Classifying Transformer Winding Deformation Fault Types and Degrees Using FRA Based on Support Vector Machine

JIANGNAN LIU[1], ZHONGYONG ZHAO[1], CHAO TANG[1], CHENGUO YAO[2], (Member, IEEE), CHENGXIANG LI[2], AND SYED ISLAM[3], (Fellow, IEEE)

[1]College of Engineering and Technology, Southwest University, Chongqing 400716, China
[2]State Key Laboratory of Power Transmission Equipment and System Security and New Technology, Chongqing University, Chongqing 400044, China
[3]School of Science Engineering and Information Technology, Federation University Australia, Ballarat, VIC 3353, Australia

Corresponding author: Zhongyong Zhao (zhaozy1988@swu.edu.cn)

ABSTRACT As an important part of power system, power transformer plays an irreplaceable role in the process of power transmission. Diagnosis of transformer's failure is of significance to maintain its safe and stable operation. Frequency response analysis (FRA) has been widely accepted as an effective tool for winding deformation fault diagnosis, which is one of the common failures for power transformers. However, there is no standard and reliable code for FRA interpretation as so far. In this paper, support vector machine (SVM) is combined with FRA to diagnose transformer faults. Furthermore, advanced optimization algorithms are also applied to improve the performance of models. A series of winding fault emulating experiments were carried out on an actual model transformer, the key features are extracted from measured FRA data, and the diagnostic model is trained and obtained, to arrive at an outcome for classifying the fault types and degrees of winding deformation faults with satisfactory accuracy. The diagnostic results indicate that this method has potential to be an intelligent, standardized, accurate and powerful tool.

INDEX TERMS Transformer, winding faults, FRA, SVM.

I. INTRODUCTION

Large power transformers constitute very expensive and vital components in electric power systems [1]. The reliability of power transformers, which are critical core equipment in power transmission and distribution systems, dictates the safe and reliable performance of the entire electrical system [2]. Most importantly, the stable and safe operation of power transformer is of significance to the normal operation of power system. The faults of power transformers will have serious impact on the safety of power grids [3]. It is necessary to pay attention to higher reliability of power transformers, due to possible failures of an electric system and because of their cost [4].

Due to above reasons, condition monitoring of transformer operation status has attracted more and more attention throughout the world [5]. Winding mechanical deformation

The associate editor coordinating the review of this article and approving it for publication was Canbing Li.

faults are one of common fault types of transformers. At present, many winding mechanical fault diagnosis methods are proposed theoretically and practically, for instance, the short-circuit impedance (SCI) method based on the principle of short-circuit impedance measurement, the low-voltage impulse (LVI) method based on the principle of signal analysis, the frequency response analysis (FRA) method, *etc*. In SCI method, measured SCI of a phase winding is compared to the value that appears at the nameplate or factory test results [6]. In LVI proposed by Lech, W. and Tyminski, L. in 1966 [7], time domain signals of winding before and after fault are compared and analyzed to present the information of winding deformation.

Above all, FRA method has been widely accepted because it's economic, accurate, simple and fast. FRA, firstly introduced by Dick and Erven [8], has adopted sweep frequency sinuous signal to excite transformer windings and measured the response signal in the frequency domain to construct a frequency response signature [9]. The transformer winding

is proven to be equivalent of an electrical network consisting of resistance, capacitance and inductance in high frequency range, and its frequency response signature can represent the status of the winding [10]. The frequency response signature of the transformer after the factory production can be regarded as the mostly original, standard and healthy signature of the transformer, which is frequently called fingerprint. By comparing the frequency response curve of the subsequent transformer in fault status with its fingerprint, the specific fault type and severity of windings could be diagnosed [11].

Frequency response signatures of transformers with the same fault type always present the similar characteristics and patterns. On this basis, after comparing frequency response signatures with the fingerprint, the common characteristics of the certain fault type can be extracted. Based on this principle, it is only need to extract the different characteristics of frequency response signature with different faults; as a result, the types of transformer winding mechanical faults can be diagnosed. The same is true for the diagnosis of fault degree. However, as so far, there is still no standard and reliable code for interpretation of FRA signature [12], [13], in which, the analysis of FRA mainly relies on visual inspection or mathematical calculation. The diagnosis result by visual inspection is easily affected by the subjectivity of personnel, while the winding fault type and degree are not easily recognized by simple mathematical calculation.

Recently, artificial intelligence (AI) has developed as an advanced technique which was successfully used in many fields. Some relevant researches on the application of FRA and AI algorithm in transformer fault diagnosis have been successively conducted. For instance, A. J. Ghanizadeh and G. B. Gharehpetian trained the neural network classifiers by processing the FRA data [14]. Bigdeli used SVM to diagnose transformer winding fault on the basis of transfer function (TF) [15]. Zhao combined SVM with impulse frequency response analysis (IFRA) to diagnose transformer fault [16]. Deng classified the transformer winding deformation based on SVM and finite element analysis (FEA) [17]. However, in reference [14], the neural network is not fit for the small sample data of transformer winding faults, sometimes it is easy to fall into the local optimal solution, which leads to premature maturity. What's more, the required data for training and testing are obtained by the simulation in this study, instead of the FRA data measured in the actual transformers. In reference [15], only a few feature quantities have been extracted and applied to the fault diagnosis, which can be further improved. Reference [16] only diagnoses the fault types, without further classifying and predicting the degree of fault. In reference [17], the identification of winding fault is based on the short circuit impedance method, not the FRA signature; what's more, the data for training are obtained by FEA. The optimization algorithm of SVM parameters can still be further improved.

In view of above background, this study proposes the identification of actual transformer winding deformation faults by combining FRA and SVM. SVM is popular and powerful due to its unique advantages in solving small sample, nonlinear and high-dimensional pattern recognition problems; it has good generalization ability in the case of limited samples. This is of great practical significance for the fault diagnosis and prediction of power transformer, which, the identification of winding deformation faults is always the problem of sample shortage and high nonlinearity between the fault phenomena and fault reasons [18].

In this study, the SVM model is trained by FRA data. The parameters of the SVM model are optimized by the particle swarm optimization (PSO) algorithm. The characteristics of transformer FRA signatures with different fault types and degrees can be obtained by the SVM model, which is used for fault classification. Besides, some characteristic features are introduced by statistical analysis of FRA signatures. Both the fault types and degrees are comprehensively discriminated. In addition, to compare the proposed method with the current state of the art of SVM on discriminating FRA signatures, two common parameter optimization algorithms-the grid search algorithm and genetic algorithm are also discussed in this study.

The rest of this paper is organized as follows. Firstly, the experimental setup and the experimental results are introduced in the Section II. The key tools and procedures used in this study are described in the Section III. The classifying results and analysis are presented in the Section IV. Refer to Section V for the details of conclusions of this study and further research work. All the data tables supporting this article can be seen in Section VI.

II. EXPERIMENTAL SETUP AND RESULT
A. EXPERIMENTAL SETUP
In this paper, a specially manufactured model transformer is adopted to perform all experiments. Detailed parameters of the model transformer are shown in Tab. 1. Three common transformer faults are simulated by experiments, namely, disk space variation (DSV), inter-disk short circuit (SC) and radial deformation (RD). More information about the model transformer and the detailed experimental setup can be found in [16], [19], [20].

The DSV fault has been revealed that the capacitance parameter dominates the effect and this fault could be emulated by changing the inter-disk capacitance parameter [11]. The influence of DSV fault on transformer is simulated by connecting several disks with paralleled capacitors. The magnitude of capacitance indicates the degree of fault, including 50 pF, 67 pF, 100 pF, 200 pF, 400 pF, 600 pF and 800 pF. The influence of SC fault on transformer is simulated by shorting the connectors between the adjacent disks. The transformer manufacturer also produced some windings with variable deformations, which are used to replace the middle 10-disk windings to simulate the RD faults [16]. The image of the winding RD fault is shown in Figure 1. In Figure 1 (a), d represents the amount of RD, which is a variable, θ represents the angle which is fixed at 45°, the ratio of d and the
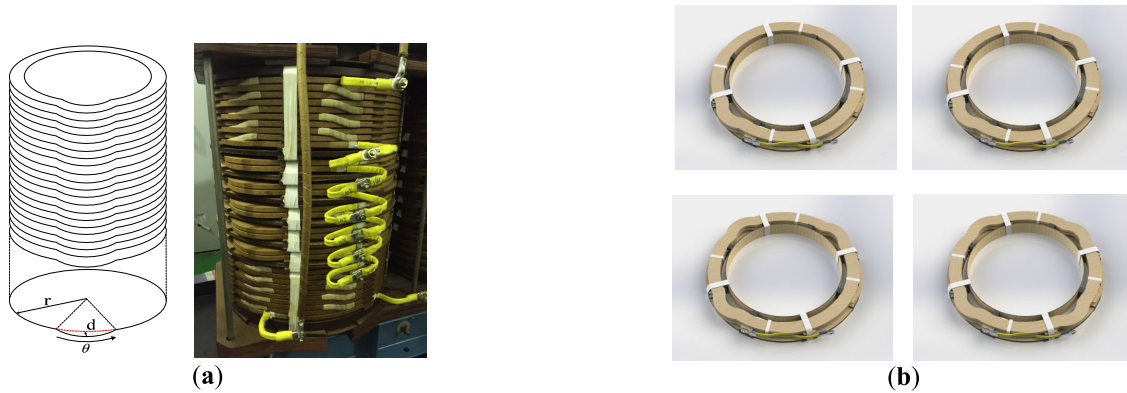
FIGURE 1. Diagrammatic sketch of winding RD fault: (a) diagram and image of RD; (b) 3D visualization of winding RD with faults manufactured at different directions.

TABLE 1. Design specifications of specifically manufactured model transformer.

Parameters		Nominal value
Rated voltage (kV)		10/0.4
Capacity (kVA)		400
Rated current (A)		23/577
Frequency (Hz)		50
Number of phase		3
Connection type		Ynyn0
Tank (mm)		1705*740*1415
HV winding	Outer radius (mm)	250
	Inner radius (mm)	210.5
	Height (mm)	519
LV winding	Outer radius (mm)	174.5
	Inner radius (mm)	158
	Height (mm)	363
Iron core	Yoke radius (mm)	150
	Yoke height (mm)	1190
	Yoke length (mm)	1390

winding radius r are set to be 3%, 5%, 7% and 10% to emulate the different degree of RDs produced at one direction [20]. There are also other RD fault windings in which the faults are manufactured at different directions, but the ratio of d and r is fixed at 5%, as shown in Figure 1(b) [16].

Figure 2 (a) shows the schematic diagram of transformer experiment and Figure 2 (b) indicates the image of measurements of the research.

All fault emulated experiments are performed using FRA end to end open circuit connection. Two sets of FRA data under transformer healthy status are measured and taken as the standard FRA signatures. The reason why two sets of FRA data are chosen as the fingerprints is that the measurement errors were taken into account. In the process of FRA measurement, there are a number of unavoidable factors that can interfere with the measurement results to varying degrees.

Taking two sets of FRA data as the fingerprints can weaken the influence of measurement error on the diagnosis results to a certain extent. The two sets of data were named normal 1 and normal 2.

B. EXPERIMENTAL RESULT

In the case of SC fault, 15 groups of different frequency response data were measured. In the case of DSV fault, 21 groups of different FRA data were obtained. RD fault experiments were carried out with 18 different conditions. Figure 3 $\sim$ 5 show the comparison between the new measured FRA traces and the fingerprint of transformer with different fault status, including variable fault degrees and locations, respectively.

III. APPLICATION OF SVM TO DIAGNOSE WINDING FAULT

A. BRIEF INTRODUCTION TO THE KEY TOOLS

SVM has been in existence for a couple of decades, but it has been developed as a powerful tool since it was firstly proposed by Vapnik in the field of machine learning. The application of SVM in fault diagnosis has attracted an increasing attention in recent years due to its good classification performance. The SVM is based on the statistical learning theory [21]. Generally speaking, it is a two-classification model. Its basic model is described as a linear classifier with the largest spacing in feature space. Namely, the learning strategy of the SVM is to maximize the spacing, which can be ultimately transformed into a convex quadratic programming problem.

SVM is a classification method based on structural risk minimization (SRM) criterion. The learning strategy is to maximize the interval, and the solution of the optimal hyperplane can ultimately be transformed into the solution of a convex quadratic function problem [22]. It has particular advantages of solving the classification problem of small-scale samples, non-linearity, and high-dimension. According to Statistical Learning Theory (SLT), which is the theory of finite sample statistics put forward by Vapnik *et al* [23]–[26],
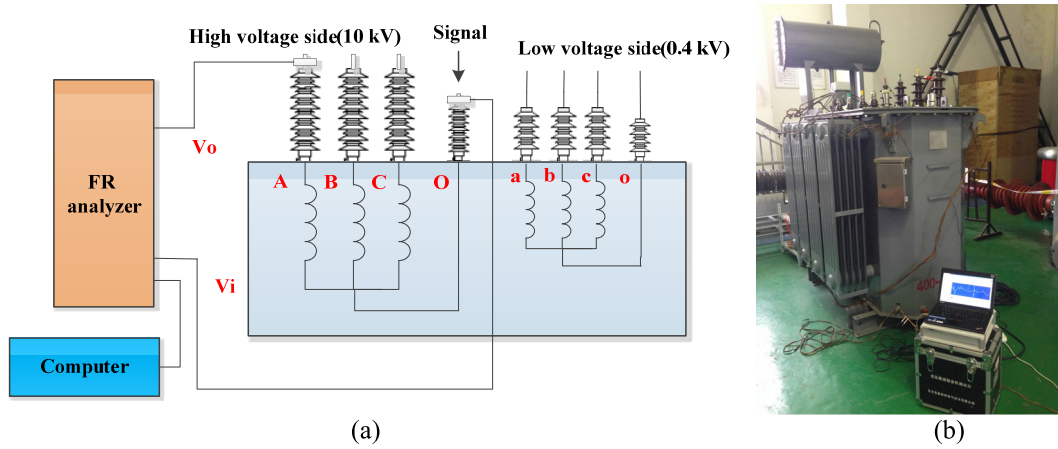
FIGURE 2. Image of measurements of the research: (a) schematic diagram of transformer experiment; (b) image of the testing ground.

FIGURE 3. Visualization of DSV Fault Experiment Data (the solid line is fingerprint and the dotted line is new trace which corresponds to different fault status).

FIGURE 5. Visualization of RD Fault Experiment Data(the solid line is fingerprint and the dotted line is new trace which corresponds to different fault degree).
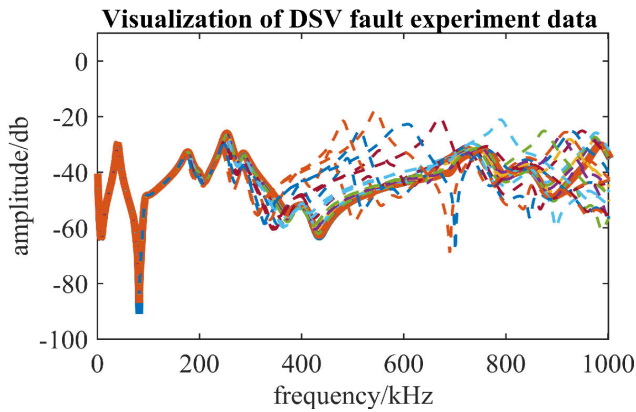

FIGURE 4. Visualization of SC Fault Experiment Data (the solid line is fingerprint and the dotted line is new trace which corresponds to different fault status).
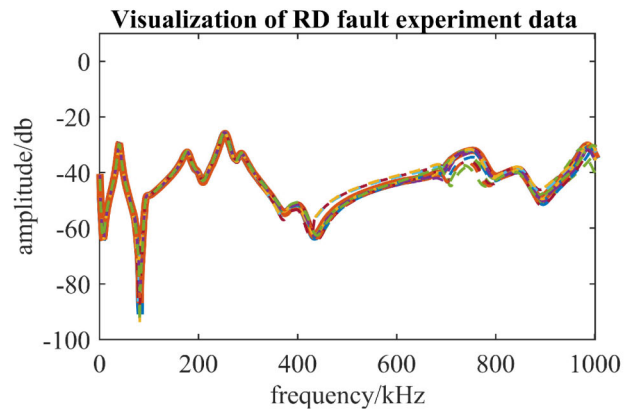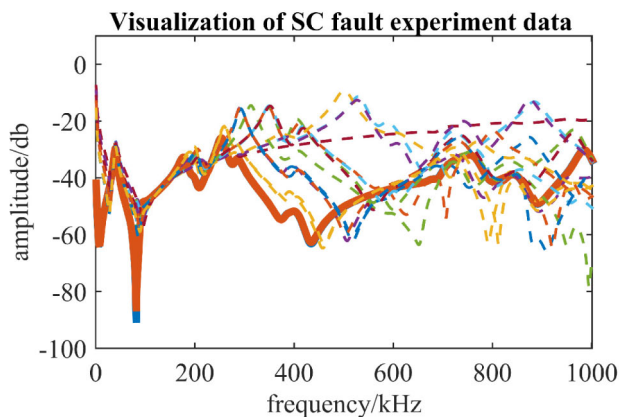
the so-called structural risk minimization is to reduce the Vapnik-Chervonenkis (VC) dimension of the learning machine while ensure the classification accuracy rate (empirical risk), so as to control the expected risk of the learning machine on the whole sample set [25]. The relationship between the expected risk and the empirical risk is shown by:

$$R_{exp} \leq R_{emp} + \Phi\left(\frac{n}{h}\right) \tag{1}$$

where R_{exp} and R_{emp} represent the expected risk and the empirical risk (the classification accuracy rate), and Φ is the confidence interval that is related to the number of samples and the VC dimension, where n is the number of samples and h represent the VC dimension.

The confidence interval Φ in the above formula decreases monotonously with the increase of ratio of n to h. In order to make the model has better generalization ability, it is necessary to reduce the VC dimension while minimizing the empirical error. Define the classification interval as ρ. At the ρ interval, the VC dimension h of the hyperplane sets satisfies the following relationship:

$$h = f\left(\frac{1}{\rho^2}\right) \tag{2}$$

where f is a monotonic increasing function. It can be seen from Equation (2) that h is inversely proportional to the square of ρ. Thus, maximizing the sample interval can minimize the VC dimension to make the generalization ability of the model stronger.

In summary, the principle of SVM is to maximize the interval between two classes of samples in the process of sample classification, so that the training model has strong generalization ability.

The most widely used radial basis function (RBF) is selected as the kernel function of the SVM algorithm in this study. The expression of the function is as follows:

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{g^2}\right) \quad (3)$$

The LibSVM provides complete functions for training and testing of SVM models [27].

The parameters of the SVM model are optimized by the PSO algorithm. The PSO algorithm was firstly proposed by Eberhart and Kennedy in 1995 [28], [29]. Its basic concept stems from the study of foraging behavior of birds. The PSO algorithm emulates the behavioral characteristics of this biological population and is used to solve the optimization problem.

The initialization of the PSO algorithm is a group of random particles, and then the optimal solution is found through multiple iterations. In each iteration, the particle updates itself through the optimal solution it calculates (namely the individual optimum) and the optimal solution currently calculated by the entire population (namely the global optimum). The formula for updating the velocity and position of particles is introduced by Equation (4) and Equation (5).

$$V_i = W * V_i + C_1 * R_1 * (E_{best} - P_i)$$
$$+ C_2 * R_2 * (G_{best} - P_i) \quad (4)$$
$$P_i = P_i + V_i \quad (5)$$

where P_i and V_i are the position and velocity of the i th particle, respectively. C_1 and C_2 are acceleration constants, also known as learning rates. W is an inertia constant, and R_1 and R_2 are random numbers in the range of 0 to 1. E_{best} represents the individual optimum and G_{best} represents the global optimum.

PSO algorithm can be used in many fields because of its superior optimization performance, which plays an important role in swarm intelligence algorithm.

B. STRUCTURE OF DATA SETS

Extracting the representative fault features is an important prerequisite for accurately discriminating the various types of faults. At present, the mainstream feature extraction methods from FRA signatures are divided into two categories, the statistical indicators and FRA signature waveform features [30]. In this study the feature extraction based on statistical indicators is used to quantify the difference between the fingerprint and the measured new trace. Calculating statistical indicators

is easy and fast, and the noise has little influence on the calculation results [30].

The data used for training and testing of SVM model are experimental FRA data of transformer as mentioned above. The following Tab. 2 shows the eight different mathematical features [31]–[33] of the measured data extracted by a series of mathematical operations between the new trace and the fingerprint, the calculate expression of the features is also presented. These mathematical features include the Correlation Coefficient (CC), the Euclidean Distance (ED), the maximum of difference (MAX), the Integral of Absolute difference (IA),

TABLE 2. Extracted mathematical index.

Feature	Definition	Formula
τ_1	Correlation Coefficient	$\dfrac{\sum\limits_{i=1}^{n} X_i Y_i}{\sqrt{\sum\limits_{i=1}^{n}[X_i]^2 \sum\limits_{i=1}^{n}[Y_i]^2}}$
τ_2	Euclidean Distance	$\sqrt{\sum\limits_{i=1}^{n}(Y_i - X_i)^2}$
τ_3	Maximum of difference	$\max(Y_i - X_i)$
τ_4	Sum Squared Error	$\dfrac{1}{n}\sum\limits_{i=1}^{n}(Y_i - X_i)^2$
τ_5	Sum Squared Ratio Error	$\dfrac{1}{n}\sum\limits_{i=1}^{n}(\dfrac{Y_i}{X_i} - 1)^2$
τ_6	Integral of Absolute difference	$\int |Y(f) - X(f)| df$
τ_7	Sum Squared Max-Min Ratio Error	$\dfrac{1}{n}\sum\limits_{i=1}^{n}(\dfrac{\max(Y_i, X_i)}{\min(Y_i, X_i)} - 1)^2$
τ_8	Root Mean Square Error	$\sqrt{\dfrac{1}{n}\sum\limits_{i=1}^{n}\left(\dfrac{|Y_i| - |X_i|}{\dfrac{1}{n}\sum\limits_{i=1}^{n}|X_i|}\right)^2}$

the Sum Squared Error (SSE), the Sum Squared Ratio Error (SSRE), the Sum Squared Max-Min Ratio Error (SSMMRE) and Root Mean Square Error (RMSE).

The reasons why these 8 features are chosen can be seen below. CC describes whether the relationship between two independent variables is close. To a certain extent, it is to describe the similarity. ED is a commonly used definition of distance, which is the actual distance between two points in m-dimensional space. Its geometric meaning is clear. The correlation coefficient and Euclidean Distance are the two frequently used features of FRA signature for pattern recognition. MAX represents the maximum range of data changes. IA determines the gap between the new trace and the fingerprint. Intuitively speaking, it is the sum of the areas enclosed by two curves. The actual error can be reflected by this index. Additionally, SSE and RMSE are similar, which explain the dispersion between finger print and the new trace in different ways, and measure the deviation between them. At last, SSRE and SSMMRE are similar, which highlight the impact of relatively large errors and weaken the impact of relatively small errors. In practical situations, the transformer is most likely diagnosed as a healthy state when the curve is slightly offset. However, when there is a large deviation in the curves, the transformer may be identified as a fault state. The characteristics of SSRE and SSMMRE conform to this common fault diagnosis criterion.

The form of data set is 111-by-8, where 111 represents 111 groups of data and 8 is the feature number. Additionally, it's not difficult to understand that the form of data label matrix is 111-by-1. In other words, each group of data contains eight components. The details are shown in the Tab. 3 and Expression (6).

$$\begin{bmatrix} \tau_{1,1} & \tau_{1,2} & \cdots & \tau_{1,8} \\ \tau_{2,1} & \tau_{2,2} & \cdots & \tau_{2,8} \\ \vdots & \vdots & \ddots & \vdots \\ \tau_{i,1} & \tau_{i,2} & \tau_{i,j} & \tau_{i,8} \\ \vdots & \vdots & \ddots & \vdots \\ \tau_{110,1} & \tau_{110,2} & \cdots & \tau_{110,8} \\ \tau_{111,1} & \tau_{111,2} & \cdots & \tau_{111,8} \end{bmatrix} \quad (6)$$

Expression (6) shows the detail structure of the data set, where $\tau_{i,j}$ represents the j th feature of the i th data group.

When transformer fault types are classified, 1-3 rows of data label matrix were set to 1 to represent health status, and 4-45 rows of data label matrix were set to 2 to represent DSV fault, 46-75 rows of data label matrix were set to 3 to represent SC fault, 76-111 rows of data label matrix were set to 4 to represent RD fault. 4,6,8,10…rows of data set were taken as training set, and the corresponding rows of data label matrix were taken as training label matrix. Accordingly, 5,7,9,11…rows of data set were taken as testing set, and the corresponding rows of data label matrix were taken as testing label matrix.

When classifying the degree of fault, because of the difficulty in quantifying the degree of the fault in the case of

TABLE 3. Structure of data set and details of each group.

The Group Number Of Data Set	The Status Of Winding	Reference Standard Data
1~3	Normal	Normal1,Normal2
4~24	DSV	Normal1
25~45	DSV	Normal2
46~60	SC	Normal1
61~75	SC	Normal2
76~93	RD	Normal1
94~111	RD	Normal2

SC fault, the SVM algorithm is only used to classify the cases of DSV fault and RD fault. In the case of DSV fault, 4-24 rows of data set were taken as the training set and 25-45 rows of data set were taken as the testing set. Data label matrix can be determined by the fault level of corresponding rows. In addition, considering that the experimental FRA data are limited in the case of RD fault, only a few samples were taken randomly as the testing set and the rest as the training set.

C. PROCEDURE OF PROPOSED METHOD
The procedure of the proposed method is depicted in Figure 6. The original data is used to plot the corresponding FRA signatures, and then the features of the curves are extracted as the input of SVM. The input is divided into two parts, one
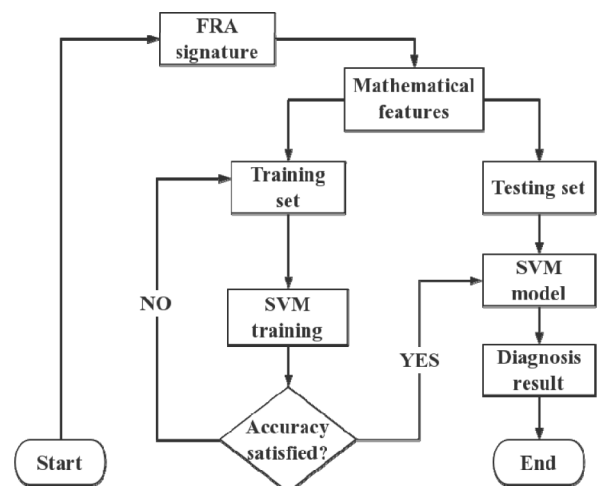

FIGURE 6. Block diagram of the proposed method.

for training and the other for testing. Only when the training accuracy meets the requirement, the model completes the training and is used for testing, and finally the diagnostic results are obtained.

D. RESULT OF PROPOSED METHOD FOR FAULT DIAGNOSIS

After training the SVM model, the model is tested with the testing set. The detail result of model testing is shown in following Tab. 4-10.

TABLE 4. Classification results for DSV.

No.	The Result of Classification
1	True
2	True
3	True
...	...
20	True
21	True
Overall Classification Accuracy Rate For DSV: 100% (21/21).	

TABLE 5. Classification results for SC.

No.	The Result of Classification
1	True
2	True
3	True
...	...
14	True
15	True
Overall Classification Accuracy Rate For SC: 100% (15/15).	

TABLE 6. Classification results for RD.

No.	The Result of Classification
1	True
2	True
3	False
...	...
12	False
...	...
17	True
18	True
Overall Classification Accuracy Rate For RD: 88.89% (16/18).	

When the aim of the classifier is fault types, half of the data set is extracted as a training set and the other half as a testing set. In order to make the SVM model to identify the various fault types, the training set learned by the model contains samples of various fault types, DSV, SC and RD. The

TABLE 7. Summary of diagnostic results about fault types.

Fault Type	Classification Accuracy Rate
DSV	100% (21/21)
SC	100% (15/15)
RD	88.89% (16/18)
Comprehensive Classification Accuracy Rate For Fault Type:96.30%(52/54)	

TABLE 8. Classification results for the level of DSV.

No.	The Actual Level of Fault(pF)	The Result of Classification(pF)
1	50	50
2	67	67
3	100	100
...	...	...
9	67	50
...	...	...
20	600	600
21	800	800
Overall Classification Accuracy Rate For The Level Of DSV: 95.24% (20/21).		

TABLE 9. Classification results for a typical testing set for the level of RD.

No.	The Actual Level of Fault	Diagnostic Result
1	3%	5%
2	5%	5%
3	7%	7%
4	10%	10%
5	5%	5%
The Average Classification Accuracy Rate: 80%(4/5)		

training accuracy of the model is 100% (54/54). The testing results of various fault types are given in Tab. 4-6.

Table 7 is a summary of the testing results from Tab. 4 to Tab. 6. From Table 7, it is clear that the proposed method is capable of discriminating fault types. The comprehensive classification accuracy can reach up to 96.30%. In particular, DSV and SC fault can be identified without error.

When the aim of the classifier is fault level, two groups of experiments were carried out, namely, DSV level classification and RD level classification.

TABLE 10. Processing result of 1000 groups of randomized trials for discriminating RD fault degree in 3 times.

No.	Diagnostic Accuracy Rate
1	68.36%
2	68.6%
3	68.02%
Average	68.33%

When it comes to DSV level classification, similarly, the DSV fault data was divided into two equal parts, one for training and another for testing. The training accuracy of model is 100% (21/21). The result of classifying the levels of DSV faults are presented in Tab. 8.

As shown in Tab. 8, basically, the fault degree of DSV can be accurately identified by the model. It is remarkable that in the only wrong set of data, there is not much difference between what the model predicts (67pF) and what it actually is (50pF). Therefore, the model is capable of detecting the fault degrees of DSV fault.

When classifying the degree of RD fault, testing set and training set are randomly extracted from data set. The following tables are the classification results of a typical training set and its testing set. The diagnostic accuracy rate varies with the selection of testing sets and training sets. Anyway, it is not difficult to conclude that the classification accuracy rate must obey the statistical law. Based on this idea, 1000 groups of randomized trials were conducted for 3 times. In each of the randomized trials, 5 groups of FRA data of RD fault are selected randomly from the data set and taken as the testing set, and the selected testing set cover all of the fault degree. The rest of data are taken as the training set and cover all fault degrees, too. The average training accuracy of the above 3000 random trials equals to 96.79%. The diagnostic accuracy rates of experimental data are listed in Tab. 10. It can be concluded that the diagnostic accuracy rate for discriminating the winding RD fault degree is around 70% by the statistical law. The experimental data are visualized in Figure 7.

E. COMPARISON OF PARAMETER OPTIMIZATION ALGORITHMS BETWEEN PROPOSED METHOD AND CURRENT LITERATURE

The penalty coefficient and kernel function parameters (C, g) of SVM model have a crucial impact on the performance of the model. In other words, the quality of model parameters determines the performance of the model. In this study, except for the PSO algorithms, the other two frequently used algorithms - the grid search and genetic algorithm are also used to optimize the (C, g).

The traditional parameter optimization method in the most of existing literatures is the grid search algorithm. In fact, the grid search algorithm is essentially an enumeration method. In a given interval, many small mesh intervals are
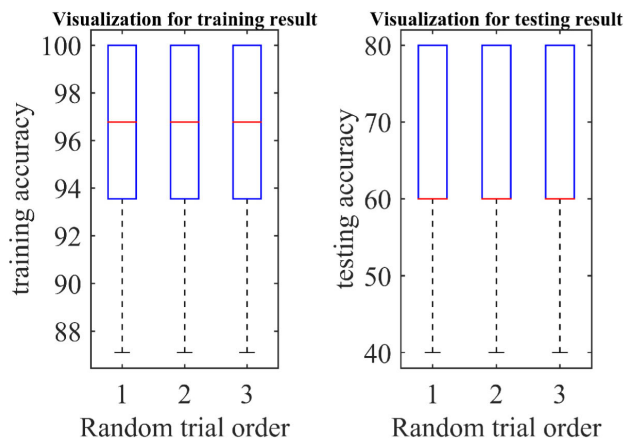

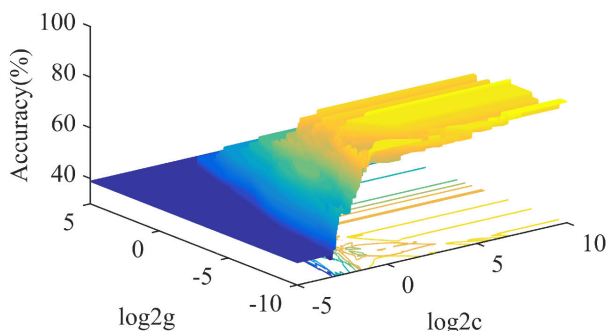**FIGURE 7.** Visualization of randomized trails result.

FIGURE 8. Result diagram of grid search algorithm.

divided according to the given step size. By calculating whether the parameters in each mesh interval can reduce the training error of the model, the optimal parameters can be determined.

In addition, both of genetic algorithm and PSO algorithm are popular at the moment. They are quick and accurate to obtain the optimal parameters with greater probability.

In order to illustrate the differences of these three algorithms in optimizing parameters for the SVM model, taking the classification of fault types as an example, each method is debugged three times to calculate the most appropriate parameters. The debugging results of each optimization algorithm method are shown in the Tab. 11. Similar results were obtained for classification of fault degree.

Figure 8 ～ 10 show the result diagrams of above parameter optimization methods, respectively.

When the grid search algorithm is used to optimize the SVM parameters, the parameters are first searched in a large interval. The search interval and step size are then manually adjusted referring to the results of parameter optimization. Thus, this method is often time-consuming in the early search stage and the human intervention is needed. When genetic algorithm is used to optimize the parameters,

TABLE 11. The differences between three methods of parameter optimization.

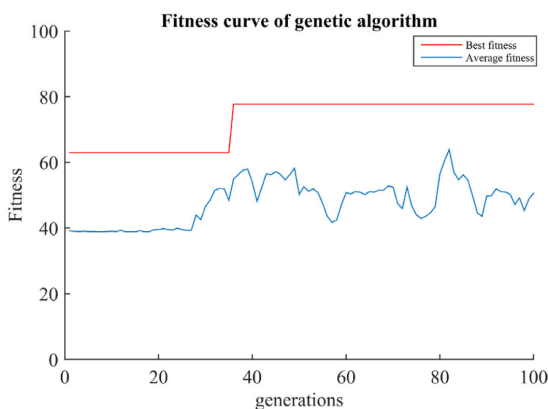| Method of parameter optimization	No.	Debugging Result	
		Elapsed time	Testing accuracy
grid search algorithm	1	77.6195 seconds	92.59%
	2	24.5956 seconds	92.59%
	3	1.1404 seconds	92.59%
genetic algorithm	1	14.5244 seconds	96.30%
	2	14.8636 seconds	92.59%
	3	7.4793 seconds	96.30%
PSO algorithm	1	2.2693 seconds	96.30%
	2	2.4790 seconds	96.30%
	3	2.2284 seconds	96.30%

FIGURE 9. Result diagram of genetic algorithm.

genetic algorithm involves operations such as crossover and mutation, which has strong randomness. Occasionally, the appropriate parameters cannot even be found. In most cases, genetic algorithm can converge to an approximate optimal solution as quickly as possible. When PSO is used to optimize the SVM parameters, the convergence speed of the algorithm is fastest among the three methods, and the quality of the solution is the best.

In summary, among the three methods of model parameter searching algorithms, the traditional grid search algorithm is far inferior to the current popular genetic algorithm and PSO.
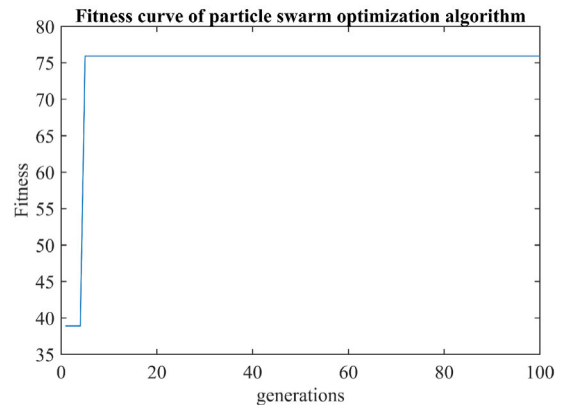

FIGURE 10. Result diagram of PSO algorithm.

Among them, the PSO algorithm is the best, which demonstrates the advantages of the proposed method.

IV. CONCLUSION

In this paper, a method of combining SVM with FRA has been evaluated for discriminating fault types and degrees of transformer deformation faults. A series of experiments were carried out on an actual transformer. Eight features are extracted from FRA raw data by mathematical statistical calculation. FRA Data is used to train SVM model, and then PSO algorithm is used to optimize SVM model parameters. The data processing result shows that the comprehensive accuracy rate of the diagnosis reaches up to 96.3% when the classifier is used to discriminate the fault types of winding DSV, SC and RD. When the trained classifier is used to classify the fault degree of DSV, the accuracy rate reaches up to 95.24%. In addition, the accuracy rate is around 70% when the classifier is applied to discriminate the fault degree of RD. The applicability and effectiveness of the method has been already demonstrated, the comparison result of the proposed method with other two parameter optimized methods demonstrates the advantages of PSO method. Overall, the combination of SVM and FRA has reached significant potential for transformer winding deformation fault diagnosis. In further research, fault location will be analyzed. Phase information of frequency response can also be considered for fault diagnosis. It is also practical and meaningful to use massive data of FRA to solve the actual fault diagnosis problems when more reliable SVM models are trained and obtained.

APPENDIX
See Tables 1–11.

REFERENCES

[1] C. Jettanasen and A. Ngaopitakkul, "A novel probabilistic neural network-based algorithm for classifying internal fault in transformer windings," *Ieej Trans. Electr. Electron. Eng.*, vol. 8, no. 2, pp. 123–131, Mar. 2013.

[2] F. Yuan, J. Guo, Z. Xiao, B. Zeng, W. Zhu, and S. Huang, "A transformer fault diagnosis model based on chemical reaction optimization and twin support vector machine," *Energies*, vol. 12, no. 5, p. 960, Mar. 2019.

[3] B. Lee, J.-W. Park, P. A. Crossley, and Y. C. Kang, "Induced voltages ratio-based algorithm for fault detection, and faulted phase and winding identification of a three-winding power transformer," *Energies*, vol. 7, no. 9, pp. 6031–6049, Sep. 2014.

[4] M. Gutten, D. Korenciak, M. Kucera, R. Janura, A. Glowacz, and E. Kantoch, "Frequency and time fault diagnosis methods of power transformers," *Meas. Sci. Rev.*, vol. 18, no. 4, pp. 162–167, Aug. 2018.

[5] H. Kaixing and P. Zaiping, "Vibration model of power transformer under short-circuit condition," in *Proc. Int. Conf. Elect. Mach. Syst.*, Nov. 2009, pp. 1–5.

[6] M. Bagheri, M. S. Naderi, and T. Blackburn, "Advanced transformer winding deformation diagnosis: Moving from off-line to on-line," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 19, no. 6, pp. 1860–1870, Dec. 2012.

[7] W. Lech and L. Tyminski, "Detecting transformer winding damage-the low voltage impulse method," *Elect. Rev.*, vol. 179, no. 21, pp. 768–772, Jan. 1966.

[8] E. P. Dick and C. C. Erven, "Transformer diagnostic testing by frequency response analysis," *IEEE Trans. Power App. Syst.*, vol. PAS-97, no. 6, pp. 2144–2153, Nov. 2007.

[9] S. Alsuhaibani, Y. Khan, A. Beroual, and N. H. Malik, "A review of frequency response analysis methods for power transformer diagnostics," *Energies*, vol. 9, no. 11, p. 879, Oct. 2016.

[10] W. H. Portilla, G. A. Mayor, J. P. Guerra, and C. Gonzalez-Garcia, "Detection of transformer faults using frequency-response traces in the low-frequency bandwidth," *IEEE Trans. Ind. Electron.*, vol. 61, no. 9, pp. 4971–4978, Sep. 2014.

[11] A. Abu-Siada, N. Hashemnia, S. Islam, and M. A. S. Masoum, "Understanding power transformer frequency response analysis signatures," *IEEE Elect. Insul. Mag.*, vol. 29, no. 3, pp. 48–56, May 2013.

[12] N. Hashemnia, A. Abu-Siada, and S. Islam, "Improved power transformer winding fault detection using FRA diagnostics—Part 1: Axial displacement simulation," *IEEE Trans. Dielectr. Elect. Insul.*, vol. 22, no. 1, pp. 556–563, Feb. 2015.

[13] S. M. Saleh, O. Gouda, and S. H. El-Hoshy, "Proposed diagnostic methodology using the cross-correlation coefficient factor technique for power transformer fault identification," *IET Electr. Power Appl.*, vol. 11, no. 3, pp. 412–422, Mar. 2017.

[14] A. J. Ghanizadeh and G. B. Gharehpetian, "ANN and cross-correlation based features for discrimination between electrical and mechanical defects and their localization in transformer winding," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 21, no. 5, pp. 2374–2382, Oct. 2014.

[15] M. Bigdeli, M. Vakilian, and E. Rahimpour, "Transformer winding faults classification based on transfer function analysis by support vector machine," *IET Electr. Power Appl.*, vol. 6, no. 5, pp. 268–276, May 2012.

[16] Z. Y. Zhao, C. Tang, Q. Zhou, L. N. Xu, Y. G. Gui, and C. G. Yao, "Identification of power transformer winding mechanical fault types based on online IFRA by support vector machine," *Energies*, vol. 10, no. 12, p. 2022, Nov. 2017.

[17] D. E. Xiangli, X. I. Haiyuan, and X. I. Xiaofu, "Classification method of transformer winding deformation based on SVM and finite element analysis," (in Chinese), *Proc. CSEE*, vol. 35, no. 22, pp. 5778–5786, 2015.

[18] L. Sima, "Research on methods of fault diagnosis and prognosis for power transformer based on improved support vector machines," (in Chinese), Ph.D. dissertation, School Elect. Eng., WuHan Univ., WuHan, China, 2012.

[19] Z. Y. Zhao, C. Tang, C. Yao, Q. Zhou, L. Xu, Y. Gui, and S. Islam, "Improved method to obtain the online impulse frequency response signature of a power transformer by multi scale complex CWT," *IEEE Access*, vol. 6, pp. 48934–48945, 2018.

[20] Z. Zhao, C. Yao, C. Li, and S. Islam, "Detection of power transformer winding deformation using improved FRA based on binary morphology and extreme point variation," *IEEE Trans. Ind. Electron.*, vol. 65, no. 4, pp. 3509–3519, Apr. 2018.

[21] K. Hong, L. Wang, and S. Xu, "A variational mode decomposition approach for degradation assessment of power transformer windings," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 4, pp. 1221–1229, Apr. 2019.

[22] A. Christmann and I. Steinwart, "Support vector machines," *Inf. Sci. Statist.*, vol. 158, no. 18, pp. 1–28, 2008.

[23] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proc. 5th Annu. Workshop Comput. Learn. Theory*, vol. 5, Jul. 2008, pp. 144–152.

[24] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[25] V. N. Vapnik, *Statistical Learning Theory*. New York, NY, USA: Wiley, 2009.

[26] V. Blanz, B. Schölkopf, H. Bülthoff, C. Burges, V. Vapnik, and T. Vetter, "Comparison of view-based object recognition algorithms using realistic 3D models," in *Proc. Int. Conf. Artif. Neural Netw.*, vol. 1112, 1996, pp. 251–256.

[27] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, Apr. 2011, Art. no. 27.

[28] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proc. 6th Int. Symp. Micro Mach. Hum. Sci.*, Oct. 1995, pp. 39–43.

[29] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. Int. Conf. Neural Netw. (ICNN)*, vol. 4, 1995, pp. 1942–1948.

[30] H. Lin, "Transformer fault diagnosis algorithm based on frequency response data," Ph.D. dissertation, South China Univ. Technol., Guangzhou, China 2015.

[31] M. H. Samimi and S. Tenbohlen, "FRA interpretation using numerical indices: State-of-the-art," *Int. J. Electr. Power Energy Syst.*, vol. 89, pp. 115–125, Jul. 2017.

[32] N.-K. Wesley, S. Bhandari, A. Subramaniam, M. Bagheri, and S. K. Panda, "Evaluation of statistical interpretation methods for frequency response analysis based winding fault detection of transformers," in *Proc. IEEE Int. Conf. Sustain. Energy Technol. (ICSET)*, Nov. 2017, pp. 36–41.

[33] E. Rahimpour, M. Jabbari, and S. Tenbohlen, "Mathematical comparison methods to assess transfer functions of transformers to detect different types of mechanical faults," *IEEE Trans. Power Del.*, vol. 25, no. 4, pp. 2544–2555, Oct. 2010.

JIANGNAN LIU was born in Ganzhou, Jiangxi, China, in 1999. He is currently pursuing the bachelor's degree with the Department of Electrical Engineering, College of Engineering and Technology, Southwest University, Chongqing, China. His areas of research include condition monitoring and fault diagnosing for power transformer, and application of artificial intelligence.

ZHONGYONG ZHAO was born in Guangyuan, Sichuan, China. He received the B.Sc. and Ph.D. degrees from Chongqing University, Chongqing, China, in 2011 and 2017, respectively, all in electrical engineering. He received the scholarship from the China Scholarship Council to enable him to attend a joint-training Ph.D. program with Curtin University, Perth, Western Australia, from 2015 to 2016. He is currently an Associate Professor with the College of Engineering and Technology, Southwest University, Chongqing, China. His areas of research include condition monitoring and fault diagnosing for HV apparatus, and pulsed power technology.
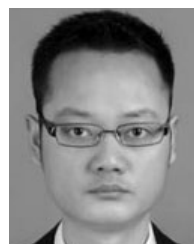
CHAO TANG was born in Sichuan, China, in 1981. He received the M.S. and Ph.D. degrees in electrical engineering from Chongqing University, China, in 2007 and 2010, respectively. As a Ph.D. student (2008–2009), and as a Visiting Scholar (2013–2013) and (2015–2016), he studied with the Tony Davies High Voltage Laboratory, University of Southampton, U.K., doing some researches on the dielectric response characteristics and space charge behaviors of oil-paper insulation. He is currently a Professor with the College of Engineering Technology, Southwest University, China. His research interests include mainly in the field of on-line monitoring of insulation conditions and fault diagnosis for high-voltage equipment.

CHENGUO YAO (M'08) was born in Nanchong, Sichuan, China. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from Chongqing University, Chongqing, China, in 1997, 2000, and 2003, respectively. He became a Professor with the School of Electrical Engineering, Chongqing University, in 2007. He has been a Visiting Scholar with the Old Dominion University, Norfolk, VA, USA, from 2017 to 2018. His research interests include online monitoring of insulation condition and insulation fault diagnosis for HV apparatus, pulsed power technology, and its application in biomedical engineering.

CHENGXIANG LI was born in Shandong, China. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from Chongqing University, Chongqing, China, in 2002, 2005, and 2011, respectively, where he is currently an Associate Professor with the School of Electrical Engineering. His areas of research include pulse power technology and its application in biomedical engineering, and online monitoring of insulation condition for HV apparatus.

SYED ISLAM received the B.Sc. degree in electrical engineering from the Bangladesh University of Engineering and Technology, Bangladesh, in 1979, and the M.Sc. and Ph.D. degrees in electrical power engineering from the King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia, in 1983 and 1988, respectively.

He is currently the Dean of the School of Science Engineering and Information Technology, Federation University, Australia. He received the Dean's medallion for research at Curtin University, in 1999. He received the IEEE T Burke Haye's Faculty Recognition Award, in 2000. He received the Curtin University Inaugural Award for Research Development, in 2012. He received the Sir John Madsen medal for Best Electrical Engineering Paper in Australia, in 2011 and 2014, respectively. He has published over 300 technical papers in his area of expertise. His research interests include in condition monitoring of transformers, wind energy conversion, and smart power systems. He has been a keynote speaker and invited speaker at many international workshops and conferences. He has been a Visiting Professor with the Shanghai University of Electrical Power, China.

Prof. Islam is a member of the Steering Committee of the Australian Power Institute and a member of the WA EESA Board. He is a Fellow of the Engineers Australia, a Fellow of the IET, and a Chartered Engineer in the United Kingdom. He is a founding Editor of the IEEE Transactions on Sustainable Energy and an Associate Editor of the *IET Renewable Power Generation*. He was the Guest Editor-in-Chief for the IEEE Transactions on Sustainable Energy Special Issue on Variable Power Generation Integration into Grid.

• • •