International Journal of Innovative Computing, Information and Control
Volume **21**, Number **2**, April **2025**

ICIC International ©2025 ISSN 1349-4198
pp. **389–406**

A HYBRID MODEL/DATA-DRIVEN METHOD FOR OPEN-CIRCUIT FAULT DIAGNOSIS IN NPC THREE-LEVEL INVERTERS

Weilin Yang[1], Chao Zhang[1], Dezhi Xu[2,3,*] and Yujian Ye[2]

[1]School of Internet of Things Engineering
Jiangnan University
No. 1800, Lihu Avenue, Wuxi 214122, P. R. China
wlyang@jiangnan.edu.cn; 6221915028@stu.jiangnan.edu.cn

[2]School of Electrical Engineering
[3]Engineering Research Center of Electrical Transport Technology, Ministry of Education
Southeast University
No. 2, Sipailou, Xuanwu District, Nanjing 210096, P. R. China
yeyujian@seu.edu.cn
*Corresponding author: xudezhi@seu.edu.cn

Received August 2024; revised November 2024

Abstract. *In this paper, the diagnosis and location issue of open-circuit faults in neutral-point-clamped three-level inverters is analyzed. A hybrid method based on model-based and data-driven fault diagnosis is proposed for the insulated gate bipolar transistors open-circuit faults. First, the three-phase current residual is attained by subtracting the true value of the three-phase current generated by the inverter with the estimated three-phase current generated by the state estimator. Second, the Park's vector modulus and wavelet transformation algorithms are utilized for normalization of three-phase current residuals and feature analysis. Then, the 11 fault features of current residuals are extracted and the datasets of fault feature are established. Finally, the fault feature samples are utilized to train the random forest model to achieve state classification. The proposed method can improve the diagnostic accuracy compared with traditional fault diagnosis methods. The effectiveness and the robustness of this method under various conditions are validated by experimental results.*

Keywords: Neutral-point-clamped three-level inverter, Hybrid faults diagnosis, Model-based method, Data-driven method, Random forest

1. **Introduction.** As the increasing consumption of conventional energy sources such as oil and coal, new energy technology has become a research hotspot [1]. Coupled with the global recognizes the importance of reducing energy consumption. The new energy grid connection represented by photovoltaic power generation needs to be realized by inverters [2]. Due to the important role that inverters play in the area of new energy, the demand for high-power inverters is increasing [3]. Neutral-point-clamped (NPC) three-level inverters are widely applied by virtue of merits of flexible control, high voltage resistance, lower harmonic distortion, and output current waveform tending towards sine wave [4]. The NPC three-level inverter has more power switches in contrast with traditional two-level inverters, resulting in more intricate circuits and diminished stability [5]. The insulated gate bipolar transistors (IGBTs) are operated as power switches in inverters, working under the high operating frequency for a long time. Among all faults of the inverter, the probability of IGBT failures is extremely high [6].

DOI: 10.24507/ijicic.21.02.389

The primary IGBT faults in three-level inverters are open-circuit and short-circuit faults. In the case of short-circuit faults, the three-phase current experiences an instantaneous surge, and the power supply is immediately disconnected by the protection circuit in the system with the intention of protecting the entire system. On the contrary, open-circuit faults lead to distortion in load current and abnormal operation of switching devices, resulting in reduced performance of the inverter system. If not addressed promptly, it can lead to secondary failures in other devices, thereby impacting overall system functionality. Therefore, this paper focuses on the study of open-circuit faults in inverters.

Since the significant losses caused by open-circuit faults, numerous experts and scholars have dedicated efforts towards researching on inverter open-circuit fault diagnosis. Many open-circuit fault diagnosis methods have been suggested, which are broadly categorized into three types: model-based, signal-based and data-driven methods.

Model-based fault diagnosis methods require accurate system models that can effectively utilize system information to provide reliable diagnostic results. However, it heavily relies on system model and parameters precision [7]. In Huang et al. [8], a fault diagnosis method of permanent magnet synchronous motor (PMSM) is proposed, which takes the change of normalized cost function as the fault diagnosis indicator to realize fault identification and location. In Huang et al. [9], this paper combines the advantages of model predictive control (MPC) and mixed logical dynamic model (MLDM) to diagnose the PMSM open-circuit fault. In Liu et al. [10], a fault estimation control method based on hierarchical structure is proposed, and the state of the system is estimated by combining observers in a decentralized manner. In Huang et al. [11], the fault diagnosis method of inverters in complex condition is realized by using the evaluation variable and position information of reversed trend comprehensive assessment-based algorithm. In Yong et al. [12], the online fault diagnosis algorithm based on the characteristic phase and amplitude of the current residual for the voltage-source inverters are proposed. However, the threshold value setting requires adjustment as the load changes, which is difficult to set as a constant and has limited applicability. Thus, the accuracy of fault diagnosis algorithms is hard to ensure.

Signal-based fault diagnosis methods require prior knowledge of the system for fault diagnosis, which is less reliant on models [13]. However, complete prior knowledge is necessary to ensure accurate diagnosis, as noise and other factors can easily affect its results. In Huang et al. [14], the virtual mirrors are constructed with fault information from different angles and the feature data is processed by cross comparison processing. The inverter fault diagnosis is realized by using the joint decision method. In Tang et al. [15], the random forest (RF) algorithm is used to scale and optimize the current signal which are extracted under different fault modes.

Data-driven fault diagnosis method exhibits excellent capabilities in data processing, portability, and multiple fault handling capabilities, while being independent of models and prior knowledge. However, it necessitates a lot of historical data for training, resulting in high computational costs and lengthy execution times [7]. In Sonawane and Patil [16], feature extraction from three-phase current, voltage, torque, and speed is carried out by track and hunt meta heuristic, and the deep neural network (DNN) is used to identify and locate faults. In Yao et al. [17], the parameters of the extreme learning machine (ELM) fault diagnosis model are optimized using whale optimization algorithm (WOA) to identify and locate faults. In Wang et al. [18], the fault diagnosis method based on extended deep belief network (EDBN) is proposed to get the most out of the valuable information in the raw data and its good fault classification performance is verified by experiments.

In order to enhance excellent diagnosis performance, researchers have proposed schemes that integrate various fault diagnostic technologies. In Sánchez et al. [19], a hybrid method based on model and data processing is applied to fault diagnosis in NPC inverters. In Xia et al. [20], a hybrid fault diagnosis algorithm based on model and data-driven method is proposed. The data-driven prediction model is generated using ELM. Signal prediction and residual generation diagnostic methods were proposed based on the model. In Wang et al. [21], an adaptive disturbance suppression based fault-tolerant control method is proposed, which combines real-time model identification and nonlinear dynamic inversion, and uses adaptive disturbance suppression controller to realize fault-tolerant control. In Kou et al. [22], a combination of data-driven and signal-based methods is implemented for fault diagnosis. Concordia transform is used to extract features, followed by construction of an RF classifier for online fault diagnosis. Therefore, this paper proposed a hybrid fault diagnosis method for the NPC inverter. By adopting both advantages of model-based and data-driven methods, the proposed hybrid fault diagnosis method has good adaptability as well as higher diagnostic accuracy and robustness. The diagnostic accuracy rates of simulation and StarSim HIL experiments are 98.5% and 99.6%, respectively. The main contributions of this study are discussed next.

1) The three-phase current residual is utilized as the original dataset, and its features are employed to train the data-driven model. Compared to model-based methods, it avoids the manual adjustment of the residual threshold under different loads and improves the universality of the diagnosis method.

2) Sample data is pre-processed using wavelet transformation (WT) and Park's vector modulus (PVM) normalization algorithm. The statistical characteristics of three-phase current residual signals are extracted to enhance distinctiveness of data features.

3) The random forest is trained using the fault feature extracted from residual currents. Compared to traditional data-driven methods, this method offers higher accuracy with less data, improved diagnostic accuracy, as well as anti-interference and robustness.

The structure of the article is as follows. The current fault characteristics of the NPC three-level inverter are presented in Section 2. The hybrid fault diagnosis method is described in Section 3. The hybrid fault diagnosis algorithm is verified through simulation and experiment in Section 4 and Section 5. The conclusions are given in Section 6.

2. **Basic Theory.**

2.1. **Working principle of the NPC three-level inverter.**

The topology of the NPC three-level inverter is detailed in Figure 1. C_1 and C_2 are the two equivalent capacitors on the DC side, which have the functions of buffering energy and dividing voltage. The NPC three-level inverter consists of phase-A, phase-B and phase-C bridge legs. Each bridge leg includes four IGBTs (S_{x1}-S_{x4}), four freewheeling diodes (D_{x1}-D_{x4}), two clamping diodes (D_{x5}-D_{x6}), where $x = a, b, c$ represents phase-A, phase-B and phase-C bridge legs. Two clamping diodes provide a path for the current in the O state of the inverter circuit, which are used to clamp the DC side voltage. The three-phase load current is represented by i_A, i_B and i_C. The load is resistance and inductance.

The NPC three-level inverter has three work modes (P, O, N). Because the three-level inverter circuit has the characteristics of three-phase symmetrical distribution, the phase-A bridge leg is analyzed as an example. Figure 2 shows current flow paths according to work modes, where the solid line represents the current flowing from the DC side to the load side, defined as i_A^+, and the dashed line represents the current flowing from the load side to the DC side, defined as i_A^-. If the inverter works in P status, when the current

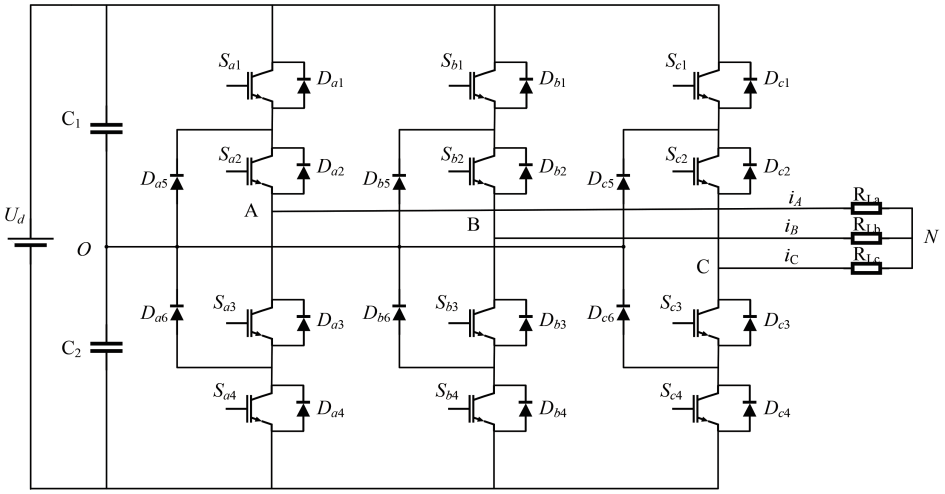


FIGURE 1. Topology of the NPC three-level inverter

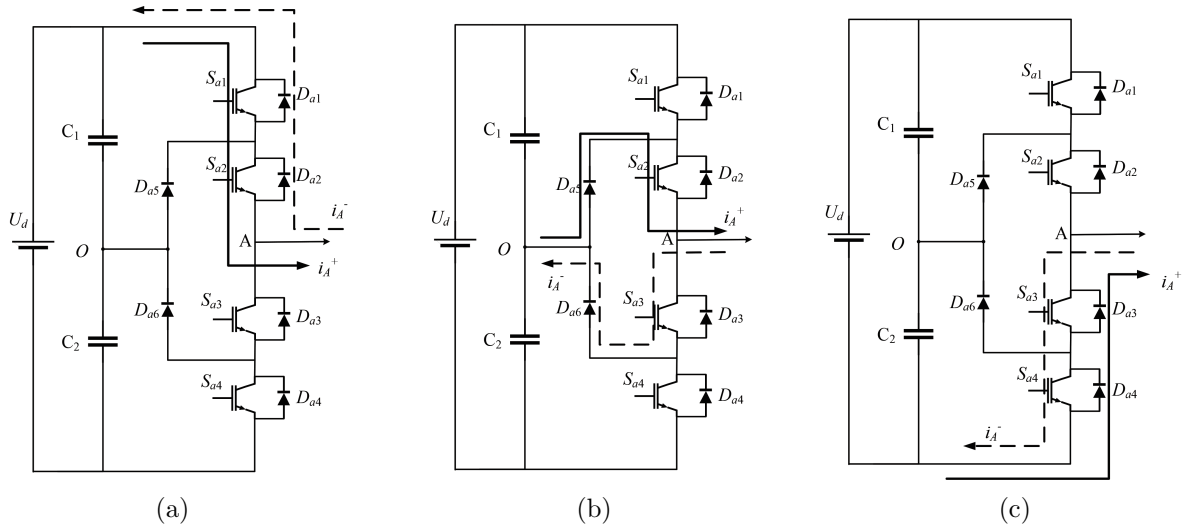


FIGURE 2. Current flow paths according to work modes: (a) P status; (b) O status; (c) N status

is positive, the current flows to the load side through S_{a1} and S_{a2}; when the current is negative, the current flows to the DC side through D_{a1} and D_{a2}.

2.2. **Fault features analysis of the NPC three-level inverter.** When the inverter works normally, the three-phase current signals are sinewave, and each phase differs by 120°. The single-tube open fault of phase-A bridge leg is analyzed as an example. Figure 3 shows signals of three-phase current at phase-A single-tube fault conditions. When S_{a1} open fault occurs, the positive half-period amplitude of the phase-A current decreases to half of its normal value, while the negative half-period remains essentially unchanged. When S_{a2} open fault occurs, the positive half-period amplitude of the phase-A current becomes zero, resulting in phase-A deficiency, while the negative half-period remains basically unaffected. The current waveform features between S_{a4} and S_{a1} (S_{a3} and S_{a2}) are opposite to each other. It can be concluded that different positions of IGBTs exhibit distinct current features when open faults occur.

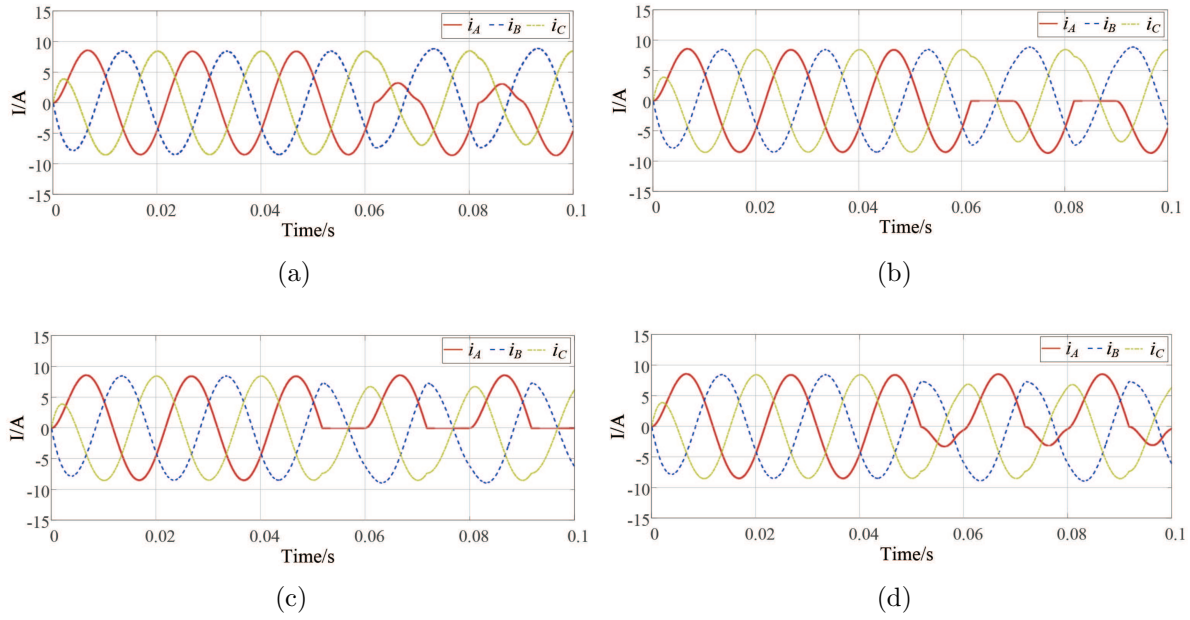


FIGURE 3. Signals of three-phase current at single-tube open fault of phase-A bridge leg conditions: (a) S_{a1} open fault; (b) S_{a2} open fault; (c) S_{a3} open fault; (d) S_{a4} open fault

3. Fault Diagnosis Algorithm Based on a Hybrid Model-Based and Data-Driven Method.

3.1. **Modeling of the NPC three-level inverter.** According to the inverter topology, the three-phase voltage equation can be obtained as follows:

$$\begin{cases} U_{AN} = Ri_A + L\dfrac{di_A}{dt} \\ U_{BN} = Ri_B + L\dfrac{di_B}{dt} \\ U_{CN} = Ri_C + L\dfrac{di_C}{dt} \end{cases} \tag{1}$$

where R is resistance, L is inductance, U_{AN}, U_{BN}, U_{CN} is three-phase phase voltage and i_A, i_B, i_C is three-phase phase current. When the inverter works normally, there is

$$U_{AN} + U_{BN} + U_{CN} = 0 \tag{2}$$

As per Kirchhoff's law, the following results are obtained:

$$\begin{cases} U_{AN} = U_{AO} + U_{ON} \\ U_{BN} = U_{BO} + U_{ON} \\ U_{CN} = U_{CO} + U_{ON} \end{cases} \tag{3}$$

where U_{AO}, U_{BO}, U_{CO} is three-phase terminal voltage, U_{ON} is the voltage between DC neutral point O and load neutral point.

Combining (2) and (3), the following result is as follows:

$$U_{ON} = -\frac{1}{3}\left(U_{AO} + U_{BO} + U_{CO}\right) \tag{4}$$

Combining (4) and (3), the relationship between terminal voltage and phase voltage is obtained:

$$\begin{bmatrix} U_{AN} \\ U_{BN} \\ U_{CN} \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} U_{AO} \\ U_{BO} \\ U_{CO} \end{bmatrix} \tag{5}$$

Combining (1) to (5), there are

$$\begin{bmatrix} L\dfrac{di_A}{dt} \\ L\dfrac{di_B}{dt} \\ L\dfrac{di_C}{dt} \end{bmatrix} = -R \begin{bmatrix} i_A \\ i_B \\ i_C \end{bmatrix} + \frac{1}{3} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} U_{AO} \\ U_{BO} \\ U_{CO} \end{bmatrix} \tag{6}$$

The phase-B bridge leg is as an example to analyze the relationship between various variables in the normal operation of the inverter ($S_{bi} = 1$ when S_{bi} is ON and $S_{bi} = 0$ when S_{bi} is OFF, $i = 1, 2, 3, 4$), and the expression of terminal voltage U_{BO} can be obtained as follows:

$$U_{BO} = \frac{1}{2} U_{dc} \left(S_{b1} S_{b2} \overline{S_{b3} S_{b4}} - \overline{S_{b1} S_{b2}} S_{b3} S_{b4} \right) \tag{7}$$

Then, there are

$$\begin{cases} U_{AO} = \dfrac{1}{2} U_{dc} \left(S_{a1} S_{a2} \overline{S_{a3} S_{a4}} - \overline{S_{a1} S_{a2}} S_{a3} S_{a4} \right) \\ U_{BO} = \dfrac{1}{2} U_{dc} \left(S_{b1} S_{b2} \overline{S_{b3} S_{b4}} - \overline{S_{b1} S_{b2}} S_{b3} S_{b4} \right) \\ U_{CO} = \dfrac{1}{2} U_{dc} \left(S_{c1} S_{c2} \overline{S_{c3} S_{c4}} - \overline{S_{c1} S_{c2}} S_{c3} S_{c4} \right) \end{cases} \tag{8}$$

According to $\lambda_m = \left(S_{m1} S_{m2} \overline{S_{m3} S_{m4}} - \overline{S_{m1} S_{m2}} S_{m3} S_{m4} \right)$ $(m = A, B, C)$ and combining (5), the following results are as follows:

$$\begin{bmatrix} U_{AN} \\ U_{BN} \\ U_{CN} \end{bmatrix} = \frac{U_{dc}}{6} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} \lambda_A \\ \lambda_B \\ \lambda_C \end{bmatrix} \tag{9}$$

Combining (6) and (9), the state space model vector expression of the NPC three-level inverter can be obtained:

$$\dot{x} = Ax + B\lambda \tag{10}$$

where $x = [i_A \ i_B \ i_C]^T$; $\lambda = [\lambda_A \ \lambda_B \ \lambda_C]^T$; $A = \mathrm{diag}(-R/L, -R/L, -R/L)$; $B = U_{dc}/6L[2, -1, -1; -1, 2, -1; -1, -1, 2]$.

When the inverter has a single-tube open fault, the S_{b1} open fault of phase-B bridge leg is analyzed as an example. The expression of U'_{BO} can be obtained as follows:

$$U'_{BO} = \frac{1}{2} U_{dc} \left(\overline{\delta_b \overline{S_{b1}}} S_{b2} \overline{S_{b3} S_{b4}} - \overline{S_{b1} S_{b2}} S_{b3} S_{b4} \right) \tag{11}$$

where δ_b means the polarities of phase-B current ($\delta_b = 1$ when $i_B > 0$ and $\delta_b = 0$ when $i_B < 0$). According to $\lambda'_B = \left(\overline{\delta_b \overline{S_{b1}}} S_{b2} \overline{S_{b3} S_{b4}} - \overline{S_{b1} S_{b2}} S_{b3} S_{b4} \right)$, the following results are as follows:

$$U'_{BO} = \frac{1}{2} U_{dc} \lambda'_B \tag{12}$$

Combining (5), there are

$$\begin{bmatrix} U_{AN} \\ U'_{BN} \\ U_{CN} \end{bmatrix} = \frac{U_{dc}}{6} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} \lambda_A \\ \lambda'_B \\ \lambda_C \end{bmatrix} \tag{13}$$

When an open fault occurs in the NPC three-level inverter, compared with the normal condition, its operating state is deviated, and different IGBTs open faults have different deviations. Therefore, the state estimator of inverter system in the normal situation is constructed. The three-phase current residual is attained through subtracting the true value of the three-phase current generated by the inverter with the estimated three-phase current generated by the state estimator.

The designed state estimator under normal condition is

$$\dot{\hat{x}} = A\hat{x} + B\lambda \tag{14}$$

The residual current equation is

$$\Delta\dot{x} = A\Delta x + B\left(\lambda' - \lambda\right) = A\Delta x + B\Delta\lambda \tag{15}$$

When the IGBT works normally, the estimated value of the state estimator equals the actual current value ($\Delta x = 0, \Delta\lambda = 0$) and the current residual is 0. The S_{b1} open fault of phase-B bridge leg is analyzed as an example, $S_{b1} = 0$, the current residual state equation of the inverter is as follows:

$$\begin{bmatrix} \Delta\dot{i}_A \\ \Delta\dot{i}_B \\ \Delta\dot{i}_C \end{bmatrix} = A \begin{bmatrix} \Delta i_A \\ \Delta i_B \\ \Delta i_C \end{bmatrix} + \frac{U_{dc}}{6L} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ \Delta\lambda_B \\ 0 \end{bmatrix} \tag{16}$$

Among $\Delta\lambda_B = \left(\overline{\delta_b S_{b1}} S_{b2} \overline{S_{b3} S_{b4}} - S_{b1} S_{b2} \overline{S_{b3} S_{b4}}\right)$ and assuming the initial value of the current residual is 0, there are

$$\begin{cases} \Delta i_A = -\dfrac{U_{dc} \cdot \Delta\lambda_B}{6R}\left(1 - e^{-\frac{R}{L}t}\right) \\ \Delta i_B = \dfrac{2U_{dc} \cdot \Delta\lambda_B}{6R}\left(1 - e^{-\frac{R}{L}t}\right) \\ \Delta i_C = -\dfrac{U_{dc} \cdot \Delta\lambda_B}{6R}\left(1 - e^{-\frac{R}{L}t}\right) \end{cases} \tag{17}$$

According to (17), when the S_{b1} of phase-B bridge leg is open faulty, the current residual of phase-B is $\Delta i_B \leq 0$, the current residual of phase-A and phase-C is $\Delta i_A = \Delta i_C \geq 0$. The current residual of phase-B is twice as large as the current residual of phase-A and phase-C, and the relationship between three-phase is

$$\Delta i_B = -2\Delta i_A = -2\Delta i_C \leq 0 \tag{18}$$

The current residual information for single-tube open fault of phase-B bridge leg is listed in Table 1. In this paper, the signals of the current residual are utilized as the data-driven raw data.

3.2. **Data preprocessing and feature extraction.** During the normal operation, the signals of three-phase load current are given as

$$\begin{cases} i_A = I_m \sin wt \\ i_B = I_m \sin(wt - 2/3\pi) \\ i_C = I_m \sin(wt + 2/3\pi) \end{cases} \tag{19}$$

where I_m is the current maximum amplitude, i_A, i_B, i_C are the instantaneous values of the current in phase-A, phase-B, and phase-C, and w is the current frequency. The load is made independent by three-phase current normalization, and the transient caused by load changes is suppressed with PVM. First, the three-phase current (i_A, i_B, i_C) is transformed into two-phase current (i_d, i_q):

$$\begin{bmatrix} i_d \\ i_q \end{bmatrix} = \sqrt{\frac{2}{3}} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} i_A \\ i_B \\ i_C \end{bmatrix} \tag{20}$$

TABLE 1. The current residual information for single-tube open fault of phase-B bridge leg

State	Fault IGBT	The current residual relationship
Normal	–	$\Delta i_A = \Delta i_B = \Delta i_C = 0$
Signal-tube open fault	S_{b1}	$\Delta i_B = -2\Delta i_A = -2\Delta i_C \leq 0$
	S_{b2}	$\Delta i_B = -2\Delta i_A = -2\Delta i_C \leq 0$
	S_{b3}	$\Delta i_B = -2\Delta i_A = -2\Delta i_C \geq 0$
	S_{b4}	$\Delta i_B = -2\Delta i_A = -2\Delta i_C \geq 0$

The PVM calculation formula is

$$\left|\overline{i_{\text{park}}}\right| = \sqrt{{i_d}^2 + {i_q}^2} \tag{21}$$

The normalized three-phase current is as follows [23]:

$$i_{k\text{park}} = \frac{i_T}{\left|\overline{i_{\text{park}}}\right|} \tag{22}$$

where $T = A$, B or C and $i_{k\text{park}}$ is the normalized three-phase current.

In signal fault data processing, WT is a common time-frequency analysis method, which can well display the overall state of the signal, observe the details of the signal and characterize its local characteristics [24]. In this paper, WT is used to decompose the signals of the three-phase current residual and PVM is used to normalize it, thus completing the signals of the three-phase current residual processing. The 11 faulty features of three-phase current residual signals in Table 2 [25] are computed for the normalized signals.

By simulating the single-tube open faults, the current residual sample dataset containing the normal state and 12 types of single-tube open fault is obtained. In addition, each phase of the current residual signal can extract 11 faulty features, with a total of three phases. Therefore, each group of samples has 33 faulty features. Thus, the faulty feature dataset based on three-phase current residual signals is established.

3.3. **Random forest.** RF is used to diagnose the open fault of the NPC inverter [26]. RF has good predictive classification ability and is easy to implement. As a flexible machine learning algorithm, RF is widely used in the field of fault classification [27]. Multiple decision trees are established through random sampling with playback, and the final classification result is determined by the maximum voting in the sample leaf nodes of the decision tree [28].

TABLE 2. The faulty features of three-phase current residual signals

Number	Feature	Expression
1	Max value	$X_1 = \max(x(i))$
2	Min value	$X_2 = \min(x(i))$
3	Range	$X_3 = \max(\lvert x(i)\rvert) - \min(\lvert x(i)\rvert)$
4	Mean value	$X_4 = \left(\frac{1}{n}\right)\sum_{i=1}^{n} x(i)$
5	Standard deviation	$X_5 = \sqrt{\frac{\sum_{i=1}^{n}(x(i)-X_4)^2}{(n-1)}}$
6	Kurtosis	$X_6 = \frac{\sum_{i=1}^{n}(x(i)-X_4)^4}{\left((n-1){X_5}^4\right)} - 3$
7	Skewness	$X_7 = \frac{\sum_{i=1}^{n}(x(i)-X_4)^3}{\left((n-1){X_5}^3\right)}$
8	Crest factor	$X_8 = \frac{\max(\lvert x(i)\rvert)}{\sqrt{(1/n)\sum_{i=1}^{n} x(i)^2}}$
9	Impulse factor	$X_9 = \frac{\max(\lvert x(i)\rvert)}{\sqrt{(1/n)\sum_{i=1}^{n} \lvert x(i)\rvert}}$
10	Shape factor	$X_{10} = \frac{\sqrt{(1/n)\sum_{i=1}^{n} x(i)^2}}{(1/n)\sum_{i=1}^{n} \lvert x(i)\rvert}$
11	Latitude factor	$X_{11} = \frac{\max(\lvert x(i)\rvert)}{((1/n)\sum_{i=1}^{n} \lvert x(i)\rvert)}$

3.4. **Hybrid fault diagnosis model framework.** The framework of the hybrid model fault diagnosis includes current residual generation, data preprocessing, feature extraction and fault identification to diagnose inverter open-circuit faults. The four steps are as follows.

Step 1. Current Residual Generation. Subtracting the true value of the three-phase current generated by the inverter with the estimated three-phase current generated by the state estimator is used to obtain the current residuals.

Step 2. Data Preprocessing. WT is used to decompose the signals of the three-phase current residual and PVM is used to normalize it.

Step 3. Feature Extraction. Combined with the signals of three-phase current residual fault characteristics table, the 11 faulty features are calculated.

Step 4. Fault Identification. The datasets of faulty feature are segmented into training set and testing set, and fault diagnosis model based on RF is constructed to realize fault identification and location.

4. Simulation Results.

4.1. **Simulation setup.** In an effort to verify the effectiveness of the proposed method, the simulation is carried out in the environment of MATLAB/Simulink. The NPC three-level inverter is controlled by Space Vector Pulse Width Modulation (SVPWM). The

single-tube open faults are studied in the paper. There are 13 kinds of inverter (normal state and open faults of S_{a1}-S_{a4}, S_{b1}-S_{b4}, S_{c1}-S_{c4}). Different combinations of resistors and inductors were selected to simulate load changes, and there were 5 groups of loads: 1) R = 5.0 Ω, L = 12 mH; 2) R = 5.0 Ω, L = 13 mH; 3) R = 5.0 Ω, L = 14 mH; 4) R = 5.0 Ω, L = 20 mH; 5) R = 5.0 Ω, L = 30 mH. In each group of loads, take one electrical period as a group and extract 10 groups of current residual signals. There were 650 ($5 \times 13 \times 10$) groups of current residual signals to construct dataset Case A.

4.2. **Results and discussion.** For load of R = 5.0 Ω and L = 12 mH, the NPC three-level inverter works normally. Figure 4 shows the real value and estimated value of the three-phase current at normal state. Figure 4(c) shows that the three-phase current residual is zero.

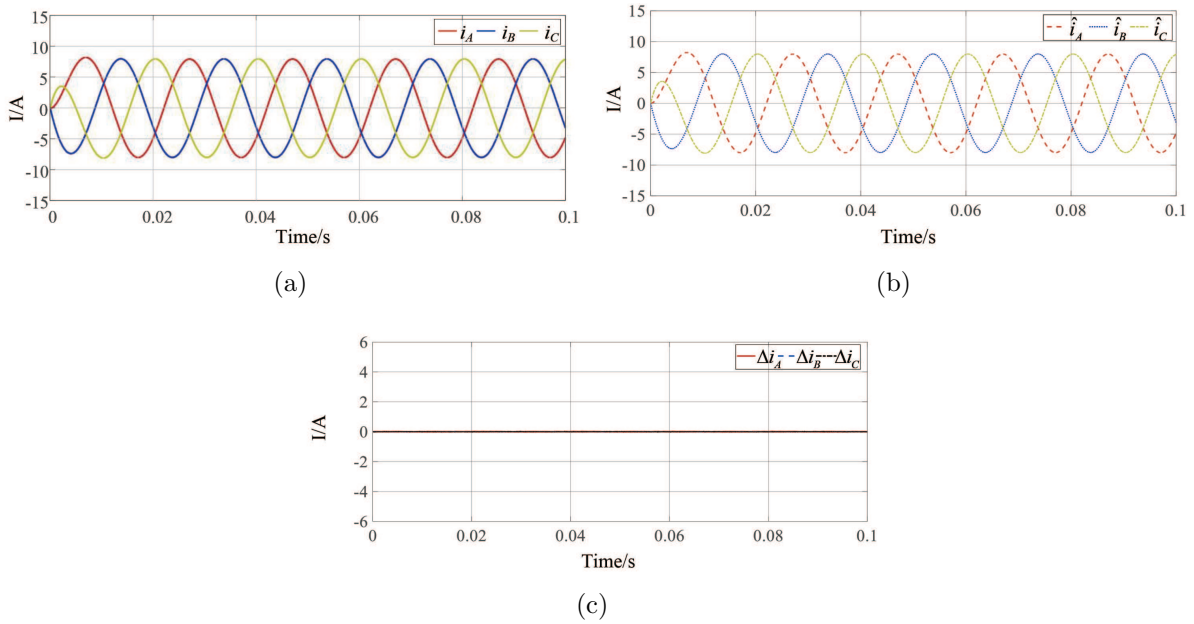


FIGURE 4. Normal state: (a) The real value of the three-phase current; (b) the estimated value of the three-phase current; (c) the three-phase current residual

The four IGBTs of phase-B bridge leg are set to have open faults. The three-phase current residuals of S_{b1}, S_{b2}, S_{b3}, S_{b4} open fault of phase-B bridge leg are shown Figure 5, which meets the current residual information table of the above analysis.

The 33 (3×11) features of each sample are taken as the input vector. In the dataset Case A, 450 groups of samples were randomly selected as the training set, and 200 groups of samples were selected as the testing set.

The number of decision trees in the RF model affects the accuracy of fault diagnosis and training time. If decision trees are too small, the accuracy is affected. If decision trees increase too much, the training time also increases. Training time and diagnostic accuracy for different numbers of decision trees in Case A are shown in Figure 6. The training time increases with the number. As the number gradually increases from 50 to 200, the accuracy rate increases accordingly, and at 200, the diagnostic accuracy reaches 98.5%. As the number gradually increases from 300 to 500, the accuracy declines. To sum up, the best number of decision trees in the RF model is 200.

The confusion matrix of results on datasets of Case A is shown in Figure 7. The row represents the actual state, and the column represents the predicted state. In the confusion

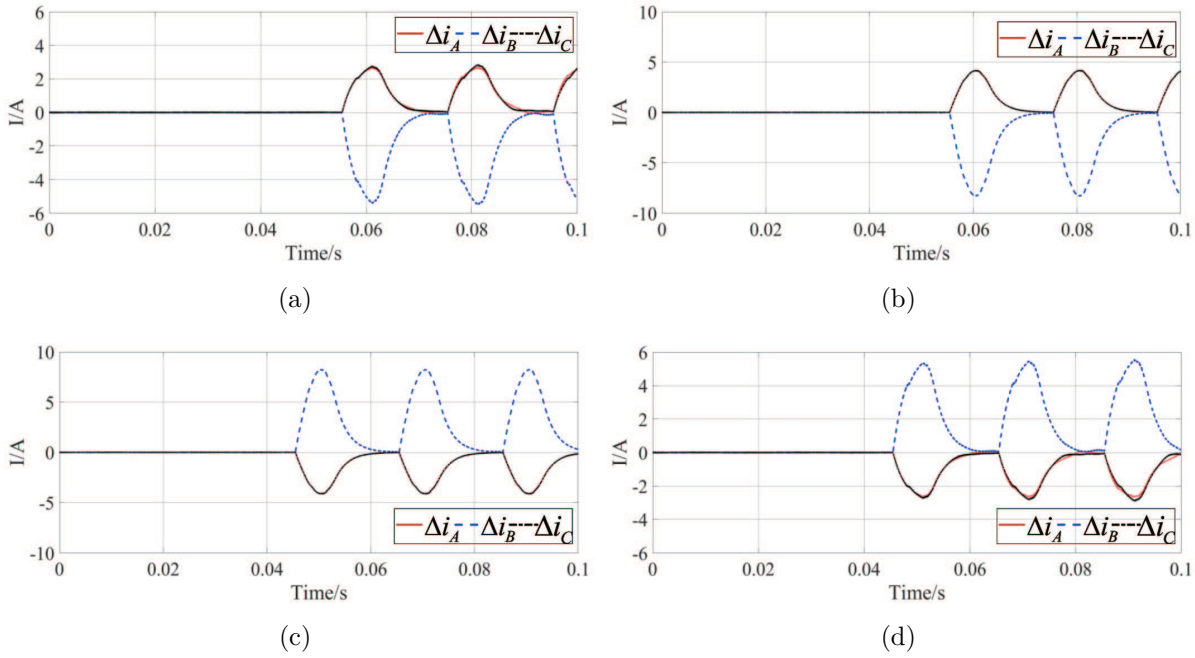


FIGURE 5. Single-tube open fault of phase-B bridge leg: (a) The three-phase current residual of S_{b1} open fault; (b) the three-phase current residual of S_{b2} open fault; (c) the three-phase current residual of S_{b3} open fault; (d) the three-phase current residual of S_{b4} open fault

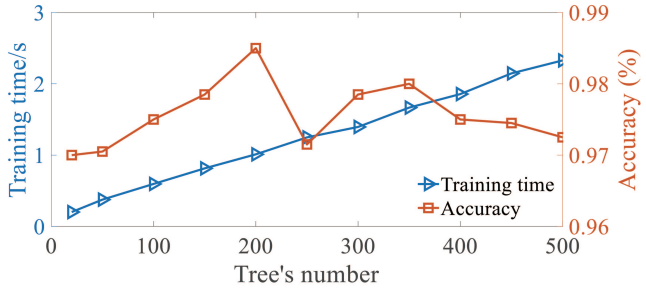


FIGURE 6. Training time and diagnostic accuracy for different numbers of decision trees in Case A

matrix, diagonal elements stand for the number of samples that the model correctly classifies for each state, while off-diagonal elements stand for incorrectly classified samples. The results show that the accuracy of normal state, four IGBTs in phase-A, S_{b3}, S_{b4}, S_{c1}, S_{c2} and S_{c3} is 100%. The accuracy of S_{b1} is the lowest, 91.7%, which is misclassified as S_{b2}. The overall accuracy of Case A is 98.5%.

4.3. **Comparison of fault diagnosis methods under different noise conditions.** In an effort to prove the practicability of the hybrid fault diagnosis method, different Gaussian white noises are added to the original signal. Noise interference with SNR of 20 dB, 10 dB and 1 dB is injected into the testing sample data. In an effort to prove the superiority of the proposed method, it is in comparison with traditional machine learning methods (ELM and SVM). At the same time, compared with the data-driven method, the three-phase current is utilized as the original signal for fault diagnosis. Three-phase current-WT-RF, three-phase current residual-RF, three-phase current residual-WT-ELM,

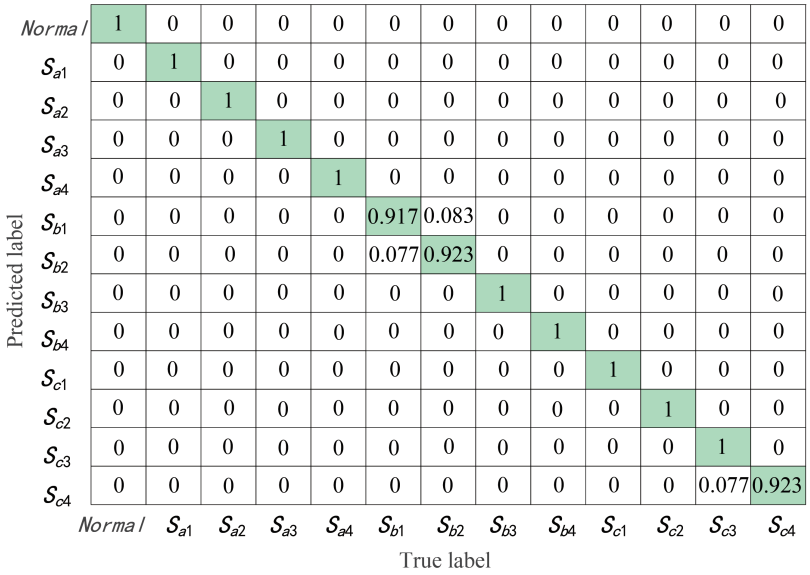


FIGURE 7. Confusion matrix of results on datasets of Case A

TABLE 3. Accuracy of fault identification under noise conditions in Case A

Original dataset	Feature extraction	Diagnosis method	Accuracy rate			
			Normal	SNR = 20 dB	SNR = 10 dB	SNR = 1 dB
The current residual	/	RF	97%	96.5%	95%	93%
The current		RF	98%	97.5%	94%	90.5%
The current residual	WT	ELM	95%	93%	92%	89.5%
		SVM	97.5%	94%	93.5%	92.5%
		RF	**98.5%**	**98%**	**96.5%**	**94%**

three-phase current residual-WT-SVM and three-phase current residual-WT-RF are compared under three SNR conditions, and the accuracy of fault identification is shown in Table 3. The noise has a great influence on diagnostic accuracy, and the comparison experiment under the three noise conditions shows that the hybrid method based on model-based and data-driven fault diagnosis has the highest accuracy and good robustness.

5. **Hardware in the Loop Experiment.**

5.1. **Experiment platform introduction.** To further prove the effectiveness of the hybrid fault diagnosis method, StarSim MT3200 test platform is used to conduct real-time simulation experiments. Figure 8 shows the experimental platform, which includes host computer, oscilloscope, interface box, StarSim MT3200 and dSPACE1202. The NPC three-level inverter circuit is constructed by a real-time hardware in-the-loop system. The control algorithm model is completed by dSPACE1202. The real-time simulator and controller are connected through actual physical I/O, which transmit signals.

5.2. **Experiment setup.** The normal state and 12 kinds of single-tube open-circuit fault state are studied using the experimental platform. The three-phase current residual signals could be obtained by experiments of hardware circuits with different loads. Four different loads were used for the experiment: 1) R = 5.0 Ω, L = 12 mH; 2) R = 5.0 Ω, L = 20 mH; 3) R = 5.5 Ω, L = 12 mH; 4) R = 5.5 Ω, L = 20 mH. In each group of loads, take one electrical period as a group and extract 20 groups of current residual signals. A total

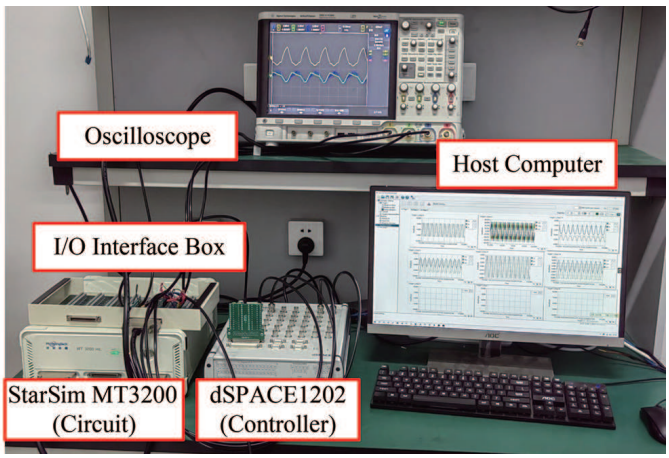


FIGURE 8. Experimental platform based on StarSim HIL

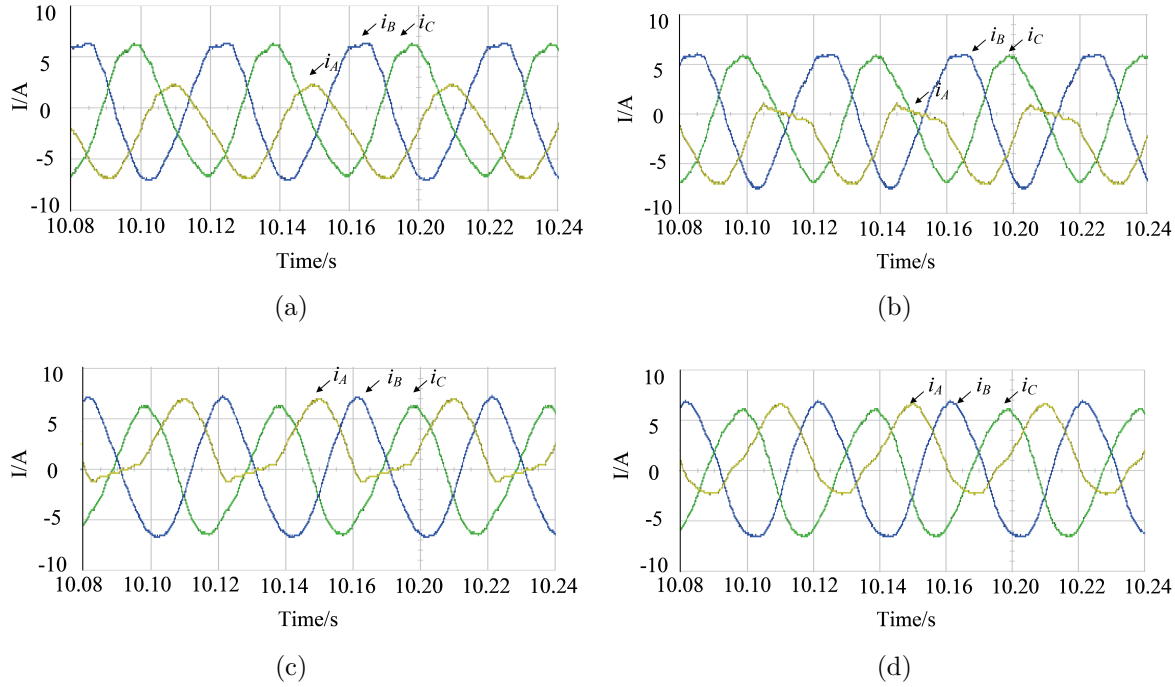


FIGURE 9. The three-phase current waveform at single-tube open fault of phase-A bridge leg conditions: (a) S_{a1} open fault; (b) S_{a2} open fault; (c) S_{a3} open fault; (d) S_{a4} open fault

of 1040 (52×20) three-phase current residual signals were utilized to construct dataset Case B.

5.3. **Results and discussion.** For load of R = 5.5 Ω and L = 12 mH, the single-tube open fault of phase-A bridge leg conditions are taken as examples. The three-phase current waveform is shown in Figure 9.

In the dataset Case B, 740 groups of samples were randomly selected as the training set, and 300 groups of samples were selected as the testing set.

Training time and diagnostic accuracy for different numbers of decision trees in Case B are shown in Figure 10. If the number is too small, the diagnostic accuracy is influenced due to the underfitting. If the number is too large which leads to the overfitting problem,

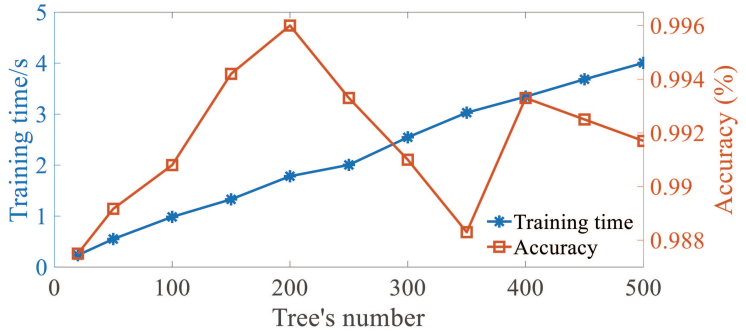


FIGURE 10. Training time and diagnostic accuracy for different numbers of decision trees in Case B

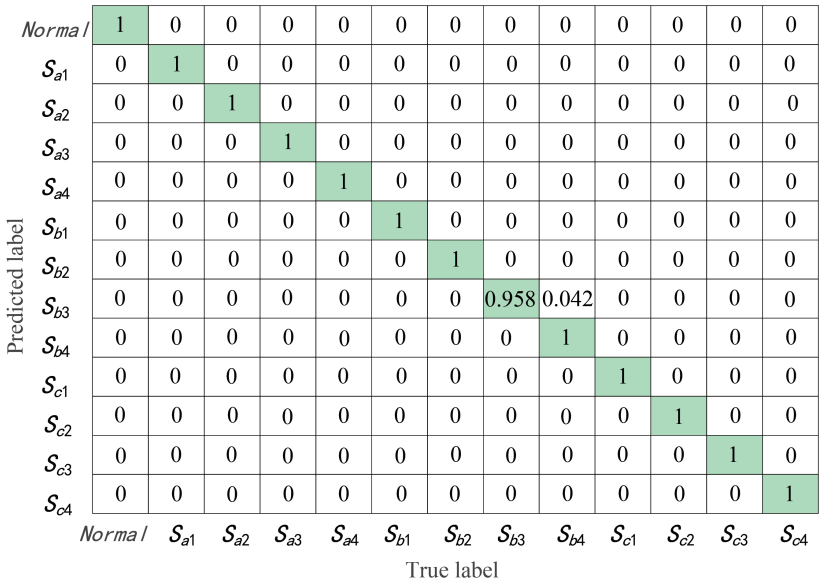


FIGURE 11. Confusion matrix of results on datasets of Case B

the diagnostic accuracy decreases. The diagnostic accuracy is positively correlated with the number of decision trees between 10 and 200, and at 200, the diagnostic accuracy reaches the maximum of 99.6%. When the number exceeds 250, the diagnostic accuracy does not increase. In the experiment, the best number of decision trees is 200.

Figure 11 shows the confusion matrix of results on datasets of Case B. The results show that the accuracy of normal state, four IGBTs in phase-A, S_{b1}, S_{b2}, S_{b4} and four IGBTs in phase-C is 100%. The accuracy of S_{b3} is the worst, 95.8%, and all of them are misclassified as S_{b4}. The normal state has 100% accuracy and no other fault states are misclassified as normal state, which indicates that the proposed method can precisely distinguish whether there is any failure occurring in Case B. The overall accuracy of Case B is 99.6%.

Table 4 shows the diagnostic accuracy under different noise conditions in Case B. As can be seen from Table 4, under the same conditions of feature extraction method, by comparing the source of the original dataset, we find that the hybrid diagnosis method has higher accuracy than the model-based method under normal and different noise conditions. In the experiment of the StarSim HIL platform, the proposed method still has the highest accuracy and good robustness.

TABLE 4. Accuracy of fault identification under noise conditions in Case B

Original dataset	Feature extraction	Diagnosis method	Accuracy rate			
			Normal	SNR = 20 dB	SNR = 10 dB	SNR = 1 dB
The current residual	/	RF	98.3%	97.6%	96.3%	93.3%
The current		RF	99%	98.3%	97.3%	93.6%
The current residual	WT	ELM	97.3%	95%	93.3%	92%
		SVM	98.6%	97.3%	95.6%	93%
		RF	**99.6%**	**98.6%**	**98%**	**94.3%**

6. **Conclusions.** In this paper, a hybrid model/data-driven method for the IGBTs open-circuit fault diagnosis is proposed. The current residuals in different fault types at different load conditions are collected as the original sample dataset. The three-phase current residuals are decomposed by WT. Then, the 33 fault features of current residuals are extracted which is normalized by PVM. Use data-driven method based on RF method to train the failure diagnostic model. The diagnostic accuracy rates of simulation and StarSim HIL experiments are 98.5% and 99.6%, respectively.

The results indicate that the proposed method has good diagnostic performance under different noise conditions compared with other machine learning methods. Compared with the model-based method, the proposed method can identify and locate IGBT faults under different loads and has good adaptability. The proposed method has higher diagnostic accuracy than the traditional data-driven method, and has good robustness. The main research work in the future is to discover more fault features and improve the diagnostic accuracy.

Acknowledgment. This work is partially supported by the National Natural Science Foundation of China (Grant No. 62222307, No. 61973140) and the Natural Science Foundation of Jiangsu Province (Grant No. BK20211235). The authors also appreciatively acknowledge the helpful suggestions and comments of the reviewers, which have enhanced the presentation.

REFERENCES

[1] H. Wang, J. Lu, Z. Wu and M. Yu, Diffusion path of leading technology in new energy industry based on fsQCA – Take China's solar energy industry as an example, *Technology Analysis & Strategic Management*, pp.1-16, 2023.

[2] S. Luo, F. Wu and K. Zhao, Modified single-carrier multilevel SPWM and online efficiency enhancement for single-phase asymmetrical NPC grid-connected inverter, *IEEE Transactions on Industrial Informatics*, vol.16, no.5, pp.3157-3167, 2019.

[3] P. Kakosimos and H. Abu-Rub, Predictive control of a grid-tied cascaded full-bridge NPC inverter for reducing high-frequency common-mode voltage components, *IEEE Transactions on Industrial Informatics*, vol.14, no.6, pp.2385-2394, 2017.

[4] U.-M. Choi, J.-S. Lee, F. Blaabjerg and K.-B. Lee, Open-circuit fault diagnosis and fault-tolerant control for a grid-connected NPC inverter, *IEEE Transactions on Power Electronics*, vol.31, no.10, pp.7234-7247, 2016.

[5] J. Qiu, Y. He, C. Lei, Q. Jiao and J. Liu, An improved LMSVM method for leakage current suppression and neutral-point voltage control in transformerless NPC three-level inverters, *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol.10, no.3, pp.3100-3113, 2021.

[6] V. Smet, F. Forest, J.-J. Huselstein, F. Richardeau, Z. Khatir, S. Lefebvre and M. Berkani, Ageing and failure modes of IGBT modules in high-temperature power cycling, *IEEE Transactions on Industrial Electronics*, vol.58, no.10, pp.4931-4941, 2011.

[7] J. Liang, K. Zhang, A. Al-Durra, S. Muyeen and D. Zhou, A state-of-the-art review on wind power converter fault diagnosis, *Energy Reports*, vol.8, pp.5341-5369, 2022.

[8] W. Huang, J. Du, W. Hua, W. Lu, K. Bi, Y. Zhu and Q. Fan, Current-based open-circuit fault diagnosis for PMSM drives with model predictive control, *IEEE Transactions on Power Electronics*, vol.36, no.9, pp.10695-10704, 2021.

[9] W. Huang, J. Du, W. Hua, K. Bi and Q. Fan, A hybrid model-based diagnosis approach for open-switch faults in PMSM drives, *IEEE Transactions on Power Electronics*, vol.37, no.4, pp.3728-3732, 2022.

[10] C. Liu, B. Jiang, R. J. Patton and K. Zhang, Hierarchical-structure-based fault estimation and fault-tolerant control for multiagent systems, *IEEE Transactions on Control of Network Systems*, vol.6, no.2, pp.586-597, 2018.

[11] Z. Huang, L. Dong and X. Yao, Reversed trend comprehensive assessment-based inverter fault diagnosis under various complex interferences in microgrid, *Measurement*, 114301, 2024.

[12] C. Yong, J.-J. Zhang and Z.-Y. Chen, Current observer-based online open-switch fault diagnosis for voltage-source inverter, *ISA Transactions*, vol.99, pp.445-453, 2020.

[13] Z. Gao and X. Liu, An overview on fault diagnosis, prognosis and resilient control for wind turbine systems, *Processes*, vol.9, no.2, 300, 2021.

[14] Z. Huang, Z. Wang and C. Song, Complementary virtual mirror fault diagnosis method for microgrid inverter, *IEEE Transactions on Industrial Informatics*, vol.17, no.11, pp.7279-7290, 2021.

[15] S. Tang, H. Wang, W. Wang and C. Liu, A fault diagnosis method for active power factor correction power supply based on seagull algorithm optimized kernel-based extreme learning machine, *International Journal of Circuit Theory and Applications*, vol.52, no.3, pp.1116-1135, 2024.

[16] V. R. Sonawane and S. B. Patil, Track and hunt metaheuristic based deep neural network based fault diagnosis model for the voltage source inverter under varying load conditions, *Advances in Engineering Software*, vol.177, 103414, 2023.

[17] F. Yao, Q. Liu, B. Ji, W. Zhang, Y. Wu and Y. Tang, Open circuit fault diagnosis of three-phase inverter based on SR-WOA-ELM, *International Journal of Circuit Theory and Applications*, vol.52, no.6, pp.2786-2802, 2024.

[18] Y. Wang, Z. Pan, X. Yuan, C. Yang and W. Gui, A novel deep learning based fault diagnosis approach for chemical process with extended deep belief network, *ISA Transactions*, vol.96, pp.457-467, 2020.

[19] J. A. P. Sánchez, D. U. Campos-Delgado, D. R. Espinoza-Trejo, A. A. Valdez-Fernández and C. H. De Angelo, Fault diagnosis in grid-connected PV NPC inverters by a model-based and data processing combined approach, *IET Power Electronics*, vol.12, no.12, pp.3254-3264, 2019.

[20] Y. Xia, Y. Xu and B. Gou, Current sensor fault diagnosis and fault-tolerant control for single-phase PWM rectifier based on a hybrid model-based and data-driven method, *IET Power Electronics*, vol.13, no.18, pp.4150-4157, 2020.

[21] X. Wang, S. Wang, S. An and H. Liu, An adaptive disturbance suppression based fault-tolerant control approach against the control surface faults, *International Journal of Innovative Computing, Information and Control*, vol.19, no.1, pp.213-227, 2023.

[22] L. Kou, C. Liu, G.-W. Cai, J.-N. Zhou, Q.-D. Yuan and S.-M. Pang, Fault diagnosis for open-circuit faults in NPC inverter based on knowledge-driven and data-driven approaches, *IET Power Electronics*, vol.13, no.6, pp.1236-1245, 2020.

[23] J.-H. Park, D.-H. Kim, S.-S. Kim, D.-J. Lee and M.-G. Chun, C-ANFIS based fault diagnosis for voltage-fed PWM motor drive systems, *IEEE Annual Meeting of the Fuzzy Information, Processing of NAFIPS'04*, Banff, AB, Canada, vol.1, pp.379-383, 2004.

[24] R. Yan, Z. Shang, H. Xu, J. Wen, Z. Zhao, X. Chen and R. X. Gao, Wavelet transform for rotary machine fault diagnosis: 10 years revisited, *Mechanical Systems and Signal Processing*, vol.200, 110545, 2023.

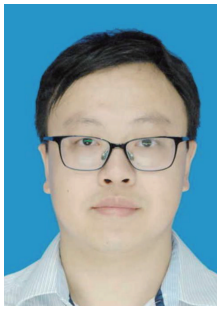
[25] S. Liu, X. Qian, H. Wan, Z. Ye, S. Wu and X. Ren, NPC three-level inverter open-circuit fault diagnosis based on adaptive electrical period partition and random forest, *Journal of Sensors*, vol.2020, pp.1-18, 2020.

[26] L. Breiman, Random forests, *Machine Learning*, vol.45, pp.5-32, 2001.

[27] K. Dhibi, M. Mansouri, M. Hajji, K. Bouzrara, H. Nounou and M. Nounou, A novel hybrid methodology for fault diagnosis of wind energy conversion systems, *Energy Reports*, vol.9, pp.5362-5371, 2023.

[28] S. Yang, P. Yang, H. Yu, J. Bai, W. Feng, Y. Su and Y. Si, A 2DCNN-RF model for offshore wind turbine high-speed bearing-fault diagnosis under noisy environment, *Energies*, vol.15, no.9, 3340, 2022.

Author Biography



Weilin Yang received his B.Eng. degree in Machine Design & Manufacture and Their Automation from University of Science and Technology of China, Hefei, China, in 2009, and the Ph.D. degree in Mechanical Engineering from City University of Hong Kong, Hong Kong SAR in 2013.

He was a postdoctoral researcher at Masdar Institute of Science and Technology (now Khalifa University), Abu Dhabi, UAE, 2013-2016. He was a research engineer of General Electric (GE) Global Research, Shanghai, 2016-2017. He joined Jiangnan University in July 2017, where he is currently an Associate Professor. His research interests include modeling and control of energy systems, robust model predictive control, and data-driven control.

Chao Zhang received his B.S. degree in Electrical Engineering and Automation from Jiangnan University, Wuxi, China, in 2022. He is currently pursuing the M.S. degree in Electrical Engineering with Jiangnan University, Wuxi, China.

His current research interests include machine learning, hybrid systems and fault diagnosis.

Dezhi Xu received the Ph.D. degree in Control Theory and Control Engineering from Nanjing University of Aeronautics and Astronautics, China, in 2013.

He was a Visiting Fellow with the Department of Biomedical Engineering, City University of Hong Kong, China, from 2018 to 2019. He is currently a Professor and Doctoral Supervisor with the Southeast University. His research interests include data-driven control, fault diagnosis and fault-tolerant control, multi-agent systems and cyber-physical systems, technologies of renewable energy, motor control, and smart grid. Dr. Xu was supported by the National Natural Science Fund for Excellent Young Scientists Fund Program in 2022. He was a recipient of the First Class Prize of Science and Technology Progression from the China General Chamber of Commerce in 2016, and the Best Young Scholar of Jiangnan University in 2022. He was a Guest Editor for the International Journal of Innovative Computing, Information and Control and the Electric Power. He currently serves as an Editorial Board Member for the International Journal of Innovative Computing, Information and Control, the Electric Power, the Electrotechnical Application and the Electrical Engineering. He is a Committee Member of the Association of Energy Internet, and Trusted Control in Chinese Association of Automation (CAA), and the Energy Storage in China Renewable Energy Society (CRES).

Yujian Ye received the B.Eng. (Hons) degree in Electrical and Electronic Engineering from Northumbria University, Newcastle Upon Tyne, U.K., in 2011, and the M.Sc. degree with Distinction in Control Systems and the Ph.D. degree from Imperial College London, London, U.K., in 2013 and 2017, respectively. He performed Postdoctoral research also with Imperial College London from 2016 to 2020, and then joined Southeast University, Nanjing, China with Associate Professorship in 2021.

He is currently a Professor with Young Endowed Chair Honor with the School of Electrical Engineering at Southeast University and an Honorary Lecturer at Imperial College London. His current research interests include development and application of novel data analytics and artificial intelligence techniques in low-carbon energy-transportation-information systems modeling, analysis and control, and optimization of economics of power system operation and planning. He serves as the Associate Editor of several prestigious and reputable international journals, including IEEE Transactions on Smart Grid, IEEE Transactions on Industrial Informatics and IEEE Transactions on Industry Applications. He also serves as the Subject Editor of "Emerging Technologies in Smart Grids" for IET Smart Grid, and a Young Editorial Board Member of Applied Energy and cross-disciplinary journal Nexu.