WILEY | Hindawi

Research Article

A Method for Parameter Identification of Distribution Network Equipment Based on Sequential Model-Based Optimization

Bin Li [ID],[1] **Jia Yang Ma** [ID],[1] **Ke Hu** [ID],[2] **Shi He Xu** [ID],[3] **Hao Jiao** [ID],[4] **Jin Ming Chen** [ID],[4] **and Wei Liu** [ID][4]

[1]*Nanjing Institute of Technology, Nanjing 211167, Jiangsu, China*
[2]*Chongqing University of Posts and Telecommunications, Chongqing 400065, China*
[3]*University of Science and Technology of China, Hefei 230027, Anhui, China*
[4]*State Grid Jiangsu Electric Power Co., Ltd., Research Institute, Nanjing 210096, Jiangsu, China*

Correspondence should be addressed to Ke Hu; huke@cqupt.edu.cn

We present the first model-based parameter identification method in the power distribution network to successfully achieve parameter identification directly based on sequential model-based optimization. This method is building a model with a posteriori probability to optimize an objective function. Furthermore, to achieve an efficient exploration, three different acquisition functions, i.e., random search, tree-structured Parzen estimator approach, and simulated annealing, were proposed. We applied our three models and the conventional model-free method to the actual feeder data with no adjustment of the other conditions. The experiment shows that our method achieves at least 25% and 70% improvements in accuracy and convergence speed, respectively.

1. Introduction

In the power distribution network (PDN), reliable and accurate parameters are the basis of security analysis, system control, state estimation, line loss calculation, power flow calculation, protection setting, and fault analysis [1]. Owing to the improved distribution power grid construction, new applications, such as new energy vehicles and distributed generation [2], depend on it. Thus, to ensure security and stability, it is necessary to measure the PDN effectively using real-time information. In reality, owing to the lack and limitations of real-time measuring equipment, some import parameters, such as the line resistance, line reactance, transformer resistance, transformer reactance, transformer conductance, and transformer electrical susceptance, are often difficult to obtain directly. These parameters are assumed to be static, with ideal values instead of actual values. In particular, under the influence of unstable operating conditions, environment, grounding resistance, etc. [3], it is difficult to avoid the differences in data between the distribution management system and the actual situation, which do not reflect the real-time operation status in PDN, resulting in poor parameter calculations for the distribution network.

Currently, the parameter identification research for lines and transformers is more active in the area of the power transmission network (PTN) than the PDN. Some of these studies include the theoretical formula calculation method based on the self-geometric spacing method [4] and the identification method based on field measurements of the voltage, current, power, frequency, and other network parameters using electrical instruments. Compared with other main networks, the PDN covers a larger area, and its measurement conditions are unsatisfactory. Therefore, the parameter identification methods used in the PTN may not be suitable for the PDN.

In the field of PDN, with the wide application of supervisory control and data acquisition, power management unit (PMU), and advanced metering infrastructure (AMI), some new approaches focusing on efficiency and error have been developed, such as the full-scale approach [5], PSOSR [6], normalized Lagrange multiplier (NLM) test [7], finite-

time algorithm (FTA) [8], residual method, sensitivity analysis method, Lagrange multiplier method [9], and Heffron–Phillips method [10]. With the development of machine learning, some new approaches have been proposed, such as a method that considers distance space [11], particle swarm optimization (PSO) algorithm [12], interior point method [13], ensemble Kalman filtering [14], evolutionary strategies [15], estimation using synchrophasor data [16], PSCAD simulation [17], and deep learning [18]. Most of these methods are based on lack of original data of the line in PDN measurement devices; consequently, they inevitably exhibit some deviations. It has been proven that these methods are effective with simulation data and have strict requirements for the precision of voltage and voltage phase angle. In addition, they require data regarding the voltage, current, active power, and reactive power on the high- and low-voltage sides. Therefore, several measuring devices should be placed in the grid; otherwise, only a part of the parameters can be obtained. To solve these problems, identification equations for the equipment parameters in the PDN are established, and it builds relationships between our target (parameters to be identified, such as the line resistance, line reactance, transformer resistance, transformer reactance, transformer conductance, and transformer electrical susceptance) and raw data (data can be obtained easily, such as the active power, reactive power, voltage on the high side, and voltage on the low side). Thus, the problem reverts to how to solve node equations, and two model-free methods, such as the least squares method (LS) [19] and the latest research MCMC [20] method, are applied to solve these equations. Based on node equations, these two methods work only based on the limited real-time measurement data collected from the end layer of the grid, and the parameters in the entire grid can be calculated with different types of nodes [21]. However, these two methods both have some disadvantages. The LS method assumes that the parameter space is approximately convex and a satisfactory local extremum cannot be easily attained; hence, this method is sensitive to the initial value, and the calculation process converges at an unsatisfactory result easily [20]. Moreover, the whole process of parameter calculations is time consuming, and it increases exponentially with increasing scale of the grid. The MCMC method can obtain high-precision parameter values without phase angle information; additionally, this method is insensitive to the initial values and exhibits fast convergence [20]. But, it has some strict requirements for convergence. For example, it performs unsatisfactory with standard node equations and needs to design an unexplainable and complex equation (named *loss function* in [20]), especially instead of node equations. In addition, to acquire excellent identification results, it requires an excess constrained parameter search space when working. Even if it performs better than LS algorithm on regular data, it is hard to predict the parameter with the outlier's value or high deviations. Thus, the MCMC method cannot be used in fault detection, and it is hard to be applied in real situation.

To sum up, the current parameter identification research in the PDN mainly faces the following problems:

(1) Due to the complexity of network structure in the PDN, there is little research progress in this field.

(2) The studies and experiences in the PTN are hard to transfer to the PDN.

(3) In the PDN, most of the studies about parameter identification are based on an idea condition that enough real-time measurement devices are placed on the whole grid. But in reality, these devices are mostly only placed at the end layer of grid. So, most of them cannot be widely applied in the current situation.

(4) Calculation based on node equations can solve the problem lacking real-time measurements, and two methods, the LS method and MCMC method, are proposed. After the analysis mentioned above, there exists many application problems in the MCMC method currently; thus, in this study, we did not make comparisons with it.

(5) The LS method is a feasible solution currently, but as mentioned above, there are disadvantages in accuracy and calculation efficiency.

In this study, we present the first model-based solution about parameter identification in the PDN and propose another feasible method based on sequential model-based global optimization (SMBO) [22]. In addition, we also research how to explore more efficiently in this model-based solution. SMBO is based on Bayes optimization [23, 24]. The principle of this theory is building a model with a posteriori probability to optimize an objective function using the existing samples in parameter identification tasks [25, 26]. In this model, samplings are regarded as being from one of the distributions such as the Gaussian distribution or the compound Gaussian distribution. Thus, the model can be represented easily, and the dimensions of the model are lower than those of most of the machine learning and deep learning models. If one existing point is from a sample point, the mean is the value of the optimized objective function at this point, and the variance is 0. The mean and variance of other unknown sample points are fitted by the posterior probability model, which may not be close to the true value. Thereafter, a collection method, such as Markov chain Monte Carlo, is used to constantly explore the value by sampling from the objective function corresponding to these unknown sample points, and then, it constantly updates the model of posterior probability using the acquisition function. Because the acquisition function can consider exploration and exploitation, points with better performance will be increasingly selected. Thus, good optimization can be achieved.

In the experiments, raw data are sampled from an actual 10 kV feeder. To verify whether our work can be applied widely, the whole identification process is with no adjustment of the other conditions, e.g., based on standard node equations and the parameter search spaces are not limited. Moreover, to improve the exploration ability to be more efficiently and obtain a higher accuracy and coverage speed, three different acquisition functions, that is, the random

search approach (RS) [27, 28], tree-structured Parzen estimator approach (TPE) [28], and simulated annealing (SA) [29], are applied in SMBO. Finally, comparisons are made between the LS method and our solutions.

2. Materials and Methods

2.1. Model of the Power Flow Calculation. In the PDN, transformers and lines are the main components. The analysis in the PDN usually assumes all feeders from a substation as the object of the system, and each feeder is regarded as a basic analysis unit. Because the line in the PDN is shorter and the voltage is lower than in other networks in the power system, the charging capacitance of the line is neglected. As Figure 1 shows, according to graph theory in power analysis, the distribution transformer is regarded as a Γ-type equivalent circuit in this study.

Considering the error and computational complexity, we assumed the three-phase balance as the premise in the calculation of the power flow. The total power was distributed evenly in each phase for calculation. The power flow calculation method is shown in Figure 2.

The active power P_d, reactive power Q_d, and voltage U_d on the high-voltage side of the transformer at bus D can be obtained directly by real-time measurement. Other basic parameters in the line and transformer, such as the transformer reactance R_d, transformer resistance X_d, transformer conductance G_d, transformer electrical susceptance B_d, line resistance R_{cd}, and line reactance X_{cd}, are important to the PDN, but are difficult to detect. These basic parameters satisfy the following formula [21]:

$$
\begin{cases}
P_d = P_{Ld} + \dfrac{P_{Ld}^2 + Q_{Ld}^2}{U_{Ld}^2} R_d^T + U_{Ld}^2 G_d^2, \\[3mm]
Q_d = Q_{Ld} + \dfrac{P_{Ld}^2 + Q_{Ld}^2}{U_{Ld}^2} X_d^T + U_{Ld}^2 B_d^2, \\[3mm]
U_d = \sqrt{\left(U_{Ld} + \Delta U_d^T\right)^2 + \left(\delta U_d^T\right)^2}.
\end{cases}
\tag{1}
$$

ΔU_d^T and δU_d^T are the longitudinal and transverse components of the transformer impedance voltage drop at bus D in V, respectively.

$$
\begin{cases}
\Delta U_d^T = \dfrac{P_{Ld} R_d^T + Q_{Ld} X_d^T}{U_{Ld}}, \\[3mm]
\delta U_d^T = \dfrac{P_{Ld} X_d^T + Q_{Ld} R_d^T}{U_{Ld}}.
\end{cases}
\tag{2}
$$

The equation of bus C can be expressed as [21]

$$
\begin{cases}
U_c = \sqrt{\left(U_d + \Delta U_{cd}^T\right)^2 + \left(\delta U_{cd}^T\right)^2}, \\[3mm]
\Delta U_{cd}^T = \dfrac{P_d R_{cd}^T + Q_d X_{cd}^T}{U_{Ld}}, \\[3mm]
\delta U_{cd}^T = \dfrac{P_d X_{cd}^T + Q_d R_{cd}^T}{U_d}.
\end{cases}
\tag{3}
$$

Thus, the final equation is [21]

$$
f_c(x) = \sqrt{\left(U_d + \Delta U_{cd}^T\right)^2 + \left(\delta U_{cd}^T\right)^2} - \sqrt{\left(U_{Lc} + \Delta U_c^T\right)^2 + \left(\delta U_c^T\right)^2}.
\tag{4}
$$

Through the abovementioned analysis, a set of node equations was obtained. Thus, the parameters in the line and transformer were calculated based on these equations and the measured power and voltage data.

2.2. SMBO. Generally, to solve the identification problem, LS is applied based on the node equations and measurement data. However, considering the parameter identification task in the PDN, there are some problems while using LS:

(1) In principle, the parameter space may be nonconvex, which is unsuitable for LS.

(2) In the PDN, there are many parameter identification equations to be solved owing to the large scale of lines and nodes. Therefore, the calculation efficiency is low.

(3) There are exponential differences at the numerical level, and the identification results are significantly affected by measurement errors.

To solve these problems, the SMBO optimization method was introduced. The SMBO method was used to solve the node equations based on the measurement data to identify the parameters in the line and transformer. During SMBO operation, a model can be trained and updated through iterations, and then, a set of objective parameters can be collected based on the current model to improve the performance of the model in the next iteration.

We consider that the training set is $\{(\Phi_1, O_1), \ldots, (\Phi_n, O_n)\}$ and $\Theta_i = \{(\Theta_{i,1}, \ldots, O_{i,d})\}$ is represented as a set of parameter configurations when training the model, where d represents the number of parameters required and O represents the corresponding observations. This model aims to predict the fitness of the current observation O_{n+1} based on a set of new parameters Θ_{n+1} (Algorithm 1).

In SMBO, each iteration can obtain a new model distribution (such as the Gauss process) based on the domain $H = (x_1, f(x_1), \ldots, x_n, f(x_n))$, which has been evaluated.
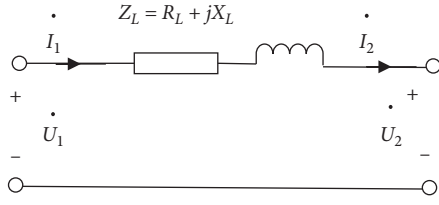
FIGURE 1: Γ-type equivalent circuit of the transformer and Π-type equivalent model of the low-voltage distribution line [20].
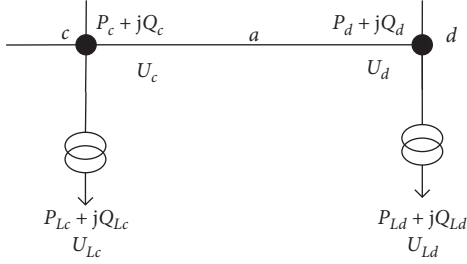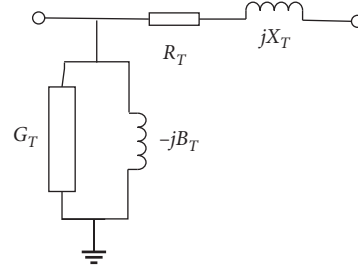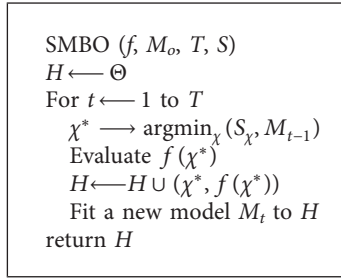

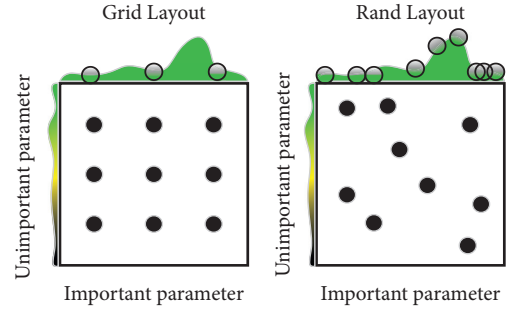FIGURE 2: Power flow calculation circuit model [20].

FIGURE 3: Schematic diagram of RS's advantage. The grid and random search of nine trials for optimizing the function $f(x, y) = g(x) + h(y) \approx g(x)$ with low effective dimensionality is shown. Above each square, $g(x)$ is shown in green, and the left of each square $h(y)$ is shown in yellow. With a grid search, nine trials only tested $g(x)$ in three distinct locations. Using a random search, all nine trials explored distinct values of g.

```
SMBO (f, Mₒ, T, S)
H ⟵ Θ
For t ⟵ 1 to T
    χ* ⟶ argmin_χ (S_χ, M_{t-1})
    Evaluate f(χ*)
    H ⟵ H ∪ (χ*, f(χ*))
    Fit a new model M_t to H
return H
```

ALGORITHM 1: Algorithm SMBO.

Furthermore, at each iteration, a new object x^* ($x^* = \text{argmin}_{x_n} f(x_n)$) is obtained based on the new model distribution. Thereafter, the evaluation function is applied, and $(x^*, f(x^*))$ are added to domain H. Finally, the model distribution and a new x^* are both updated. The entire process can be expressed as shown in the following chart, where f represents the evaluation function and M represents the model distribution.

In this study, three different exploration methods for finding parameters in SMBO, namely, RS, TPE, and SA, were selected for comparison with the conventional LS. Three different optimization functions are briefly introduced below.

2.3. Random Search.

Bergstra et al. proved that random search is more efficient in the parameter identification field than the grid search in finite iterations [27]. Although the results of random search may be different, the experiments prove that random search is better than grid search, especially when the iteration step is finite [27]. Figure 3 illustrates how point grids and uniformly random point sets differ in coping with low effective dimensionality.

2.4. Simulated Annealing.

SA is a random optimization algorithm based on the Monte Carlo theory. SA is regarded as a greedy algorithm; however, what makes it different from a standard greedy process is that SA also obeys a random policy and introduces random actions with a low probability during the search process. Thus, it has the ability to prevent the case of local optima.

2.5. Tree-structured Parzen Estimator Approach.

TPE improves the traditional form of the model distribution, $p(y \mid x)$ to $p(x|y)$ and $p(y)$. For $p(x \mid y)$, the TPE is defined as follows:

$$p(x \mid y) = l(x), \quad if \ y < y,$$
$$p(x \mid y) = g(x), \quad if \ y \geq y^*, \tag{5}$$

where $l(x)$ represents the loss function of the current $x^{(i)}$ when $f(x^{(i)}) < y^*$ and $g(x)$ is the opposite case. Based on domain $B = \{x^{(i)}, \ldots, x^{(k)}\} \in H$, which is calculated by the k-means algorithm, TPE can be regarded as a prior Gaussian distribution, where each $x^{(i)} \in B$. For each discrete $x^{(i)}$, the prior distribution is p_i; thus, the posterior distribution is $N \times p_i + C_i$, where N represents the sample number and C_i represents the probability in domain B. The entire optimization process is based on the expected improvement in EI [26].

$$EI_{y^*}(x) := \int_{-\infty}^{y^*} (y^* - y) p(y \mid x) = \int_{-\infty}^{y^*} (y^* - y) \frac{p(x|y)p(y)}{p(x)} dy. \tag{6}$$

In TPE, $\gamma = p(y < y^*)$ and $p(\text{x}) = \int_R p(x|y)p(y)dy = \gamma l(x) + (1 - \gamma)g(x)$; thus,

$$\int_{-\infty}^{y^*} (y^* - y) p(x|y)p(y)dy = l(x) \int_{-\infty}^{y^*} (y^* - y)p(y)dy$$
$$= \gamma y^* l(x) - l(x) \int_{-\infty}^{y^*} p(y)dy. \tag{7}$$

Consequently,

$$EI_{y^*}(x) = \frac{\gamma y^* l(x) - l(x) \int_{-\infty}^{y^*} p(y)dy}{\gamma l(x) + (1 - \gamma)g(x)}. \tag{8}$$

From equation (8), more points that perform better are obtained according to the improved $(g(x)/l(x))$ and an optimal x^* with the maximum EI value.

3. Experiment

3.1. Preparation. For the experiment, the whole schedule is as shown in Figure 4.

(1) Establishing node equations: the terminal node equations are introduced in equations (1)–(4)

(2) Sampling data

(3) Parameter identification: at this step, five sets of parameters will be obtained

  (a) Four sets of parameters were obtained using different identification algorithms. Among them, one set was from the contrast group and conducted the LS method; the other three sets were from the experimental group and conducted the three different acquisition functions, respectively.

  (b) The remained one set is the static parameters.

(4) Result comparison

  (a) As shown in Figure 4, the identification error is evaluated by the difference between U_{algo} and U_{std}, where the U_{algo} represents the high-voltage side calculated by the identification parameters and U_{std} is the sampled true value of the high-voltage side

  (b) Convergence speed among different identification processes

In the process of sampling data, an actual 10 kV feeder was selected for the calculations. The 10 kV feeder is composed of a transformer (S11-M-400/10) and eight overhead transmission lines. The specific topology is illustrated in Figure 5.

The whole process, including exploration and solving node equations, has low cost for calculation, and the dimension of the studied model is not high, so ordinary research computers can finish this work.

3.2. Convergence. SMBO is a model-based method, and it works obeying the metaheuristic optimization. It has been verified that three different explorations have the ability to reach an ideal result close to the global optimized solution, but we must overcome the randomness in identification results. To solve this problem, we focus on the standard precision of these parameters and think that randomness can be ignored when the parameter update quantity is less than the standard precision in the following iterations.

The static parameters of the standard 10 kV feeder network are as follows.

3.2.1. Line

(1) Line resistance: $R_{cd} = 0.1263 \, (\Omega/\text{km})$

(2) Line reactance: $X_{cd} = 0.1665 \, (\Omega/\text{km})$

3.2.2. Transformer

(1) Transformer resistance: $X_d = 10.000 \, (\Omega/ \text{km})$

(2) Transformer reactance: $R_d = 2.825 \, (\Omega/ \text{km})$

(3) Transformer conductance: $G_d = 5.7 \times 10^{-6} \, (\text{S})$

(4) Transformer electrical susceptance: $B_d = 3.2 \times 10^{-5} \, (\text{S})$

To sum up, the precisions and convergence conditions of all identified parameters are summarized in Table 1.

Thus, when the difference between successive identified parameters is less than the corresponding precision (in other words, on convergence conditions), we think it achieves convergence and the impact of randomness can be neglected. In this study, we observe the updated error in following 10 iterations to ensure convergence.

3.3. Raw Data Descriptions. Without loss of generality, raw data collected on 1 January 2020 were selected randomly. The data were collected using SCADA, and the sampling period was 15 min. Figure 6 shows the three-phase first section voltage (U_A, U_B, and U_C) on the high-voltage side, Figure 7 shows the three-phase voltage (u_a, u_b, and u_c) on the low-voltage side, Figure 8 shows the three-phase active power (P_a, P_b, and P_c) on the low-voltage side, and Figure 9 shows the three-phase reactive power (Q_a, Q_b, and Q_c).

From Figure 6, it can be seen that the high-voltage side (marked as U_c in Figure 2) in the sampling data obeys the three-phase balance; thus, it satisfies the requirements of solving node equations and parameter identification. From Figure 7 to Figure 9, the trend of change in active power, reactive power, and low-voltage side (marked as P_d, Q_d, and U_d in Figure 2, respectively) is consistent. Thus, it is verified
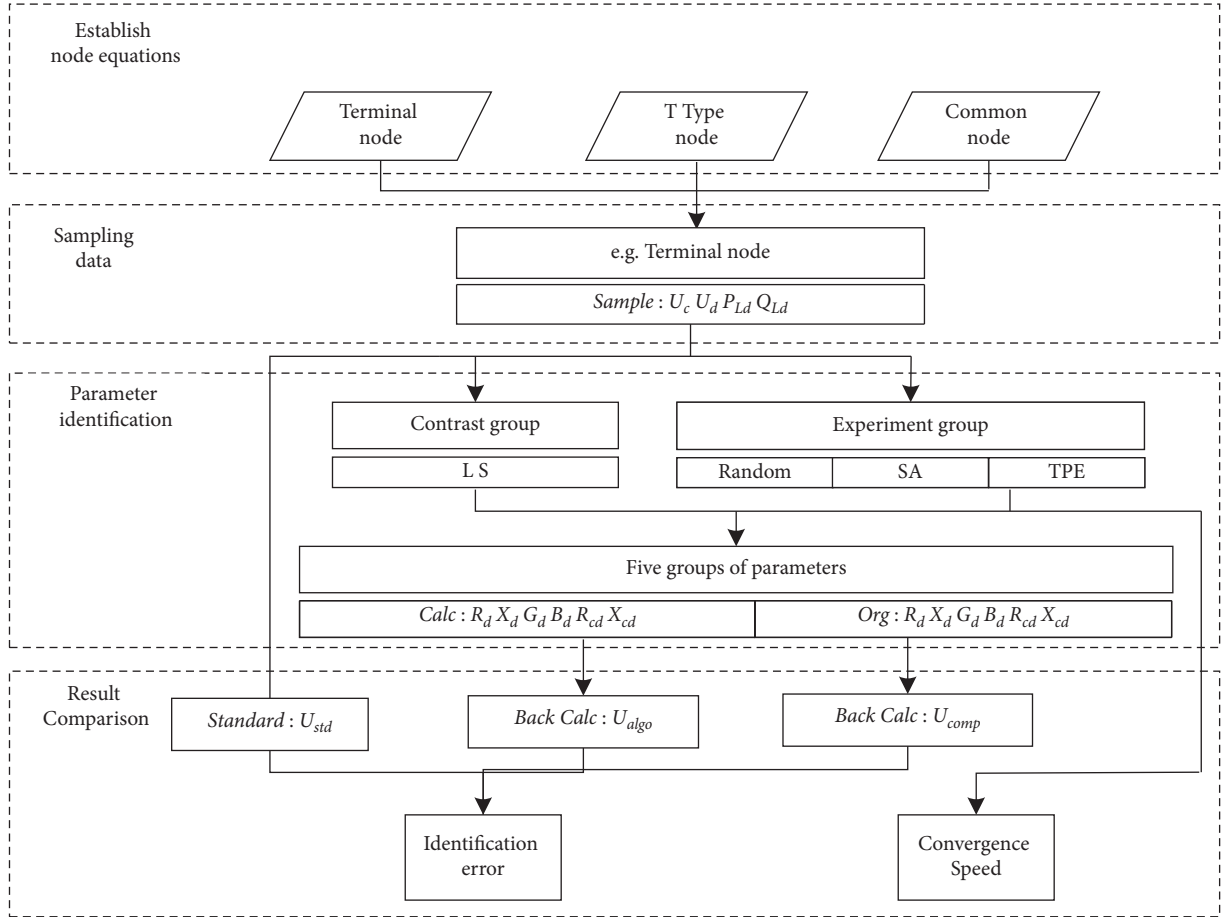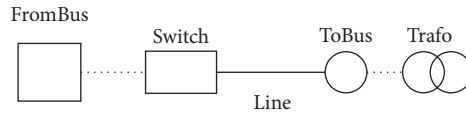
FIGURE 4: Schematic diagram of the experiment.

FIGURE 5: Ten-kilovolt feeder topology [20].

TABLE 1: Precisions and convergence conditions of all parameters.

	R_{cd}	X_{cd}	X_d	R_d	G_d	B_d
Precisions	1×10^{-4}	1×10^{-4}	1×10^{-3}	1×10^{-3}	1×10^{-6}	1×10^{-5}
Convergence conditions	1×10^{-5}	1×10^{-5}	1×10^{-4}	1×10^{-4}	1×10^{-7}	1×10^{-6}

that the data are stable and can be used to perform parameter identification.

4. Results

After parameter identification, one set of parameters in the contrast group based on LS is obtained. As shown in Figure 4, a back calculation was conducted, and three sets of parameters in the experimental group were obtained based on three different acquisition functions.

Because there is no real-time and accurate data of the line parameters, U_{std} is the standard voltage value per unit used for the voltage data, and U_{algo} is the high-voltage side value calculated with the identification parameter value. The identification results can be described by the difference between U_{std} and U_{algo}. In addition, the producer (factory) provides each line and transformer an original parameter value as reference. Therefore, these parameters are suitable to be regarded as the baseline, and we can also calculate the voltage values U_{comp} with these original values.

Figure 10 shows the identification parameter calculated using LS, Figure 11 shows the identification parameter calculated by RS, Figure 12 shows the identification
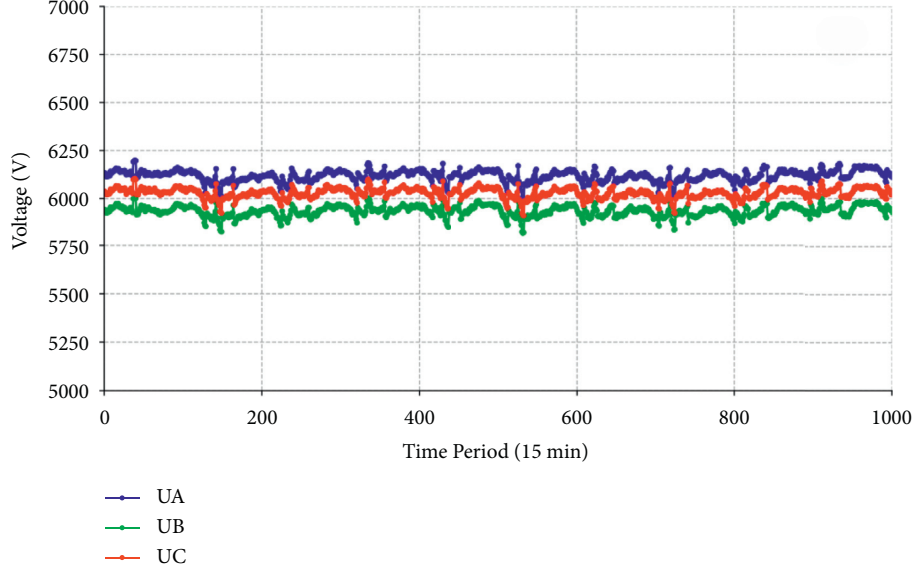
FIGURE 6: U_A, U_B, and U_C on the high-voltage side.
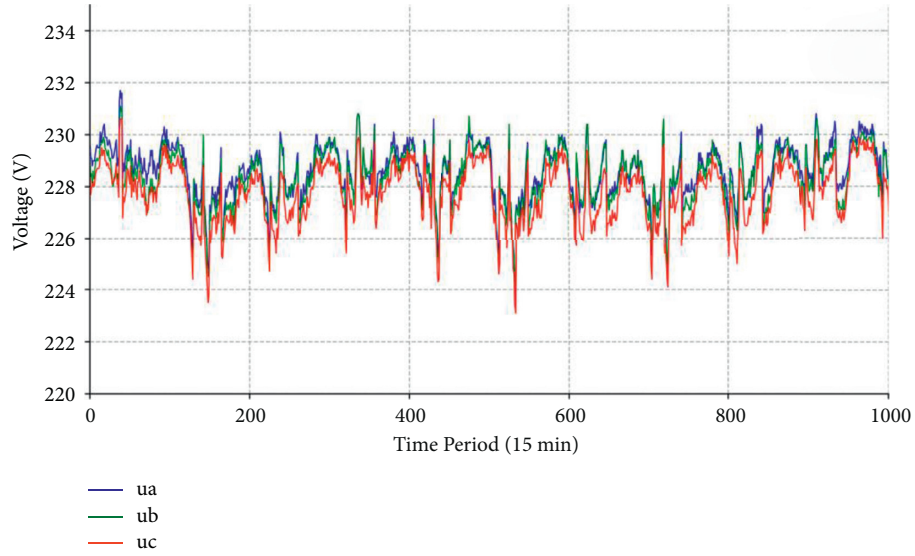
FIGURE 7: u_a, u_b, and u_c on the low-voltage side.

parameter calculated by SA, and Figure 13 shows the identification parameter calculated by TPE.

In Figures 10–13, the gap between the blue line U_{std} and the green line U_{algo} represents the difference between the identification parameters and the true value. The gap in the LS method is larger than that in the other three SMBO methods. To show the difference in fitting error, the four metrics, that is, the mean absolute error (MAE), roost-mean-square error (RMSE), mean absolute percentage error (MAPE), and symmetric mean absolute percentage error (SMAPE), are applied, which are defined as follows:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |\widehat{y}_i - y_i|,$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{\widehat{y}_i - y_i}{y_i} \right|,$$

$$SMAPE = \frac{100\%}{n} \sum_{i=1}^{n} \frac{|\widehat{y}_i - y_i|}{\left((|\widehat{y}_i| + |y_i|)/2 \right)}.$$
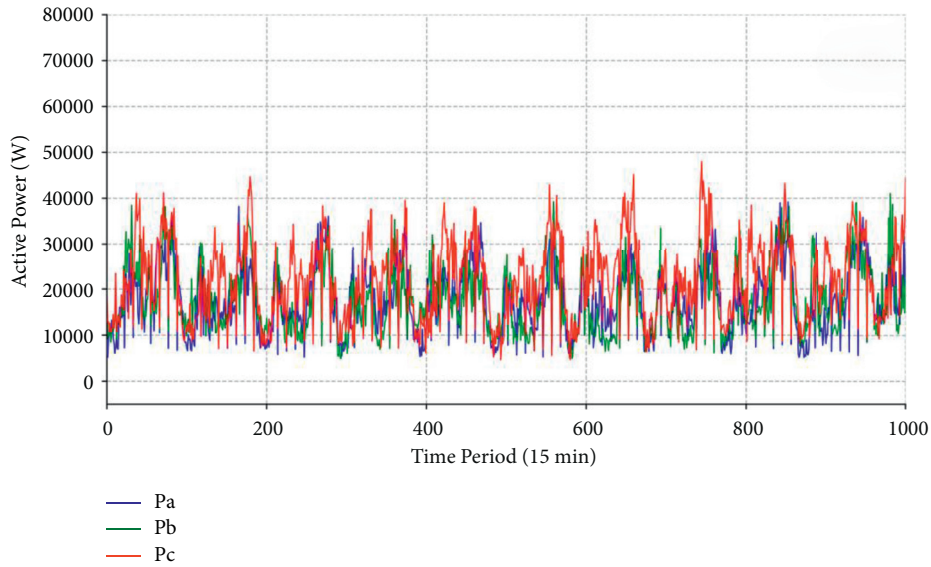
(9)

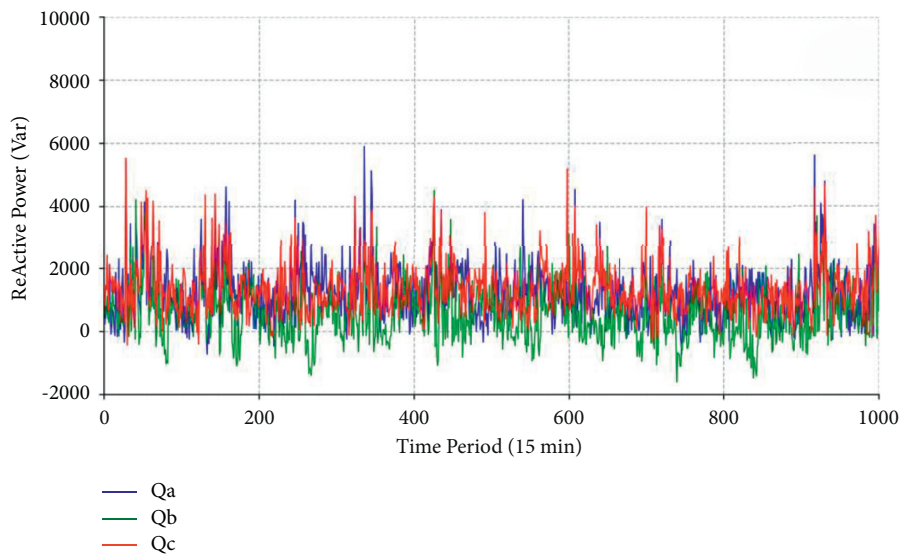FIGURE 8: P_a, P_b, and P_c on the low-voltage side.

FIGURE 9: Q_a, Q_b, and Q_c on the low-voltage side.

The overall performance of all the methods is shown in Table 2. The *Comp* group represents the parameter from the original parameter value; the other four groups include one traditional LS method and three methods (RS, SA, and TPE) based on SMBO. The convergence speed (*Converg spd*) was recorded, and all four corresponding methods were tested at the same calculation plat and same amount of data for calculation during 1000 iterations.

From Table 2, it can be seen that all the four different methods based on the algorithms (LS, RS, SA, and TPE) yield better results than those based on the *Comp* group, which demonstrates that the original parameter value is unreliable

and it is necessary to perform parameter identification. In addition, three SMBO methods perform better than LS on the abovementioned four metrics, indicating that the identification results based on SMBO are closer to the real values than the results based on LS. In the experimental group, the TPE method performed the best in all four metrics for fitting error, while the RS method performed the worst, which indicates that, in the parameter exploration process, a model based on an updating distribution is better than random.

Furthermore, from the column of *Converg spd*, the convergence speed of all three SMBO methods is
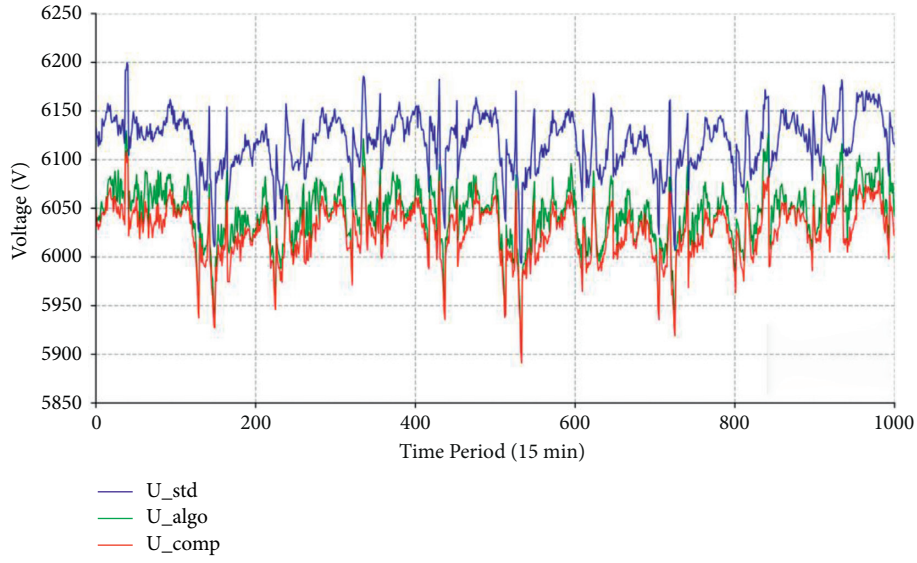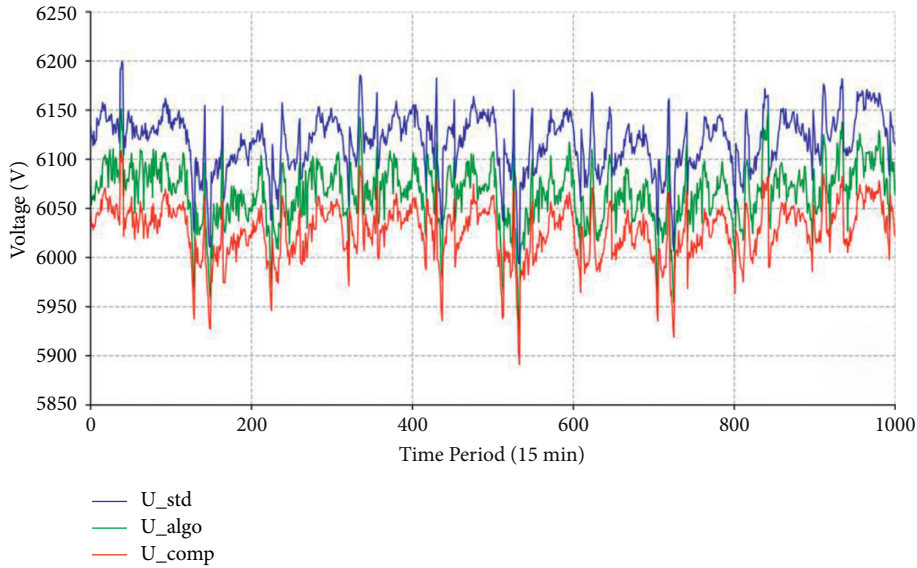
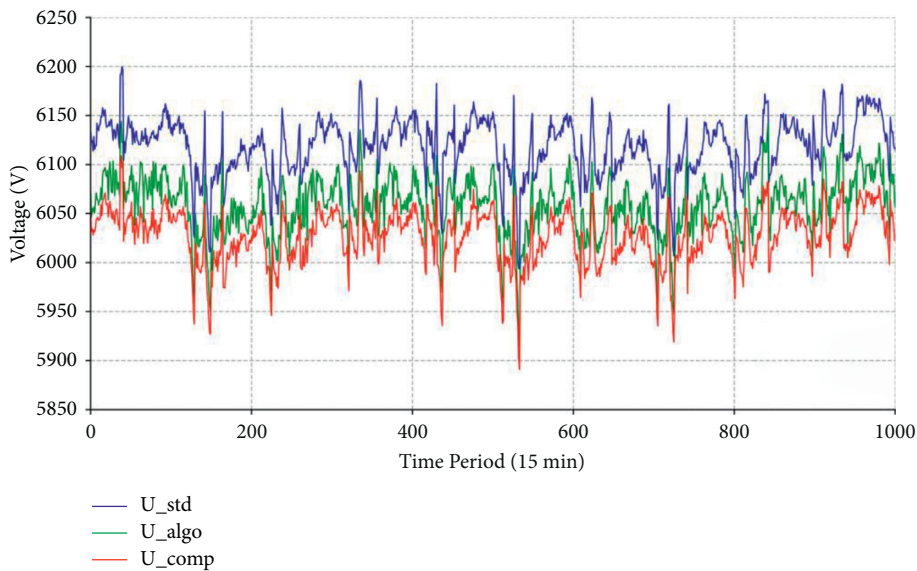FIGURE 10: Result calculated by LS.

FIGURE 11: Result calculated by RS.

FIGURE 12: Result calculated by SA.
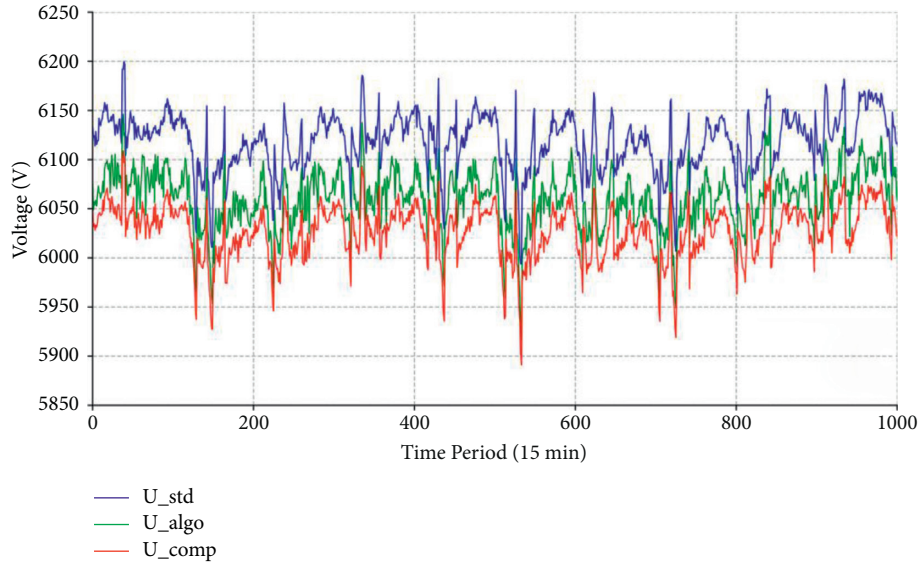
FIGURE 13: Result calculated by TPE.

TABLE 2: Metrics summary of all methods.

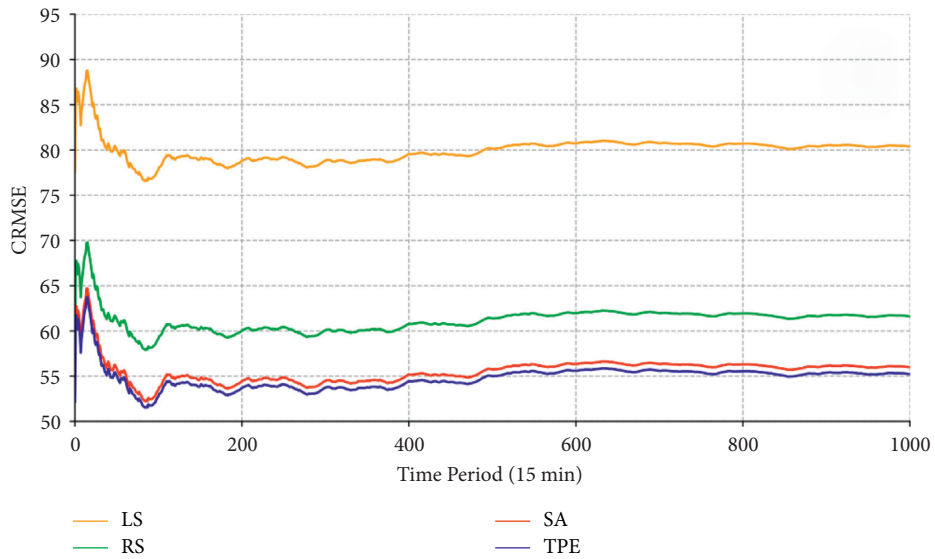|	MAE	RMSE	MAPE	SMAPE	Converg spd (s)
Comp	90.524	90.820	1.479	1.490	—
LS	76.908	78.163	1.257	1.265	1498
RS	57.791	59.451	0.944	0.949	311
SA	51.736	53.757	0.845	0.849	408
TPE	51.056	53.046	0.834	0.838	420

FIGURE 14: CRMSE curve with different parameter identification methods.

TABLE 3: Hypothesis testing results.

Gr. A	Gr. B SMBO			
P val	LS	RS	SA	TPE
LS	1.000	<0.01	<0.01	<0.01
RS	<0.01	1.000	<0.01	<0.01
SA	<0.01	<0.01	1.000	<0.01
TPE	<0.01	<0.01	<0.01	1.000

significantly faster than that of LS. In the experimental group, the RS method runs the fastest when convergence occurs because the calculation cost of the RS method is the lowest.

In addition, to compare the effects of different identification methods, the cumulative root-mean-square error (CRMSE) was adopted in this study.

For an arbitrary point at order m in the samples,

$$CRMSE^{(m)} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(u_{std}^{(i)} - u_{algo}^{(i)}\right)}, \tag{10}$$

where u_{std} represents the true value and u_{algo} represents the value from one of the identification methods. According to the definition, CRMSE represents the dynamic value error well with the incremental sampling sequence.

From Figure 14, there is a gap between the LS curve and the three SMBO curves for the whole sample sequence. Furthermore, in the group of the three SMBO curves, there is a gap between the RS curve and the other two curves.

Hypothesis testing is one of the methods used to determine whether there are significant differences in statistics between the two groups (e.g., groups A and B) of samples. By observing p, a significant difference between these two groups of samples was obtained.

(1) If $p > 0.05$, there were no significant differences. There was no significant difference in the mean and distribution between groups A and B. (2) If $p \leq 0.05$, it indicated a significant difference. There was a significant difference in the mean and distribution between groups A and B.

The testing results are in Table 3.

From Table 3, it can be seen that compared with the LS method, the p value of all three SMBO methods is less than 0.01, indicating that the results based on SMBO are all superior to those of the LS method.

In addition, the performances of the three SMBO-based methods are different. After hypothesis testing, it was found that the TPE method performed best in the experiments; the result of SA was similar to that of TPE, and the RS method performed worse than the other two methods.

5. Conclusions

In the PDN, accurate values of parameters in the line and transformer are important, but it is difficult to perform real-time measurements. In this study, we focused on the parameter exploration process and presented a method based on SMBO. In addition, three different exploration solutions,

RS, SA, and TPE, were introduced. Experiments were conducted using the known partial measurement data sampled from an actual 10 kV feeder.

The experimental results demonstrated that the SMBO performs better than LS in terms of accuracy and convergence speed. The TPE method performs best in terms of accuracy, and RS achieves the fastest convergence speed. Our method achieves at least 25% and 70% improvements in accuracy and convergence speed, respectively. Thus, our method can satisfy the requirements of the parameter identification task of the PDN.

However, some problems remain, such as lesser calculated high-voltage values than the true measurement values in Figures 10–13. Therefore, intensive work is required to design the last equation in node equations.

Data Availability

The data used to support the findings of this study are currently under embargo while the research findings are commercialized. Requests for data, 6/12 months after publication of this article, will be considered by the corresponding author.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

References

[1] L. Jianlin, M. Huimeng, and H. Dong, "Present development condition and trends of energy storage technology in the integration of distributed," *Renewable Energy Transactions of China Electrotechnical Society*, vol. 31, no. 14, pp. 1–10+20, 2016.

[2] C. Wang, "Challenges on coordinated planning of smart distribution networks driven by source-network-load evolution," *Proceedings of the CSEE*, vol. 40, no. 8, pp. 2385–2396, 2020.

[3] P. M. Anderson, *Analysis of Faulted Power Systems*, Wiley-IEEE Press, New York, NY, USA, 1995.

[4] C. Wang, "Wave propagation in overhead wires with ground return," *Bell System Technical*, vol. 5, pp. 539–554, 1926.

[5] M. Lave, M. J. Reno, and J. Peppanen, "Distribution system parameter and topology estimation applied to resolve low-voltage circuits on three real distribution feeders," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 3, pp. 1585–1592, 2019.

[6] Z. Wang, C. Hu, W. Wang, X. Kong, and W. Zhang, "A prognostics-based spare part ordering and system replacement policy for a deteriorating system subjected to a random lead time," *International Journal of Production Research*, vol. 53, no. 15, 2015.

[7] Y. Lin and A. Abur, "Strategic use of synchronized phasor measurements to improve network parameter error

detection," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 5281–5290, 2018.

[8] A. Arnaldo, S. Rangel, and A. Carlos, "Finite–time identification of the Thévenin equivalent parameters in power grids," *International Journal of Electrical Power & Energy Systems*, vol. 116, Article ID 105534, 2020.

[9] B. V. Krishna and N. P. Srinivasu, "A direct method for distribution system load flow solutions," *International Journal of Engineering and Advanced Technology*, vol. 8, no. 6s3, 2019.

[10] B. Zaker, G. B. Gharehpetian, and M. Karrari, "Small signal equivalent model of synchronous generator-based grid-connected microgrid using improved Heffron-Phillips model," *International Journal of Electrical Power & Energy Systems*, vol. 108, pp. 263–270, 2019.

[11] Q. Yan and M. Zheng, "A network parameter error identification method considering distance space," *Electric Power Engineering Technology*, vol. 34, no. 05, pp. 25–28, 2015, in chinese.

[12] S. Zhang, "Power system state estimation based on particle swarm optimization algorithm," *Power System Protection and Control*, vol. 38, no. 22, pp. 86–89, 2010, in chinese.

[13] M. F. Zhou, "A hybrid strategy based on interior point method and DPSO for electricity transmission network," *Electric Power Engineering Technology*, vol. 33, no. 01, pp. 22–25, 2014, in chinese.

[14] C. Carquex, C. Rosenberg, and K. Bhattacharya, "State estimation in power distribution systems based on ensemble kalman filtering," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6600–6610, 2018.

[15] W. do Nascimento Sepulchro, L. Frizera Encarnacao, and M. Brunoro, "Harmonic distortion and power flow state estimation for distribution systems based on evolutionary strategies," *IEEE Latin America Transactions*, vol. 13, no. 9, pp. 3066–3071, 2015.

[16] J. Song, E. Dall'Anese, A. Simonetto, and H. Zhu, "Dynamic distribution state estimation using synchrophasor data," *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 821–831, 2020.

[17] M. A. Zhibin, S. Zhang, and L. I. Junge, *A Single-phase Grounding Fault Section Location Method for Power Distribution Network Based on Parameter Identification*, Guangdong Electric Power, Guangdong, China, 2019.

[18] M. Mahmud, M. S. Kaiser, A. Hussain, and S. Vassanelli, "Applications of deep learning and reinforcement learning to biological data," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 6, pp. 2063–2079, 2018.

[19] K. Yuan and Z. Wei, "Power line parameter identification based on multi-innovation least square algorithm," *Electric Power Engineering Technology*, vol. 39, no. 04, pp. 55–160, 2020, in chinese.

[20] H. Wang, H. Jiao, J. Chen, and W. Liu, "Parameter identification for a power distribution network based on MCMC algorithm," *IEEE Access*, vol. 9, pp. 104154–104161, 2021.

[21] K. Wu and H. H. Yuan, "A method for Parameter identification of distribution network equipment based on least square method," *Energy and Energy Conservation*, vol. 05, pp. 58–63, 2019.

[22] F. Hutter, H. H. Hoos, and K. Leyton-Brown, "Sequential model-based optimization for general algorithm configuration," *Lecture Notes in Computer Science*, vol. 6683, pp. 507–523, 2011.

[23] M. Pelikan, *Bayesian Optimization Algorithm*, Springer, Berlin, Heidelberg, 2005.

[24] F. Matthias, "Springenberg jost tobias. Hutter. Frank. Initializing bayesian hyperparameter optimization via meta-learning," *Proceedings of the National Conference on Artificial Intelligence*, vol. 2, pp. 1128–1135, 2015, June 1.

[25] F. Hutter, H. H. Hoos, and T. Stützle, "Automatic algorithm configuration based on local search," *AAAI*, vol. 7, pp. 1152–1157, 2007.

[26] S. Jasper, L. Hugo, and P. Adams Ryan, "Practical Bayesian optimization of machine learning algorithms," *Advances in Neural Information Processing Systems*, vol. 4, pp. 2951–2959, 2012.

[27] B. James, "Bengio yoshua. Random search for hyper-parameter optimization," *Journal of Machine Learning Research, February*, vol. 13, pp. 281–305, 2012.

[28] J. Bergstra, "Algorithms for hyper-parameter optimization," *Advances in Neural Information Processing Systems*, vol. 11, 2011.

[29] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science (New York, N.Y.)*, vol. 220, pp. 671–680, 1983.