Online Fault Diagnosis for Single-Phase PWM Rectifier Using Data-Driven Method

Kun ZHANG, Bin GOU, and Xiaoyun FENG

Abstract—**In power electronic traction transformer, the failure of the single-phase PWM rectifier will lead to performance degradation of the system. Thus, a feasible data-driven method is proposed to realize online fault diagnosis of single-phase PWM rectifier in this paper. The principle of the data-driven method is to construct a signal predictor based on historic database utilizing the nonlinear autoregressive exogenous (NARX) model with a randomized learning algorithm named extreme learning machine (ELM). Besides, the ensemble method is employed to improve the prediction accuracy and robustness against load fluctuation. In online diagnosis, the predictor and sensor operate simultaneously and their residual is generated. Afterward, the fault detection is conducted by comparing the residual with fault threshold and the fault classification is completed based on system fault symptoms and fault residuals analysis. Several hardware-in-loop tests are implemented to verify the applicability of the proposed diagnosis method. Test results show that this data-driven method is effective to perform the online fault diagnosis with fast fault detection speed and high classification accuracy, and robust against load fluctuation.**

Index Terms—**Ensemble learning, extreme learning machine (ELM), nonlinear autoregressive exogenous (NARX), online fault diagnosis, single-phase PWM rectifier.**

I. Introduction

As the high-speed traction system is developing towards the targets of high efficiency, energy-saving and light weight, power electronic traction transformer (PETT) is considered as a superior choice compared to traditional power frequency transformer [1]. However, due to the complex cascade structure and the fragility of the large number of power components and sensors, the internal part and subunit of PETT are prone to come across malfunction [2]. As shown in Fig. 1, as the basic unit of the input stage of PETT, single-phase PWM rectifier plays a vital role in the operation of PETT. The failure of single-phase PWM rectifier, including sensor faults and IGBT open-circuit faults, may result in a severe deterioration or even shutdown of the whole system. Therefore, it is urgently essential to investigate in fault online diagnosis for the single-phase PWM rectifier.

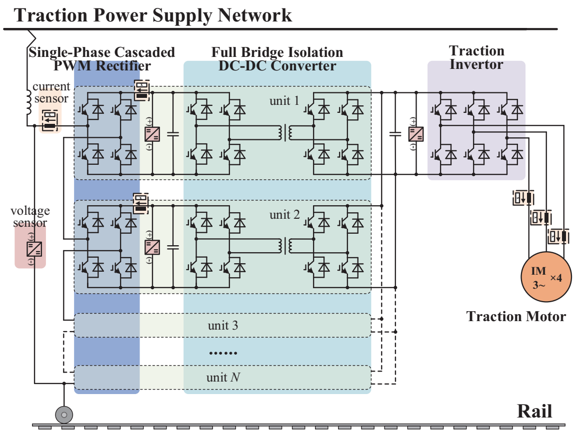


Fig. 1. Structure of PETT.

Generally, fault diagnosis methods can be categorized into the classes of model-based, signal-based, and knowledge-based methods [3]. For model-based methods, the main idea is to build an analytical model describing the relationship among system variables and indicating the system status, based on which the system signals can be estimated, and then fault diagnosis schemes are developed and implemented. The existing fault diagnosis methods for single-phase PWM rectifier are mostly model-based methods [4]–[7]. [4] proposes a fast model-based method for open-switch fault diagnosis of the single-phase PWM rectifier, based on the mixed logical dynamic (MLD) model and residual generation. Also based on MLD model, [5] utilizes the change rate of current residual for different IGBT faults to realize the fault detection of single-phase cascaded PWM rectifier. [6] proposes a method to determine an unknown input observer with minimum sensitivity to disturbance and maximum to the faults when achieving the fault detection for single-phase PWM rectifier. These model-based methods have salient advantage in diagnosis speed but suffer from parameter uncertainties and difficulty of complex system modeling. On the other hand, signal-based methods do not need the mathematical system model but the signal symptoms and features extracted by signal processing techniques [8]–[10]. The fault diagnosis is then executed by checking the signal patterns and features. This method has high diagnostic accuracy but requires large computational quantities or even additional hardware and performs poorly in the case of loads fluctuating.

Manuscript received December 30, 2021; revised February 18, 2022; accepted March 2, 2022. Date of publication March 30, 2022; date of current version March 25, 2022. This paper was presented in part at the 1st IEEE International Power Electronics and Application Symposium (PEAS), Shanghai, China, November 2021. *(Corresponding author: Kun Zhang.)*

All authors are with the School of Electrical Engineering, Southwest Jiaotong University, Chengdu 611731, China (e-mail: kunz@my.swjtu.edu.cn).

Digital Object Identifier 10.24295/CPSSTPEA.2022.00005

The knowledge-based method is also known as the data-driven method because what it needs is not a priori system model or signal symptom, but a historic database, which is more available in complicated industrial processes. In data-driven methods, the principle is to apply intelligent algorithms to extract the underlying knowledge (i.e. the mapping relationship among system variables) from a historic database, enabling the data-driven approach to solve problems with unknown system structure and parameters or complicated models [11]–[14].

In many research and cases, the diagnostic problems are mostly formulated as pattern recognition problems and data-driven method is considered as a classification tool [3], which suffers from a large computational burden, unbalanced fault dataset, and unfavorable diagnosis speed, making related on-line applications more difficult to realize. To overcome such challenges and realize the online fault diagnosis with fast diagnostic speed, a data-driven signal predictor is designed in this paper, the principle of which is derived from the model-based method. But unlike the model-based approach, the data-driven predictor is constructed based on historic database rather than the system mathematical model.

First, to establish the mapping relationship among system variables including grid voltage and current, DC-link voltage and switch command signals, a random learning algorithm named ELM is selected, which has fast convergence speed, favorable generalization ability, and computationally efficient mechanism [15]. Besides, a useful data structure based on nonlinear autoregressive exogenous (NARX) is utilized to improve the signal prediction performance [16]. Additionally, the ensemble learning method is employed to fortify the robustness against loads fluctuating. As for diagnosis decision-making, a hybrid classifier based on the fault residual symptoms combined with the fault performance of system variables is designed. The data-driven predictor and the physical sensor work simultaneously and their outputs are compared in real-time, generating the current residual, based on which the fault detection is conducted online. Furthermore, the specific type of fault is identified based on the fault system symptoms and fault residual analysis.

The rest of this paper is organized as follows: Section II briefly describes the system of single-phase PWM rectifier. Section III introduces the general framework of the proposed method. Section IV gives the fault analysis of the rectifier and shows the fault diagnosis decision-making mechanism. Section V presents the experimental results and analysis, and finally, a general conclusion of this paper is given in Section VI.

II. System Description

A typical single-phase PWM rectifier is shown as Fig. 2, where u_N and i_N are the grid voltage and current, L_N and R_N are the grid leakage inductance and resistance, L_s and C_s are the series resonant circuit inductance and capacitance, C_d is the dc-link capacitance. R_L is the load resistance. u_d and i_L refer to dc-link voltage and current.

For data-driven methods, the feature vector and target vector are significant for the methodology performance. To choose suitable feature and target vectors of the data-driven predictor and establish their mapping relationship, the dependence of system variables of single-phase PWM rectifier should be confirmed. According to the research of [17], the mathematical state-space model of single-phase PWM rectifier can be constructed as:

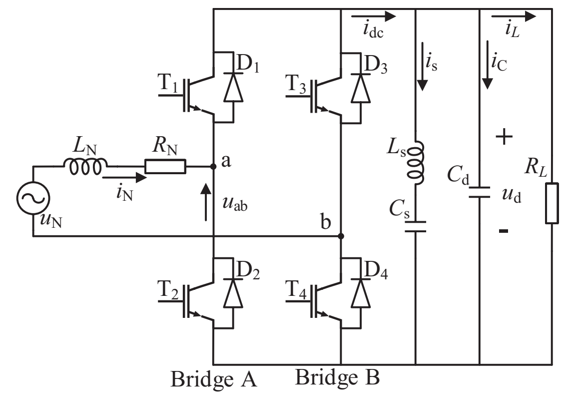


Fig. 2. Topology of single-phase PWM rectifier.

$$\dot{i}_N = -\frac{R_N}{L_N} i_N + \frac{1}{L_N} u_N - \frac{1}{L_N}(S_a - S_b) u_d \tag{1}$$

where S_a and S_b are the switch gate signal function of two bridges, as shown in the following:

$$S = \begin{cases} 1, & \text{the upper switch conducts} \\ -1, & \text{the lower switch conducts} \end{cases} \tag{2}$$

According to the mathematical model, the system variable grid-side current i_N is related to grid-side voltage u_N, dc-link voltage u_d, and switch command signals s_1–s_4. Therefore, i_N is selected as the target vector. u_N, u_d, and s_1–s_4 are chosen as the feature vectors. The principle of the data-driven method is to learn from the historic database of the rectifier and build the nonlinear mapping relationship between the feature vectors and target vector using an intelligent algorithm. After that, a predictor is built to emulate the value of grid-side current i_N utilizing the real-time sampling signals.

III. Proposed Methodology

A. General Framework

The general framework of the proposed data-driven diagnosis method can be briefly illustrated as Fig. 3, the whole process is divided into offline stage and online stage. At the offline training stage, the sample data of the rectifier working on healthy state, including grid-side voltage u_N, dc-link voltage u_d, switch command signals s_1–s_4 and grid-side current i_N, are collected to build a historic database for training. Then the historic time-series data is constructed as NARX structure and ensemble ELM algorithms are applied to learn the mapping relationship f between the feature vectors and the desired

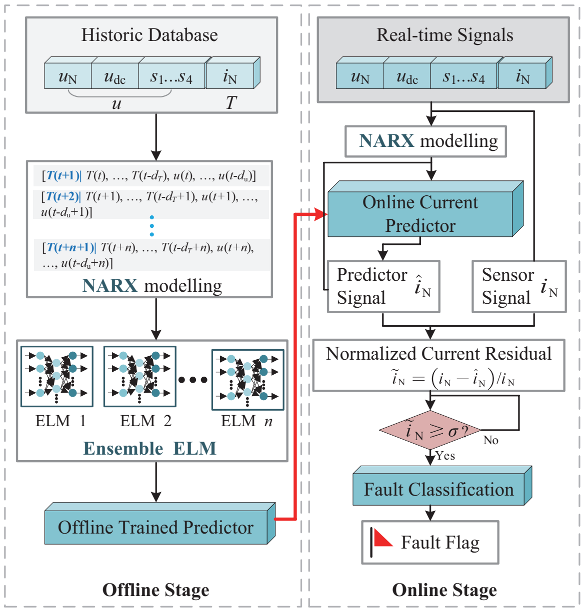


Fig. 3. Overall flowchart of proposed method.

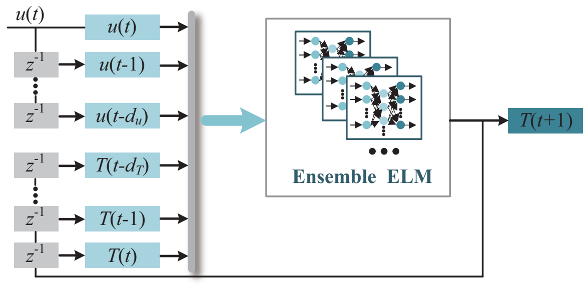


Fig. 4. NARX structure.

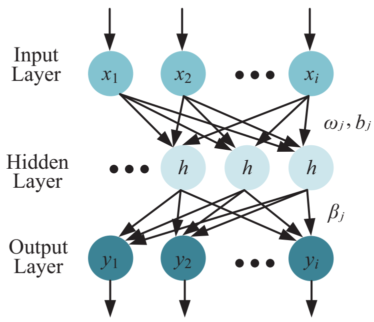


Fig. 5. ELM structure.

target. After the offline learning, the predictive model, which can be regarded as a digital emulator indicating the healthy operational mode of the rectifier, is built for the online process. At the online stage, the real-time sampled data of the rectifier reshaped as NARX structure are delivered to the current predictor. The predictor and the physical grid-side current sensor are operating simultaneously and generating the residual in real-time. Once the residual value overtakes the threshold, the fault detection flag will have a change and thus the fault is detected. Furthermore, based on the fault current residual and various performance of the rectifier when different faults occur, the specific type of the fault can be recognized in detail by the hybrid classifier and a fault classification flag will be output.

B. Current Prediction Based on NARX Structure

NARX is a resultful structure for modeling the nonlinear dynamic system and dealing with time-series data, which can realize an effective prediction by its autoregressive topology, as shown in Fig. 4. NARX can be mathematically formulated as:

$$T(t+1) = f[T(t),...,T(t-d_T), u(t),..., u(t-d_u)] \quad (3)$$

where T, u are the target and feature vectors, d_T, d_u refer to a time-delayed step. Generally, the nonlinear mapping f is established by the intelligent algorithm.

Based on the NARX model merged with the rectifier mathematical expression, the predictive function for i_N can be formed as:

$$i_N = f(i_{NH}, u_N, u_d, s) \quad (4)$$

where i_{NH} is the past grid-side current value, s is the switch command signal vector. By integrating the function (4) with (3), the predictor structure can be represented as follows:

$$\hat{i}_N(t+1) = f\begin{bmatrix} [i_N(t),\cdots,i_N(t-d_T),u_N(t),\cdots,u_N(t-d_u)] \\ u_d(t),\cdots,u_d(t-d_u),s(t),\cdots,s(t-d_u) \end{bmatrix} \quad (5)$$

Based on this NARX model, the predicted value of grid-side current is calculated iteratively every sampling interval and the prediction is processed accurately and effectively.

In [18], it has been demonstrated that the NARX model has not only a powerful computing capacity as good as a fully connected RNN but also a better performance in terms of learning capability and convergence speed, which is desirable in online applications.

C. Extreme Learning Machine

According to (3) and (5), to realize the signal prediction based on the NARX model, it is essential to select a suitable learning algorithm to extract the mapping relationship between the inputs and the output. Here ELM learning algorithm gets the nod for its unique merits of fast learning speed and favorable generalization capacity. Proposed in [15] by Huang, ELM is an advanced algorithm for dealing with the single hidden layer feedforward neural network. Different from conventional algorithms, the input weights and bias of the network in ELM are randomly produced, with the output weights calculated by Moore-Penrose pseudoinverse.

A typical structure of the ELM network is constructed as Fig. 5. For a dataset comprising N' instances (x_i, y_i), where $x_i \in R^n$ and $y_i \in R^m$, the mapping relationship between the input layer x and the output layer y of the ELM network can be mathematically formulated as:

$$y_i = \sum_{j=1}^{N} \beta_j \cdot h(\omega_j \cdot x_i + b_j), \ i = 1, 2, \cdots, N' \quad (6)$$

where ω_j is the input layer weight vector connected to j^{th} node of the hidden layer, h is the activation function, β_j is the hidden layer weight connected to the output layer, b_j and N refer to the bias vector and the number of hidden nodes. And (6) can be simplified as:

$$Y = H\beta \tag{7}$$

where H refers to the output matrix of the hidden layer, which converts the input space of J-dimension to N-dimension output space. In the learning process of the algorithm, the input weight ω and bias b are generated stochastically, while β is calculated by Moore-Penrose inverse shown as follows:

$$\beta = H^{\dagger}Y \tag{8}$$

$$H^{\dagger} = H^{T}(HH^{T})^{-1} \tag{9}$$

D. Ensemble ELM

As a randomized learning algorithm, ELM enjoys the advantages of fast learning speed and favorable generalization capacity but suffers from undesirable robustness against disturbance, which means that although ELM is suitable for online fault diagnosis in single-phase PWM rectifier, it performs badly once the load of the rectifier fluctuates suddenly. To address this problem, an effective approach named ensemble learning is adopted, the critical idea of which is to generate a group of individual learners as an ensemble and deliver the output with a combination strategy.

Firstly ELM is selected as the base learning algorithm, and a set of individual learners are produced based on ELM, denoted as $\{h_{\mathrm{E}1}, h_{\mathrm{E}2}, ..., h_{\mathrm{E}p}\}$, p is the number of learners. Then the training dataset, which is comprised of time-series data $D\{(x_1, y_1), (x_2, y_2), ..., (x_t, y_t), ...\}$, is imported to every single ELM. Afterward, the outputs of every single ELM are combined by a combination strategy and thus final output is generated. Generally, averaging is employed as the combination strategy in the regression task. Moreover, due to the individual learners are all based on the same learning algorithm ELM, simple averaging is adopted, which can be formulated as follows:

$$\overline{T} = \text{mean}\sum_{i=1}^{p} h_{\mathrm{E}i}[D(x_t, y_t)] \tag{10}$$

where $\overline{T}$ is the final output of the ensemble learner. To summarize, the ensemble rule can be formed as follows:

ELM Ensemble Rule

Input : Individual Learners, h_{E};
Training Dataset, $D = \{(x, y)\}$;
Number of Learners, N;
Process:
1: For $i = 1, 2, \dots, p$
2: $T_i = h_{\mathrm{E}i}[D(x_t, y_t)]$
3: End For
Output : $\overline{T}(t) = \text{mean}\,(\sum_{i=1}^{p} h_{\mathrm{E}i})$

It is worth mentioning that this ensemble rule is conducted at every iteration step of the prediction process, as shown in Fig. 3. Integrating (3) and (10), the NARX structure with ensemble learning can be formulated as follows:

$$\overline{T}(t+1) = \text{mean}\sum_{i=1}^{p} h_{\mathrm{E}i}\Big(\overline{T}(t), ..., \overline{T}(t-d_T), u(t), ..., u(t-d_u)\Big) \tag{11}$$

The integrating of individual learners can bring great benefits in two aspects [19]. From the perspective of statistics, since the hypothesis space of learning tasks is generally large, there may be multiple hypotheses that achieve the same performance on the training set, which may cause the case that the single learner with poor generalization capacity is put to use. But by virtue of the ensemble approach, this risk can be decreased. From another point of the presentation, the true hypothesis of some learning tasks may stay beyond the hypothesis space considered by the adopted learning algorithm. However, the hypothesis space can be broadened by integrating numerous learners, and a better approximation can be achieved in this way. Benefiting from such advantages, the ensemble learner has better performance and superior generalization capacity than the single learner, and the robustness of the proposed predictor can be improved.

IV. Fault Analysis and Decision-Making Mechanism

For single-phase PWM rectifier, the switch IGBTs are easily to come to open-circuit fault because of high power stresses, aging, and unpredicted operational condition. Besides, the sensors installed to accomplish the closed-loop control strategy are also prone to malfunction due to device aging and surrounding interference. Therefore, regarding to IGBT open-circuit fault and sensor fault, the fault analysis is given and the decision-making mechanism of fault diagnosis is designed based on a hybrid classifier in this section.

A. IGBT Open-Circuit Fault Analysis

The direction of current flowing into the rectifier is defined as the reference direction. When the open-circuit fault occurs in switch T_1, the positive current keeps unchanged as there exists another available path with freewheeling diode D_1, as shown in Fig. 6(a), where the yellow arrow is the current that keeps unchanged and the red arrow is the current that will change. But the negative current flowing through T_1 originally has to change its path to the bottom part of bridge A by D_2, which will lead to the reduction of both the negative part of i_{N} and the amplitude of u_{d}. And the extent of the reduction is related to the load R_L. The topologies of other switch faults are shown in Fig. 6(b), (c), and (d) and the fault analysis is the same again.

As the system performance of T_4 open-circuit fault is the same as T_1, they are identified as one type of IGBT open-circuit fault. Similarly, the open-circuit fault of T_2 and T_3 are treated as another type.

B. Sensor Fault Analysis

Based on the characteristics and the performance of various

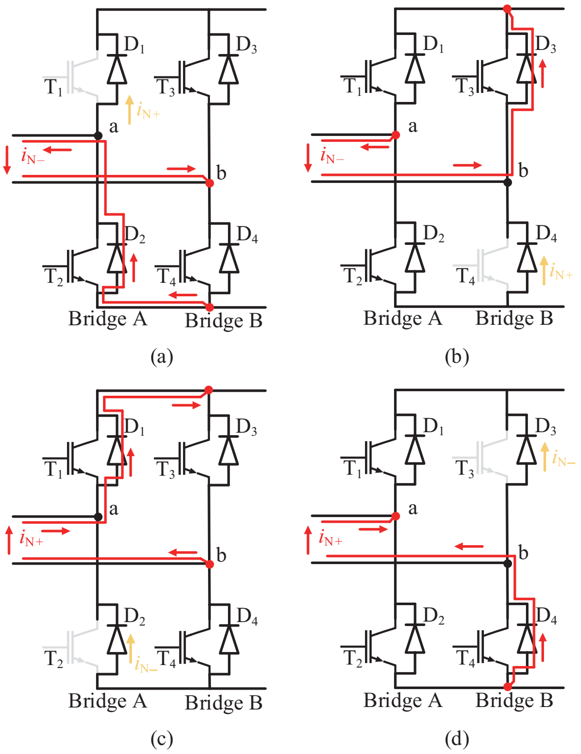


Fig. 6. Topologies of IGBT switch open-circuit fault in single-phase PWM rectifier. (a) T_1 fails, (b) T_4 fails, (c) T_2 fails, (D) T_3 fails.

faults, sensor faults are generally categorized into four types: gain fault, drift fault, stuck fault, and noise fault, which can be formulated as follows:

$$i(t) = \begin{cases} d(t), & 0 \leqslant t \leqslant t_{\mathrm{F}} \\ A \cdot d(t), & t \geqslant t_{\mathrm{F}} \end{cases} \tag{12}$$

$$i(t) = \begin{cases} d(t), & 0 \leqslant t \leqslant t_{\mathrm{F}} \\ d(t) + K_1, & t \geqslant t_{\mathrm{F}} \end{cases} \tag{13}$$

$$i(t) = \begin{cases} d(t), & 0 \leqslant t \leqslant t_{\mathrm{F}} \\ K_2, & t \geqslant t_{\mathrm{F}} \end{cases} \tag{14}$$

$$i(t) = \begin{cases} d(t), & 0 \leqslant t \leqslant t_{\mathrm{F}} \\ d(t) + n(t), & t \geqslant t_{\mathrm{F}} \end{cases} \tag{15}$$

where $i(t)$ is the real output of the sensor, $d(t)$ is the desired output in normal operation, K_1, K_2 and A are constant values. $n(t)$ is a time sequence of Gaussian white noise. t_{F} refers to the time when the fault occurs. According to the above-mentioned analysis, the faults of single-phase PWM rectifier can be classified into six types, as shown in Table I.

C. Fault Diagnosis Decision-Making Mechanism

To satisfy the requirements of both diagnosis time and diagnosis accuracy, a two-stage decision-making mechanism including fault detection and fault classification is designed for fault diagnosis.

For the fault detection step, the principle derived from model-based methods is to calculate the residual between predicted signal and real-time sampled sensor signal and compare it with the fault threshold, which takes a very short time. At the online stage, the data-driven predictor and sensor are working simultaneously, and the current residual can be calculated as:

TABLE I
FAULT TYPE AND FAULT LABELS

Fault Type	Fault Label
sensor gain fault	1
sensor drift fault	2
sensor stuck fault	3
sensor noise fault	4
T_1/T_4 open-circuit	5
T_2/T_3 open-circuit	6

$$\tilde{i}_{\mathrm{N}} = \frac{i_{\mathrm{N}} - \hat{i}_{\mathrm{N}}}{i_{\mathrm{N}}} \tag{16}$$

where i_{N} refers to the real sample value of the grid-side current sensor. Then the fault detection rule can be defined as follows:

$$\begin{cases} \left|\tilde{i}_{\mathrm{N}}\right| > \sigma, & \text{fault occurs} \\ \left|\tilde{i}_{\mathrm{N}}\right| < \sigma, & \text{no fault} \end{cases} \tag{17}$$

where σ is the fault detection threshold. After that, the fault diagnosis will process to the fault classification stage.

The occurrence of different types of faults will cause a different impact on rectifier system signals, which means that the system fault symptom is corresponding to the type of fault. In our previous work [2], the fault classification is completed by the logical judgment module integrating the fault symptom of system signals based on IF-THEN rules developed from prior knowledge. One example of the logical judgment rule is:

Antecedent: $\left|\tilde{i}_{\mathrm{N}}\right| > \sigma$ AND the grid-side current increases AND the dc-link voltage decreases.

Consequent: Then the fault is sensor gain fault.

Moreover, aiming at all six types of faults of the rectifier, a fault classification rule table is constructed based on related fault knowledge basement, as presented in Table II.

Although the classifier constructed based on the fault symptom of the system signal can fulfill the fault classification with high accuracy [2], there still exists the risk of misdiagnosis if the load of the rectifier has a fluctuation. To alleviate such misdiagnosis risk, the fault residual is taken into consideration. As shown in Fig. 7, the current residual between predictor signal and sensor signal has diverse changes when the different fault occurs. According to the waveform characteristic and distribution of these fault residuals, several descriptive statistics including variance, integral, kurtosis, and skewness are adopted to the analysis of fault residuals, as shown in Fig. 8. It is worth mentioning that kurtosis is a measure of steepness of the probability distribution of a real-valued random variable and skewness is a measure of asymmetry. Based on the diversity

TABLE II
FAULT CLASSIFICATION RULES BASED ON SYSTEM FAULT SYMPTOM

Fault Type	Fault Symptom
Sensor gain fault	Grid current increases and dc-link voltage decreases
Sensor drift fault	Grid current stays the same and dc-link voltage ripple increases
Sensor stuck fault	Grid current remains at a constant value
Sensor noise fault	Grid current increases and dc-link voltage stays the same
T_1/T_4 open-circuit	Positive part of grid current remains but negative part decreases
T_2/T_3 open-circuit	Negative part of grid current remains but positive part decreases

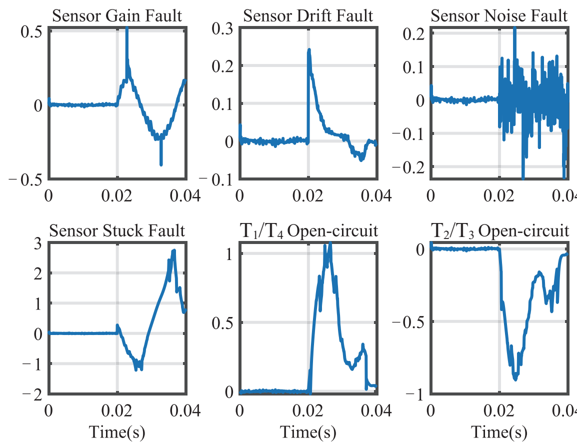


Fig. 7. Waveform of different type of fault residual in single-phase PWM rectifier.

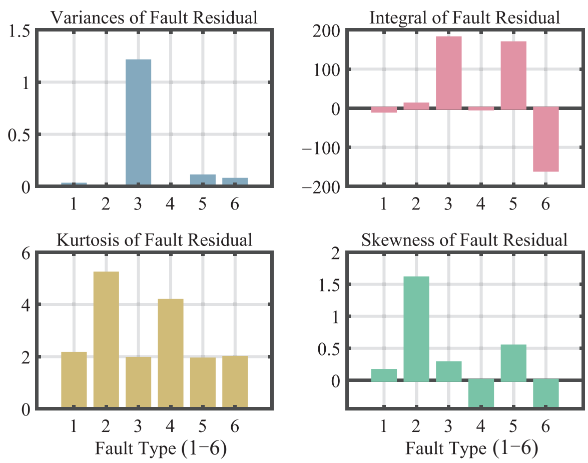


Fig. 8. The variance, integral, kurtosis, and skewness of fault residuals. Fault type: 1, sensor gain fault; 2, sensor drift fault; 3, sensor stuck fault; 4, sensor noise fault; 5, T_1/T_4 open-circuit fault; 6, T_2/T_3 open-circuit fault.

and feature of different fault residuals in these descriptive statistics, the fault classifier can be built as Fig. 9, where $v = 1$, $s = 1$, $i_+ = 100$, $i_- = -100$, $k = 3$.

Eventually, the fault classification is completed by the hybrid classifier combining system fault symptom and fault residual analysis. When a fault is detected, the classifier based on

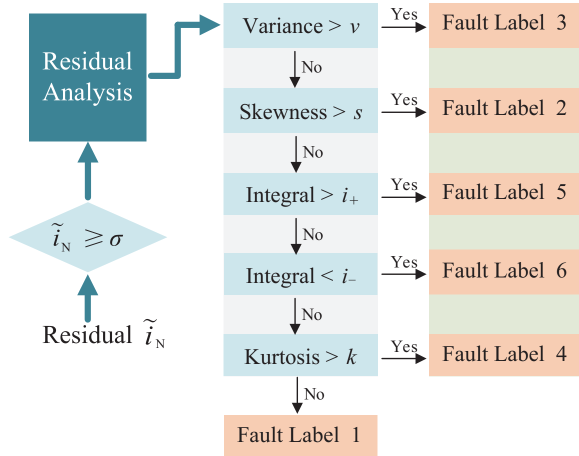


Fig. 9. Fault classifier based on fault residual.

system fault symptom and the classifier based on fault residual analysis work simultaneously and the fault label is output only if their classification results are consistent. In this way, the misdiagnosis risk of the load fluctuating case can be reduced and the fault classification accuracy and effectiveness can be guaranteed.

V. EXPERIMENTAL VALIDATION

A. Offline Modeling and Training

From the perspective of error-ambiguity decomposition [20], the higher the accuracy and diversity of the individual learners, the better the ensemble, which can be formulated as follows:

$$E = \overline{E} - \overline{A} \tag{18}$$

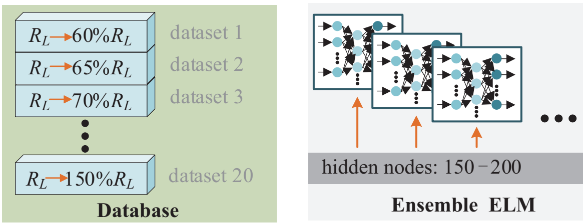
where E is the ensemble generalization error. $\overline{E}$ and $\overline{A}$ refer to the weighted mean of the generalization error and ambiguity of individual learners, which can be formed as follows:

$$\overline{E} = \sum_{i=1}^{p} \omega_i E_i \ ; \ \overline{A} = \sum_{i=1}^{p} \omega_i A_i \tag{19}$$

where p is the number of individual learners, and ω_i is the weight. As the ambiguity of individual learners, A_i represents the diversity of individual learners in samples.

Therefore, the performance of the ensemble learner can be optimized by increasing the accuracy and diversity of individual learners. On the one hand, the accuracy is guaranteed by applying ELM as the base learning algorithm. On the other hand, to increase the diversity of individual learners, the common idea is to introduce randomness in the learning process by perturbing data samples, input attributes, and algorithm parameters.

As a result, the data sample disturbance and algorithm parameters disturbance are adopted to increase the diversity of the ensemble ELMs learner, as shown in Fig. 10. On the offline stage, considering the case of load fluctuating, the dataset is



(a) (b)

Fig. 10. Increasing the diversity of ensemble ELM in offline modeling and training. (a) data sample disturbance, (b) algorithm parameters disturbance.

derived from the rectifier working at rated load R_L changing to different load, varying from 60% R_L to 150% R_L, as shown in Fig. 10(a). Then the training dataset of every single ELM is randomly selected in the database. Besides, the randomness is also introduced into the algorithm parameters in the modeling process by setting different hidden neuron nodes number. According to the survey in [21], the range of hidden node number is set as [150, 200], as shown in Fig. 10(b).

B. Online Experimental Testing

To validate the practicability and effectiveness of the designed diagnosis method for practical application, several hardware-in-loop tests are carried out. The experimental system parameters are listed in Table III. As presented in Fig. 11, the experimental platform consists of an AC power, a single-phase PWM rectifier circuit, an RT Box simulator, and an upper computer. The control strategy of the rectifier and the diagnosis algorithm is conducted by RT Box with corresponding software in the upper computer.

The predictor model is built and trained in the upper computer so that the grid-side current prediction and fault diagnosis can be realized by the RT Box simulator connected to the computer. Additionally, due to the convenience of the RT Box simulator, the real-time testing results can be observed and saved in the upper computer without using an oscilloscope.

Based on the fault analysis in section IV, four kinds of sensor faults and two IGBT open-circuit faults are simulated in the rectifier to verify the applicability of the proposed data-driven diagnosis method. For sensor gain fault, the gain factor A is set as 1.1. And for sensor drift fault, the drift factor K_1 is set as 4, which is 0.2 times the grid current.

Fig. 12 shows the test result of the rectifier operating in normal condition with load fluctuation. In this case, the predictor signal follows the sensor signal closely and their normalized residual keeps lower than 0.04 even after the moment that the load changes from R_L to 135% R_L (the fluctuating range taken into consideration is set as 60% to 150%). Therefore, the fault threshold σ is set as 0.04, which means that no fault occurs as long as the normalized residual staying lower than it. Moreover, there is no risk of misdiagnosis even if the load has a fluctuation.

Fig. 13(a)–(f) presents test results of the rectifier working with different types of fault occurrence, including sensor gain fault, drift fault, stuck fault, and noise fault, and T_1/T_4 open

TABLE III
PARAMETERS OF SINGLE PHASE PWM RECTIFIER

Symbol	Parameter	Value
u_N	grid voltage (RMS)	60 V
L_N	equivalent filter inductor	3 mH
R_s	parasitic resistor	0.1 Ω
u_d	reference DC-link voltage	120 V
C_d	DC-link capacitor	2.8 mF
R_L	rated equivalent load	20 Ω
f_s	switching frequency	1 kHz
T_c	control cycle	50 μs
f_N	grid frequency	50 Hz

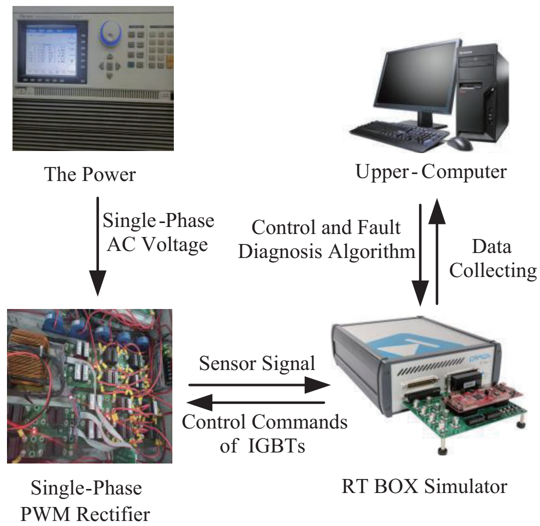


Fig. 11. Experimental platform.

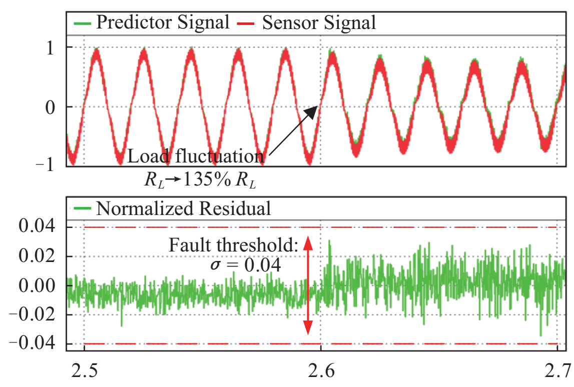


Fig. 12. Test result of the rectifier working at normal condition with load fluctuating.

circuit fault, T_2/T_3 open circuit fault. It can be seen that the normalized residual keeps lower than the fault threshold and the fault detection label and classification label stay at zero when the rectifier is operating at normal conditions. But once a sensor fault or IGBT open-circuit fault occurs, the normalized residual will have a change or even a large deviation compared to normal operation immediately and exceed the fault threshold (i.e. $\tilde{i}_N > \sigma$), and the fault detection label will be set to 0.5 within 1 ms, indicating a fault occurring in the system.

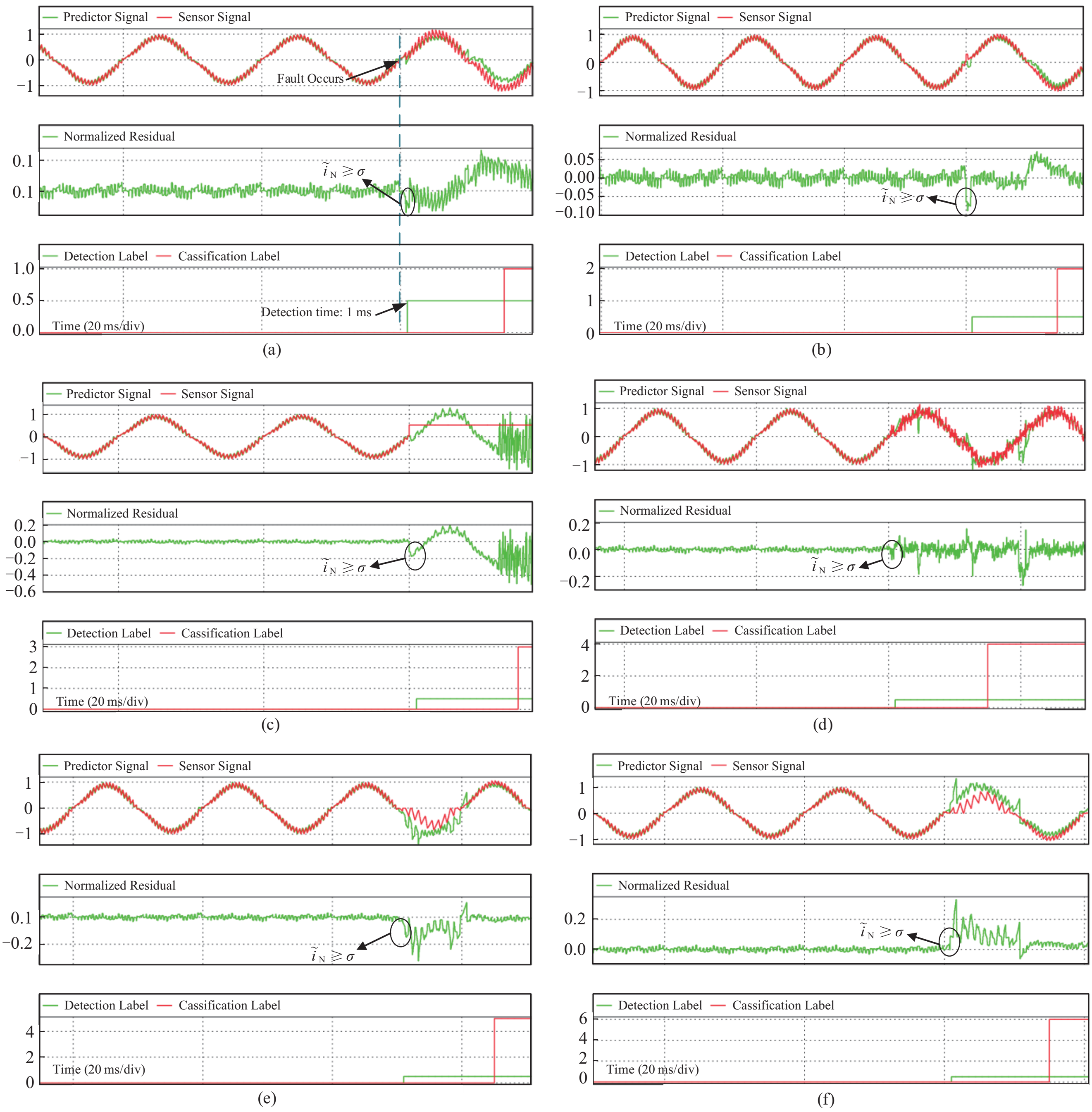


Fig. 13. Test results of the rectifier working with fault occurrence. (a) sensor gain fault, (b) sensor drift fault, (c) sensor stuck fault, (d) sensor noise fault, (e) T_1/T_4 open-circuit fault, (f) T_2/T_3 open-circuit fault.

Furthermore, in about 15 ms (i.e. 3/4 fundamental period time) after the fault occurs, the corresponding classification label varying from 1 to 6 will be output and thus the fault type is classified in detail and located accurately.

According to these test results, the proposed data-driven method has been validated as an effective and practical approach for the online fault diagnosis in single-phase PWM rectifier, which can realize the signal estimation in a data-driven way. Moreover, the proposed method can perform the online fault diagnosis with fast fault detection speed and high fault classification accuracy based on the detection-classification mechanism.

VI. Conclusion

This paper proposes a data-driven method dealing with online fault diagnosis problems of the single-phase PWM rectifier, which is the essential subunit of the cascaded power electrical traction transformer. The critical idea of this data-driven approach is to construct a signal predictor applying the NARX model with ensemble ELMs. The predictor and the real sensor operate simultaneously and generate a residual. Then fault detection is conducted by comparing the residual with fault threshold, and fault classification is completed by a hybrid classifier combining the system fault symptoms with fault residual analysis. The experimental test results demonstrate that the proposed data-driven method is feasible and effective to realize the online fault diagnosis of single-phase PWM rectifier with fast diagnosis speed and high accuracy, and is robust against load fluctuation.

References

[1] J. Zhang, J. Liu, S. Zhong, J. Yang, N. Zhao, and T. Q. Zheng, "A power electronic traction transformer configuration with low-voltage IGBTs for onboard traction application," in *IEEE Transactions on Power Electronics*, vol. 34, no. 9, pp. 8453–8467, Sept. 2019.

[2] K. Zhang and B. Gou, "A data-driven method for online fault diagnosis in single-phase PWM rectifier," in *Proceedings of* 2021 *IEEE 1st International Power Electronics and Application Symposium (PEAS)*, Shanghai, China, Nov. 2021.

[3] Z. Gao, C. Cecati, and S. X. Ding, "A survey of fault diagnosis and fault-tolerant techniques—Part II: Fault diagnosis with knowledge-based and hybrid/active approaches," in *IEEE Transactions on Industrial Electronics*, vol. 62, no. 6, pp. 3768–3774, Jun. 2015.

[4] B. Gou, X. Ge, S. Wang, X. Feng, J. B. Kuo, and T. G. Habetler, "An open-switch fault diagnosis method for single-phase PWM rectifier using a model-based approach in high-speed railway electrical traction drive system," in *IEEE Transactions on Power Electronics*, vol. 31, no. 5, pp. 3816–3826, May 2016.

[5] D. Xie and X. Ge, "Residual-changing-rate based open-circuit fault diagnosis for a single-phase cascaded H-bridge rectifier," in *Transactions of China Electrotechnical Society,* vol. 33, no. 16, pp. 3822–3834, Aug. 2018.

[6] N. Egone, L. Ma, and N. Qin, "Faults detection and identification based on robust residual generation for a single-phase PWM rectifier," in *Proceedings of 2019 IEEE 3rd Conference on Energy Internet and Energy System Integration (EI2)*, Changsha, China, Nov. 2019, pp. 732–736.

[7] S. Zhuo, A. Gaillard, L. Xu, C. Liu, D. Paire, and F. Gao, "An observer-based switch open-circuit fault diagnosis of DC–DC converter for fuel cell application," in *IEEE Transactions on Industry Applications*, vol. 56, no. 3, pp. 3159–3167, May/Jun. 2020.

[8] Q. T. An, L. Sun, and L. Z. Sun, "Current residual vector-based open-switch fault diagnosis of inverters in PMSM drive systems," in *IEEE Transactions on Power Electronics*, vol. 30, no. 5, pp. 2814–2927, May 2015.

[9] J. Poon, P. Jain, C. Spanos, S. K. Panda, and S. R. Sanders, "Fault prognosis for power electronics systems using adaptive parameter identification," in *IEEE Transactions on Industry Applications*, vol. 53, no. 3, pp. 2862–2870, May/Jun. 2017.

[10] F. Bento and A. J. M. Cardoso, "Open-circuit fault diagnosis and fault tolerant operation of interleaved DC-DC boost converters for homes and offices," in *IEEE Transactions on Industry Applications*, vol. 55, no. 5, pp. 4855–4864, Sept./Oct. 2019.

[11] B. Cai, Y. Zhao, H. Liu, and M. Xie, "A data-driven fault diagnosis methodology in three-phase inverters for PMSM drive systems," in *IEEE Transactions on Power Electronics*, vol. 32, no. 7, pp. 5590–5600, Jul. 2017.

[12] I. Bandyopadhyay, P. Purkait, and C. Koley, "Performance of a classifier based on time-domain features for incipient fault detection in inverter drives," in *IEEE Transactions on Industrial Informatics*, vol. 15, no. 1, pp. 3–14, Jan. 2019.

[13] Z. Huang, Z. Wang, X. Yao, and H. Zhang, "Multi-switches fault diagnosis based on small low-frequency data for voltage-source inverters of PMSM drives," in *IEEE Transactions on Power Electronics*, vol. 34, no. 7, pp. 6845–6857, Jul. 2019.

[14] Y. Xia, Y. Xu, and B. Gou, "A data-driven method for IGBT open-circuit fault diagnosis based on hybrid ensemble learning and sliding-window classification," in *IEEE Transactions on Industrial Informatics*, vol. 16, no. 8, pp. 5223–5233, Aug. 2020.

[15] G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: Theory and application," in *Neurocomputing*, vol. 70, pp. 489–501, 2006.

[16] J. M. P. Menezes, and G. A. Barreto, "Long-term time series prediction with the NARX network: An empirical evaluation," in *Neurocomputing*, vol. 71, no. 16-18, pp. 3335–3343, Oct. 2008.

[17] B. Gou, X. Ge, S. Wang, X. Feng, J. B. Kuo, and T. G. Habetler, "An open-switch fault diagnosis method for single-phase PWM rectifier using a model-based approach in high-speed railway electrical traction drive system," in *IEEE Transactions on Power Electronics*, vol. 31, no. 5, pp. 3816–3826, May 2016.

[18] H. T. Siegelmann, B. G. Horne, and C. L. Giles, "Computational capabilities of recurrent NARX neural networks," in *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 27, no. 2, pp. 208–215, Apr. 1997.

[19] T.G. Dietterich, "Ensemble methods in machine learning," in *International Workshop on Multiple Classifier Systems*, vol. 1857. Springer, Berlin, Heidelberg, Dec. 2000.

[20] A. Krogh and J. Vedelsby, "Neural network ensembles, cross validation, and active learning," in *Proceedings of Neural Information Processing Systems: Natural and Synthetic*, pp. 231–238, Denver, Colorado, USA, 1995.

[21] Y. Xia, B. Gou, and Y. Xu, "A data-driven method for current sensor fault diagnosis in single-phase PWM rectifier," in *Proceedings of 2019 9th International Conference on Power and Energy Systems (ICPES)*, Perth, WA, Australia, Dec. 2019.

Kun Zhang received the B.S. degree in electrical engineering and automation from Southwest Jiaotong University, Chengdu, China, in 2021.

He is currently pursuing the master degree in the School of Electrical Engineering, Southwest Jiaotong University. His main research interests include machine learning and fault diagnosis for power electronics.

Bin Gou received the B.S. and Ph.D. degrees in electrical engineering from Southwest Jiaotong University, Chengdu, China, in 2010 and 2016, respectively.

He is currently an Associate Professor with the School of Electrical Engineering, Southwest Jiaotong University. His research interests include data-analytics for fault diagnosis and health management of power drive systems.

Xiaoyun Feng received the B.S., M.S., and Ph.D. degrees in electrical engineering from Southwest Jiaotong University, Chengdu, China, in 1983, 1988, and 2001, respectively.

She is currently a Professor with the School of Electrical Engineering, Southwest Jiaotong University. Her research interests include fault prognostic and health management for electric traction systems.