

Article

# Fault Detection and Classification of Shunt Compensated Transmission Line Using Discrete Wavelet Transform and Naive Bayes Classifier

Elhadi Aker \*, Mohammad Lutfi Othman \*, Veerapandiyan Veerasamy, Ishak bin Aris,  
Noor Izzri Abdul Wahab and Hashim Hizam

Advanced Lightning Power and Energy Research (ALPER), Department of Electrical and Electronics Engineering, Faculty of Engineering, Universiti Putra Malaysia (UPM), 43400 UPM Serdang, Selangor, Malaysia; veerapandian220@gmail.com (V.V.); ishak\_ar@upm.edu.my (I.b.A.); izzri@upm.edu.my (N.I.A.W.); hhizam@upm.edu.my (H.H.)

\* Correspondence: hadi.aker@yahoo.com (E.A.); lutfi@upm.edu.my (M.L.O.); Tel.: +60-11-1083-6907 (E.A.); +60-1-9275-5209 (M.L.O.)

Received: 12 October 2019; Accepted: 17 November 2019; Published: 3 January 2020



**Abstract:** This paper presents the methodology to detect and identify the type of fault that occurs in the shunt compensated static synchronous compensator (STATCOM) transmission line using a combination of Discrete Wavelet Transform (DWT) and Naive Bayes (NB) classifiers. To study this, the network model is designed using Matlab/Simulink. Different types of faults, such as Line to Ground (LG), Line to Line (LL), Double Line to Ground (LLG) and the three-phase (LLLG) fault, are applied at disparate zones of the system, with and without STATCOM, considering the effect of varying fault resistance. The three-phase fault current waveforms obtained are decomposed into several levels using Daubechies (db) mother wavelet of db4 to extract the features, such as the standard deviation (SD) and energy values. Then, the extracted features are used to train the classifiers, such as Multi-Layer Perceptron Neural Network (MLP), Bayes and the Naive Bayes (NB) classifier to classify the type of fault that occurs in the system. The results obtained reveal that the proposed NB classifier outperforms in terms of accuracy rate, misclassification rate, kappa statistics, mean absolute error (MAE), root mean square error (RMSE), percentage relative absolute error (% RAE) and percentage root relative square error (% RRSE) than both MLP and the Bayes classifier.

**Keywords:** static synchronous compensator (STATCOM); discrete wavelet transform (DWT); multi-layer perceptron neural network (MLP); Bayes and Naive Bayes (NB) classifier

## 1. Introduction

Restructuring and deregulation of a power system with increases in energy demand, environmental hurdles, economic factors and right of way, forces the utilities to use the transmission lines to their thermal limits. Also, some developed countries that have surplus power generation supply the load demand through a large number of distribution companies, leading to transmission line overloading. On the other hand, the connection of renewable energies into the grid causes an unbalance in the system voltage. All of these problems can be resolved economically by enhancing the thermal stability of the line through the placement of a flexible AC transmission systems (FACTS) device into the system [1]. Generally, the shunt compensation device like the static compensator (STATCOM) is a widely used FACTS device for increasing the transmission line capability of the system. STATCOM is a parallel-connected device which controls one or more alternating current (AC) system parameters, such as system stability, power quality and voltage control via the injection and absorption of reactive power from the system by adjusting its control action [2–4].

The reliability of power system operation is affected due to occurrences of faults in the transmission lines, leading to equipment damage. In order to ensure the secure and safe operation of the power system network, it is essential to implement an effective protection scheme within the shortest time span to avoid the cascading failure of the system. This is achieved through an advanced fault classification technique that supports an effective, reliable, fast and secured way of relaying operation in the protective system [4]. A numerous study were made for the location of fault in the transmission lines as presented in the literature, only a few of these studies consider the effect of a FACTS-compensated line, and others fail to consider their effects [5–10]. The problem of over-reach and under-reach conditions due to the injection and absorption of reactive power by STATCOM into the system leads to a false tripping of the relay [11]. Therefore, the identification of a fault in the presence of the FACTS device is a crucial issue in power system protection.

Over the years, distance relay-based transmission line protection schemes were adapted for secure and reliable operation of power systems [12–14]. But the presence of series/shunt FACTS devices leads to mal-operation of this conventional relay to detect and locate the fault [15,16]. Moreover, the fault signal is non-stationary in nature, and the analysis of such a signal is a cumbersome process. Therefore, researches proposed the numerical relays based on signal processing techniques, namely Fourier Transform (FT), Fast FT, discrete FT and short time FT that are extensively used in the initial stage for the analysis of the fault signal. It is observed through rigorous analysis that FTs are not suitable for locating a time-varying fault transient signal, and also the information on the time of the occurrence of transients cannot be obtained. To cater for this limitation, S-transform-based fault locations were used for locating the time and frequency information of the fault signal. But it involves a large number of mathematical computation and calculation time that results in degrading the performance of the numerical relay [17–20].

The aforementioned drawbacks are overcome by the time–frequency-based discrete wavelet transform (DWT) approach, which is broadly used for the classification and location of faults and power quality mitigation problems such as sag and swell in the system [21]. One of the major problems with DWT is the selection of mother wavelets for particular applications. However, many works in the literature claimed that Daubechies 4 (db4) is the best suited mother wavelet for power system transient signal analysis like a fault. The detailed explanation of this is portrayed in [22]. The rapid filtering of the original signal from the noise signal with a minimum processing time makes the DWT analysis to extract the features more accurately than other signal processing methods. Because of these reasons the features are extracted using DWT analysis in this work. Then the obtained features are used to train the artificial intelligence (AI) or machine learning (ML) classifiers. Numerous computational intelligence classifiers were proposed in the literature for the location of faults in the system, such as the multilayer perceptron (MLP) neural network, support vector machine (SVM), fuzzy logic, particle swarm optimization (PSO), and so on. The Artificial Neural Network (ANN) and SVM classifiers consume large time for training, and also the efficacy of fuzzy depends upon rules framed by the expertise [6,7,13,23,24]. Besides, many different methods of classifier are proposed in the literature, ranging from a heuristic rule of thumb to formal mathematics [24]. Despite all, the proposed work uses a simple, efficient and sensitive type of a probabilistic neural network-based Naive Bayes (NB) approach for the selection of features, and to classify the type of fault in the system.

The remainder of the paper is organized as follows: Section 2 deals with the system model studied, and Section 3 portrays the proposed method of fault classification with detailed explanation about the extraction of features using DWT analysis. Section 4 describes the MLP neural network- and probabilistic network-based classifiers, such as the Bayes and NB methods to classify the fault when it occurs in the system. Section 5 presents the results and discussion of the proposed work of fault classification with comparative analysis presented in Section 6. The conclusion and future work is made in the last part of the paper.

## 2. System Model Studied

To validate the proposed method of fault detection scheme, it is necessary to acquire the field data from the real time power system network, as the real time data acquisition is a quite tedious and cumbersome process. Therefore, the system under study for fault application considers real time Libya power system data for simulation and the possibility of the occurrence of numerous faults are simulated using Matlab/Simulink. Figure 1 depicts the shunt STATCOM compensated power system model consisting of three phase supply, transmission line network, STATCOM and load. The parameters for the simulation are as follows:

**Generator:** Base MVA rating of 300 MVA, 400 kV, frequency =60 Hz, internal Resistance ( $R_{in}$ ) = 0.8929  $\Omega$ , internal source inductance ( $L_{in}$ ) = 16.56 mH, short circuit MVA rating of 100 and base kV of 300 kV.

**Transmission line:** Positive and zero sequence resistance of 0.0279  $\Omega$ /km and 0.3046  $\Omega$ /km, positive and zero sequence inductance of 0.828 mH/km and 3.820 mH/km, positive and zero sequence capacitance of 11.66  $\eta$ F/km and 7.03  $\eta$ F/km, respectively.

**STATCOM:** Voltage rating of 400 kV with 100 MVA base. The model consists of 48 pulses Voltage Source Converter (VSC).

**Load:** The system consists of an active and reactive power load of 210 MW and 150 MVar respectively.

**Circuit Breakers:** During normal operation, the breaker is considered as closed. For simulation of a fault, the fault is applied for the period of 6/60 to 7/60. The breaker resistance of 0.001  $\Omega$ , the Snubber resistances and the capacitance of 1 M $\Omega$  and infinite value, are considered for simulation study.

The transmission line length of 300 km, considered for each zone (Z1, Z2 and Z3) of line, is assumed to be 100 km. The detailed explanation of simulation parameters and STATCOM are also presented in [11]. The dataset for the training of neural networks (NN) is obtained by introducing the various faults, considering the effect of fault resistance and with/without STATCOM at different locations like 100 km, 200 km and 300 km of the mid-point compensated power system.

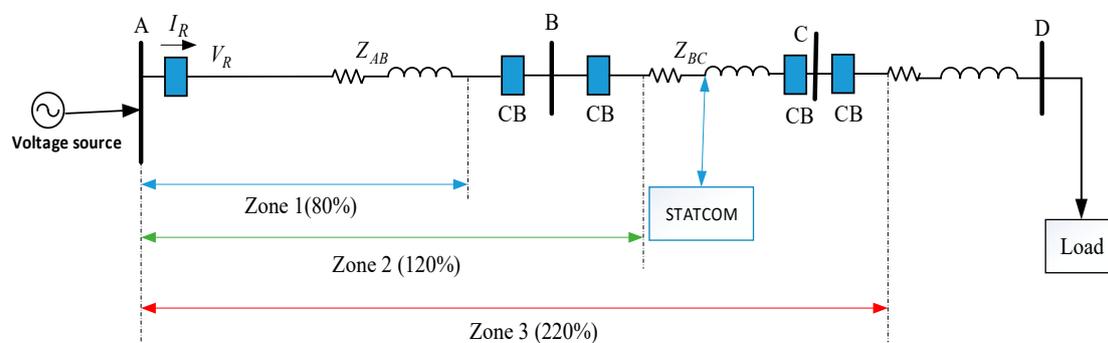


Figure 1. Power System Model.

The power system model is protected from fault by different zones of protection scheme Z1, Z2 and Z3. Thus, the relay responds to various zones of protection, and the trip signal is obtained from the intelligence relaying scheme developed using the NB classifier. In the proposed work, the percentage of distance protection relay by different zones such as Z1, Z2 and Z3 are assumed to be 80%, 120% and 220% of the total line length, respectively.

### Proposed Method of Fault Detection

This section presents the steps for the detection of faults in the power system using the NB method of classification. The detailed steps are illustrated in Figure 2, and also presented as follows:

**Step 1: Data Acquisition**—The shunt compensated power system model is simulated using Matlab/Simulink under various cases of disturbance, and the current signal is obtained for extracting the features to train the NN.

Step 2: Feature Extraction—The data for training are obtained by sampling the current signal using advanced signal processing techniques like DWT, and the features such as SD and energy values are obtained for the system with and without shunt compensation to study the effect of STATCOM compensation.

Step 3: Training Phase—In this phase, the obtained SD and energy values are acquired for different locations of faults and various values of fault resistance.

Step 4: Fault detection—Here, the trained NN is tested for the occurrence of different faults in the system, and this process is repeated for every cycle of operation.

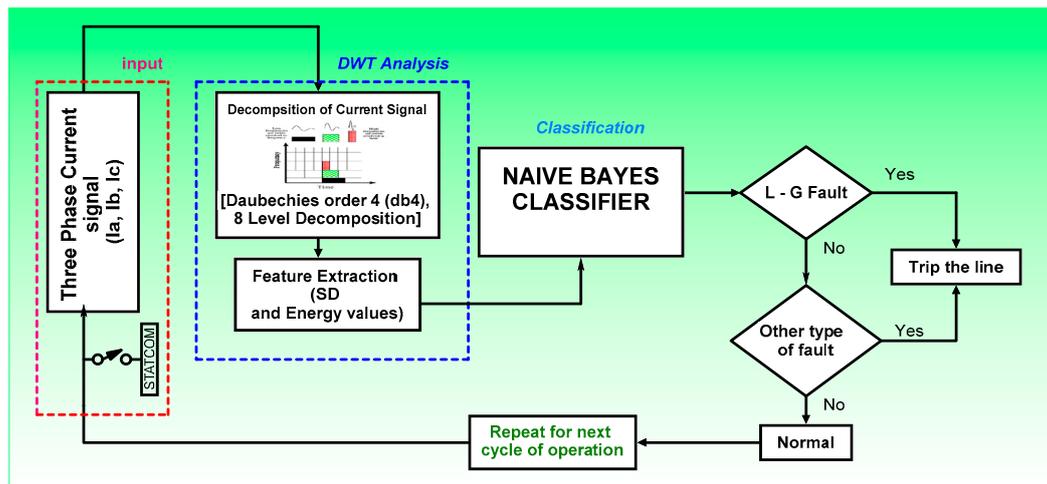


Figure 2. Proposed method of fault classification.

### 3. Feature Extraction Using Discrete Wavelet Transform

Wavelet transform (WT) is a widely used signal processing tool for analyzing the high frequency transient signal in applications like bearing fault detection, transmission line faults and power quality disturbances, namely voltage sag and swell detections, as wavelet analysis overcomes the limitations of FT by localizing the fault signal both in time and frequency domains. Fourier analysis does not provide information about the time of occurrence of the fault/disturbance in the non-stationary current/voltage waveform of the power system. In general, WT exists in two forms: a continuous and a discrete method. The latter is extensively used in the literature, due to its resolution and its applicability in real time. The detailed explanation on the application of WT in a power system is discussed in [21,22].

DWT is a significant tool that analyzes the time varying, transient signal-like faults by decomposing the min to an approximation (A) and detailed coefficients (D) through successive filtering of high-pass and low-pass signals, as depicted in Figure 3.

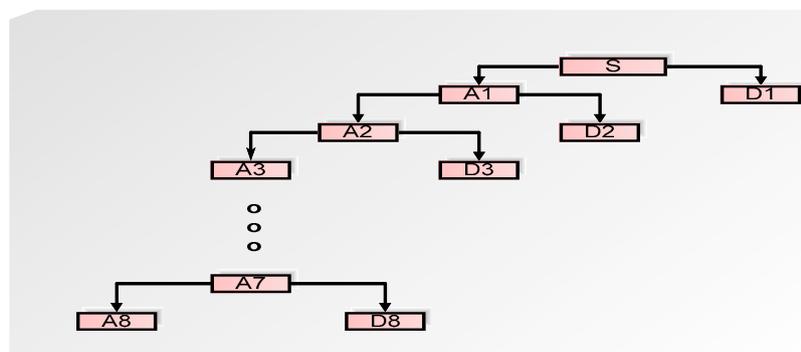


Figure 3. Discrete Wavelet Transform (DWT) Decomposition at eight levels.

As the number of decomposition levels increases, the DC noise present in the fault signal can be suppressed. In this work, a mother wavelet of db4 with eight levels is used to extract the features by sampling the current signal of one cycle with the sampling frequency of 20 kHz and 333 samples per cycle of the current waveform. Among various mother wavelets existent in literature, Daubechies (db4) has been broadly used in power system fault locations because of its ability to locate the fast transients in a low frequency sinusoidal signal. The bandwidths of each level of decomposition are presented in Table 1.

**Table 1.** Detailed Coefficient Levels Frequency Band kHz.

Detailed Coefficient Levels	Frequency Band in kHz
D1	20 to 10
D2	10 to 5
D3	5 to 2.5
D4	2.5 to 1.25
D5	1.25 to 0.625
D6	0.625 to 0.3125
D7	0.3125 to 0.15625
D8	0.15625 to 0.0781

### 3.1. Feature Extractions

The main aim of feature extraction is to provide the significant information for the classifier to classify the type of event through the features calculated, using standard deviation (SD) and energy values. The detailed information of this is discussed as follows.

#### 3.1.1. Standard Deviation (SD)

The SD is defined as the statistical measure of variation or dispersion that exists in the original signal and is given as follows,

$$SD = \sqrt{\frac{1}{n-1} \left[ \sum_{i=1}^n (x_i - \bar{x})^2 \right]} \quad (1)$$

where  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ ,  $x$  represent the data vector and  $n$  is the number of elements in  $x$ .

#### 3.1.2. Energy Value (E)

To test the effectiveness of the proposed classifier, this work uses another approach to calculate features based on the energy of the decomposed current signal. The spectral energy of the decomposed signal can be obtained using Equation (2),

$$E = \sum_{i=1}^n |x_i|^2 \quad (2)$$

where  $n$  is the number of detailed coefficient levels and  $x$  represents the data vector. To calculate the features, a moving window of one cycle of current wavelet coefficient is passed and the features are extracted for training the classifiers [25].

## 4. Fault Classifiers

This section presents Bayesian-based fault classifiers to identify and classify the type of fault that occurs in the shunt STATCOM compensated transmission lines. The comparative study is made with the conventional MLP neural network for the system with and without STATCOM. Here in this work, each fault that occurs in the system is considered as a class, and the same is used for training the neural network. The assumed classes for classification are:  $C_1$ -Normal,  $C_2$ -LG fault,  $C_3$ -LL fault,  $C_4$ -LLG fault

and C<sub>5</sub>-LLLG fault. Moreover, the effectiveness of the method is also tested for occurrence of fault at different locations of the transmission lines.

#### 4.1. Multi-Layer Perceptron (MLP) Network

Multi-Layer Perceptron (MLP) is the most widely used neural network for the identification and detection of the types of fault in a power system in the literature. MLP is a supervised feed forward network, as it requires learning the desired output to be classified. Figure 4 represents the MLP network that consists of the input ( $u_1$ ,  $u_2$  and  $u_3$ ), hidden and output layers.

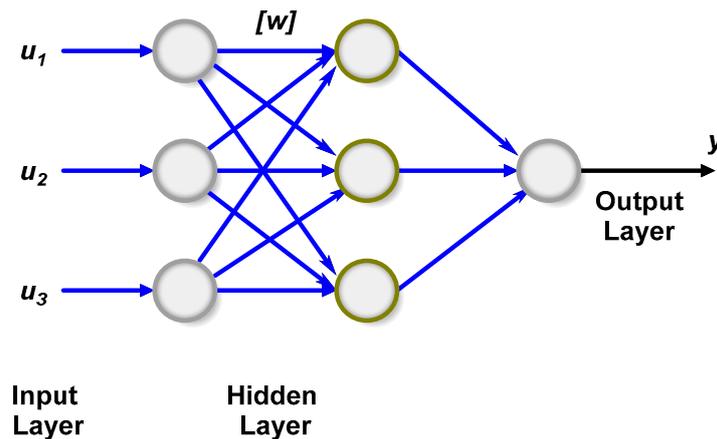


Figure 4. Perceptron (MLP) neural network.

The output [ $y$ ] of the network is the weighted sum of input neurons and is defined as,

$$y_i = W_{i_0} + \sum_{j \in \text{pred}(i)} (W_{ij} a_j) \quad (3)$$

where  $a_j$  represents the output of the previous layer neuron,  $W_{ij}$  is the weight between the  $i$ th and  $j$ th neuron, and  $W_{i_0}$  is the input bias of this neuron. In this work, the MLP network is trained using the back propagation method, and the detailed explanation is presented in [26,27].

#### 4.2. Bayes and Naive Bayes Classifiers

The conventional MLP neural network performs the classification by adjusting the weight of the network through a small penalty factor that sometimes leads to over fitting. This problem is overcome using a principle approach called Bayes theorem by the Bayesian neural network (BNN). BNN was invented by Israeli Judea Pearl in 1980s, a statistical-based, supervised classifier that determines the variable to be classified in a more way relevant to the class, by evaluating the probability of how likely is its occurrence in that class. This is achieved with the prior information obtained about the occurrence of event that takes the form of prior probability density function [28–31]. Thus, the Bayes theorem can be defined as

$$\text{Posterior probability} = \frac{\text{Class prior probability} * \text{likelihood}}{\text{Predictor prior probability}} \quad (4)$$

The simplified form can be expressed as,

$$P(C|L_1, L_2, \dots, L_n) = \frac{P(C) \cdot P(L_1, L_2, \dots, L_n | C)}{P(L_1, L_2, \dots, L_n)} \quad (5)$$

$$P(C|L) = \frac{P(C) \cdot P(L|C)}{P(L)} \quad (6)$$

where  $P(C)$  is the class probability and  $P(L|C)$  represents the likelihood of datasets  $\{L_1, L_2, \dots, L_n\}$  of variables in class  $C = [C_1, C_2, \dots, C_5]$ . The classification problem can be defined as,

$$\arg \max \left[ P(C|L) = \frac{P(C) \cdot P(L|C)}{P(L)} \right] \tag{7}$$

Here the attribute  $P(L)$  does not vary with the class and can be assumed as constant, and the above equation is approximated as,

$$\arg \max [P(C|L) = P(C) \cdot P(L|C)] \tag{8}$$

In Equations (7) and (8), the most probable output from the input arguments (data) is represented as  $\arg \max$ . It is also called the global maxima of output.

The computation burden of BNN increases as the number of likelihood terms in the class raises exponentially with the attributes  $L = \{L_1, L_2, \dots, L_n\}$ . To cater this limitation, all features in a class are assumed to be independent, and that results in the Naive Bayes (NB) classifier that reduces the number of parameters to be estimated from  $2(2n - 1)$  to  $2n$  [25,30,31]. NB is a linear classifier that divides the input data set into the training and prediction step for identifying the type of class using Bayes' theorem. In the training phase, the classifier determines the probability distribution pertaining to the features of any given class is independent. During the prediction phase, our classifier estimates the posterior probability of the test sample data belonging to a respective class. Then the method classifies the samples based on the maximum likelihood of posterior probability. The NB classifier has been widely used because of its simplicity, being easy to implement with a high accuracy and sound theoretical basis that guarantees the optimized results. The probability function defined in (8) can be rewritten with the assumption of independent features as,

$$P(C|L_1, L_2, \dots, L_n) = P(C) \cdot P(L_1|C)P(L_2|C) \dots P(L_n|C) \tag{9}$$

In this work,  $L$  is assumed to be the number of variables, i.e., the type of fault that occurs in the system. Let  $L = \{L_1, L_2, L_3, L_4, L_5\} = \{\text{Normal}, \text{LG}, \text{LL}, \text{LLG}, \text{LLL}\}$ , then  $P(L)$  denotes the probability distribution over the sesystem states, as represented in Figure 5,

$$P(L) = \{x_1, x_2, x_3, x_4, x_5\}, x_i \geq 0; \tag{10}$$

$$\sum_{i=1}^n x_i = 1 \tag{11}$$

where  $x_i$  is the probability of  $L$  for being in state  $L_i$ . The assumed probability of each disturbance is as follows:  $P(\text{Normal}) = P(L_1) = 0.2$ ,  $P(\text{LG}) = P(L_2) = 0.2$ ,  $P(\text{LL}) = P(L_3) = 0.3$ ,  $P(\text{LLG}) = P(L_4) = 0.2$  and  $P(\text{LLL}) = P(L_5) = 0.2$ .

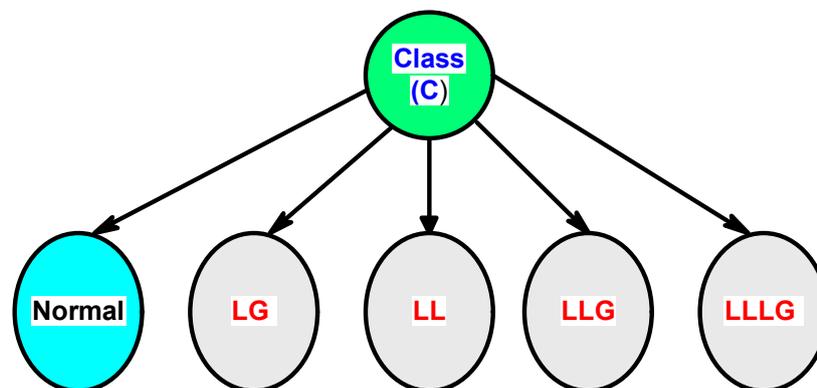


Figure 5. Naive Bayes (NB) classifier of the proposed work.

The conditional probability for the proposed work considering different possible events is portrayed in Table 2. It is seen that the classifier has  $(12 \times 5) = 60$  probabilities.

**Table 2.** Conditional Probability for the proposed work.

Cases		Normal	LG	LLG	LL	LLLG
Without STATCOM	SDA	0.33	0.89	0.5	0.49	0.34
	SDB	0.33	0.05	0.48	0.49	0.35
	SDC	0.33	0.06	0.02	0.02	0.31
	E-A	0.58	0.08	0.01	0.01	0.43
	E-B	0.23	0.89	0.5	0.51	0.24
	E-C	0.19	0.02	0.49	0.48	0.33
With STATCOM	SDA	0.33	0.22	0.12	0.13	0.34
	SDB	0.34	0.56	0.46	0.44	0.34
	SDC	0.33	0.22	0.42	0.43	0.32
	E-A	0.54	0.39	0.03	0	0.43
	E-B	0.15	0.42	0.08	0.44	0.24
	E-C	0.31	0.19	0.89	0.56	0.32

#### 4.3. Performance Indices of Classifier

The Kappa Statistic (K) is the statistical measure of classifiers that compute the constancy among the predicted type of fault and the actual type of fault, and is defined as follows,

$$K = \frac{P(\text{OF}) - P(\text{EF})}{(1 - P(\text{EF}))} \quad (12)$$

where  $P(\text{OF})$  is the probability of the observed fault,  $P(\text{EF})$  is the probability of the predicted type of fault. It ranges between 0 and 1.

Mean Absolute Error (MAE) and Root Mean Square Error (RMSE)—MAE is the absolute mean of the error calculated between the predicted and observed value, and is depicted as follows [21],

$$\text{MAE} = \frac{|\sum_{i=1}^n (E_P - E_O)|}{n} \quad (13)$$

RMSE is the square root of the mean of variance between the predicted and observed type of fault detected by the classifiers, and is given by,

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (E_P - E_O)^2}{n}} \quad (14)$$

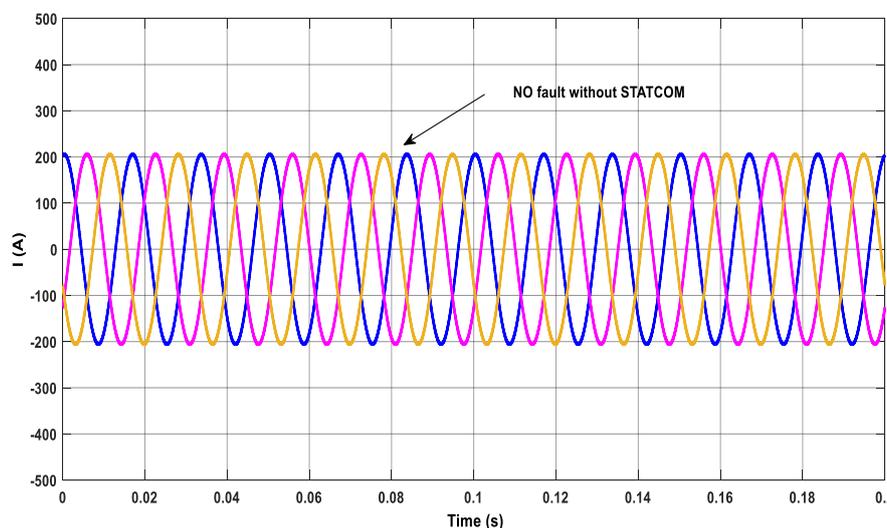
where  $E_P$  is the predicted type of fault, and  $E_O$  is the expected type of fault.

## 5. Results and Discussion

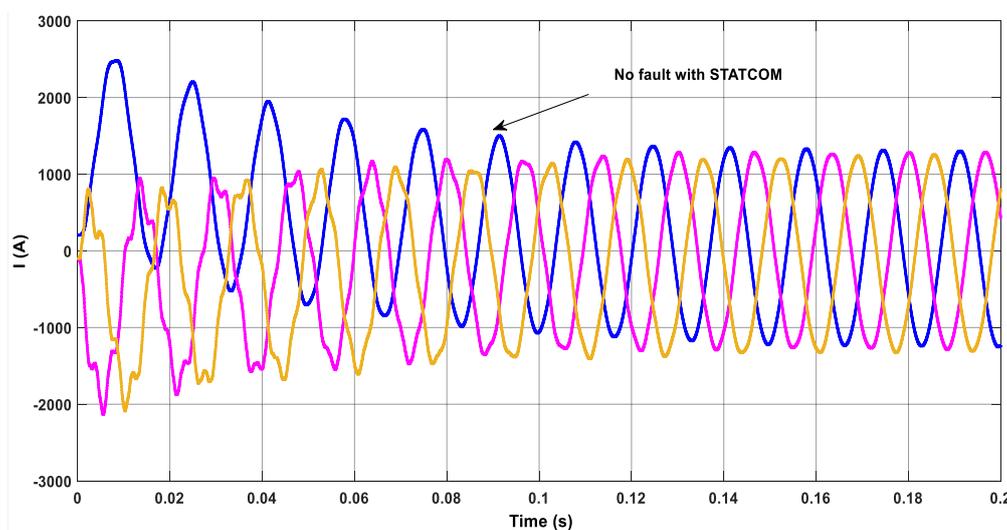
This section describes the simulation of a proposed probabilistic NB-based classifier to classify the fault and the location of the fault in a transmission line. The effect of the probabilistic classifier is studied for the transmission line with and without compensations. The simulation is carried out for the power system model depicted in Figure 1, and various plausible faults such as LG, LL, LLG and LLLG in the system, considering the variation in fault resistances. The simulation is carried out for a time period of one cycle, and the fault is applied during 0.1 to 0.12 s. Figures 6 and 7 depict the three-phase current waveform of the system without and with STATCOM, respectively. The minimum and maximum values of the peak magnitude of this three phase current signal are captured for the system with and without compensation that are illustrated in Tables 3 and 4. It is seen from the results, the magnitude of current signal increases for the system with STATCOM device, and the same is

presented in the form of a waveform; for the case of the LG fault in the system with and without STATCOM, and these are portrayed in Figures 8 and 9 respectively.

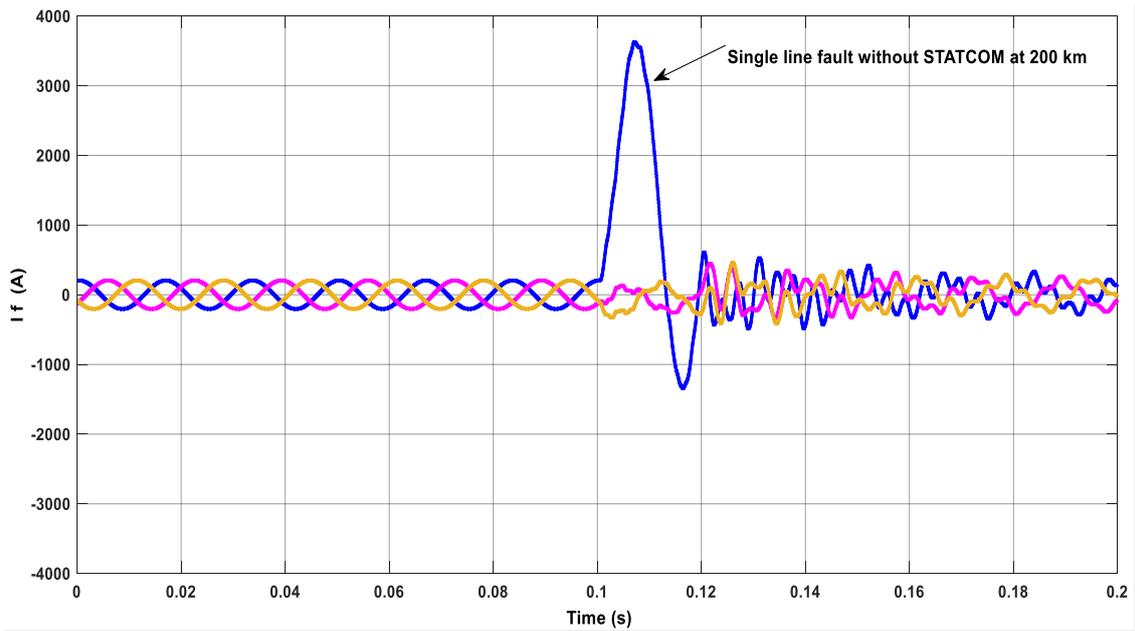
Then the current signal obtained for various cases of fault is analyzed using the db4 mother wavelet of DWT analysis with eight level coefficients to extract the features, such as SD and energy values for training the classifiers. Figures 10 and 11 represent the DWT analysis of current waveform under the normal operation of the system without and with STATCOM, respectively. In general, the magnitude of coefficients is high for the compensated system compared to the uncompensated system. Figures 12 and 13 portray the DWT analysis of the LG fault current waveform considering without and with STATCOM, respectively. Also, it is observed that the coefficients of the detailed coefficient are low when a fault occurs after the location of the STATCOM (at 150 km) device. This effect is due to the STATCOM, where the system fault current reduces as the distance of the fault increases from the fault location point. Tables 5 and 6 represent the extracted features (SD and energy values) for training the classifiers. The trained classifiers are tested with the test data, and the type of fault that occurs in the system is detected by the classifiers. The performance of the classifier for classification of various faults in the system for cases with and without STATCOM, using the features of SD and energy values, are presented as different cases, as discussed in forthcoming subsections.



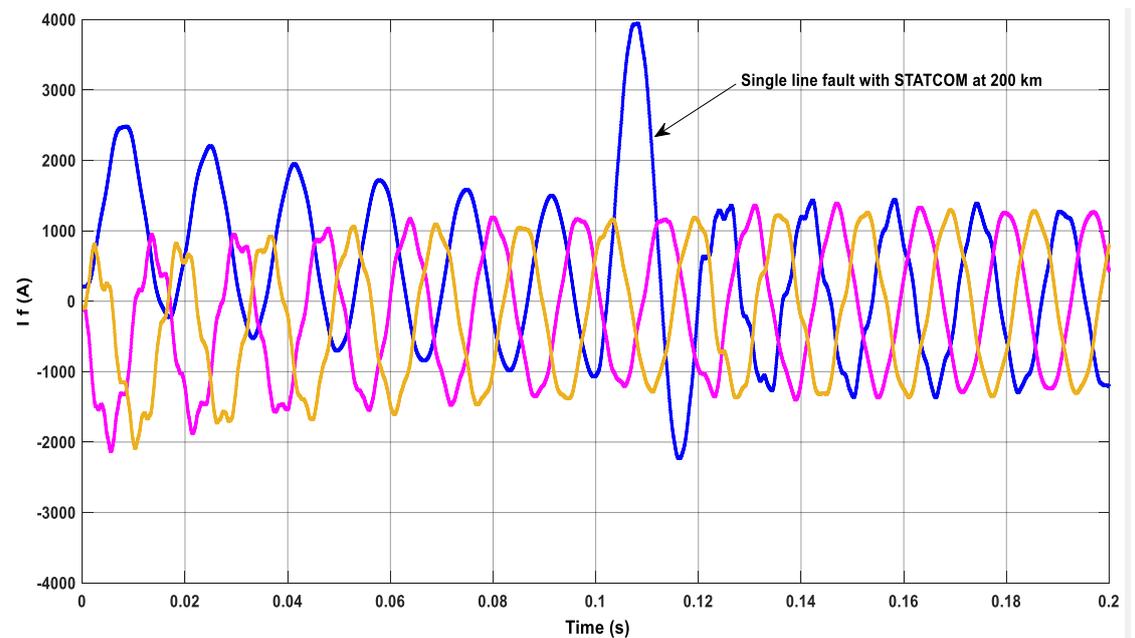
**Figure 6.** Three-phase current waveform under normal condition without STATCOM compensation.



**Figure 7.** Three-phase current waveform under normal conditions with midpoint compensation.



**Figure 8.** Three-phase current waveform during a Line to Ground (LG) fault (at 200 km) in Phase A without STATCOM compensation.



**Figure 9.** Three-phase current waveform during the LG fault (at 200 km) in Phase A with STATCOM compensation.

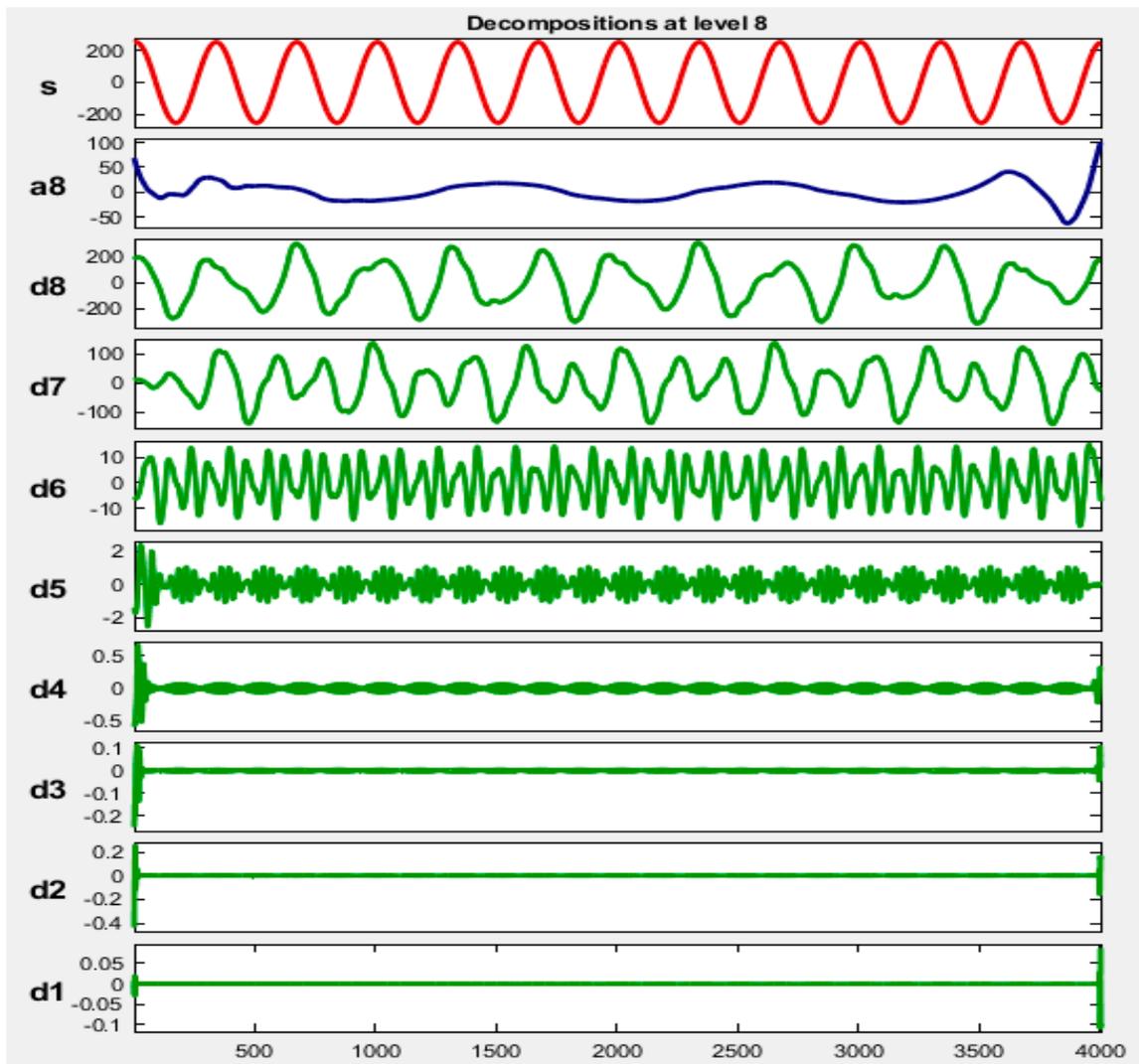


Figure 10. Discrete Wavelet Transform (DWT) analysis of Phase A under normal conditions without compensation.

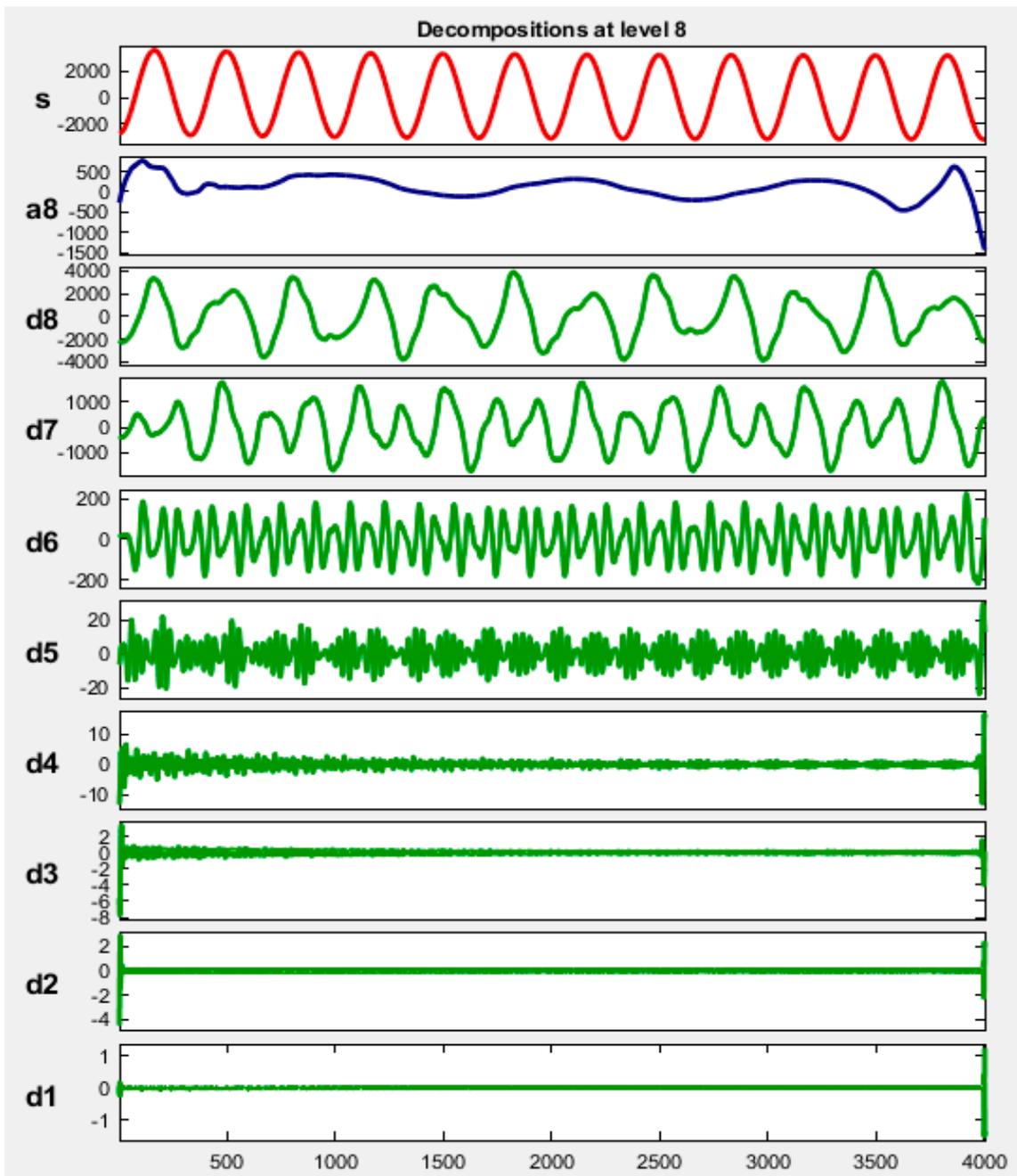


Figure 11. DWT analysis of Phase A under normal conditions with STATCOM compensation.

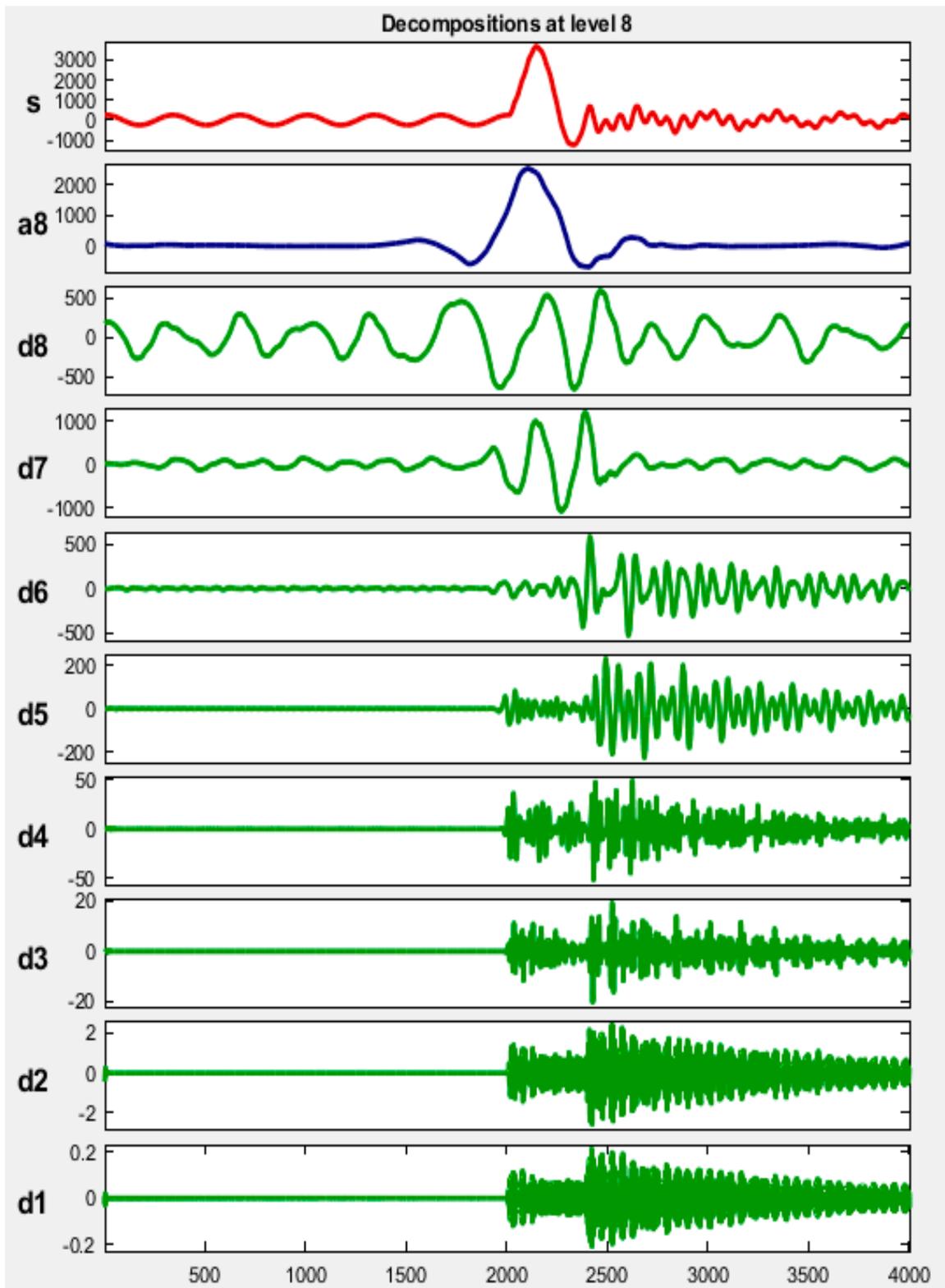


Figure 12. DWT analysis of Phase A during the LG fault without STATCOM compensation.

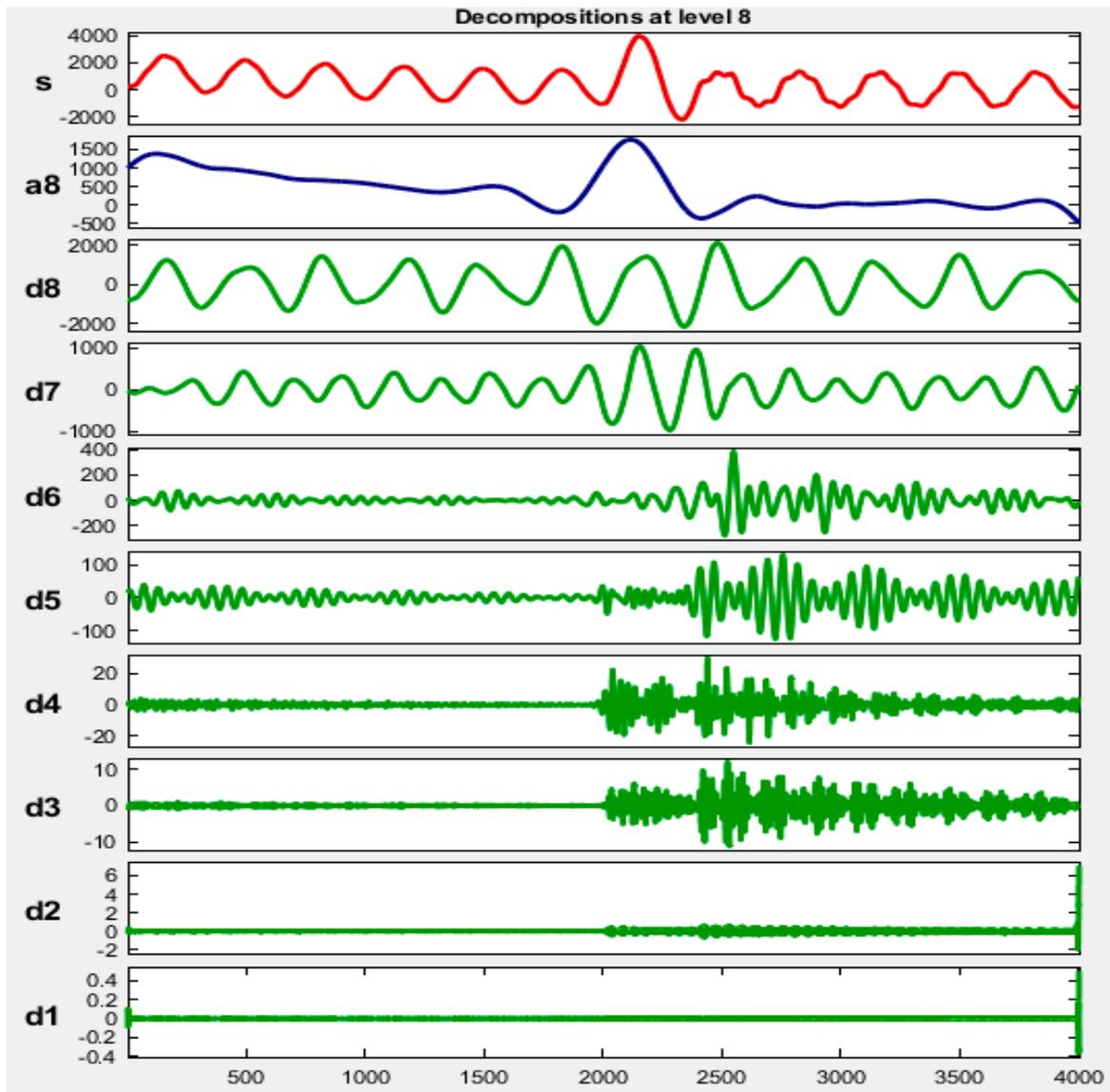


Figure 13. DWT analysis of Phase A during the LG fault with STATCOM compensation.

**Table 3.** Current magnitude during normal conditions and faults at different locations without the STATCOM-compensation.

Fault Distance	Type of Fault	Without STATCOM					
		Minimum Current			Maximum Current		
		I a kA	I b kA	I c kA	I a kA	I b kA	I c kA
100 km	No fault	−0.205	−0.205	−0.205	0.205	0.205	0.205
	LG	−2.57	−0.34	−0.46	6.95	0.28	0.25
	LL	−4.11	−12.5	−0.25	12.6	4.05	0.25
	LLG	−4.19	−12.0	−0.71	1.34	4.3	0.65
	LLLG	−3.88	−12.0	−12.4	1.52	6.76	4.3
200 km	LG	−1.23	−0.27	−0.39	3.67	0.19	0.18
	LL	−2.19	−7.01	−0.25	7.06	2.16	0.25
	LLG	−2.1	−6.78	−0.45	7.56	2.34	0.38
	LLLG	−1.97	−7.06	−7.16	8.32	3.78	2.82
300 km	LG	−0.78	−0.294	−0.37	2.49	0.185	0.19
	LL	−1.56	−4.78	−0.25	4.93	1.47	0.25
	LLG	−1.51	−4.85	−0.51	5.08	1.62	0.37
	LLLG	−1.31	−5.17	−4.97	5.72	2.62	2.16

**Table 4.** Current magnitude during normal conditions and faults at different locations with STATCOM-compensation.

Fault Distance	Type of Fault	With STATCOM					
		Minimum Current			Maximum Current		
		I a kA	I b kA	I c kA	I a kA	I b kA	I c kA
100 km	No f	−1.22	−2.07	−2.08	2.51	1.26	1.21
	LG	−3.36	−1.04	1.17	6.95	1.23	0.8
	LL	−4.57	−11.7	−1.24	11.8	4.58	1.07
	LLG	−4.74	−11.4	−1.3	1.26	4.82	1.18
	LLLG	−4.57	−11.5	−1.1.9	1.43	7.02	4.91
200 km	LG	−2.20	−1.12	−1.23	3.97	1.23	1.08
	LL	−2.80	−6.3	−1.25	6.38	2.71	1.07
	LLG	−2.85	−6.25	−1.36	6.76	2.99	1.09
	LLLG	−2.72	−6.47	−6.46	4.49	4.06	3.3
300 km	LG	−1.85	−1.19	−1.28	3.18	1.22	0.84
	LL	−2.22	−4.56	−1.27	4.61	2.22	1.07
	LLG	−2.33	−4.61	−1.38	4.88	2.41	1.17
	LLLG	−2.22	−4.84	−4.79	5.32	3.24	2.68

**Table 5.** Standard deviation (SD)-based feature values for classification.

Condition	Type of Fault	Location km	Without STATCOM			With STATCOM		
			SD-A ( $\times 10^3$ )	SD-B ( $\times 10^3$ )	SD-C ( $\times 10^3$ )	SD-A ( $\times 10^3$ )	SD-B ( $\times 10^3$ )	SD-C ( $\times 10^3$ )
Normal	No fault	100	0.177	0.177	0.177	0.875	0.877	0.866
		200	0.177	0.177	0.177	0.875	0.877	0.866
		300	0.177	0.177	0.0177	0.875	0.877	0.866
LG	AG	100	3.087	0.166	0.204	3.394	0.8	0.797
		200	1.582	0.154	0.196	2.046	0.825	0.817
		300	1.058	0.145	0.19	1.674	0.851	0.835
	BG	100	0.3	3.17	0.267	0.793	3.49	0.835
		200	0.245	1.63	0.198	0.821	2.11	0.836
		300	0.238	1.1	0.196	0.838	1.72	0.859
	CG	100	0.263	0.299	2.66	0.854	0.811	3.305
		200	0.193	0.243	1.37	0.852	0.831	1.888
		300	0.193	0.238	0.921	0.874	0.849	1.569
LLG	ABG	100	5.865	5.65	2.82	5.81	5.69	0.766
		200	3.158	3.03	2.06	3.188	3.14	0.803
		300	2.14	2.15	2.05	2.357	2.33	0.832
	BCG	100	0.188	5.65	4.99	0.755	5.67	5.12
		200	0.17	3.06	2.71	0.799	3.14	2.87
		300	0.161	2.09	1.84	0.834	2.35	2.16
	CAG	100	5.108	0.287	5.15	5.247	0.759	5.21
		200	2.749	0.203	2.79	2.932	0.794	2.9
		300	1.842	0.202	1.87	2.202	0.833	2.17
LL	AB	100	5.723	5.67	0.177	5.633	5.67	0.838
		200	3.097	3.04	0.177	3.085	3.11	0.842
		300	2.105	2.05	0.177	2.279	2.3	0.849
	BC	100	0.177	5.255	5.691	0.851	5.281	5.245
		200	0.177	2.868	2.832	0.856	2.944	2.905
		300	0.177	1.964	1.929	0.86	2.204	2.164
	CA	100	4.998	0.177	5.06	5.112	0.846	5.04
		200	2.693	0.177	2.75	2.85	0.852	2.8
		300	1.8	0.177	1.86	2.131	0.858	2.09
LLLG	ABCG	100	6.254	6.48	5.69	6.224	6.43	5.75
		200	3.368	3.51	3.1	3.397	3.37	3.19
		300	2.263	2.39	2.1	2.493	2.58	2.36

**Table 6.** Energy-based feature values for classification.

Condition	Type of fault	Location km	Without STATCOM			With STATCOM		
			E-A ( $\times 10^8$ )	E-B ( $\times 10^8$ )	E-C ( $\times 10^8$ )	E-A ( $\times 10^8$ )	E-B ( $\times 10^8$ )	E-C ( $\times 10^8$ )
Normal	No fault	100	1.25	0.49	0.4	22.7	6.26	13.1
		200	1.25	0.49	0.4	22.7	6.26	13.1
		300	1.25	0.49	0.4	22.7	6.26	13.1
LG	AG	100	96.4	0.56	0.51	128	5.36	11.4
		200	25.9	0.56	0.51	56.5	5.62	12.1
		300	12	0.51	0.46	42.7	5.85	12.3
	BG	100	1.64	57.1	0.51	21.3	70.7	13.2
		200	1.44	15.3	0.41	25.8	27.5	12.3
		300	1.5	7.08	0.37	22.3	18.8	13
	CG	100	1.39	0.76	72.9	21.7	5.74	97.1
		200	1.33	0.6	18.8	21.2	6.22	38.6
		300	1.18	0.63	8.47	22.1	6.11	28.8
LLG	ABG	100	301	223	0.71	307	214	11.3
		200	87.1	65	0.5	105	64	12.8
		300	41.8	30.4	0.45	63.8	34.3	13.6
	BCG	100	1.36	184	179	20.8	185	200
		200	1.2	54.6	53.4	21.3	58.1	67
		300	1.18	25	22.7	22.1	32.8	44.4
	CAG	100	318	0.73	313	326	5.17	305
		200	94.6	0.51	93	106	5.09	94.9
		300	41.6	0.52	41.2	66.9	5.53	56.5
LL	AB	100	255	254	4.05	265	234	12.8
		200	74.7	73.9	0.4	92.6	68.3	12.9
		300	35.6	35	0.4	56.8	35.8	12.9
	BC	100	1.24	174	169	22.4	170	186
		200	1.24	53	50.2	22.4	49.5	62.4
		300	1.24	23.5	22.3	22.4	30.2	40.8
	CA	100	308	0.5	312	314	5.8	300
		200	91.5	0.49	93.4	103	5.87	91.7
		300	40.5	0.49	41.3	65.5	5.98	54.3
LLLG	ABCG	100	425	241	315	414	236	315
		200	125	70.5	94.7	130	71.9	97
		300	57.5	33.3	40.7	76.6	38.6	59.1

Case-1: In this study, the transmission fault classification and identification in a transmission network is done without STATCOM. Table 7 presents the confusion matrix for classification of different states of the system, such as Normal, LG, LLG, LL and LLLG fault. Here, the fault in the system is classified using the SD values obtained by the DWT analysis for different types of fault occurring at the distances of 100 km, 200 km and 300 km of an overhead transmission line, and this is given in Table 5. Then these data are used for training the neural network and the classification results obtained are presented in Table 8.

The result shows that the proposed NB method of classifier is more accurate compared to the MLP and Bayes methods of classification. Moreover, the % misclassification rate of the proposed method is 0%, whereas the rate is 20% and 80% for the MLP and Bayes approaches of classification, respectively. The MLP method of classification fails to detect the LLG type of fault, and on the other hand, the Bayes method fails to classify all types of fault and whose performance is inferior compared to other methods.

It is inferred from Figure 14 and Table 8 that the NB classifier is the most significant method to classify the various types of fault in the system compared to all other methods.

**Table 7.** Confusion Matrix for Classification.

Classes	C1	C2	C3	C4	C5	System State
C1	1	0	0	0	0	Normal
C2	0	1	0	0	0	LG
C3	0	0	1	0	0	LLG
C4	0	0	0	1	0	LL
C5	0	0	0	0	1	LLLG

Case-2: Here in this study, the classification and identification of the fault is done without STATCOM, as in case-1. But in this case, instead of SD values, the energy values obtained from DWT analysis for different types of faults occurring at various distances of 100 km, 200 km and 300 km has been taken for training the network, which is illustrated in Table 6. The results obtained reveal that the NB method of classification is better than the other two methods, such as MLP and Bayes classifiers. Figure 14 represents the % accuracy rate of the proposed method is 100%, whereas this is 60% and 20% for MLP and Bayes networks, respectively. The MLP method of classification fails to detect LG and LLG faults, whilst the Bayes classifier is unable to detect all types of faults. It is seen that the propounded NB has a 0% misclassification rate, the MLP has 40% and the Bayes method has 80% of the misclassification rate, as depicted in Table 8.

Case-3: This case is similar to case-1, but in this study the STATCOM is connected at the midpoint of the transmission line, and the occurrence of faults at different locations such as 100 km, 200 km and 300 km, are studied. The SD values obtained are used to train the network, like the case-1, and shown in Table 5. It is observed from Table 8 that the proposed NB classifier performance is more predominant in terms of accuracy and % misclassification rate compared to the MLP and Bayes methods of classification, and is also shown in Figure 14. The Bayes method fails to identify all types of fault, except when the system is operating in normal conditions and MLP method fails to detect the LLG type of fault as with case-1. It is inferred from the results that both the MLP and Bayes classifier performances are the same for the transmission line involving with and without STATCOM, and the proffered NB method classifier outperforms compared to these approaches.

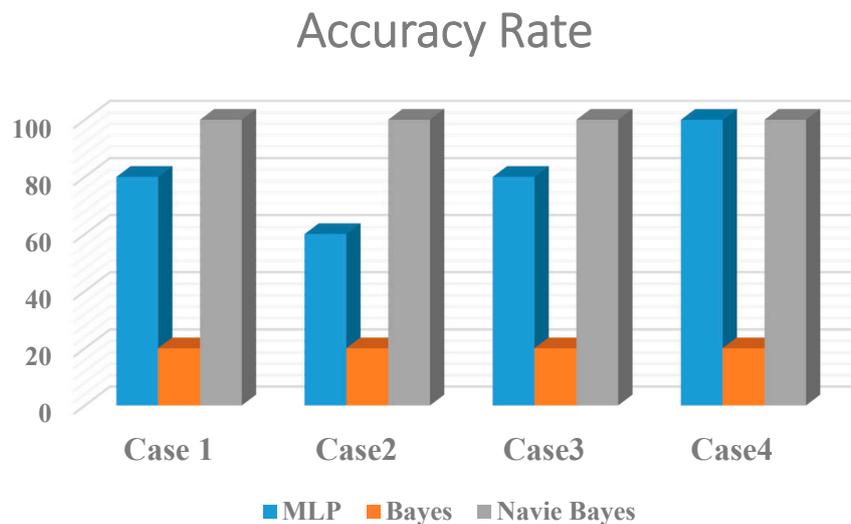
**Table 8.** Classifiers' Accuracy and Misclassification Rate.

Cases	Accuracy Rate			Misclassification Rate					
	MLP	Bayes	Naïve Bayes	MLP		Bayes		Naive Bayes	
				% Rate	Type of Fault	Rate	Type of Fault	Rate	Type of Fault
Case-1	80	20	100	20	C3	80	C2-C5	0	0
Case-2	60	20	100	40	C2-C3	80	C2-C5	0	0
Case-3	80	20	100	20	C3	80	C2-C5	0	0
Case-4	100	20	100	0	0	80	C2-C5	0	0

Case-4: This case is analogous to case-2, with the incorporation of STATCOM connected at the midpoint of the transmission line for supporting the reactive power and to improve the voltage profile of the system performance. In this context, the energy values obtained from DWT analysis for different types of faults at various distances of 100 km, 200 km and 300 km has been used for training the network and this is portrayed in Table 6.

Figure 14 represents the proposed NB classifier is very efficient compared to the MLP and Bayes methods. The % accuracy of NB and MLP are 100%, but the Bayes method is only 20% accurate. On the flipside, the % misclassification rate is 0% for NB and the MLP method, and it is 80% for the Bayes approach. It is deduced from the results, the proffered NB classifier gives accurate results for all

cases and its performance is significantly predominant than the MLP and Bayes method as depicted in Table 8.



**Figure 14.** Accuracy rate of the classifiers.

#### Performance Evaluation of Classifiers

The robustness of the classifier is evaluated by various performance indices, such as Kappa Statistics (KS), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Percentage Relative Absolute Error (%RAE) and Percentage Root Relative Square Error (%RRSE) for classifiers, namely Bayes, MLP and the NB approach. Firstly, the KS index for various classifiers is presented in Table 9 and Figure 15. The result shows that the indices are '1' for the proposed NB classifier for all the cases and the values lies in the range of 0.5–1 for the MLP classifier (for various cases) and is almost '0' for the Bayes method of classification. It is inferred from the KS index, the proffered method of classifier outperforms for various cases compared to the other classifiers. Secondly, the MAE is less than 0.1 for the proposed classifier, whereas the value lies in the range of 0.1–0.3 for the MLP method, and it is greater than 0.3 for the Bayes approach under various cases. Moreover, the RMSE is also less than 0.1 for the NB method, and the value lies in the range of 0.2–0.4 for MLP, and it is almost 0.4 for the Bayes classifier for case-1 to case-4. It is seen that the indices such as MAE and RMSE are comparatively very low, as shown in Figures 16 and 17 for the intended NB method of classifier than other approaches presented, proving that the proposed classifier is more robust and efficient.

Lastly, the %RAE and %RRSE is proven to be significantly less for the propounded NB method compared to the MLP and Bayes classifiers, as depicted in Table 10 and Figure 18. It is observed that the results outperform for all the cases by the NB approach rather than the MLP and Bayes classifier methods.

**Table 9.** Performance comparison of various Classifiers.

	Kappa Statistics			MAE			RMSE		
	MLP	Bayes	Naive Bayes	MLP	Bayes	Naive Bayes	MLP	Bayes	Naive Bayes
Case-1	0.75	0	1	0.159	0.32	0.025	0.236	0.4	0.088
Case-2	0.5	0	1	0.201	0.32	0	0.292	0.4	0
Case-3	0.75	0	1	0.172	0.32	0.033	0.248	0.4	0.097
Case-4	1	0	1	0.155	0.32	0	0.227	0.4	0

## Kappa Statistics

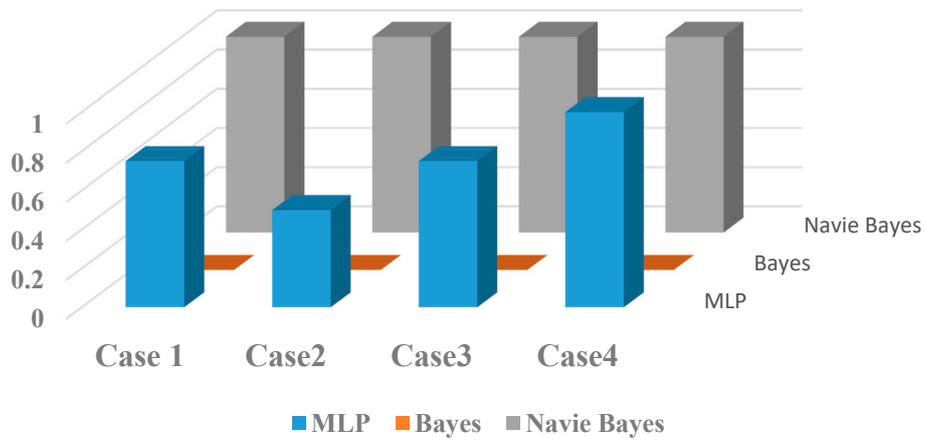


Figure 15. KappaStatistics comparison of various classifiers.

## MAE

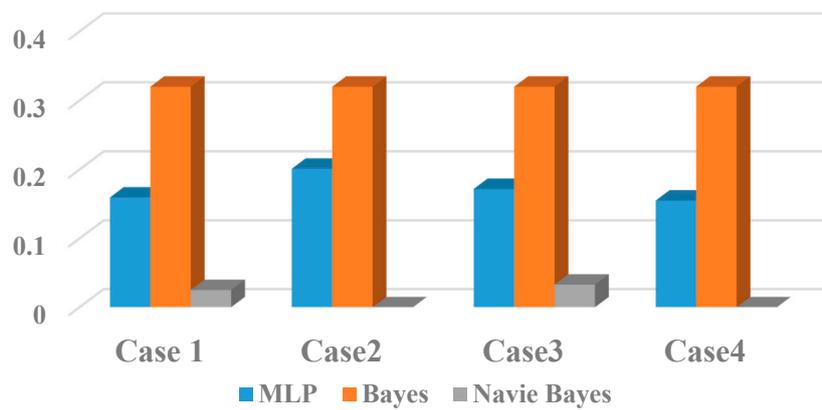


Figure 16. MAE comparison of various classifiers.

## RMSE

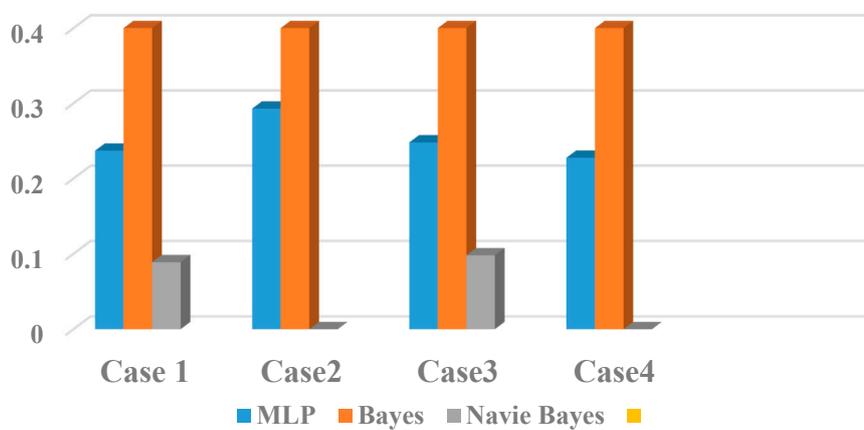
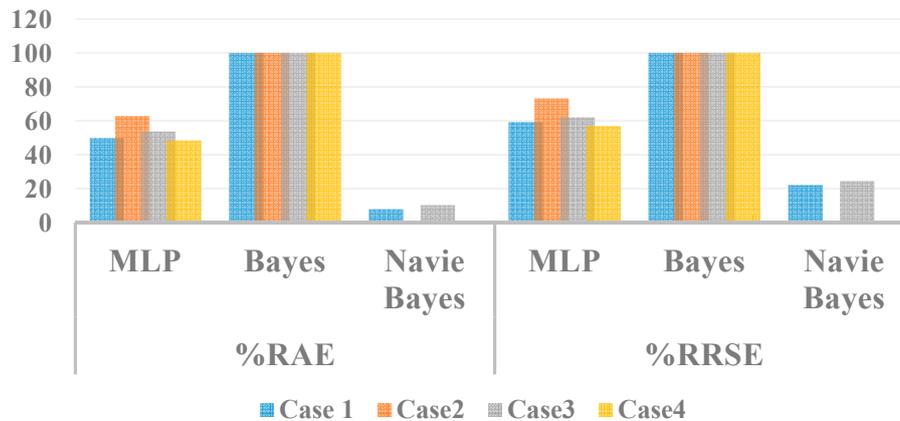


Figure 17. RMSE comparison of various classifiers.

## Relative Error



**Figure 18.** Percentage relative absolute error (% RAE) comparison of various classifiers.

**Table 10.** Percentage root relative square error (% RRSE) comparison of various classifiers.

Cases	%RAE			%RRSE		
	MLP	Bayes	Naive Bayes	MLP	Bayes	Naive Bayes
Case-1	49.89	100	7.85	59.23	100	22.21
Case-2	62.86	100	0	73.22	100	0
Case-3	53.74	100	10.29	62	100	24.46
Case-4	48.46	100	0	56.89	100	0

## 6. Comparative Analysis

This section describes the comparative analysis of various power system fault classification methods portrayed in the literature works, summarized considering the significant features of % accuracy, and illustrated in Table 11. The comparative performance of the % accuracy of the proposed NB method is made with other existing methods. The results indicate that all of the techniques have an error in identifying the type of faults because of the over-reach/under-reach of relay, due to presence of STATCOM in the transmission line. However, few literatures fail to compare the performance of the classifier without compensation.

**Table 11.** Performance comparison with literature work.

Authors	Methods	Type of Fault Considered					Fault Resistance	STATCOM	%Accuracy
		LG	LL	LLG	LLL	LLLG			
Singh. A.R [1]	Synchronized Measurements	√	√	√	√	√	√	√	99.6
Ghazizadeh A. [3]	Synchronized Measurements	√	√	√	×	√	×	√	99.07
Mishra. S.K [4]	DWT	√	√	√	×	×	√	√	-
Gupta. O.H [8]	Superimposed sequence components-based integrated impedance (SSCII).	√	√	√	√	√	√	SVC	-
Albasri. F.A [10]	Impedance Measurements	√	√	×	√	×	×	√	-
Hussain. S [15]	Unsynchronized Measurements	√	√	×	√	×	√	√	99
Proposed Work	DWT & NB	√	√	√	×	√	√	√	100

“√” represents the occurrence of fault, “×” represents the fault type is not considered for classification.

But in this paper, the identification and classification is done considering with and without STATCOM by the proffered NB classifier, and also comparison is made with MLP and the Bayes neural network. It is observed that the NB method of classification outperforms to give superior results for both the cases of system model with and without STATCOM. Also, the performance indices of the

classifier, such as Kappa statistics, MAE, RMSE, %RAE and %RRSE are evaluated to show the accuracy of presented classifiers, which is elsewhere presented in the literature.

## 7. Conclusions

This paper presents a novel probabilistic-based Naive Bayes approach to locate the fault in a shunt STATCOM compensated transmission line. In this work, a high voltage power system model of 400 kV has been simulated using MATLAB/Simulink, and various faults such as LG, LL, DLG and LLLG, are applied. The current waveform obtained under different cases of normal and fault cases are analyzed using DWT to extract the features for locating the type of fault. The fault current signal are sampled with different band of frequencies that depict the 1st, 2nd, 3rd, 4th, 5th, 6th, 7th and 8th level of the detailed coefficient and its approximation coefficient at the 8th level. The SD and Energy values have been obtained for different faults with a fault resistance of 0.001  $\Omega$ . The obtained features are used to train the classifiers to classify the type of fault. The obtained results showed that the proposed NB classifier outperforms to give a 100% accuracy rate in the case of with and without STATCOM. On the flipside, the MLP method gives an average accuracy rate of 80%, with Bayes of 20%. It also inferred from the performance indices such as Kappa statistics, MAE, %RAE and %RRSE, that the proffered NB approach gives the predominant result compared to the MLP and Bayes classifier method.

In the proposed work, though the system is specific, it is subjected to fault occurrence for various scenarios of being with and without STATCOM. Also, to test the robustness of the proposed classifier, two different features, such as standard deviation (SD) and energy values, are taken for the system with and without STATCOM. In all of the four cases presented in this paper, the proffered method of classifier outperforms to give better results than other classifiers. This claim proves even the system considered is the same, but the trained data features have the ability to give better results for various cases. So, this method can also give better results for other power system models, too. Further, the location of the fault and detection of its zones of occurrence considering with and without STATCOM is the future scope of the work.

Moreover, the presented work with the Internet of Things (IoT) paves the way for the smart relaying scheme that helps the utility to locate and isolate the faulty section from the healthy part of power system, thereby minimizing the cascading failures of the system.

**Author Contributions:** E.A. and M.L.O. proposed the main idea and performs the simulation of the work; V.V. and N.I.A.W. provided sources and wrote the paper; proof reading and final drafting was done by I.b.A. and H.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** The research received funding from the Universiti Putra Malaysia named Geran Putra Berimpak with ref. UPM/800-3/3/1/GPB/2019/9671700.

**Acknowledgments:** The authors are thankful to Center for Advanced Lightning Power and Energy Research (ALPER) and Universiti Putra Malaysia to carry out this research.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Singh, A.R.; Patne, N.R.; Kale, V.S. Synchronized measurement based an adaptive distance relaying scheme for STATCOM compensated transmission line. *Measurement* **2018**, *116*, 96–105. [[CrossRef](#)]
2. Mirzaei, M.; Vahidi, B.; Hosseinian, S.H. Accurate fault location and faulted section determination based on deep learning for a parallel-compensated three-terminal transmission line. *IET Gener. Transm. Distrib.* **2019**, *13*, 2770–2778. [[CrossRef](#)]
3. Ghazizadeh-Ahsaee, M.; Sadeh, J. Accurate fault location algorithm for transmission lines in the presence of shunt-connected flexible AC transmission system devices. *IET Gener. Transm. Distrib.* **2012**, *6*, 247. [[CrossRef](#)]
4. Mishra, S.K.; Tripathy, L.N.; Swain, S.C. DWT approach based differential relaying scheme for single circuit and double circuit transmission line protection including STATCOM. *Ain Shams Eng. J.* **2019**, *10*, 93–102. [[CrossRef](#)]

5. Chen, K.; Huang, C.; He, J. Fault detection, classification and location for transmission lines and distribution systems: A review on the methods. *High Volt.* **2016**, *1*, 25–33. [[CrossRef](#)]
6. Veerasamy, V.; Abdul Wahab, N.I.; Ramachandran, R.; Thirumeni, M.; Subramanian, C.; Othman, M.L.; Hizam, H. High-impedance fault detection in medium-voltage distribution network using computational intelligence-based classifiers. *Neural Comput. Appl.* **2019**, *31*, 9127–9143. [[CrossRef](#)]
7. Veerasamy, V.; Abdul Wahab, N.I.; Ramachandran, R.; Mansoor, M.; Thirumeni, M.; Othman, M.L. High impedance fault detection in medium voltage distribution network using discrete wavelet transform and adaptive neuro-fuzzy inference system. *Energies* **2018**, *11*, 3330. [[CrossRef](#)]
8. Gupta, O.H.; Tripathy, M. An innovative pilot relaying scheme for shunt-compensated line. *IEEE Trans. Power Deliv.* **2015**, *30*, 1439–1448. [[CrossRef](#)]
9. Mehrjerdi, H.; Ghorbani, A. Adaptive algorithm for transmission line protection in the presence of UPFC. *Int. J. Electr. Power Energy Syst.* **2017**, *91*, 10–19. [[CrossRef](#)]
10. Albasri, F.A.; Sidhu, T.S.; Varma, R.K. Performance comparison of distance protection schemes for shunt-FACTS compensated transmission lines. *IEEE Trans. Power Deliv.* **2007**, *22*, 2116–2125. [[CrossRef](#)]
11. Aker, E.E.; Othman, M.L.; Aris, I.; Wahab, N.I.A.; Hizam, H.; Emmanuel, O. Adverse impact of STATCOM on the performance of distance relay. *Indones. J. Electr. Eng. Comput. Sci.* **2017**, *6*, 528–536. [[CrossRef](#)]
12. Othman, M.L.; Aris, I.; Wahab, N.I.A. Modeling and simulation of the industrial numerical distance relay aimed at knowledge discovery in resident event reporting. *Simulation* **2014**, *90*, 660–686. [[CrossRef](#)]
13. Othman, M.L.; Aris, I.; Othman, M.R.; Osman, H. Rough-Set-and-Genetic-Algorithm based data mining and Rule Quality Measure to hypothesize distance protective relay operation characteristics from relay event report. *Int. J. Electr. Power Energy Syst.* **2011**, *33*, 1437–1456. [[CrossRef](#)]
14. Othman, M.L.; Aris, I.; Othman, M.R.; Osman, H. Rough-Set-based timing characteristic analyses of distance protective relay. *Appl. Soft Comput. J.* **2012**, *12*, 2053–2062. [[CrossRef](#)]
15. Hussain, S.; Osman, A.H. Fault location on series and shunt compensated lines using unsynchronized measurements. *Electr. Power Syst. Res.* **2014**, *116*, 166–173. [[CrossRef](#)]
16. Kumar, B.; Yadav, A.; Pazoki, M. Impedance differential plane for fault detection and faulty phase identification of FACTS compensated transmission line. *Int. Trans. Electr. Energy Syst.* **2019**, *29*, e2804. [[CrossRef](#)]
17. Megahed, A.I.; Moussa, A.M.; Bayoumy, A.E. Usage of wavelet transform in the protection of series-compensated transmission lines. *IEEE Trans. Power Deliv.* **2006**, *21*, 1213–1221. [[CrossRef](#)]
18. Patel, B. A new FDOST entropy based intelligent digital relaying for detection, classification and localization of faults on the hybrid transmission line. *Electr. Power Syst. Res.* **2018**, *157*, 39–47. [[CrossRef](#)]
19. Krishnanand, K.R.; Dash, P.K. A new real-time fast discrete S-transform for cross-differential protection of shunt-compensated power systems. *IEEE Trans. Power Deliv.* **2013**, *28*, 402–410.
20. Dash, P.K.; Das, S.; Moirangthem, J. Distance protection of shunt compensated transmission line using a sparse S-transform. *IET Gener. Transm. Distrib.* **2015**, *9*, 1264–1274. [[CrossRef](#)]
21. Wilkinson, W.A.; Cox, M.D. Discrete wavelet analysis of power system transients. *IEEE Trans. Power Syst.* **1996**, *11*, 2038–2044. [[CrossRef](#)]
22. Osman, A.H.; Malik, O.P. Protection of Parallel Transmission Lines Using Wavelet Transform. *IEEE Trans. Power Deliv.* **2004**, *19*, 49–55. [[CrossRef](#)]
23. Devasahayam, V.; Veluchamy, M. An enhanced ACO and PSO based fault identification and rectification approaches for FACTS devices. *Int. Trans. Electr. Energy Syst.* **2017**, *27*, e2344. [[CrossRef](#)]
24. Dash, P.K.; Samantaray, S.R.; Panda, G. Fault classification and section identification of an advanced series-compensated transmission line using support vector machine. *IEEE Trans. Power Deliv.* **2007**, *22*, 67–73. [[CrossRef](#)]
25. Hindarto, H.; Muntasa, A. Wavelet sub-band energy for feature extraction of electro encephalo graph (EEG) signals. *J. Eng. Sci. Technol.* **2019**, *14*, 578–588.
26. Patil, K.; Jadhav, N. Multi-Layer Perceptron Classifier and Paillier Encryption Scheme for Friend Recommendation System. In Proceedings of the 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA), Pune, India, 17–18 August 2017; pp. 1–5. [[CrossRef](#)]
27. Bebarta, D.K.; Rout, A.K.; Biswal, B.; Dash, P.K. Efficient Prediction of Stock Market Indices using Adaptive Neural Network. In Proceedings of the Second International Conference on Soft Computing for Problem Solving (SocProS 2012), Jaipur, India, 28–30 December 2012; Springer: Jaipur, India, 2014; Volume 236, ISBN 978-81-322-1601-8.

28. Zhang, N.; Wu, L.; Yang, J.; Guan, Y. Naive bayes bearing fault diagnosis based on enhanced independence of data. *Sensors* **2018**, *18*, 463. [[CrossRef](#)]
29. Youn, E.; Jeong, M.K. Class dependent feature scaling method using naive Bayes classifier for text datamining. *Pattern Recognit. Lett.* **2009**, *30*, 477–485. [[CrossRef](#)]
30. Addin, O.; Sapuan, S.M.; Mahdi, E.; Othman, M. A Naïve-Bayes classifier for damage detection in engineering materials. *Mater. Des.* **2007**, *28*, 2379–2386. [[CrossRef](#)]
31. Muralidharan, V.; Sugumaran, V. A comparative study of Naïve Bayes classifier and Bayes net classifier for fault diagnosis of monoblock centrifugal pump using wavelet analysis. *Appl. Soft Comput. J.* **2012**, *12*, 2023–2029. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).