Electrical Power and Energy Systems 142 (2022) 108310

Contents lists available at ScienceDirect

International Journal of Electrical Power and Energy Systems

journal homepage: www.elsevier.com/locate/ijepes

Semi AI-based protection element for MMC-MTDC using local-measurements

Ning Tong [a], Zhenjie Tang [a], Yu Wang [a], Chun Sing Lai [a,b,*], Loi Lei Lai [a,*]

[a] Department of Electrical Engineering, School of Automation, Guangdong University of Technology, Guangzhou 510006, China
[b] Brunel Interdisciplinary Power Systems Research Centre, Department of Electronic and Electrical Engineering, Brunel University London, London UB8 3PH, UK

ARTICLE INFO

Keywords:
MMC-MTDC
Traveling wave
Artificial intelligence
Mains protection

ABSTRACT

The multi-terminal HVDC system based on the modular multilevel converter (MMC-MTDC) is a promising technique for flexible power transmissions to multiple regions. As such a system is quite sensitive to DC faults, there is an acute need to propose a protection element that can trip the local DC circuit breaker (CB) within several milliseconds once there is an internal DC line fault. However, the existing main protection scheme faces a dilemma balancing selectivity and sensitivity. To solve this problem, a novel semi artificial-intelligence (AI) based protection element is proposed, including a start-up criterion and a fault-identification criterion. The start-up criterion is based on the propagation characteristics of the initial fault-induced surge. To enhance the real-time performance of the protection element, it will not trip the fault-identification process unless the fault is identified as a forward one. The fault-identification criterion is based on artificial intelligence (AI), and further determines whether the forward fault is internal, which only works if the start-up criterion trips. Simulation results indicate that the proposed protection element has satisfactory speed, sensitivity, and selectivity against internal DC faults and is quite secure under external fault conditions. The impact of disturbances, such as the white noise, abnormal samplings, etc., on the security of the proposed protection element is also discussed.

1. Introduction

Modular multilevel converter-based multi-terminal high-voltage direct-current transmission systems (MMC-MTDC) have great flexibility for long-distance bulk-power transmission, integrations of renewable energy, connections of asynchronous power grids, etc. [1–3]. To build a cost-effective MMC-MTDC, a large number of half-bridge MMCs (HBMMCs) are employed, which makes the system quite vulnerable to DC faults. With the development of the high-speed DC Circuit Breaker (CB) [4], faults in the MMC-MTDC are allowed to be cleared without a permanent loss of power transmission. For this purpose, a protection element should be developed to trip the DC CB before the blocking of the HBMMCs, which is merely 5–6 ms in most scenarios [5,6].

Protection elements using both the local measurement and the remote-end measurements, such as the current differential protection, show outstanding performance in conventional AC systems and several MMC-MTDCs with very short DC transmission lines, such as the Nan'ao MMC-MTDC project [7]. However, such a type of protection element has limited application perspectives on long-distance DC transmission lines because of the limitations of the communication volume [8] and the long time-delay that required to avoid the abnormal differential current due to the discharge of the shunt capacitance in external fault scenarios [9,10]. From recent advances, Reference [11] proposes a pilot protection element based on the polarities of the superimposed current in both terminals using the wavelet transform. In Reference [12], a pilot protection element is developed based on the arrival-time-difference of the forward traveling wave and the backward traveling wave. The time difference is almost zero in internal-fault scenarios while relatively high under external-fault conditions. Some have also proposed methods based on the S-transform [13], the cosine similarity [14], etc. In the above works, tremendous contributions have been made to eliminate the use of remote end measurements, communication, and computation. However, they cannot work if there is a communication failure.

Due to the above limitations, there is an acute need for a fast

* Corresponding authors at: Department of Electrical Engineering, School of Automation, Guangdong University of Technology, Guangzhou, 510006, China (L.L. Lai). Brunel Interdisciplinary Power Systems Research Centre, Department of Electronic and Electrical Engineering, Brunel University London, London, UB8 3PH, UK (C.S. Lai).

E-mail addresses: tongning@gdut.edu.cn (N. Tong), tzj789@qq.com (Z. Tang), yuwang@gdut.edu.cn (Y. Wang), chunsing.lai@brunel.ac.uk (C.S. Lai), l.l.lai@ieee.org, l.l.lai@gdut.edu.cn (L.L. Lai).

https://doi.org/10.1016/j.ijepes.2022.108310
Received 12 December 2021; Received in revised form 21 March 2022; Accepted 2 May 2022
Available online 11 May 2022
0142-0615/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Nomenclature

Abbreviations

MMC	Modular Multilevel Converter
MTDC	Multi-Terminal Direct Current
HBMMC	Half-Bridge Modular Multilevel Converter
DC	Direct Current
CB	Circuit Breaker
AI	Artificial Intelligence
HVDC	High-Voltage DC systems
CLR	Current-Limiting Reactor
SNR	Signal-Noise Ratio
CNN	Convolutional Neural Network
AMF	Adaptive Median Filtering

Symbols

W_f/E_{1f}	The wave-front/energy of forward surges
W_b/E_{1b}	The wave-front/energy of backward surges
Z_1	The aerial-modal surge impedance
$\Delta u_1/\Delta i_1$	The aerial-modal superimposed voltage/current
E_{norm}	The surge energy threshold
E_{set1}	The surge energy setting value
r_{set}	The surge energy reliability coefficient
c_{set}	The surge energy ratio confidence coefficient
S_0	The neighborhood of AMF with a center of x_m
M_{min}	The minimum in a neighborhood
M_{max}	The maximum in a neighborhood
M_{med}	The median in a neighborhood
M_1/M_2	The median substruction coefficient
N_1/N_2	The median comparison coefficient
$\mathbf{Y}/\mathbf{Y}^*$	The original/z-score normalized data set
μ/σ	The mean/standard deviation
U_a/U_a^*	Local/normalized aerial-modal voltage phasor
U_z/U_z^*	Local/normalized zero-modal voltage phasor
I_a/I_a^*	Local/normalized aerial-modal current phasor
I_z/I_z^*	Local/normalized zero-modal current phasor
$\mathbf{C}/\mathbf{C}^*$	Local/normalized feeding data set for CNN
$X_{ij}^{(m,n)}/x_{ij}^{(m,n)}$	Convolution kernel/value of location (m,n) in the i-th convolutional layer, j-th feature map
P/Q	The height/width of convolution kernels
$u(w,w)$	A square window for pooling
$x_i^{w\times w}$	The selected pooling area
$\mathbf{V}_F$	The flatten vector
$\mathbf{V}_{FC}$	The output vector of fully connected layer
$w/b/\lambda$	The weight/bias/connection coefficient vector
h/H	The neuron/maximal neurons in *softmax* layer
T_{total}	The total time to isolate a DC fault
t_s	The delay for surge energy calculation
t_u	The delay of the start-up criterion
t_p	The time window of the proposed protection
t_{cb}	The delay for DC CB to interrupt a DC fault
t_{al}	The additional delay for protection hardware
r	The radius of median filtering
R	The maximum order of AMF
f_p	The sampling frequency

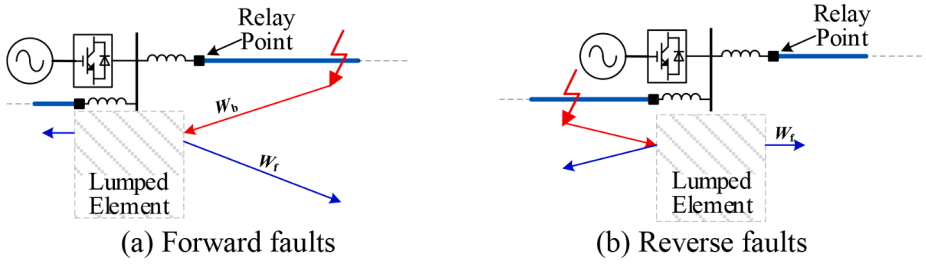


Fig. 1. Traveling wave propagation with reflection/refraction.

protection element based merely on local measurements. In this field, many have made significant contributions. In [15], a high-speed traveling wave protection element is proposed by using very short time windows and high-frequency components, and the theory has been successfully applied in the Zhangbei four-terminal MMC-MTDC. In [16], a setting-less local measurement-based protection element is proposed using model recognition, based on which internal forward faults, external forward faults, and external reverse faults are classified. In [17], boundary characteristics are employed to identify internal faults using the wavelet transform. An improved local measurement-based protection element is proposed by comparing the attenuation rate of initial voltage traveling waves in both the internal and external fault scenarios in [18], and the theory is also based on the configuration of boundary components, i.e., the fault current-limiting reactor (CLR) on the DC line-terminals. Protection elements based on the harmonic current [19], high/low frequency energy ratio [20], transverse differential current [21], nature frequency [22], and similar boundary characteristics [23,24] are also proposed and show great performance. However, the universal shortfall is that the configuration of the CLR has a significant impact on the performance of the protection element, whereas the value of CLRs is not set to meet the sensitivity requirement of the protection but to make a compromise between the current-limiting capability and the dynamic behavior of the MMC-MTDC [25]. Taking the Zhangbei Four-terminal MMC-MTDC as an example, the typical value of CLRs is no more than 150 mH. As long as a certain algorithm is employed to extract the fault feature from local measurements artificially, there should be a balance between sensitivity and selectivity. Besides, the impact of disturbances, such as the white noise and abnormal samplings, on the performance of the protection element merits further investigation.

Artificial intelligence (AI) and deep learning have witnessed significant development in recent years. As a popular form of deep learning, the convolutional neural network (CNN) has made notable contributions in fields like pattern recognition, face recognition, etc. In light of pioneer works, AI-based techniques are applied to identify grounding faults [26], locate different types of faults in the HVDC with extremely high resistive tolerance [27], etc. Arranging the measurement series into two-dimensional images can help achieve satisfactory performance. However, speed is quite critical to apply AI-based techniques to the MMC-MTDC because the protection element has to isolate DC faults within just a few milliseconds. For this reason, the framework of the AI should be as simple as possible to meet the requirement for high real-time

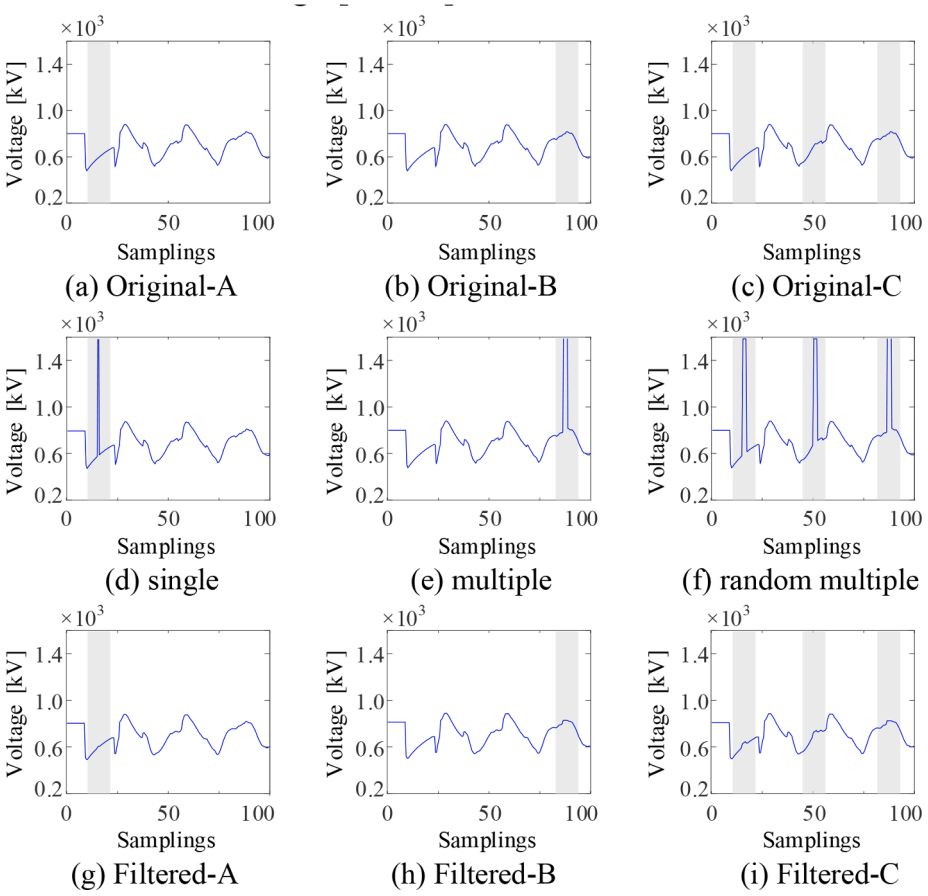


Fig. 2. AMF against different numbers of abnormal samplings.

performance.

The contribution of this paper is to propose a semi AI-based protection element using merely the local measurement. The proposed protection element should feature: 1) high sensitivity – to protect 500-kV systems, a resistive tolerance of 300 Ω is needed under the single-line-to-ground fault condition [28], regardless of the value of the line CLR; 2) fast speed – the total time cost from the fault-arrival to the complete fault-interruption by DC CB should be<5 ms; 3) high security – the operation of the proposed protection element should be immune to disturbances like white noises and abnormal samplings; and 4) high real-time performance – the structure, scenario, setting principle, etc. of the AI should be simple enough and have industrial implementation merits.

To achieve this, a measurement-based start-up criterion is proposed to determine the initial fault direction, and reverse faults are firstly excluded. Then, an AI-based fault-identification criterion using CNN is proposed to determine whether the forward fault is internal or external. Besides, adaptive median filtering (AMF) is introduced to make the proposed protection element robust to different types of disturbances. By the above means, the real-time performance of the proposed protection element is enhanced because: 1) the proposed protection element is semi AI-based since the fault-identification criterion does not work until the start-up criterion detects a forward fault, thus reducing the number of fault scenarios that have to be included in the training set; and 2) by erasing disturbances like the abnormal samplings using the AMF, the training set can be further simplified since there is no need to consider abnormal scenarios.

The remainder of this paper is organized as follows. In Section 2, theoretical studies are conducted to describe the fundaments of the start-up criterion and the AI-based fault-identification criterion. In Section 3, the methodology is illustrated. Flowcharts and sequence diagrams are made to show the behavior of the proposed semi AI-based protection element under different fault scenarios. In Section 4, simulation cases are performed to evaluate the security and sensitivity of the proposed protection element. The performance of the above two criteria is examined under both pole-to-pole faults and single-line-to-ground faults. Section 5 quantitatively discusses the merit of the proposed method to work as a main protection element, comparing with several state-of-the-art methods, and the impact of AMF on the accuracy of the fault identification. Finally, contributions are concluded in Section 5.

2. Theoretical analysis

2.1. Start-up criterion

Unlike the scenario in an AC system, for a DC system in the normal operating condition, electrical quantities like voltage, current, etc. are constants in an ideal situation when harmonics and noises are not considered. For this reason, a start-up criterion should be developed to calibrate the arrival of a disturbance since it is difficult to normalize a series of constant values and then use them as the input of a CNN. As shown in Fig. 1(a), assume that a fault on the DC line is a forward one for the protection element equipped at the relay point. Due to the discontinuity of surge impedance at the line terminal, a reflected forward surge W_f is generated from the initial backward surge W_b.

As the aerial-modal surge will appear in both the single-line-to-ground fault and the line-to-line fault, the following criterion is employed to represent the arrival of W_b:

$$E_{1b} > E_{set1} \tag{1}$$

where E_{1b} is the accumulated energy of the backward aerial-modal surge. The calculation of E_{1b} is given by Equation (2), where t_s is the

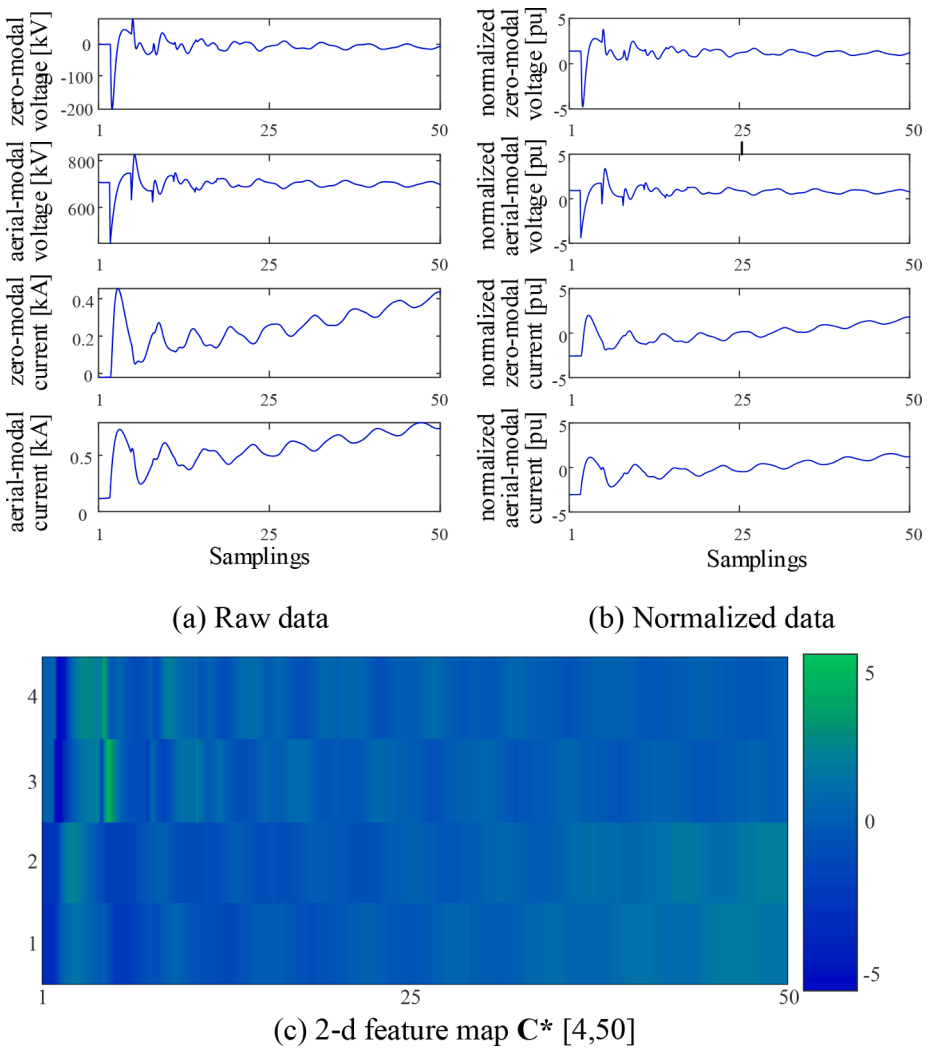


Fig. 3. Input layer data normalized and constructed.

time window; Z_1 is the aerial-modal surge impedance; Δu_1 and Δi_1 are the superimposed aerial-modal voltage and the superimposed aerial-modal current, respectively. E_{set1} is the setting value of E_{1b}.

$$E_{1b} = \int_{t-t_s}^{t} |\Delta u_1(t) - Z_1 \Delta i_1(t)| dt \tag{2}$$

Using the above criterion, forward faults can be detected because W_b represents the initial arrival of a fault-induced surge. However, if the fault is a reverse one, as shown in Fig. 1(b), the initial arrival of the fault-induced surge should only have a forward component because the initial surge is a refracted one from the reverse fault. If Equation (1) is still employed to detect the reverse fault, the arrival time of the fault-induced surge will be inaccurate. For this reason, Equation (3) is used to determine the arrival of W_f:

$$E_{1f} > E_{set1} \tag{3}$$

where E_{1f} is the accumulated energy of the forward aerial-modal surge calculated by Equation (4).

$$E_{1f} = \int_{t-t_s}^{t} |\Delta u_1(t) + Z_1 \Delta i_1(t)| dt \tag{4}$$

Combing Equations (1) and (3), a fault can be detected once any of them is true. To make them sensitive enough against all types of disturbances, E_{set1} should be set low enough. In this paper, E_{set1} is set according to the following equation:

$$E_{set1} = E_{norm} \cdot r_{set} \tag{5}$$

where E_{norm} is a threshold to exclude normal system operating conditions such as the unblocking/blocking of the DC system, the fast change of the DC power, etc. As a reliability coefficient, r_{set} is set higher than 1 to ensure reliability. However, the direction of the fault cannot be determined by this means as E_{set1} is set low enough for higher sensitivity. Thus, there are plenty of scenarios when Equations (1) and (3) are true at the same time. To prevent this, the ratio of surge energy is employed to obtain a determined fault direction. Note that in Fig. 1(a), the forward surge is generated by the reflection of the initial surge. For this reason, the calculated energy by Equation (2) will be higher than that by (4). Thus, within a short time window after the arrival of a forward fault, we have $0 < E_{1f}/E_{1b} < 1$. In contrast, under reverse faults, the calculated energy E_{1b} for the backward surge W_{1b} is almost zero because there is no backward surge at the time of fault-arrivals. For this reason, E_{1f}/E_{1b} will be pretty high, which is ideally nearly infinite. The energy ratio E_{1f}/E_{1b} is multiplied by a confidence coefficient c_{set} (higher than 1) to improve the sensitivity. For the following equations, the fault is determined to be a forward one if Equation (6) is true or reverse if Equation (7) is true.

$$0 < E_{1f}/E_{1b} < 1 \times c_{set} \tag{6}$$

$$E_{1f}/E_{1b} \geqslant 1 \times c_{set} \tag{7}$$

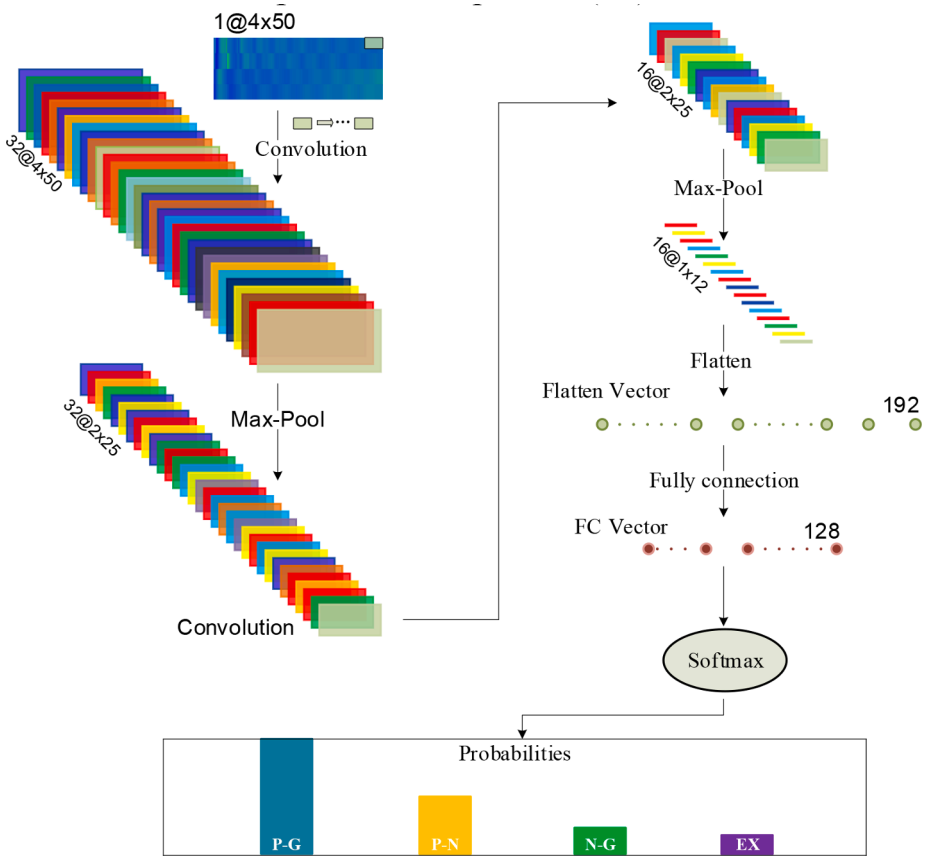


Fig. 4. CNN framework.

2.2. Data pre-processing

If a fault is determined to be a forward one, a series of local measurements will be recorded from the fault arrival time to the end of the time window. The obtained data will be fed into the CNN for further identification. The pre-processing of the input data for the CNN includes the adaptive median filtering (AMF) and the data normalization, as follows:

1) The AMF

Conventionally, in the procedures of sampling, measurement, converting, etc., two types of disturbances may be included, which are: 1) singular points caused by abnormal samplings, bit errors, packet losses, etc.; and 2) Gaussian background noises superimposed with the measurements. For protection elements, the former one is more challenging.

The core philosophy of the median filtering is to compare the value of each point in a certain range of neighborhood, calculate the median value as a new value, and then replace the former value with the calculated median value if there is a singular point located at the center of this neighborhood. For conventional median filtering algorithm, the neighborhood is set to: $S_0 \in (x_{m\text{-}r}, x_{m\text{-}r+1}, \ldots, x_{m\text{-}1}, x_m, x_{m+1}, \ldots, x_{m+r\text{-}1}, x_{m+r})$, where x_m is the center point of S_0. The indicator r is a preset-constant. Within this neighborhood, the maximum M_{max}, the minimum M_{min}, and the median M_{med} are calculated. For each x_m, Equation (8) is checked. If true, then output the original value x_m. Otherwise, the corresponding point is replaced by M_{med}.

$$M_{min} < x_m < M_{max} \tag{8}$$

The AMF is introduced according to the following steps:

a) First, a default range of neighborhood is set to $S_0 \in (x_{m\text{-}i}, x_{m\text{-}i+1}, \ldots, x_{m\text{-}1}, x_m, x_{m+1}, \ldots, x_{m+i\text{-}1}, x_{m+i})$, where x_m is the center point of S_0. The indicator i is initialized to $i = 0$. Within this neighborhood, the maximum M_{max}, the minimum M_{min}, and the median M_{med} are calculated, based on which two coefficients M_1 and M_2 are calculated:

$$\begin{cases} M_1 = M_{med} - M_{min} \\ M_2 = M_{med} - M_{max} \end{cases} \tag{9}$$

b) Second, substitute M_1 and M_2 into Equation (10). Go to step c) if the result is true. Otherwise, replace i with $i + 1$ (let the maximum value of i be R) and the neighborhood S_0 with the new one. Then, go to step a) and repeat the above process.

$$(M_1 > 0) \& (M_2 < 0) \tag{10}$$

c) Finally, compare the center point x_m with the maximum M_{max} and the minimum M_{min} according to Equations (11) and (12). Output the center point x_m if Equation (12) is true. Otherwise, output the median M_{med}.

$$\begin{cases} N_1 = x_m - M_{min} \\ N_2 = x_m - M_{max} \end{cases} \tag{11}$$

$$(N_1 > 0) \& (N_2 < 0) \tag{12}$$

For an illustrative purpose, the performance of the AMF against different numbers of abnormal samplings is shown in Fig. 2, where subfigures (a), (d), and (g) show the performance against a single abnormal sampling, subfigures (b), (e), and (h) show the performance against three continuous abnormal samplings, and subfigures (c), (f), and (i) show the performance against randomly-distributed multiple

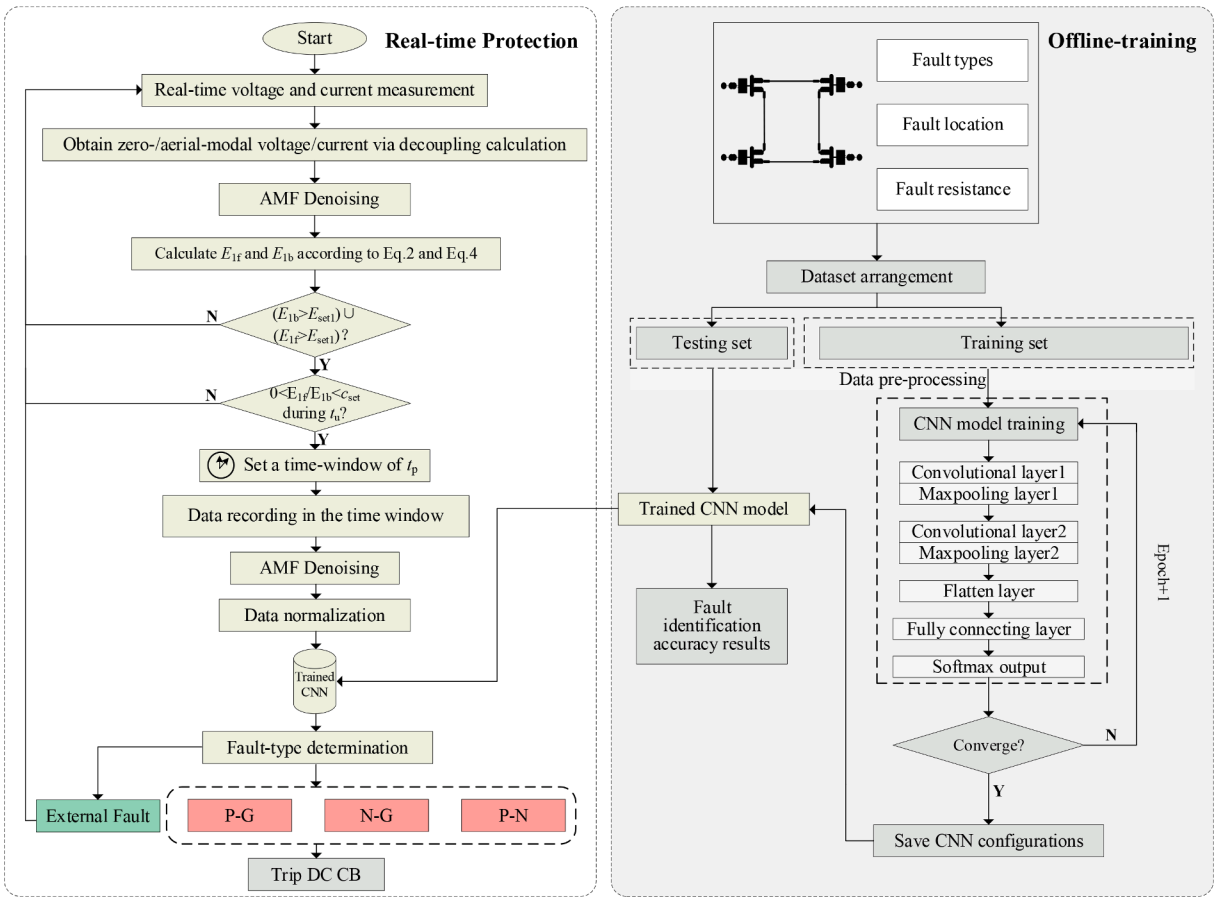


Fig. 5. The flowchart of the proposed protection.

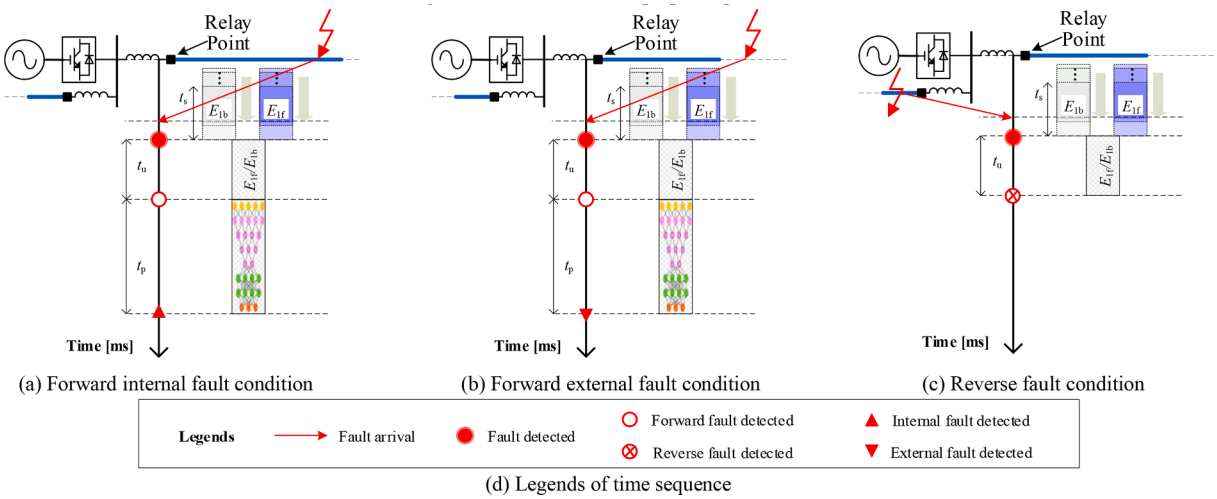


Fig. 6. Time sequence of the proposed protection.

abnormal samplings. As demonstrated, in each case, the abnormal samplings are erased evidently. Meanwhile, the feature of the original waveform can be kept, no matter what type of abnormal sampling is added to the original measurements.

According to the above, to eliminate a total of h abnormal samplings, the maximum order of the median filtering algorithm should be at least $2h+1$. However, the graphic feature of the waveform will be distorted if the order is too high, and this may leave negative impacts on the fault identification of the CNN in the next stage [29,30].

2) The data normalization

Below are the reasons why the data should be normalized before being fed into CNN. First, both the voltage and the current measurements are used as input of the CNN, while the dimension and the magnitude of the two are not the same. If the original samplings are fed into the CNN directly without normalization, the voltage measurements, which have higher magnitude, will have higher weights, and this can cause a sharp decline in the characteristics of the current measurements.

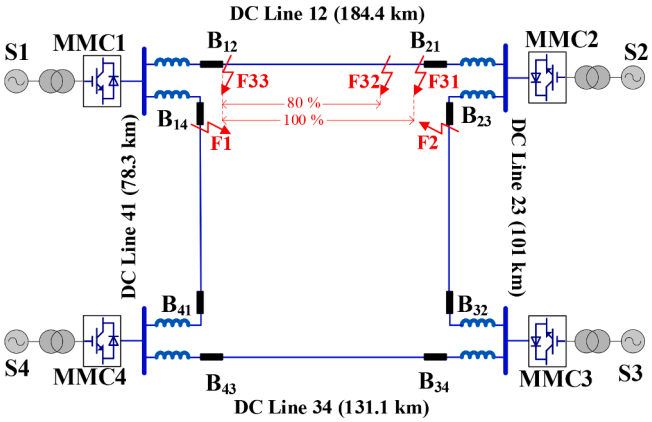


Fig. 7. MMC-MTDC construction.

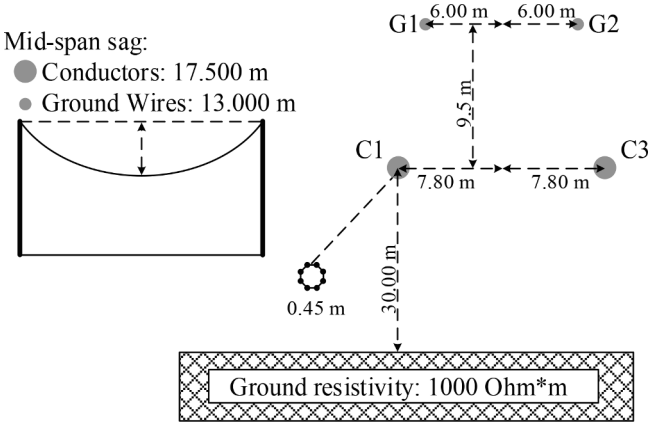


Fig. 8. The profile of DC line frequency-dependent model.

Table 1
Arrangement of the training set.

Parameter	Parameter scope	Amount
Fault type	P-G, N-G, P-N, EF	4
Location [km]	[0, 10, 247.4]	25
Fault resistance [Ω]	[0,10,300]	31

Note: [a, n, b] in parameter scope means the minimal and the maximal values are taken into consideration are a and b, and the interval is n. The total number of values is (b-a)/n + 1.

Second, with a series of normalized data, the gradient descent algorithm's optimization process can be accelerated.

In this paper, the z-score method uses Equation (13), where **Y** is the original matrix of samplings. μ stands for the mean value of the original data, and σ denotes the standard deviation of the original data. **Y*** is the matrix of samplings after normalization. Based on this, the normalized data has a mean value of 0 and a standard deviation of 1. The normalization operation to the samplings will not start until the start-up criterion is met to identify the arrival of a backward surge, indicating that there is a forward fault. For this reason, there will be no situation where the constant DC voltage, current, and other physical quantities are normalized, which will cause Equation (13) to lose its physical meaning.

$$\boldsymbol{Y}^{*} = (\boldsymbol{Y} - \mu)/\sigma \tag{13}$$

2.3. Fault identification based on CNN

After an occurrence of a fault, fault-induced surges will propagate in the MMC-MTDC with various electrical quantities such as voltage measurements, current measurements, etc. These fault components have abundant fault characteristics and are the keys to fault identification. To avoid the couplings between the positive pole and the negative pole, the post-fault aerial-modal voltage, the aerial-modal current, the zero-modal voltage, and the zero-modal current measurements are employed as inputs fed into the CNN. Conventionally, CNN is widely used in two-dimensional image recognition. Arranging the data into rectangles can facilitate the scanning of the convolution kernel, improve the local connection characteristics of CNN, and extract local features. $\mathbf{U}_a$, $\mathbf{U}_z$, $\mathbf{I}_a$, and $\mathbf{I}_z$ are the phasors of the aerial-modal voltage, the zero-modal voltage, the aerial-modal current, and the zero-modal current, respectively. Matrix **C** is used to arrange the measurements according to Equation (14). In the above equation, t_p is the length of the time window for the feeding data of CNN, and f_p is the sampling rate. In each row, the total number of samplings is $t_p{*}f_p$. Finally, for each row of data in matrix **C,** normalization is conducted according to Equation (13), which forms the final matrix **C*** in Equation (15) that will be fed into the CNN.

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{U}_a \\ \boldsymbol{U}_z \\ \boldsymbol{I}_a \\ \boldsymbol{I}_z \end{bmatrix} = \begin{bmatrix} U^{a}_{m}, U^{a}_{m+1}, \cdots\cdots U^{a}_{m+t_p f_p-1} \\ U^{z}_{m}, U^{z}_{m+1}, \cdots\cdots U^{z}_{m+t_p f_p-1} \\ I^{a}_{m}, I^{a}_{m+1}, \cdots\cdots I^{a}_{m+t_p f_p-1} \\ I^{z}_{m}, I^{z}_{m+1}, \cdots\cdots I^{z}_{m+t_p f_p-1} \end{bmatrix} \tag{14}$$

$$\boldsymbol{C}^{*} = \begin{bmatrix} \boldsymbol{U}^{*}_{a} \\ \boldsymbol{U}^{*}_{z} \\ \boldsymbol{I}^{*}_{a} \\ \boldsymbol{I}^{*}_{a} \end{bmatrix} = \begin{bmatrix} (\boldsymbol{U}_a - \mu_{ua})/\sigma_{ua} \\ (\boldsymbol{U}_z - \mu_{uz})/\sigma_{uz} \\ (\boldsymbol{I}_a - \mu_{ia})/\sigma_{ia} \\ (\boldsymbol{I}_z - \mu_{iz})/\sigma_{iz} \end{bmatrix} = \begin{bmatrix} U^{a*}_{m}, U^{a*}_{m+1}, \cdots\cdots U^{a*}_{m+t_p f_p-1} \\ U^{z*}_{m}, U^{z*}_{m+1}, \cdots\cdots U^{z*}_{m+t_p f_p-1} \\ I^{a*}_{m}, I^{a*}_{m+1}, \cdots\cdots I^{a*}_{m+t_p f_p-1} \\ I^{z*}_{m}, I^{z*}_{m+1}, \cdots\cdots I^{z*}_{m+t_p f_p-1} \end{bmatrix} \tag{15}$$

For instance, the arranging of the feeding data sets is shown in Fig. 3, where the sampling rate is 50 kHz and the time window is 1 ms. Thus, the total number of samplings in each channel is 50. After detecting the arrival of the fault-induced surge, a dataset is acquired, including the aforementioned four types of local measurements, as shown in Fig. 3(a). Then, the samplings in each channel are normalized according to Equation (13), as shown in Fig. 3(b), based on which a 2-dimensional feature map is generated in Fig. 3(c). A training set, composed of a large number of samples, is generated and fed into CNN to train the CNN. Each sample is assigned a corresponding label. There are four types of labels, namely the internal positive-pole-to-ground fault, the internal negative-pole-to-ground fault, the internal pole-to-pole fault, and the external fault.

The structure of the CNN in this paper is constructed as shown in Fig. 4. When a specific convolution kernel sweeps the above vector matrix, all the local features of the sample are extracted, and the corresponding feature map is formed. Therefore, using a determined number of convolution kernels, a corresponding number of feature maps will be generated, and the set of these feature maps is called a convolutional layer. The pooling layer is used to reduce the size of the matrix. Through the pooling layer, the number of nodes in the fully connected layer can be further reduced to reduce the parameters of the entire neural network. The use of the pooling layer is to speed up the calculation speed and prevent overfitting. The convolutional and pooling layers can work together to extract the feeding data features better. This work employs a two-layer convolutional layer and a two-layer pooling layer. In the flatten layer, the extracted features are laid out into flatten vectors, used as the input of the fully connected layer. Finally, a *softmax* layer works as a classifier to identify the feeding data, obtain a classification result, and output the classified label in a one-hot-coding manner.

During the convolution, the mathematical form of convolution kernel is as follows:

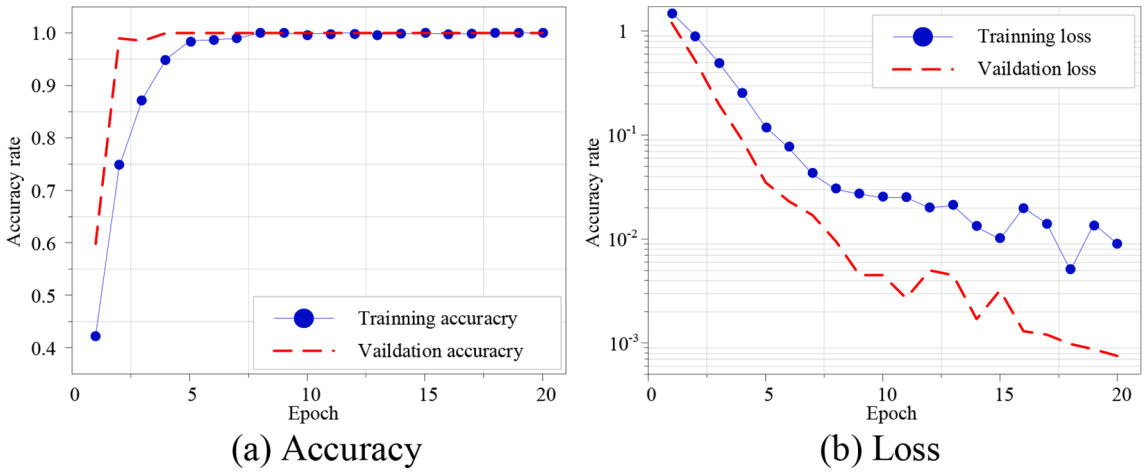


Fig. 9. Variation of accuracy and loss in each epoch.

Table 2
Study cases included in this manuscript.

Fault location	Fault resistance [Ω]	Fault type	Fault time [ms]
F1	0	P-N	0
F1	0, 150, 300	P-G	0
F2	0	P-N	0
F2	0, 150, 300	N-G	0
F31	0	P-N	3
F31	0, 150, 300	N-G	3
F32	0, 150, 300	P-G	2
F33	0, 150, 300	P-G	0

$$X_{ij}^{(m,n)} = f\left(\sum_{l}\sum_{p=0}^{P_i-1}\sum_{q=0}^{Q_i-1} x_{(i-1)l}^{(m+p)(n+q)} w_{ijl}^{pq} + b_{ij}\right) \tag{16}$$

where $x_{ij}^{(m,n)}$ represents the value of location (m,n) in the i-th convolutional layer, j-th feature map. l denotes the number of feature maps connected to the j-th feature map, i-th layer in the (i-1)-th layer. w is the weight of the l-th feature map, (i-1)-th layer, where P_i and Q_i are the height and the width of the convolution kernel. b is the bias of the j-th feature map, i-th layer. f is the activation function of the i-th layer. In this paper, the ReLU function is employed as the activation function in each layer, which is expressed as:

$$f(x) = \mathrm{ReLU}(x) = \begin{cases} x, x > 0 \\ 0, x \leqslant 0 \end{cases} \tag{17}$$

Using pooling layers after a convolution layer can help reduce the data dimension and the size of feature maps and accelerate the speed of training by increasing the generalization of the CNN model. In this paper, the Maximum pooling is applied in each pooling layer, which is given by:

$$P_i = \max_{w\times w}(x_i^{w\times w} u(w,w)) \tag{18}$$

where P_i is the output of the Maximum pooling. $u(w,w)$ is a square

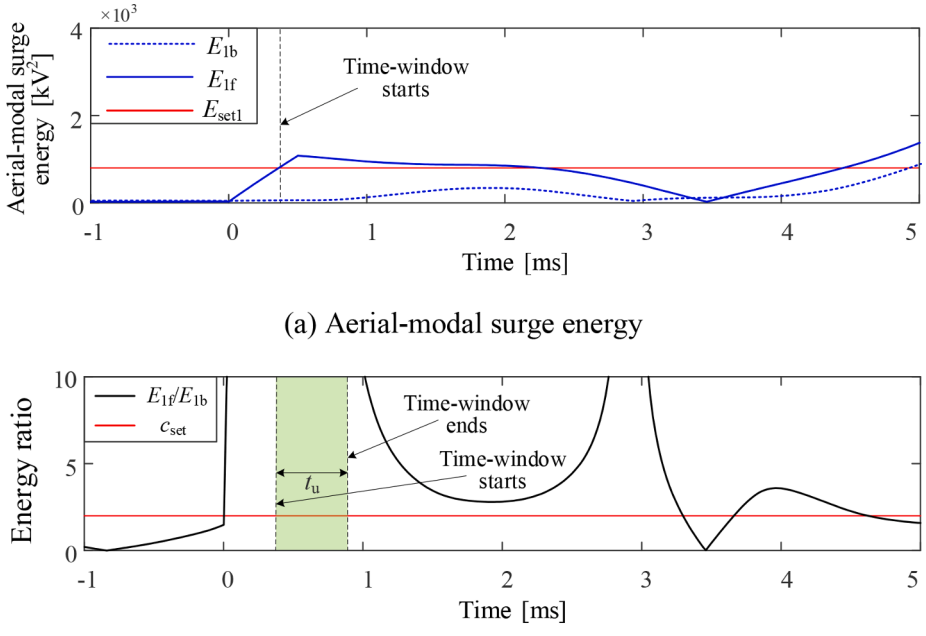


Fig. 10. Reverse faults security test.

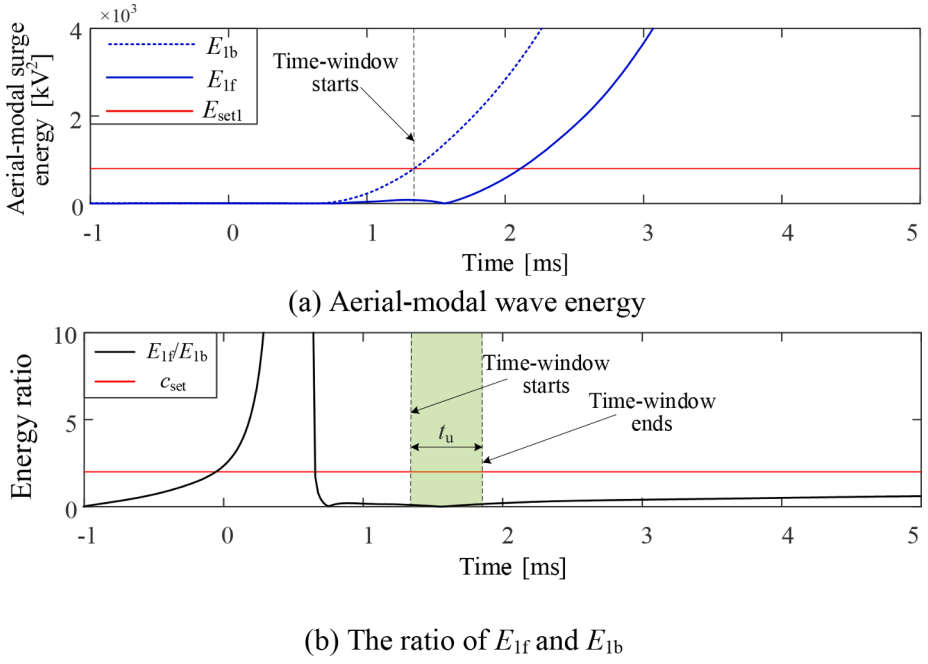


Fig. 11. Forward external faults security test.

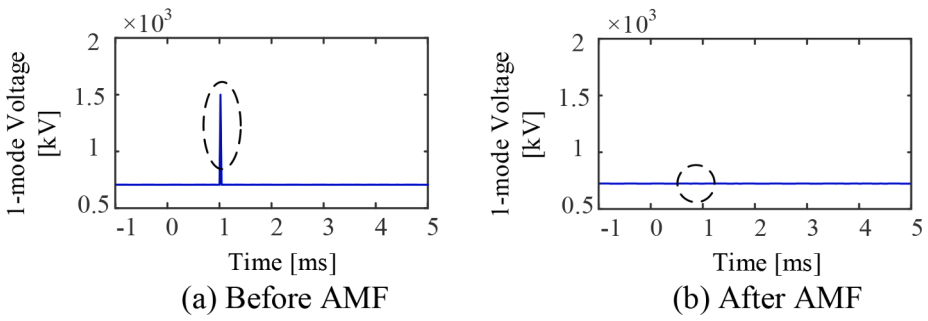


Fig. 12. An abnormal point superimposed onto the samplings.

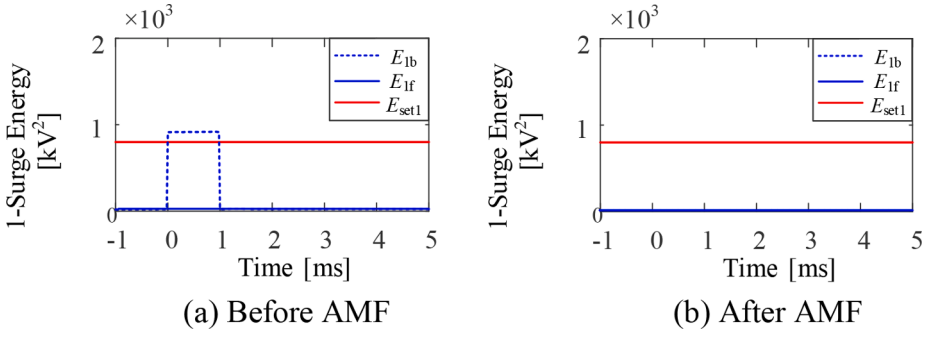


Fig. 13. Performance of the AMF against abnormal samplings.

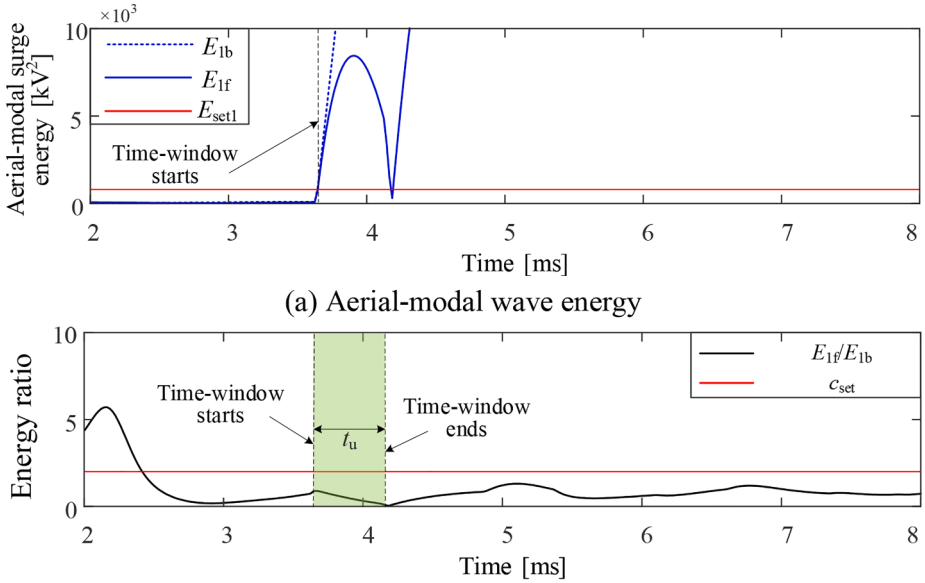
window with a mean value of 1. $x_i^{w\times w}$ is the selected pooling area, and the maximal value in this area is the output P_i. In the flatten layer, the outputs of poolings are arranged as a vector $\mathbf{V}_F$ and connected to the fully connected layer, as follows:

$$\boldsymbol{V}_{\mathrm{FC}} = f(w_k \boldsymbol{V}_F + b_k) \tag{19}$$

where $\mathbf{V}_{\mathrm{FC}}$ is the output vector of the fully connected layer. w_k is the weight of the k-th neuron. b_k is the bias of the k-th neuron. Also, the activation function of the fully connected layer is ReLU, which is expressed in Equation (17).

Finally, the vector $\mathbf{V}_{\mathrm{FC}}$ is sent to the *softmax* classification layer, where the probability of each class is calculated according to Equation (20):

$$P(T = t|\boldsymbol{V}_{FC}) = \frac{\exp(\lambda_t \boldsymbol{V}_{FC})}{\sum_{h=1}^{H} \exp(\lambda_h \boldsymbol{V}_{FC})} \tag{20}$$



(b) The ratio of E_{1f} and E_{1b}

Fig. 14. Internal faults sensitivity test.

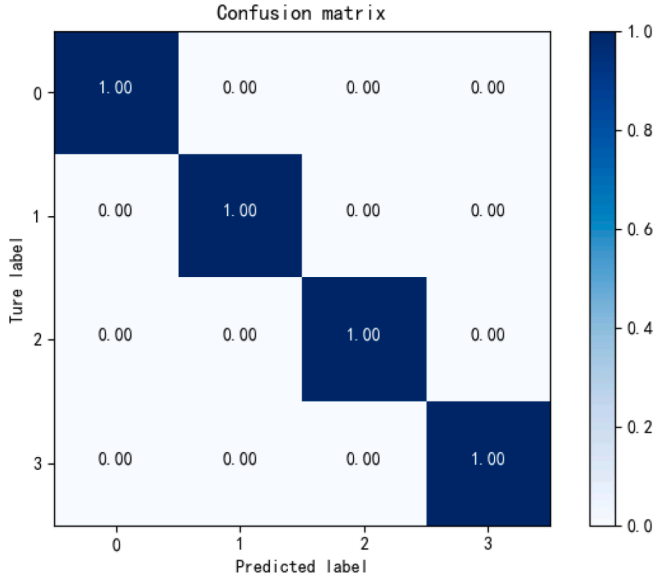


Fig. 15. Confusion matrix.

where H is the number of neurons in the *softmax* layer and is set to 4 in this paper because there are 4 types of labels. λ is the connection coefficient vector. If the output of a specific neuron is the largest one, it is most possible that the result is classified as the corresponding label. In this paper, the cross-entropy loss function (Categorical Cross-entropy) is employed in the training process, and the backpropagation algorithm is used to minimize the loss function and optimize the CNN model parameters, where the loss function is expressed by:

$$Loss = -\sum_{i=1}^{K} y_i \log(P(T=t|V_{FC})) \tag{21}$$

where K is the number of classes, and y represents the label. $P(T=t|V_{FC})$ is the output of the CNN calculated from the aforementioned *softmax*. In addition, the Adam algorithm is used to optimize the backpropagation process since it works well in terms of many engineering applications [31].

3. Methodology

The proposed protection element includes an offline and an online part, where the CNN is trained offline and applied online. The start-up criterion is also applied in an online manner. The flowchart for the

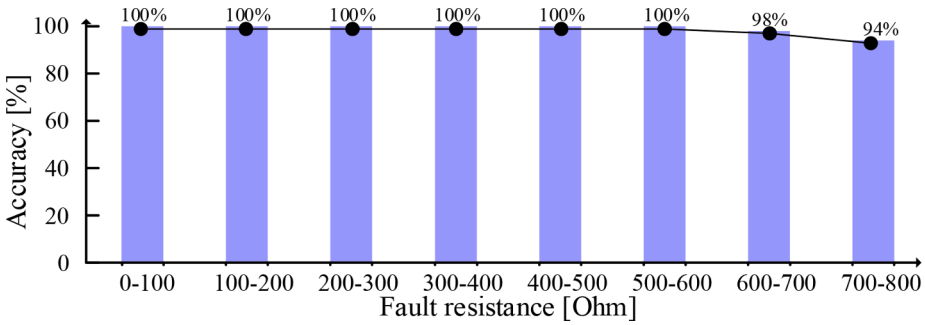


Fig. 16. Fault identification accuracy with varied fault resistances.

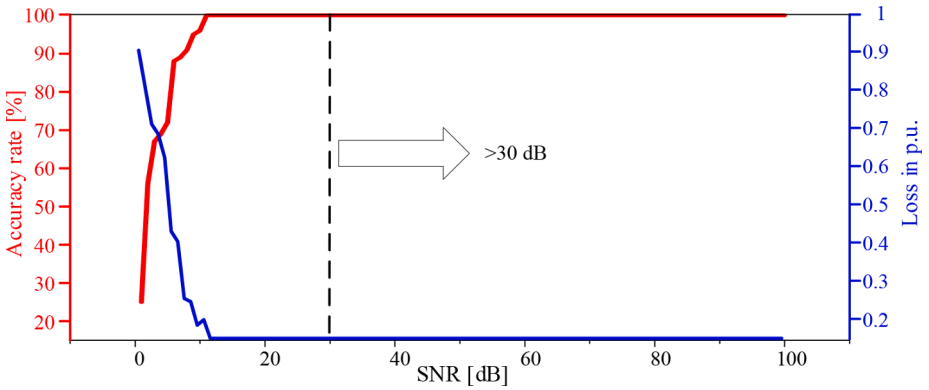


Fig. 17. Accuracies and losses under different SNRs.

methodology is shown in Fig. 5. Also, for illustrative purposes, how the proposed protection element operates under different scenarios is described in Fig. 6.

3.1. The application of the semi AI-based protection

1) Offline training part

The training set and testing set include a large number of samples, 70% of which are in the training set, and 30% are in the testing set. The training set includes different data sets (matrix **C***) from various fault locations, fault types, and fault resistances. The following issues are considered.

a) The highest fault resistance that needs to be considered in terms of engineering applications should be 100 Ω for 220-kV systems and 300 Ω for 500-kV systems.

b) To improve the accuracy of the CNN, the size of samples and the number of labels should be kept from being too large. Fortunately, the setting of the start-up criterion is for this purpose. The generated data set for each fault scenario will not be included, either for training or prediction, unless the start-up criterion trips. For instance, in the reverse fault scenario in Fig. 1(b), Equation (7) is true rather than Equation (6), indicating that the fault is a reverse one because the amplitude of the forward surge is far higher than that of the backward surge. For this reason, CNN will not work. According to the above, at least half of external fault scenarios can be excluded before the application of CNN, which not only simplifies the number of scenarios that the CNN needs to consider but also simplify the structure, setting, and inner connection of the CNN.

2) Online application part

a) For each protection element in the MMC-MTDC, voltage measurements and current measurements are recorded in real-time. Using the phase-modal decomposition equation, $\mathbf{U}_a$, $\mathbf{U}_z$, $\mathbf{I}_a$, and $\mathbf{I}_z$ are obtained. Then, after the AMF denoising, the energy of the aerial-modal backward surge and that of the aerial-modal forward surge in a short time window of t_s are obtained from Equations (2) and (4).

b) According to Equations (1) and (3), whether there is an arrival of a fault-induced initial surge is monitored in real-time. Once any of them is true, Equations (6) and (7) are calculated. In the time window of t_u, an occurring of a forward fault is detected if Equation (6) is true. Then, a data series including $\mathbf{U}_a$, $\mathbf{U}_z$, $\mathbf{I}_a$, and $\mathbf{I}_z$ in a time window of t_p is recorded. If Equation (7) stays true throughout the time window of t_u, the protection element returns to step a).

c) To eliminate noises and abnormal samplings, AMF is adopted again on the recorded measurements of $\mathbf{U}_a$, $\mathbf{U}_z$, $\mathbf{I}_a$, and $\mathbf{I}_z$ because they are measured in a totally different time window of t_p rather than the former one in the time window of t_s. Then, measurements are arranged in the form of Equation (15) with a size of $4t_p \times f_s$, which is used as the feeding data (2-d feature map **C***) of the CNN in the next step.

d) The aforementioned 2-d feature map is fed into the trained CNN. The local DC CB will be tripped if the output is an internal positive-pole-to-ground fault, internal negative-pole-to-ground fault, or internal pole-to-pole fault. Otherwise, the protection element will return to step a).

Table 3
Time-cost comparison among state-of-the-art works.

	t_s [ms]	t_u [ms]	t_p [ms]	t_{cb} [ms]	t_{al} [ms]
Reference [13]	N/A	N/A	2	2	1
Reference [32]	N/A	N/A	2	2	1
Reference [33]	N/A	N/A	2.56	2	0.44
Reference [34]	N/A	≈ 0	1	2	≈ 2
Proposed one	0.5	0.5	1	2	1

Table 4
The configuration of the simulation platform.

CPU	inter(R) Core(TM) i5-7500 cpu@ 3.40 GHz
GPU	NVIDIA GeForce 920MX
RAM	SODIMM 8 GB 2400 MHz

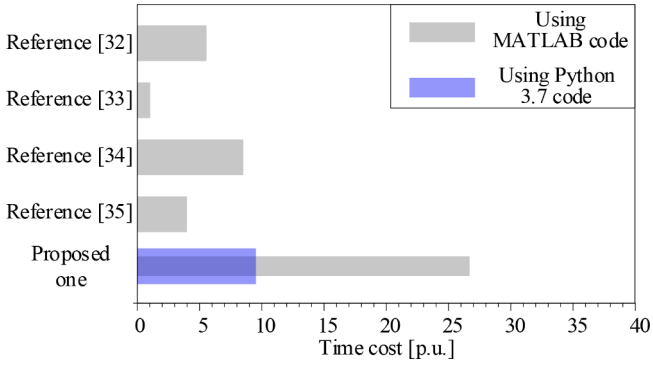


Fig. 18. The computation burden of different methods.

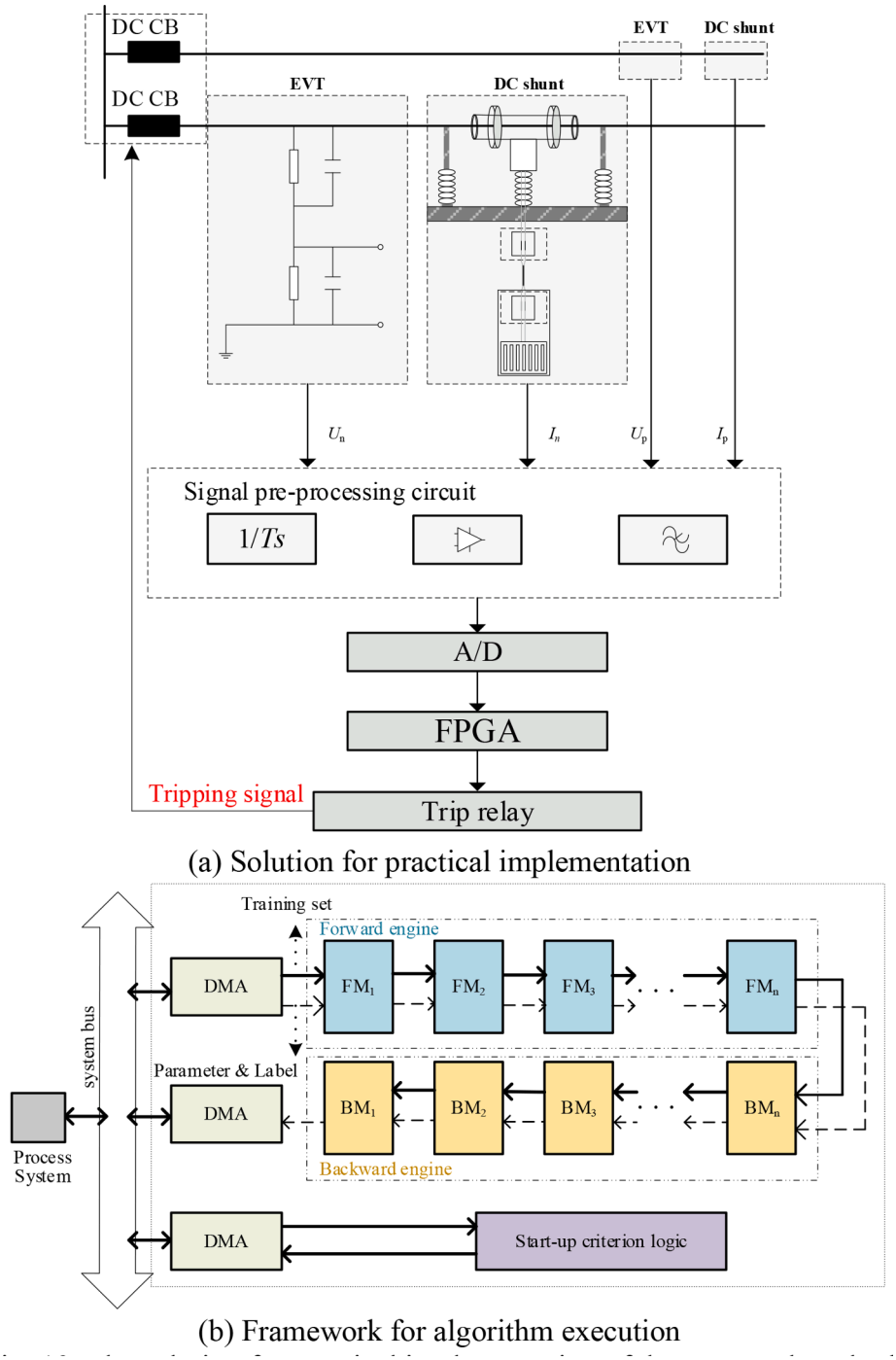


Fig. 19. The solution for practical implementation of the proposed method.

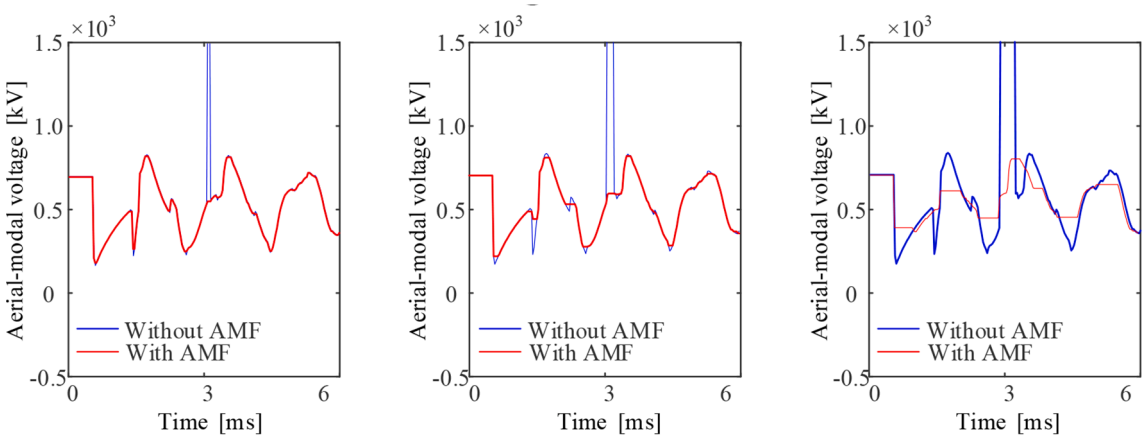


Fig. 20. Aerial-modal voltage with different continuous abnormal samplings.

Table 5
The parameters of the expanded fault set.

	Fault cases		Number of faults
Fault location	Fault type	Fault resistance [Ω]	
F_{33}, F_{32}	P-N	[1,1,9]	18
F_{31}, F_2	P-N	[0,1,10]	22
F_1	N/A	N/A	0
F_2	P-G, N-G	[0,5,300]	122
F_{33}	P-G, N-G	[5,10,295]	60
F_{32}	P-G, N-G	[5,10,295]	60
F_{31}	P-G, N-G	[0,5,300]	122

Note: [a, n, b] in parameter scope means the minimal and the maximal values are taken into consideration are a and b, and the interval is n. The total number of values is (b-a)/n + 1.

Table A1
The profile of MMC-MTDC.

System parameters	Value
Rated DC voltage [kV]	±500
Number of sub-modules	200
Current-limiting reactor [mH]	100
Arm reactance [mH]	Station-1: 61.7; Station-2: 61.7; Station-3: 30.8; Station-4: 30.8
Capacitance of sub-module [μF]	Station-1: 16292; Station-2: 16292; Station-3: 32584; Station-4: 32,584
Converter station base capacity [MW]	Station-1: 1500; Station-2: 1500; Station-3: 3000; Station-4: 3000

3.2. The operation of the semi AI-based protection

a) The logic of the protection operation against a forward internal fault is shown in Fig. 6(a). In this scenario, the initial surge that arrived at the relay location is a backward one. Once any of the two Equations (1) and (3) is true, the ratio of E_{1f}/E_{1b} will be continuously calculated in the time window of t_u. In this scenario, Equation (6) is true because the fault is a forward one. Then, the window of the CNN starts and ends after a duration of t_p. In the time window of CNN, a data set with a size of $4t_p \times f_s$ is obtained using local measurements. The data set is arranged in the form of Equation (15) and fed into CNN. Finally, the protection element trips the local DC CB.
b) In Fig. 6(b), similar to the scenario of a), the CNN will work because the fault is a forward one. However, the protection element will not trip the local DC CB because the output result of the CNN indicates an external fault.
c) As shown in Fig. 6(c), the arrived initial surge is a forward one, which means the fault is reverse. At the end of the time window of the start-up criterion, the protection element will return to the default status without tripping the fault-identification process.

According to the above, after the arrival of the fault-induced surge and the observation of a DC line fault, to isolate the fault before the blocking of HBMMCs, the total time cost should follow Equation (22):

$$T_{total} \leqslant t_s + t_u + t_p + t_{cb} + t_{al} \tag{22}$$

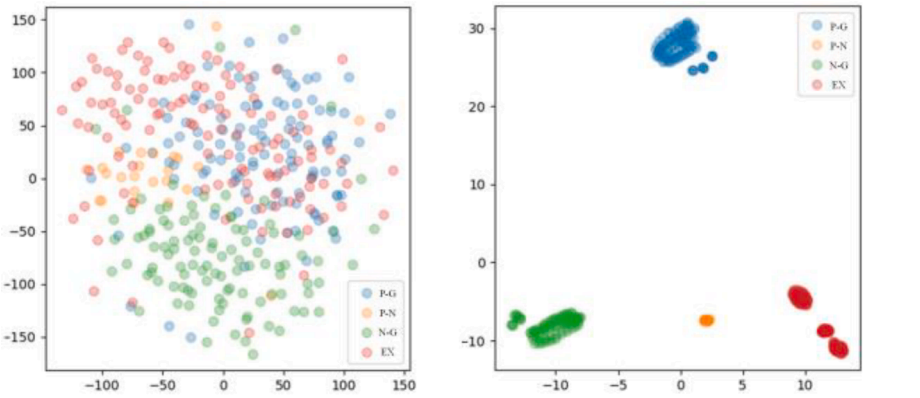
where T_{total} is the total time cost of the protection element to isolate the faulted line, which is at most the summation of the surge energy calculation delay t_s, the delay of start-up criterion time window t_u, the fault identification time delay of the CNN t_p, the typical fault interruption time of a high voltage DC CB t_{cb}, and the computation delay of the protection hardware t_{al}. To isolate the faulted line before the blocking of the HBMMC, T_{total} should be no more than 5–6 ms. In Reference [6], the transient characteristics of MMC under the pole-to-pole fault condition are analyzed, showing that the minimal blocking time of an HBMMC is about 5 ms, and this is the most severe case that should be taken into consideration. For this reason, the endurance time T_{total} is set to 5 ms. Typically, t_{cb} is regarded as 2 ms [23]. As long as we ensure that $t_s + t_u + t_p + t_{al} < 3$ ms, the speed is sufficient to meet the requirement of an MMC-MTDC.

4. Simulation studies

4.1. Protection settings

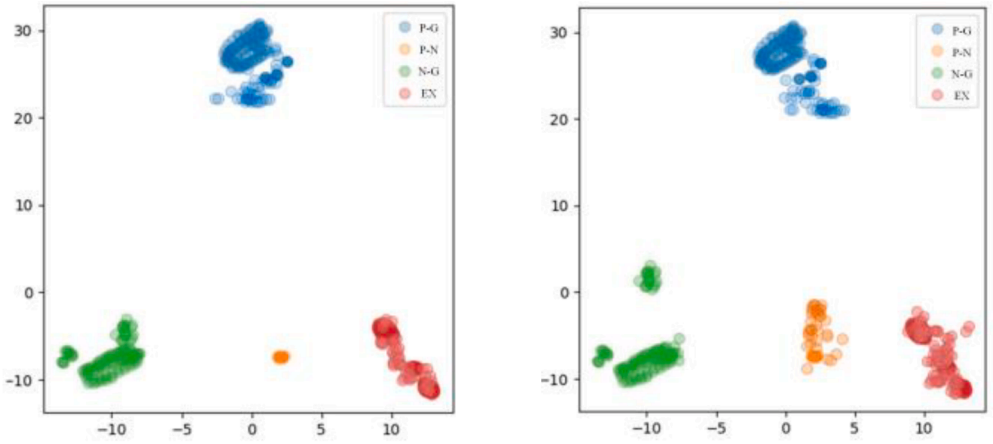
A 500-kV MMC-MTDC is built on the PSCAD/EMTDC software, as shown in Fig. 7. The profiles of the HBMMC and the system can be found in Table A1 in the appendix. The length of each DC line is 184.4 km, 101 km, 131.1 km, and 78.3 km, which is the same as the Zhangbei four-terminal MMC-MTDC. To have accurate results, the frequency-dependent model is employed for each DC line with profiles shown in Fig. 8.

According to Section 3, the constraint of $t_s + t_u + t_p + t_{al} < 3$ ms should be met if the proposed protection works as a main protection element. In this paper, both t_s (energy calculation window) and t_u (fault direction determination window) are set to 0.5 ms. t_p is set to 1 ms. Thus, theoretically, a time margin of 1 ms is left for the protection hardware to

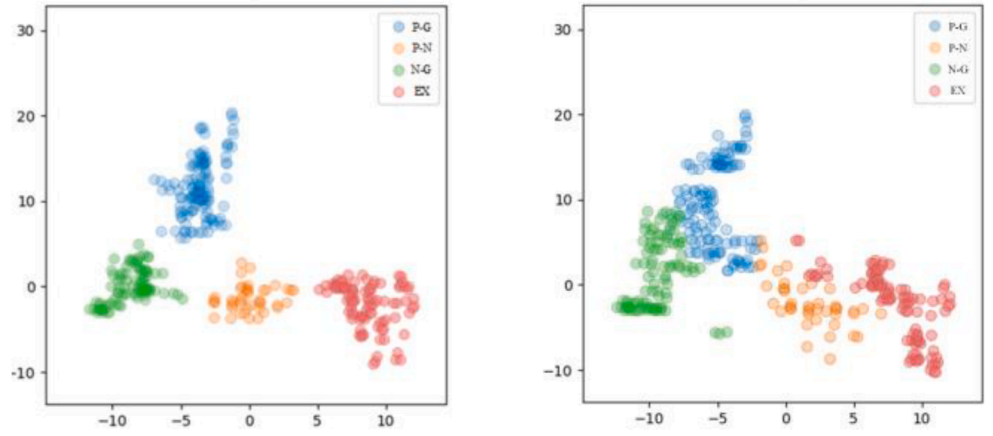


(a) Using raw data (b) Using flatten layer feature vector

Fig. 21. Performance visualization based on T-SNE.



(a) 2 abnormal points (R=5) (b) 5 abnormal points (R=11)



(c) 7 abnormal points (R=15) (d) 11 abnormal points (R=23)

Fig. 22. The accuracy rate using different orders of AMF.

compute. The sampling rate f_p is set to 50 kHz for typical industrial applications. Thus, the size of the data set fed into the CNN is 4 × 50, which is the same as the one in Fig. 3. To ensure absolute reliability, r_{set} is set to 10 in Equation (5). Via groups of simulation results, it shows that E_{norm} is no more than 80 kV^2 during the unblocking/blocking of the DC system and the fast change of the DC power. As a result, E_{set} is set to 800 kV^2. Also, to improve the sensitivity of fault direction determination according to Equations (6) and (7), c_{set} is set to 2, since the value of E_{1f}/E_{1b} can be extremely high in a reverse fault scenario while<1 in a forward fault scenario.

For illustrative purposes, the protection element at the relay location of B_{12} is taken as the main focus in this paper. As reverse faults will be excluded by the start-up criterion, which is mentioned before, the setting of the training set should only include faults that occur on DC line-12 and those on DC line-23 within the first 63 km. Considering different fault types and fault resistances, the total number of samples in the training set is 4*25*31 = 3100. Details are listed in Table 1. The total number of epochs is set to 20, and the number of feeding data is set to 200 groups in each epoch. As shown in Fig. 9, for both the training and testing sets, accuracy increases fast and approaches 100% at the end of the training, whereas the loss declines sharply.

4.2. Measurement-based start-up criterion performance in identifying forward faults

On the above platform, case studies are firstly conducted to examine the performance of the start-up criterion. As illustrated in Fig. 7, F1 represents a reverse fault. F2 is a forward external fault. F31, F32, and F33 are forward internal faults with different locations on the DC line. Also, the fault resistance, the fault type, and the fault inception time are considered in the cases, as listed in Table 2.

Case-1: Security under reverse faults.

First, the security of the proposed protection element under reverse fault conditions is assessed. Assume that a pole-to-pole bolted fault occurs at $t = 0$ ms at the backside of station MMC1 (Fault F1), reverse for relay B_{12}. The performance of the start-up criterion is shown in Fig. 10. As indicated, the start-up criterion will not begin to work until $t = 0.38$ ms. As shown by the green box in Fig. 10(b), the ratio of E_{1f}/E_{1b} is always higher than c_{set} within the 0.5-ms time window of t_u. The start-up criterion determines that the fault is a reverse one and then blocks the protection from entering the CNN identification process. As a result, security can be ensured in the reverse fault scenario.

To further assess the security of the proposed protection element under reverse fault scenarios, more cases are studied. The variation of the surge energies and energy ratio is shown in Fig. A1, Fig. A2, and Fig. A3, where the fault resistance for the grounding fault is 0 Ω, 150 Ω, and 300 Ω, respectively. According to the results, the start-up criterion will block the protection immediately after the fault occurrence, ensuring the security of the proposed protection.

Case-2: Selectivity under adjacent line faults.

Assume that a pole-to-pole bolted fault occurs at the head of the adjacent line, located at 0 km from the station MMC2 on DC Line-23 (Fault F2) when $t = 0$ ms. The performance of the start-up criterion is shown in Fig. 11. The energy of the backward surge exceeds the threshold at $t = 1.34$ ms. As illustrated, E_{1f}/E_{1b} keeps below c_{set} throughout the time window of t_u because the fault is forward, which makes Equation (6) true, and the fault is determined as forward. In additional cases shown in Fig. A4, Fig. A5, and Fig. A6, where the fault resistance for the grounding fault is 0 Ω, 150 Ω, and 300 Ω, respectively, the start-up criterion also works well. In these cases, the CNN will further determine whether the fault is internal or external, which will be discussed in detail later. Under this condition, the start-up criterion of the proposed protection element has high selectivity.

Case-3: Security against abnormal samplings.

Under the normal operating condition, electrical quantities like the voltage, current, etc., are ideally constants for DC systems and Equations (2) and (4) should be quite low. However, significant errors may be caused when calculating the above equations if errors like abnormal voltage samplings and abnormal current samplings are involved; without loss of generality, assuming that an abnormal sampling is added onto the voltage sampling, and the current sampling is kept unchanged. As a result, the abnormal sampling will be involved in the surge waveform, as shown in Fig. 12(a). The profile of the calculated surge energy is shown in Fig. 13(a). Clearly, the proposed protection element will enter into the CNN fault identification process if no countermeasure is taken, and the CNN will not output a correct result because the normal operating condition is not introduced into the training set of the CNN, partly due to the difficulty to normalize a series of constants for DC systems. In contrast, after the application of AMF, the abnormal sampling is well eliminated in Fig. 12(b). As a result, the start-up criterion will not maloperate in Fig. 13(b), and thus the CNN will not work to identify a fault scenario not included in the training set. By this means, the employed CNN does not need to work to recognize the normal operating condition, which solves the difficulty of normalization and dramatically reduces the complexity of the setting of the training set, and avoids the overfitting phenomenon that may occur during the training process.

Case-4: Sensitivity under the internal fault conditions.

To assess the performance of the start-up criterion against internal faults, several cases are studied when the fault is located at F31, F32, and F33, where F31 and F32 are remote-end faults that are not beneficial for conventional protection elements to identify. Assume that a pole-to-pole bolted fault occurs at F31, located at the end of the DC line 12 at $t = 3$ ms, as an example. As shown in Fig. 14, the backward surge energy exceeds the threshold at $t = 3.64$ ms, when the ratio of E_{1f}/E_{1b} begins to be calculated for 0.5 ms. Apparently, as shown in Fig. 14(b), the result always stays below c_{set} in this duration, indicating that this is a forward fault that the CNN should further identify. Repeated assessments are conducted for single-line-to-ground faults with various fault resistances and fault locations. In Fig. A7, Fig. A8, and Fig. A9, the internal fault is a close-in one, located at 0 km from the converter station MMC1 (Fault F33). Fig. A10, Fig. A11, and Fig. A12 represent the results of a fault at 80% of the DC line (Fault F32), but the fault type is positive-pole-to-ground rather than pole-to-pole. In Fig. A13, Fig. A14, and Fig. A15, the performance of the start-up criterion under a remote-end negative-pole-to-ground fault (Fault F31) with various fault resistances is studied. Results show that the start-up criterion can identify all forward faults with different values of fault resistances, internal or external, according to Equation (6), since the ratio of E_{1f}/E_{1b} is always below c_{set} in the time window of t_u.

4.3. AI-based fault identification criterion performance in identifying internal faults

To assess the performance of the AI-based fault identification criterion using CNN, a total of 1500 fault scenarios, which are different from those in the training set, are set and fed into the CNN. As the start-up criterion can exclude reverse faults in advance, the fault location is set to a random point on Line-12 or the first 63 km on Line-23. The fault type is randomly chosen from positive-pole-to-ground, negative-pole-to-ground, and pole-to-pole with a random fault resistance of 0–300 Ω. The results are shown in Fig. 15, where the classification results (x-axis, predicted labels) and the real results (y-axis, true labels) are compared, and the accuracy is illustrated. The value of (x, y) shows the probability of fault-type identification of a type-x fault to type-y. As indicated, no fault is misidentified as long as the fault resistance is within 300 Ω. One step further, the fault scenario is extended by setting the range of fault resistance to 0–800 Ω, where there are 820 cases for 0–300 Ω, 1060 cases for 301–500 Ω, 1120 cases for 501–800 Ω, and the result is shown

in Fig. 16. It can be seen that with the increase of the fault resistance, the accuracy of fault identification decreases from 100% to 94%. Although the established CNN has a certain error level in some specific scenarios, these scenarios are hardly seen in typical engineering application situations. For this reason, the proposed protection element can meet the requirements in terms of industrial applications.

In addition, the accuracy of the proposed CNN under different levels of white noise is studied, and Fig. 17 shows the result. It can be seen that as long as the local measurements have an SNR higher than 20 dB, the accuracy of the CNN and the correctness of the fault identification can be very close to 100%, which is adequate in terms of engineering application as the SNR of measurements are required to be higher than 30 dB. This is because the e0mployed CNN can extract the local feature of the feeding data, and the impact caused by the change of details (Gaussian noise) is weakened by a low weight in the process of convolution kernel pooling. Thus, high robustness is ensured against white noises.

5. Discussions

5.1. Speed for the proposed method to work as a main protection element in terms of engineering applications

According to Equation (22) and the analysis in Section 3, to interrupt an internal fault before the blocking of the MMC converter, the summation of time delays of t_s, t_u, t_p, and t_{al} should be no more than 3 ms. Apparently, the greater t_{al} is, the higher speed the corresponding protection element will have because a more extended time margin can be left for computation. To compare the speed of the proposed protection with other types of main protection elements, the above indicators of several state-of-the-art methods in References [32–35] are listed and compared, as shown in Table 3. As indicated, the speed of the proposed one is in the same order of magnitude. However, for the proposed protection element, there is no absolute superiority over other works in terms of speed. Meanwhile, the following issues are noteworthy:

First of all, although a time window of 0.5 ms is used to calculate the surge energy in Equations (2) and (4), E_{1f} or E_{1b} usually do not need a complete 0.5 ms to exceed the threshold of E_{set1}, leaving much more time margin for the computation delay of t_{al}. As indicated in the appendix, from Fig. A1 to Fig. A15, it is evident that the calculated surge energy will rise above the threshold immediately, which costs far$<$0.5 ms, especially in internal fault conditions. For this reason, the estimation of the surge energy calculation delay t_s is conservative. In reality, t_{al} has a range from 1 to 1.5 ms.

According to the above analysis, the computation speed of hardware is essential for the proposed method to perform as a main protection element. To analyze the problem quantitatively, the methodologies listed in Table 3 are reconducted using the same personal computer (PC), and the configuration of the PC is listed in Table 4. Taking the time delay of the method in [33] as 1p.u. as it has the simplest form of criterion, the relative time-delays of other methods, including the proposed one, are shown in Fig. 18.

It is found that the proposed method has the highest computation time–cost on the same platform. The reason is that CNN's computation burden and complexity to process fault information are more significant than that of other methods. Also, programming with Python under the same hardware configuration can significantly reduce the time cost. Python better packages and optimizes the CNN model than Matlab language. When programming with Python, the time cost can reach the same order of magnitude as the other four numerical-calculation-based criteria.

Fig. 19 shows the solution for the implementation of the proposed method, including a group of sensors, an analog/digital (A/D) transfer module, a field-programmable gate array (FPGA) based processing unit, and a trip relay to send tripping signals to the local circuit breaker, as shown in Fig. 19 (a).

For DC systems, the DC electronic voltage transformer (EVT) and the DC shunt are the common types of sensors. In Reference [36], a machine-learning framework based on Zynq series FPGA is designed, including a process system module and a programmable logic module, as shown in Fig. 19 (b). In this framework, the processing system works as the control center, communicating with the programmable logic module at a speed of up to 50 Gb/s. The fault direction calculation and the CNN computation are performed in the direct memory accesses (DMAs) embedded in the programmable logic module. Based on the above information, the processing system could renew the CNN parameters and decide whether to trip the fault-identification process according to the information from the start-up criterion. Programming with underlying languages in such a system, the required time–cost can be further reduced than using Matlab or Python on a PC.

To sum up, both the hardware configuration and the software programming are critical to the speed performance of the proposed protection element. With high-performance hardware configuration and efficient underlying-language-based programming, the proposed protection can be used as a main protection, as long as t_{al} is short enough. Otherwise, in the best case, it can be used as a back-up protection element.

5.2. Performance of fault-identification criterion considering random multiple abnormal samplings

In this paper, AMF is introduced to avoid the impact of abnormal samplings. Thus, the complexity of the established CNN and the computation burden are reduced, benefiting from a reduced size of the training set. However, to balance between the capability of keeping the original graphical feature and that against multiple abnormal samplings, there should be quantitative analysis to select the maximum order R (maximum capability against multiple abnormal samplings) of the AMF. The following issues are taken into consideration.

1) To enhance the capability against multiple abnormal samplings, the order of AMF should be high enough. According to the information in Section 2.2, multiple abnormal samplings could not be erased if R is$<$5. For this reason, we should have $R \geq 5$.
2) The maximum order of AMF should have an upper limit to prevent it from being too high. This is because, in a very short time window, the feature of the wave-front is imperative for fault identification based on CNN. Conventionally, the standard operating surge in power systems is 250/2500 μs, and a wave-front has at least 12 points. To keep the feature of wave-front from being erased, we should have $R < 25$.

To illustrate the above, a case of a positive-pole-to-ground bolted fault is conducted, and the obtained aerial-modal voltage with different numbers of abnormal samplings are shown in Fig. 20. In comparison with the filtered waveforms, it is evident that a 5-order AMF can perfectly erase the abnormal samplings in Fig. 20(a) without distorting the original feature. In Fig. 20(b), an 11-order AMF is capable of erasing more numbers of abnormal samplings, but the original waveform is

slightly distorted. In Fig. 20(c), however, the filtered waveform loses its original feature heavily because the order of AMF is too high. For this reason, an optimal R should be further determined in the range of $5 \leq R < 25$.

To quantify this, further investigation is undertaken using an expanded fault set based on Table 2. The arrangement of the expanded fault set is shown in Table 5.

Fig. 21 visualizes the distribution of fault features in the above fault set without considering the impact of abnormal samplings using t-distributed stochastic neighbor embedding (t-SNE). As shown, the four types of fault scenarios are mixed in the original feature space, whereas clustered clearly and linearly after the extraction of compressed features by the CNN. Comparing the vectors of row data with the feature vectors obtained after convolution and pooling of the CNN model, the latter shows the significant distinction in different fault types. When there are no abnormal samplings, the fault identification by the CNN has a robust feature extraction capability.

When different levels of abnormal samplings are considered, the performance of the AI-based fault-identification is shown in Fig. 22. By comparing Fig. 22 (a) with (b), it is found that although the discrimination of different fault types stays very clear for both scenarios when $R = 5$ and $R = 11$, respectively, the fault feature is lightly affected in Fig. 22 (b) because the distortion of fault features after applying a higher order of AMF. When there are 7 continuous abnormal samplings with an R of 15, the fault-identification criterion marginally works. In the scenario when there are 11 continuous abnormal points and $R = 23$, there is a sharp decline of the discrimination of fault features. In some cases, different types of faults are overlapped. This is because the transient waveform distorted too heavily due to an extremely high-order AMF. Especially in some high-resistance fault cases, the transient signal becomes a flat one, causing the mis-identification. Further studies are made to obtain the accuracy of fault-identification under different orders of AMF, and the result is shown in Fig. 22. The accuracy rate can be up to 100% as long as $R \leq 15$. To leave some space, R is set to 11 in this paper.

6. Conclusion

In this paper, a semi AI-based protection element is proposed to trip DC line faults for the MMC-MTDC. The conclusions are as follows.

1) A local measurement-based start-up criterion is employed in advance before adopting the fault identification by CNN. The configuration of the start-up criterion can: a) eliminate the negative impact of abnormal samplings with the help of the AMF, b) reduce the complexity of the CNN structure and the arrangement of the training set, and c) make the protection element more suitable for online applications because the CNN is not working unless the start-up criterion identifies a forward fault.
2) The AI-based fault-identification criterion using CNN is powerful in distinguishing between internal faults and external ones. As long as the fault resistance is under 300 Ω and the SNR of white noises is above 30 dB, very high accuracy can be achieved, which is sufficient for engineering applications. To keep the fault-identification criterion working accurately, the largest number of continuous abnormal samplings that the AMF can erase is set to 5 according to a balanced setting principle.
3) Theoretically, the speed of the proposed protection is sufficient for HBMMC-based MMC-MTDCs because both the start-up criterion and the CNN-based fault identification use very short time windows. With high-speed hardware configurations and efficient underlying-language-based programming, the proposed protection can then work as a main protection element.

CRediT authorship contribution statement

Ning Tong: Conceptualization, Methodology, Software, Writing – original draft, Funding acquisition. **Zhenjie Tang:** Methodology, Data curation, Writing – review & editing, Writing – original draft. **Yu Wang:** Writing – review & editing, Visualization. **Chun Sing Lai:** Conceptualization, Writing – review & editing, Supervision. **Loi Lei Lai:** Conceptualization, Writing – review & editing, Supervision, Funding acquisition, Resources.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported by the National Natural Science Foundation of China under Grant No. 51907069 and the Natural Science Foundation of Guangdong Province under Grant No. 2022A1515011079 and 2020A1515010766.

Data Access Statement

Data supporting this study cannot be made available due to the research data are confidential, because of the arrangement the research groups have made with the commercial partner supporting the research.

Appendix

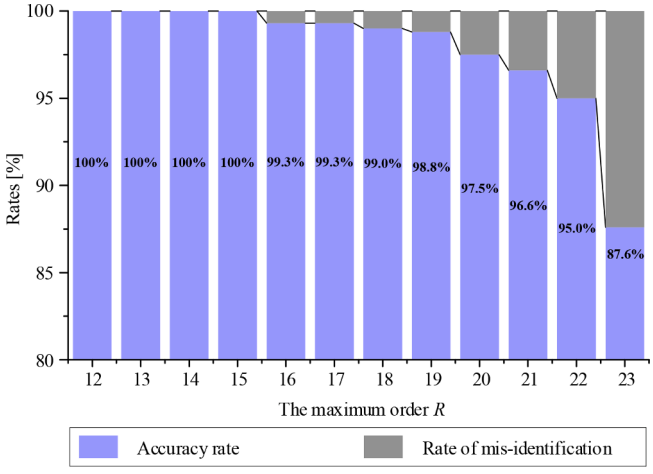


Fig. A1. Reverse, P-G, bolted fault at $t = 0$ ms.

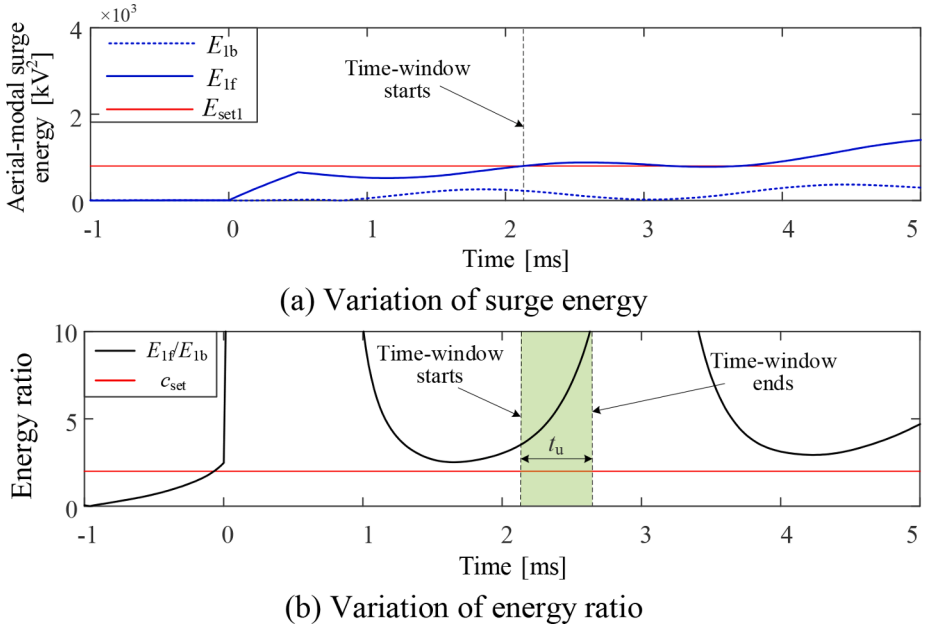


Fig. A2. Reverse, P-G, fault resistance = 150 Ω at t = 0 ms.

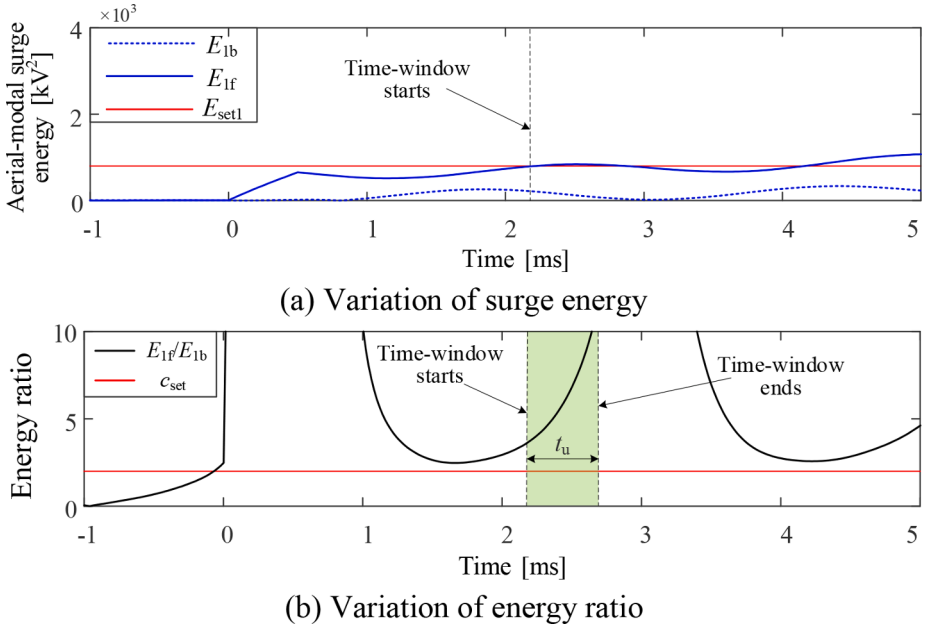


Fig. A3. Reverse, P-G, fault resistance = 300 Ω at t = 0 ms.

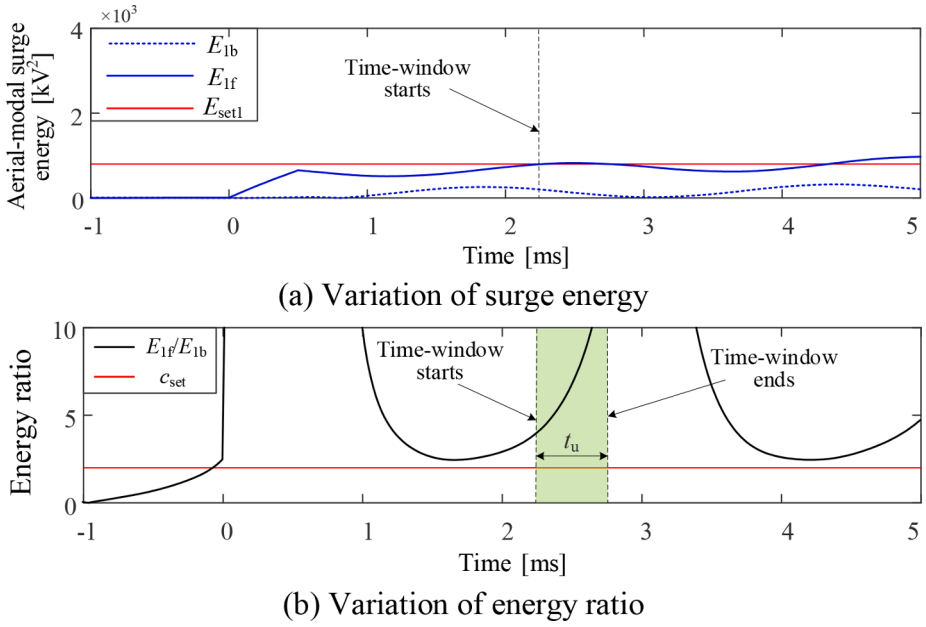


Fig. A4. Forward external, N-G, bolted fault at $t = 0$ ms.

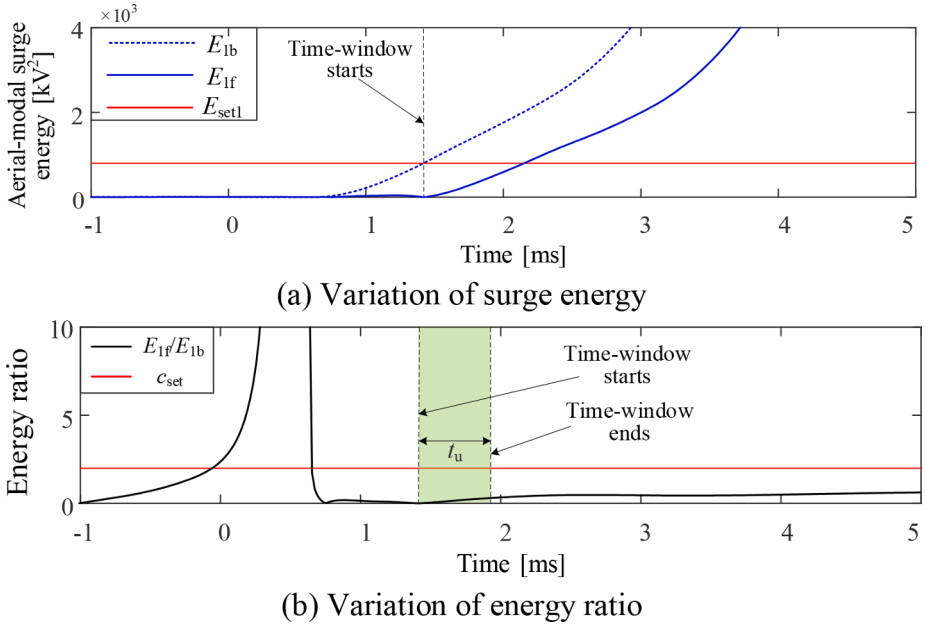


Fig. A5. Forward external, N-G, fault resistance = 150 Ω at $t = 0$ ms.

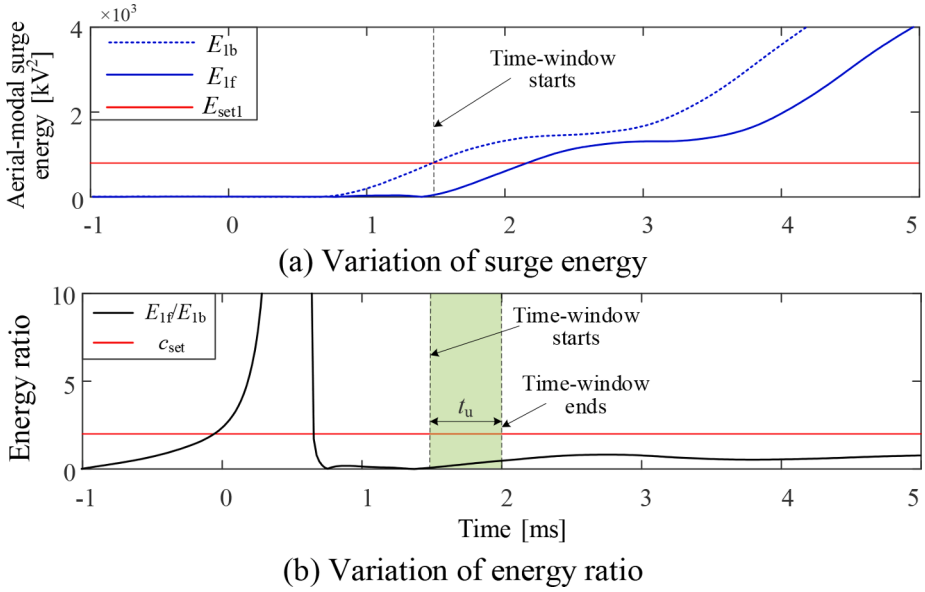


Fig. A6. Forward external, N-G, fault resistance = 300 Ω at t = 0 ms.

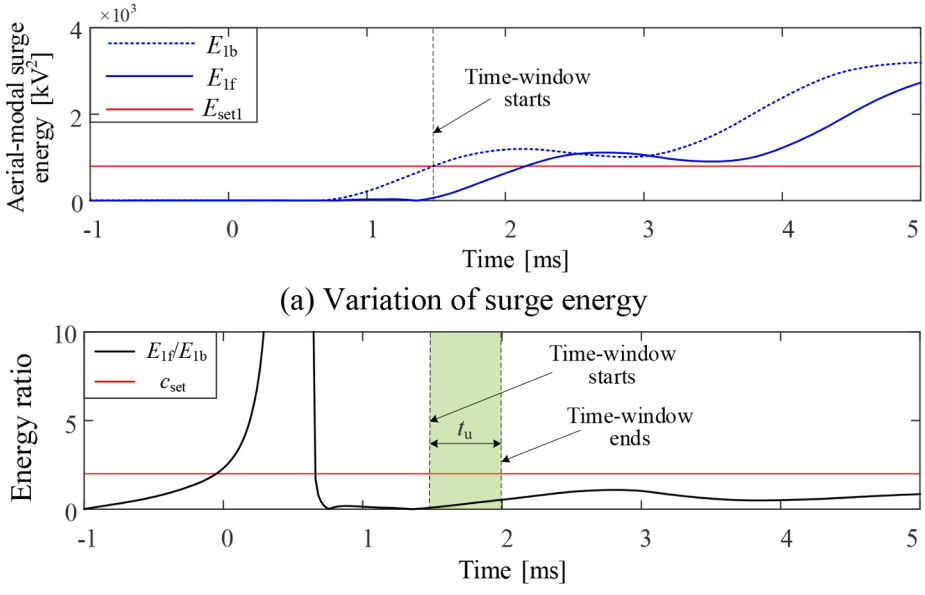


Fig. A7. Forward internal, 0%, P-G, bolted fault at t = 0 ms.

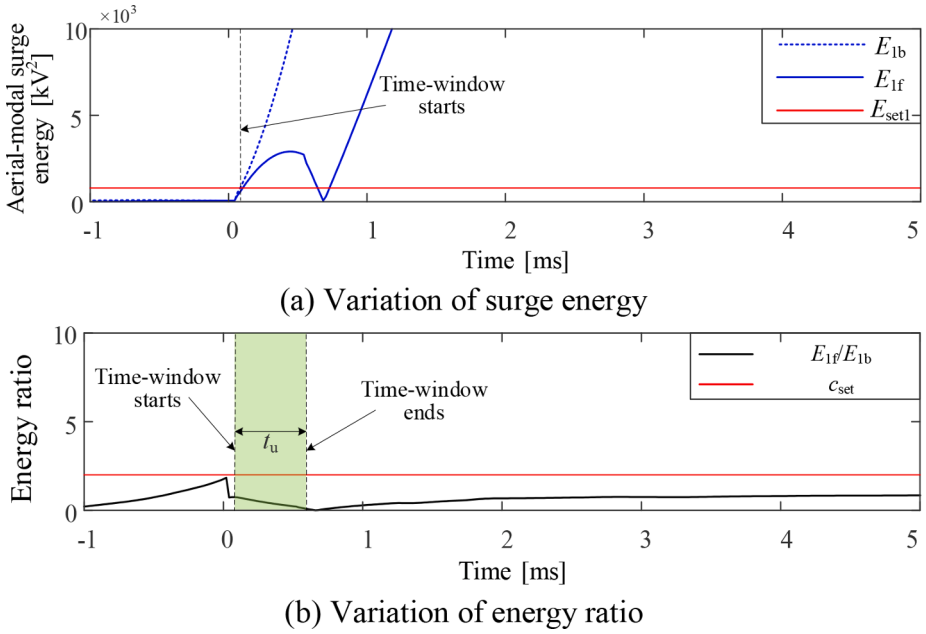


Fig. A8. Forward internal, 0%, P-G, fault resistance = 150 Ω at t = 0 ms.

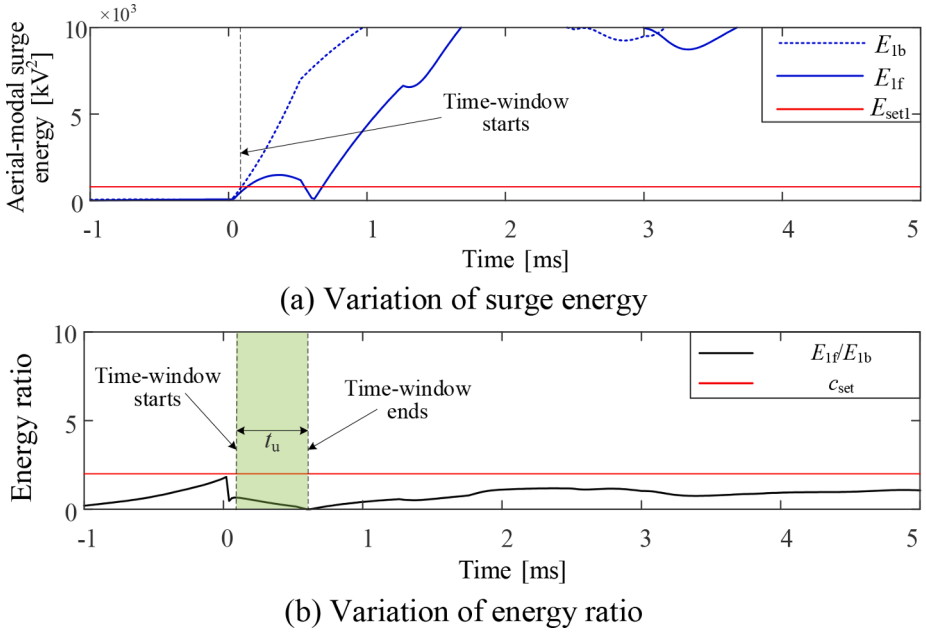


Fig. A9. Forward internal 0%, P-G, fault resistance = 300 Ω at t = 0 ms.

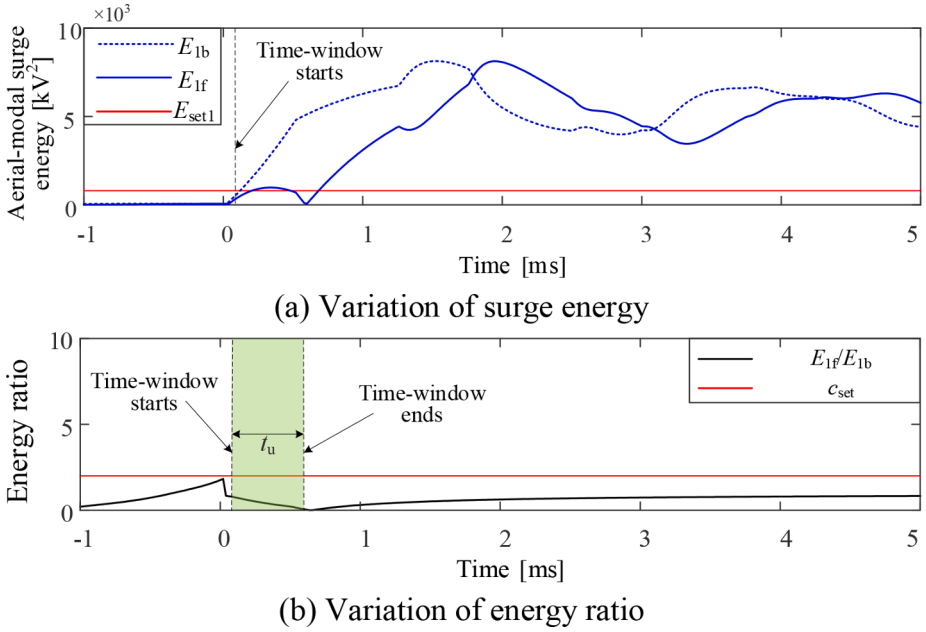


Fig. A10. Forward internal 80%, P-G, bolted fault at $t = 2$ ms.

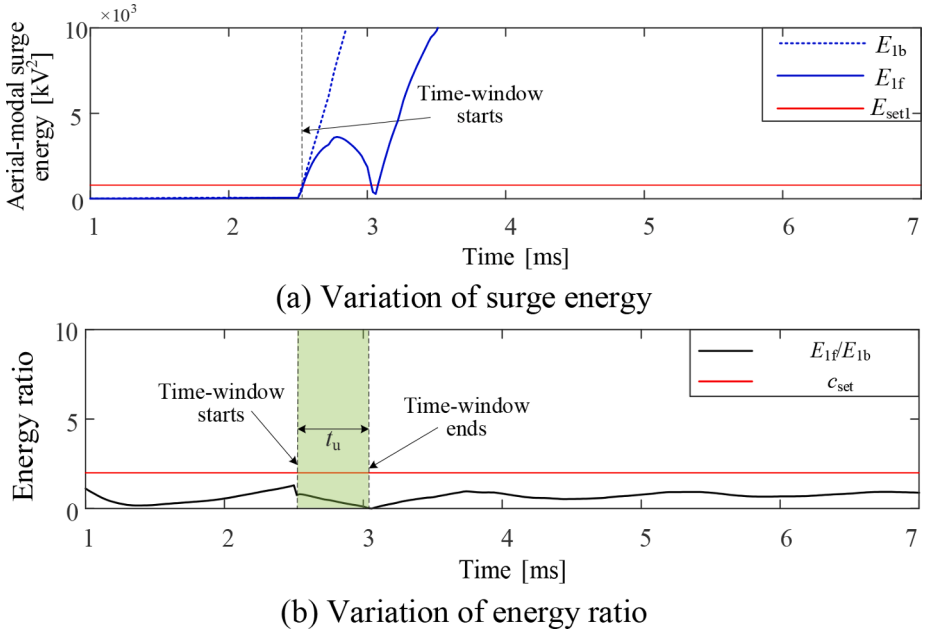


Fig. A11. Forward internal 80%, P-G, fault resistance = 150 Ω at $t = 2$ ms.

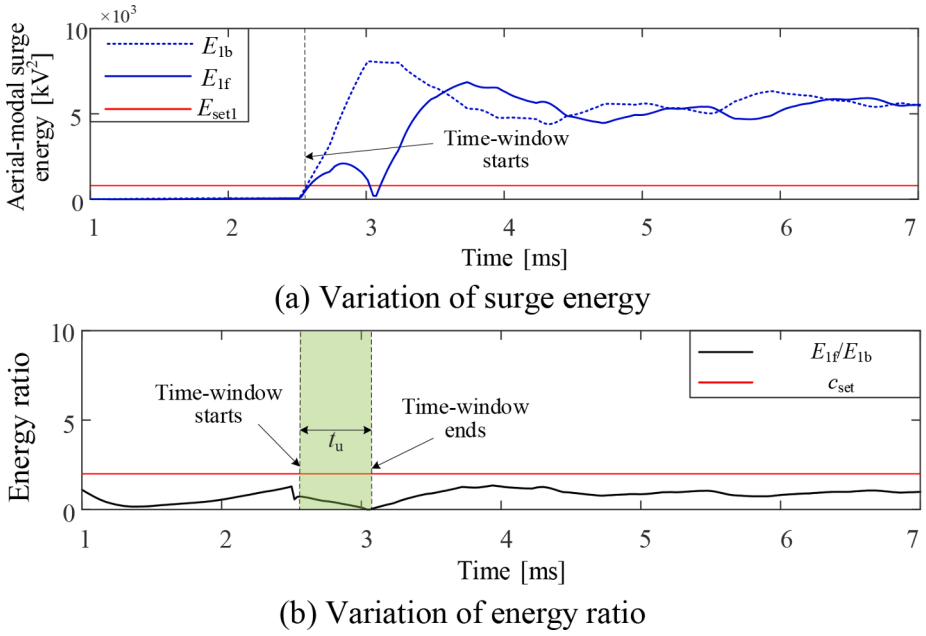


Fig. A12. Forward internal 80%, P-G, fault resistance = 300 Ω at t = 2 ms.

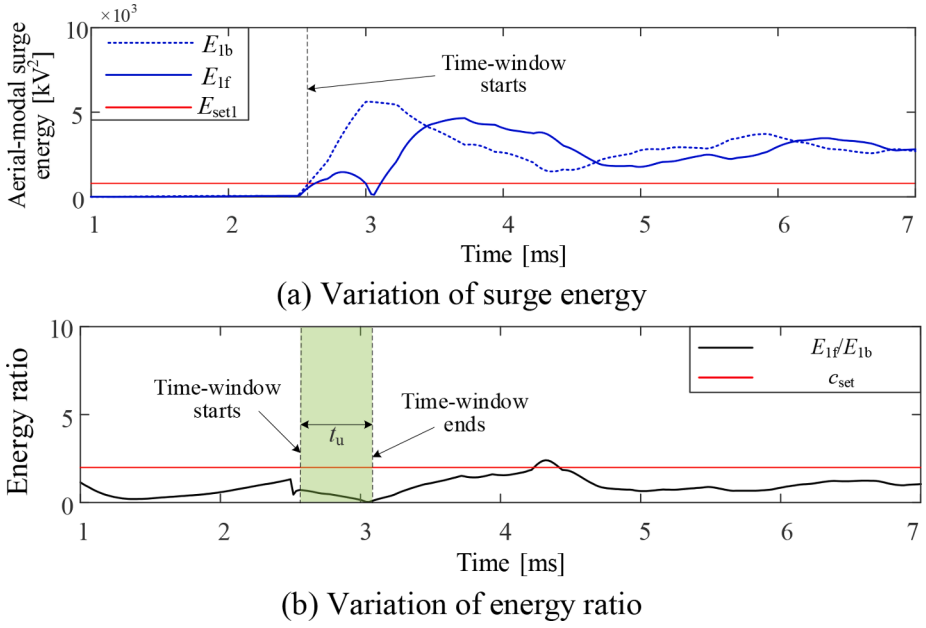


Fig. A13. Forward internal 100%, N-G, bolted fault at t = 3 ms.

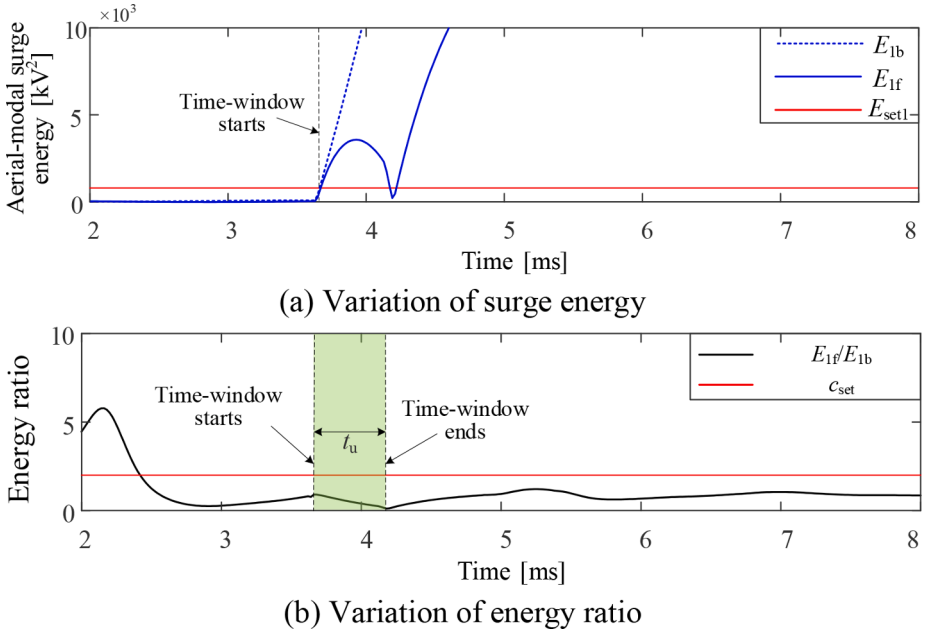


Fig. A14. Forward internal, 100%, N-G, fault resistance = 150 Ω t = 3 ms.

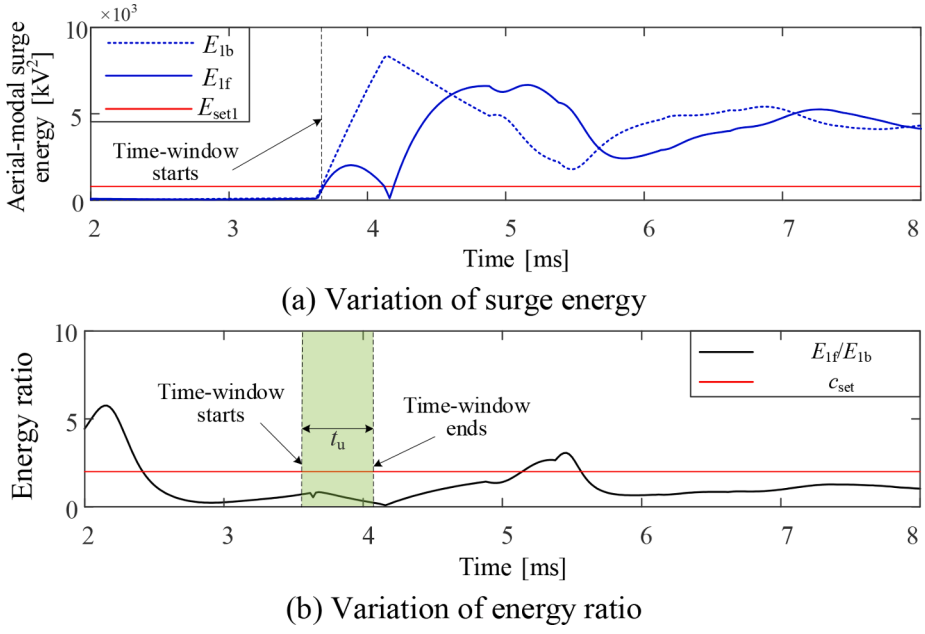


Fig. A15. Forward internal, 100%, N-G, fault resistance = 300 Ω t = 3 ms.

References

[1] Lai CS, Jia Y, Lai LL, Xu Z, McCulloch MD, Wong KP. A comprehensive review on large-scale photovoltaic system with applications of electrical energy storage. Renew Sustain Energy Rev 2017;78:439–51.

[2] Sun K, Xiao H, Pan J, Liu Y. A station-hybrid HVDC system structure and control strategies for cross-seam power transmission. IEEE Trans Power Syst 2021;36(1):379–88.

[3] Ruiwen H, Jianhua D, Lai LL. Reliability evaluation of communication-constrained protection systems using stochastic-flow network models. IEEE Trans Smart Grid 2017;9(3):2371–81.

[4] Cwikowski O, Wood A, Miller A, Barnes M, Shuttleworth R. Operating DC circuit breakers with MMC. IEEE Trans Power Delivery 2018;33(1):260–70.

[5] Suonan J, Gao S, Song G, et al. A novel fault-location method for HVDC transmission lines. IEEE Trans Power Delivery 2009;25(2):1203–9.

[6] Li B, He J, Tian J, Feng Y, Dong Y. DC fault analysis for modular multilevel converter-based system. J Mod Power Syst Clean Energy 2017;5(2):275–82.

[7] Rao H. Architecture of Nan'ao multi-terminal VSC-HVDC system and its multi-functional control. CSEE J Power Energy Syst 2015;1(1):9–18.

[8] Blond SL, Bertho R, Coury DV, Vieira JCM. Design of protection schemes for multi-terminal HVDC systems. Renew Sustain Energy Rev 2016;56:965–74.

[9] He J, Zhang M, Luo G, et al. A fault location method for flexible DC distribution network based on fault transient process. Power Syst Technol 2017;41(3):985–92.

[10] Song G, Li D, Chu X, et al. One-terminal fault location for VSC-HVDC transmission lines based on principles of parameter identification. Power System Technology 2012;36(12):94–9.

[11] Zou G, Feng Q, Huang Q, Sun C, Gao H. A fast protection scheme for VSC based multi-terminal DC grid. Int J Electr Power Energy Syst 2018;98:307–14.

[12] Jiang L, Chen Q, Huang W, Zeng Yu, Wang L, Zhao Pu. A novel directional pilot protection method for VSC-MTDC based on the initial forward and backward travelling wave head. Int J Electr Power Energy Syst 2019;109:198–206.

[13] Zhao P, Chen Q, Sun K. A novel protection method for VSC-MTDC cable based on the transient DC current using the S transform. Int J Electr Power Energy Syst 2018;97:299–308.

[14] Jia K, Wang C, Bi T, et al. Transient current waveform similarity based protection for flexible DC distribution system. IEEE Trans Ind Electron 2019;66(12):9301–11.

[15] Tang L, Dong X, Shi S, Qiu Y. A high-speed protection scheme for the DC transmission line of a MMC-HVDC grid. Electr Power Syst Res 2019;168:81–91.

[16] Zhang C, Song G, Meliopoulos APS, Dong X. Setting-less Non-unit Protection Method for DC Line Faults in VSC-MTDC Systems. IEEE Trans Ind Electron 2022;69(1):495–505.

[17] Li B, Li Y, He J, et al. A novel single-ended transient-voltage-based protection strategy for flexible DC grid. IEEE Trans Power Delivery 2019;34(5):1925–37.

[18] Zhang Y, Cong W. An improved single-ended frequency-domain-based fault detection scheme for MMC-HVDC transmission lines. Int J Electr Power Energy Syst 2021;125:106463.

[19] Zheng X, Tai N, Wu Z, et al. Harmonic current protection scheme for voltage source converter-based high-voltage direct current transmission system. IET Gener Transm Distrib 2014;8(9):1509–15.

[20] Chu X, Song G, Liang J. Analytical method of fault characteristic and non-unit protection for HVDC transmission lines. CSEE J Power Energy Syst 2016;2(4):37–43.

[21] Li S, Chen W, Yin X, et al. Protection scheme for VSC-HVDC transmission lines based on transverse differential current. IET Gener Transm Distrib 2017;11(11):2805–13.

[22] He Z, Liao K. Natural frequency-based protection scheme for voltage source converter-based high-voltage direct current transmission lines. IET Gener Transm Distrib 2015;9(13):1519–25.

[23] Sneath J, Rajapakse AD. Fault detection and interruption in an earthed HVDC grid using ROCOV and hybrid DC breakers. IEEE Trans Power Delivery 2014;31(3):973–81.

[24] Leterme W, Beerten J, Van Hertem D. Nonunit protection of HVDC grids with inductive DC cable termination. IEEE Trans Power Delivery 2015;31(2):820–8.

[25] Song G, Chu Xu, Gao S, Kang X, Jiao Z. A new whole-line quick-action protection principle for HVDC transmission lines using one-end current. IEEE Trans Power Delivery 2015;30(2):599–607.

[26] Guo M, Zeng X, Chen D, et al. Deep-learning-based earth fault detection using continuous wavelet transform and convolutional neural network in resonant grounding distribution systems. IEEE Sens J 2017;18(3):1291–300.

[27] Lan S, Chen M, Chen D. A novel HVDC double-terminal non-synchronous fault location method based on convolutional neural network. IEEE Trans Power Delivery 2019;34(3):848–57.

[28] Tong N, Lin X, Li Y, Hu Z, Jin N, Wei F, et al. Local measurement-based ultra-high-speed main protection for long distance VSC-MTDC. IEEE Trans Power Delivery 2019;34(1):353–64.

[29] Astola J, Kuosmanen P. Fundamentals of nonlinear digital filtering. CRC Press; 2020.

[30] Esakkirajan S, Veerakumar T, Subramanyam AN, PremChand CH. Removal of high density salt and pepper noise through modified decision based unsymmetric trimmed median filter. IEEE Signal Process Lett 2011;18(5):287–90.

[31] Bera S, Shrivastava VK. Analysis of various optimizers on deep convolutional neural network model in the application of hyperspectral remote sensing image classification. Int J Remote Sens 2020;41(7):2664–83.

[32] Zhao P, Chen Q, Sun K. A novel protection method for VSC-MTDC cable based on the transient DC current using the S transform. Int J Electr Power Energy Syst 2018;97:299–308.

[33] Li J, Li Y, Xiong L, et al. DC fault analysis and transient average current based fault detection for radial MTDC system. IEEE Trans Power Delivery 2019;35(3):1310–20.

[34] Li Y, Wu L, Li J, Xiong L, Zhang X, Song G, et al. DC fault detection in MTDC systems based on transient high frequency of current. IEEE Trans Power Delivery 2019;34(3):950–62.

[35] Huang Q, Zou G, Wei X, Sun C, Gao H. A non-unit line protection scheme for MMC-based multi-terminal HVDC grid. Int J Electr Power Energy Syst 2019;107:1–9.

[36] Wei X. "Research and Implementation of FPGA Accelerated Convolutional Neural Network Training," M.Sc. dissertation, School of Mechano-Electronic Engineering. Xi'an, Shaanxi Province, China: Xidian University; 2018.