In order to improve the accuracy of the two prediction methods, the data of July are divided into the training set and the test set according to the ratio of 7:3. The training set contains the data from 1st to 21st, while the test set contains the data from 22nd to 31st. In order to evaluate the accuracy of the prediction methods, the verification set is composed of the actual operating data of
ET from August 1st to 8th. The mean absolute error (MAE), mean absolute percentage error (MAPE), mean square error (MSE), and root mean square error (RMSE) are taken as the accuracy measurement indexes to compare the prediction data with the data in the verification set:

\[
MAE = \frac{1}{n} \sum_{i=1}^{n} |y - y'|, \tag{9}
\]

\[
MAPE = \frac{1}{n} \sum_{i=1}^{n} \frac{|y - y'|}{y}, \tag{10}
\]

\[
MSE = \frac{1}{n} \sum_{i=1}^{n} (y - y')^2, \tag{11}
\]

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y - y')^2}. \tag{12}
\]

Figure 7: Flowchart of BPNN parameter selection and training.

**4.4. Prediction Result.** The prediction results of the ET phase error using BPNN and the Prophet model are shown in Figures 9(a) and 9(b), respectively, and the evaluation index comparison of the two methods is shown in Table 3.

**4.5. Discussion.** As can be seen from Figure 9(a), since the overfitting of the neural network may cause worse prediction effect, the BPNN fitting curve and the actual operation curve in the prediction result are not completely matched due to the appropriate parameter settings of BPNN. In terms of prediction results, the BPNN only predicts the general trend of the ET phase error in the next 8 days. Although the MAE and RMSE indexes are not high, the daily prediction error is relatively large.

As can be seen from Figure 9(b), the Prophet model has a good effect in fitting to the ET phase error data and is highly inclusive for the mutation points and abnormal points. At the same time, as the Prophet model is very good at predicting time series with obvious periodicity, the prediction curve not only is in line with the development trend of phase error, but also has lower value of MAE and RMSE indexes. Moreover, the daily prediction error of this method is also very small, and thus, the prediction curve is very close to the actual operation curve.

Through the comparison of Figures 9(a) and 9(b), it can be seen that since the Prophet model does not exist overfitting problem, the fitting effect of the Prophet model is better than that of the BPNN method. Furthermore, compared with four evaluation indexes in Table 3, the prediction trend of the Prophet model is more accurate than that of the BPNN method. Therefore, for the ET phase error prediction, the method based on the Prophet model is significantly better than the method based on BPNN.

The Prophet model fits the time series according to (4). When the Prophet model is adopted to predict the phase error of ET, in addition to the prediction results, the development trend and periodic fluctuation of the phase error over a period of time can be obtained as well, as shown in Figure 10.

The trend component in Figure 10 represents the development and prediction trend of the ET in July, and the blue shading indicates the confidence interval of the prediction. The weekly and daily components represent the weekly fluctuations and daily fluctuations of the ET phase error. Comparing the fluctuation amplitude of each component, i.e., the trend term (−0.02, 0.04), the weekly fluctuation (−0.0025, 0.015), and the daily fluctuation (−0.03, 0.03), it can be found that the variation of the ET phase error is mainly the daily fluctuation.

In the daily component, it can be seen that the phase error is at a very high level from 9:00 PM to 8:00 AM of the next day, while at a very low level from 10:00 AM to 6:00 PM. This phenomenon may be caused by temperature variation between day and night, resulting in the difference between the error variation speed of the air-core coil current transformer and the electromagnetic current transformer, which caused significant daily fluctuations in the final phase error data. Therefore, the trend component of the metering
Table 2: Parameter settings of the Prophet model.

| Parameter                | Explanation                                      | Value       |
|--------------------------|--------------------------------------------------|-------------|
| Growth                   | The way of increasing                             | Linear      |
| Seasonality prior scale  | Control the influence of periodic change          | 10.0        |
| Change point prior scale | Control the inflection point data                 | 0.05        |
| Holidays prior scale     | Control the influence of mutation point           | 10.0        |
| Interval width           | Confidence coefficient                            | 0.95        |
| MCMC samples             | Estimated sample number of MCMC parameter         | 5000        |

Figure 8: Flowchart of Prophet model prediction.

Figure 9: ET phase error prediction results. (a) Result of BPNN. (b) Result of the Prophet model.
error obtained by the Prophet model will provide reference for follow-up research of the influence of temperature on the metering error of ET as well as forecasting the metering error.

5. Conclusions

The short-term prediction methods of ET error based on BPNN and the Prophet model are proposed in this paper. After the preprocessing of data, the prediction of 8 days is made based on the ET phase error data in July through the proposed methods, and then, the indexes of MAE, MAPE, MSE, and RMSE are adopted to compare and analyse the prediction results of the two methods. Through the case results, the following conclusions can be obtained:

(1) The fitting effect of the BPNN is negatively correlated with the prediction effect, while the Prophet model can obtain better fitting effect as well as ensures the prediction accuracy.

(2) Through the comparison of image results and quantitative analysis, it can be seen that the prediction effect of the Prophet model on the error of ET is significantly better than that of the BPNN method.

### Table 3: The comparison of fitting and prediction results.

| Intelligent algorithm | MAE     | MAPE   | MSE             | RMSE  |
|-----------------------|---------|--------|-----------------|-------|
| **BPNN**              |         |        |                 |       |
| Fitting               | 0.0116  | 1.6128 | 2.835e – 04     | 0.0168|
| Prediction            | 0.0222  | 2.2663 | 7.1379e – 04    | 0.0267|
| **The Prophet model** |         |        |                 |       |
| Fitting               | 0.0130  | 3.629  | 2.5988e – 04    | 0.0167|
| Prediction            | 0.0129  | 0.9261 | 2.2642e – 04    | 0.0150|

#### Figure 10: Trend and periodic components of ET phase error.
The periodic change of the ET error mainly depends on the daily fluctuation, which may be caused by the change of the error variation speed due to the temperature change in one day.

In summary, the Prophet model, which is good at predicting obviously periodic time series, is more suitable for short-term prediction of ET error than BPNN.

Currently, we have completed the short-term prediction of the trend of the metering error of ET through the historical metering error data. In the follow-up work, more historical operating data of ETs in different periods will be collected, including factory, operation, failure, and scraping. Thus, a multisample space of ETs can be built, so as to realize the simulation of the whole life cycle operation situation of ETs driven by data and provide reference for the operation and maintenance of ETs under the corresponding types.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the Science and Technology Project of Headquarters of State Grid Corporation of China (no. 521820170001).

References

[1] M. Gutten, D. Korenciak, M. Kucera, R. Janura, A. Glowacz, and E. Kantoch, “Frequency and time fault diagnosis methods of power transformers,” Measurement Science Review, vol. 18, no. 4, pp. 162–167, 2018.

[2] G. Rigatos and P. Siano, “Power transformers’ condition monitoring using neural modeling and the local statistical approach to fault diagnosis,” International Journal of Electrical Power & Energy Systems, vol. 80, pp. 150–159, 2016.

[3] X. Yang, W. Chen, A. Li, C. Yang, Z. Xie, and H. Dong, “BP-PNN-based methods for power transformer fault diagnosis,” Advanced Engineering Informatics, vol. 39, pp. 178–185, 2019.

[4] A. Mejia, M. Valtierra, D. Granados, J. C. Olivares-Galvan, and R. Escarela-Perez, “The application of EMD-based methods for diagnosis of winding faults in a transformer using transient and steady state currents,” Measurement, vol. 117, pp. 371–379, 2018.

[5] B. Li, J. Zhang, Y. He, and Y. Wang, “Short-term load-forecasting method based on wavelet decomposition with second-order gray neural network model combined with ADF test,” IEEE Access, vol. 5, pp. 16324–16331, 2017.

[6] S. S. M. Ghoneim, S. S. Dessouky, A. A. Elfarashkoury, and A. B. A. Sharaf, “Prediction of insulating transformer oils breakdown voltage considering barrier effect based on artificial neural networks,” Electrical Engineering, vol. 100, no. 4, pp. 2231–2242, 2018.

[7] P. Wang, X. Zhang, B. Han et al., “Prediction model for railway freight volume with GCA-genetic algorithm-generalized neural network: empirical analysis of China,” Cluster Computing, vol. 22, no. 2, pp. S4239–S4248, 2019.

[8] L. Li, G. Duan, and J. Wang, “Reserve prediction of bank outlets based on prophet framework,” Journal of Central South University (Science and Technology), vol. 50, no. 1, pp. 76–82, 2019.

[9] N. Ge, L. Sun, X. Shi et al., “Research on sales forecast of prophet-LSTM combination model,” Computer Science, vol. 46, no. 6A, pp. 446–451, 2019.

[10] N. Ge, L. Sun, X. Shi et al., “Scale prediction of AQI based on Prophet-random forest optimization model,” Environmental Pollution & Control, vol. 41, no. 7, pp. 758–766, 2019.

[11] IEC, Communication Networks and Systems in Substations-Part 9-2: Specific Communication Service Mapping (SCSM) -Sampled Values over ISO/IEC 8802-3: IEC61850-9-2, International Electrotechnical Commission, Geneva, Switzerland, 2004.

[12] V. Viswanathan and S. R. Viswanathan, Data Analysis Cookbook pp. 30–39, Packet Publishing Ltd, Birmingham, U.K., 2nd edition, 2015.

[13] M. Frigge, D. C. Hoaglin, and B. Iglewicz, “Some implementations of the boxplot,” The American Statistician, vol. 43, no. 1, pp. 50–54, 1989.

[14] B. K. Friedrich and F. Barmann, “A learning algorithm for multi-layered neural networks based on linear least squares problems,” Neural Networks, vol. 6, no. 1, pp. 127–131, 1993.

[15] X. Zhang, “Optimized character recognition algorithm based on BP artificial neural network,” Computer Engineering and Applications, vol. 48, no. 35, pp. 182–185, 2012.

[16] J. E. Dayhoff and J. M. Deleo, “Artificial neural networks,” Cancer, vol. 91, no. 8, pp. 1615–1634, 2001.

[17] S. J. Taylor and B. Letham, “Forecasting at scale,” The American Statistician, vol. 72, no. 1, pp. 37–45, 2017.

[18] S. J. Taylor and B. Letham, Prophet: Automatic Forecasting Procedure (Version 0.5), 10 Oct. 2019, https://max.book118.com/html/2017/0706120622508.shtm.

[19] D. Zhai, X. Zhang, P. Xiong et al., “Detection of ionospheric TEC anomalies based on prophet time-series forecasting model,” Earthquake, vol. 39, no. 2, pp. 46–62, 2019.

[20] M. Feng, J. Zheng, J. Ren et al., “Big data analytics and mining for effective visualization and trends forecasting of crime data,” IEEE Access, vol. 7, pp. 106111–106123, 2019.