Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 October 2019 doi:10.20944/preprints201910.0148.v1

Peer-reviewed version available at Energies 2020, 13, 243; doi:10.3390/en13010243

Article

Fault Detection and Classification of Shunt Compensated Transmission line using Discrete Wavelet Transform and Naive Bayes Classifier

Elhadi Aker[1], Mohammad Lutfi Othman[1], Veerapandiyan Veerasamy[1], Ishak Aris[1], Noor Izzri Abdul Wahab[1] and Hashim Hizam[1]

[1]Advanced Lightning and Power Energy Research (ALPER), Department of Electrical and Electronics Engineering, Faculty of Engineering, Universiti Putra Malaysia (UPM), 43400 UPM Serdang, Selangor, Malaysia; hadi.aker@yahoo.com(E.A); lutfi@upm.edu.my(M.L.O); veerapandian220@gmail.com(V.V); ishak_ar@upm.edu.my (I.A) ; izzri@upm.edu.my(N.I.A.W); hhizam@upm.edu.my (H.H).
*Correspondence: hadi.aker@yahoo.com; Tel.: (+601110836907); lutfi@upm.edu.my (M.L.O); Tel(+060192755209).

Abstract: This paper presents the methodology to detect and identify the type of fault that occurs in shunt connected static synchronous compensator (STATCOM) transmission line using a combination of Discrete Wavelet Transform (DWT) and Naive Bayes classifier. To study this, the network model is designed using Mat-lab/Simulink. The different faults such as Line to Ground (LG), Line to Line (LL), Double Line to Ground (LLG) and three-phase (LLLG) fault are applied at different zones of 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 mother wavelet of db4 to extract the features such as standard deviation and Energy values. The extracted features are used to train the classifiers such as Multi-Layer Perceptron Neural Network (MLP), Bayes and Naive Bayes (NB) classifier to classify the type of fault that occurs in the system. The results 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), relative absolute error (RAE) and root relative square error (RRSE) than MLP and 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 power system with increase in energy demand, environmental hurdles, economic factors and right of way forces the utilities to use the transmission lines to its thermal limit. Also, some developed countries that have surplus power generation supplies the load demand through large number of distribution companies leading to transmission line overloading. On the other hand, the connection of renewable energies into the grid causes unbalance in the system voltage. The utilities resolve all these problems economically by enhancing the thermal stability of the line through placement of flexible AC transmission systems (FACTS) device into the system [1]. The shunt compensation device like static compensator (STATCOM) is widely used FACTS device for increasing the transmission line capability of the system. STATCOM is a parallel connected device which controls one or more AC system parameters such as system stability, power quality and voltage control via injection and absorption of reactive power from the system by adjusting its control action [2-4]. The reliability of power system operation is affected by occurrence of fault in transmission line 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 shortest time span to avoid the cascading failure of the system. This is achieved

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 October 2019 doi:10.20944/preprints201910.0148.v1

Peer-reviewed version available at *Energies* **2020**, *13*, 243; doi:10.3390/en13010243

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 location of fault in transmission lines in the literature, only some of the study involves effect of FACTS compensated line and other fails 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 false tripping of relay [11]. Therefore, identification of fault in the presence of FACTS device is a crucial issue in power system protection.

Distance relay based transmission line protection schemes were adapted for secure and reliable operation of system [12-14]. But, the presence of series/shunt FACTS device leads to mal-operation of conventional relay to detect and locate the fault [15, 16]. Moreover, the fault signal is non-stationary in nature and the analysis of such 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 analysis of fault signal. It is observed through rigorous analysis that FTs are not suitable for locating time-varying fault transient signal and also the information on time of occurrence of transients cannot be obtained. To cater this limitation S-transform based fault location were used for locating the time and frequency information of fault signal. But it involves large number of mathematical computation and calculation time that results in degrading the performance of numerical relay [17-20].

The aforementioned drawback are overcome by the time-frequency based discrete wavelet transform (DWT) approach and is broadly used for classification and location of faults, power quality mitigation problems such as sag and swell in the system [21]. One of the major issues with DWT is selection of mother wavelets and many works in the literature on analysis of power system transients claimed that Daubechies 4 (db4) is best suited for fault analysis [22]. Because of fast filtering with less processing time makes the DWT analysis than other methods for extracting the features to train the Artificial Intelligence (AI) or machine learning (ML) classifiers in the proposed work. Also, numerous computational intelligence classifiers were proposed for location of fault in the system such as multilayer perceptron (MLP) neural network, support vector machine (SVM), fuzzy logic, particle swarm optimization(PSO) and so on. The ANN and SVM classifiers consume large time for training and the efficacy of fuzzy depends on rules framed by the expertise [6, 7, 13 23, 24.]. Also, many different methods of classifier are proposed in the literature ranging from heuristic rule of thumb to formal mathematics [24]. Despite of all, the proposed work uses a simple, efficient and sensitive type of probabilistic neural network based Naive Bayes (NB) approach for selection of features 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 classifications with detailed explanation about extraction of features using DWT analysis. Section 4 describes the MLP neural network and probabilistic network based classifiers such as Bayes and NB method to classify the fault occurs in the system. Section 5 presents the results and discussion of proposed work of fault classification with conclusion and future work 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 quite tedious and cumbersome process. Therefore, the system under study for fault application considers a real time Libya power system data for simulation and the possibility of occurrence of numerous faults are simulated using Mat lab/Simulink. Figure 1 depicts the shunt STATCOM compensated power system model and the parameters for simulation are as follows: Generator rating – 300 MVA, 400kV, 60Hz and line length of 300 km with each zones (Z1, Z2 and Z3) of line is assumed to be 100 km and load rating of 260 MVA. The detailed explanation of simulation parameters and STATCOM are presented in [11]. The dataset for training of neural networks (NN) are obtained by introducing the various fault considering effect of fault resistance and with/without STATCOM at different locations like 100km, 200km and 300 km of mid-point compensated power system.

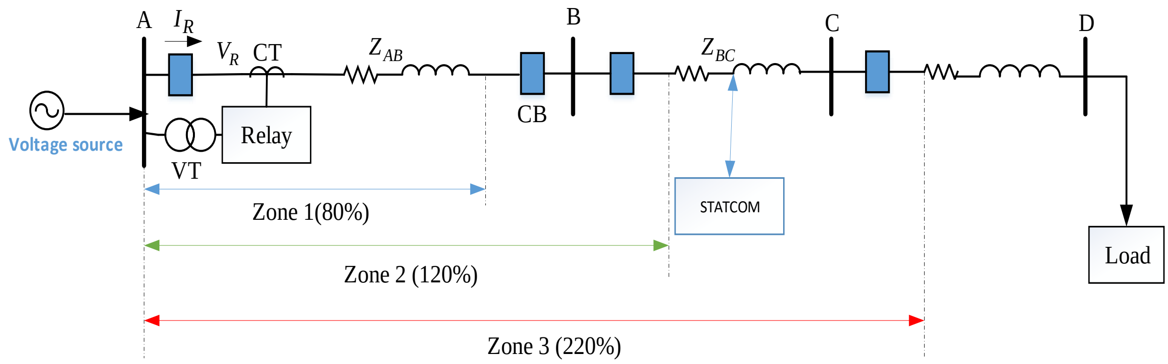


Figure 1. Libya Power System Model

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 a 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 total line length respectively.   .

2. 1. Proposed Method of Fault Detection

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

Step-1 **Data Acquisition** - The shunt compensated power system model is simulated using Mat lab/Simulink under various cases of disturbances 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 standard deviation (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 location of faults and various values of fault resistance.

Step-4 **Fault detection** – Here, the trained NN is tested for occurrence of different faults in the system and this process repeats for every cycle of operation.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 October 2019 doi:10.20944/preprints201910.0148.v1

Peer-reviewed version available at *Energies* **2020**, *13*, 243; doi:10.3390/en13010243

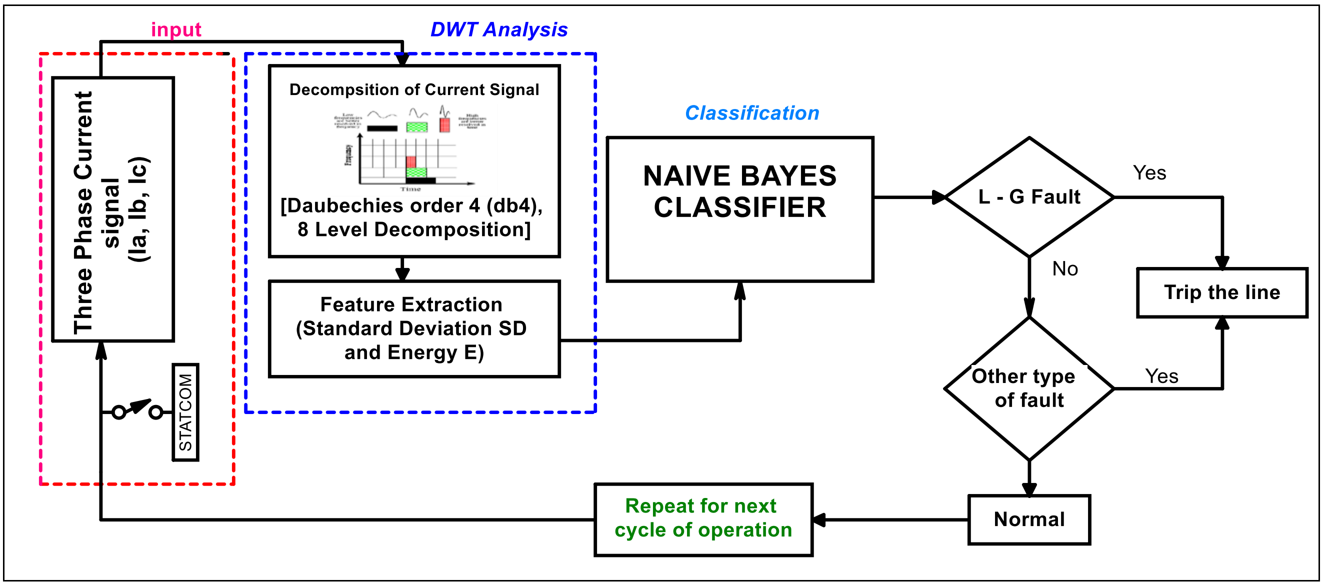


Figure 2. Proposed method of fault classification

3. Feature Extraction using Discrete Wavelet Transform

Wavelet transform (WT) have been widely used for analyzing the transient signal in ample number of applications like mechanical vibrations, image processing and also electrical power system fault detection. As wavelet analysis overcome the limitations of FT by localizing the fault signal both in time and frequency domains. As Fourier analysis, does not provide information about the time of occurrence of fault/disturbance in non-stationary current/voltage waveform of power system. In general WT exists in two forms: continuous and discrete method. The later is extensively used in the literature, due to its resolution and its applicability in real time. The detailed explanation on application of WT in power system is discussed in [21,22].

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

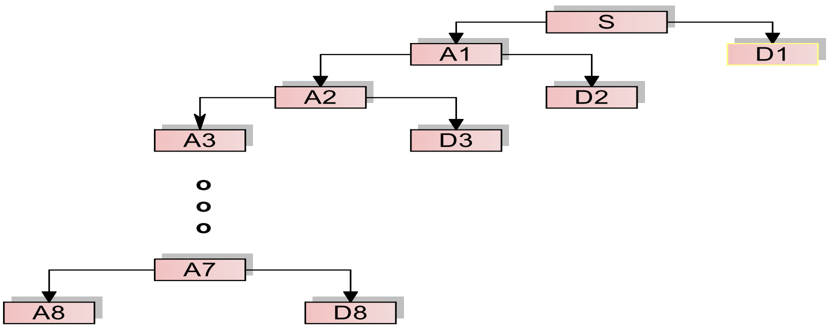


Figure 3. DWT Decomposition at eight levels

As the number of decomposition level increases, the DC noise present in the fault signal can be suppressed. In this work, an mother wavelet of Db4 with 8-level 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 current waveform. Among various mother wavelets exist in literature, Daubechies (Db4) have been broadly used in power system fault locations because of its ability to locate the fast

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 October 2019 doi:10.20944/preprints201910.0148.v1

Peer-reviewed version available at Energies 2020, 13, 243; doi:10.3390/en13010243

transients in low frequency sinusoidal signal. The bandwidth of each levels of decomposition is 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 statistical measure of how much variation or dispersion that exists in the original signal and is defined in terms of wavelet coefficient as,

$$\mathbf{SD} = \sqrt{\left\{\frac{\sum_{i=1}^{8}(\mathbf{A_8}+\mathbf{D_i})^2}{\mathbf{n}} - \left(\frac{\sum_{i=1}^{8}(\mathbf{D_8}+\mathbf{D_i})}{\mathbf{n}}\right)^2\right\}} \quad (1)$$

where n represents the number of data samples.

3.1.2 Energy Value (E)

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

$$\mathbf{E} = \sum_{i=1}^{k}[|\mathbf{D_i}|^2] + |\mathbf{A_8}|^2 \quad (2)$$

where k is the number of detailed coefficient levels. 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 [26].

4. Fault Classifiers

This section presents Bayesian based fault classifiers to identify and classify the type of fault that occurs in the shunt compensated STATCOM devices. 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 classes and the same is used for training neural network. The assumed classes for classifications are: C_1-Normal, C_2-LG fault, C_3-LL fault, C_4-LLG fault and C_5-LLLG fault. Moreover, the effectiveness of the method is also tested for occurrence of fault at different location of transmission lines.

4.1 Multi-Layer Perceptron (MLP) Network

Multi-Layer Perceptron (MLP) is the most widely used neural network for identification and detection type of fault in 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 input (u1, u2 and u3), hidden and output layer.

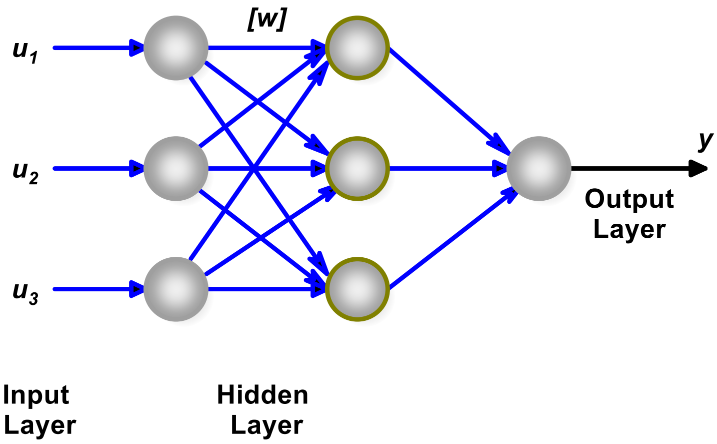


Figure 4. MLP neural network

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

$$y_i = W_{i_o} + \sum_{j \in pred(i)} (W_{ij} a_j) \tag{3}$$

where a_j represents the output of previous layer neuron, W_{ij} is the weight between i[th] and j[th] neuron and Wio is input bias of neuron. In this work, the MLP network is trained using back propagation method and the detailed explanation is presented in [27, 28].

4.2 Bayes and Naive Bayes Classifiers

The conventional MLP neural network minimizes the error of the system by adjusting the weight of the network through small penalty factor that leads to overfitting. This is avoided for any complex network through a principle approach called Bayes theorem by the Bayesian neural network (BNN). BNN is invented by Judea Pearl in 1980s, a statistical based supervised classifier that determines the variable to be classified in more relevant to the class by evaluating the probability of how likely its occurrence in that class with the prior information that takes the form prior probability density function [29]. Thus the Bayes theorem can be defined as

$$\boldsymbol{Posterior\ probability} = \frac{\boldsymbol{Class\ prior\ probability * likelihood}}{\boldsymbol{Predictor\ prior\ probability}}$$

(4)

The simplified form can be expressed as,

$$\mathrm{P}(\mathrm{C}|\mathbf{L_1}, \boldsymbol{L_2}, \ldots, \mathbf{L_n}) = \frac{\mathbf{P(C)P(L_1, \boldsymbol{L_2}, \ldots, L_n | C)}.}{\mathbf{P(L_1, \boldsymbol{L_2}, \ldots, L_n)}} \tag{5}$$

$$\boldsymbol{P(C|L)} = \frac{\boldsymbol{P(C)P(L|C)}}{\boldsymbol{P(L)}} \tag{6}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 October 2019 doi:10.20944/preprints201910.0148.v1

Peer-reviewed version available at Energies 2020, 13, 243; doi:10.3390/en13010243

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

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

Here the attributed P(L) doesn't vary with the class and can be assumed as constant and the above equation is approximated as,

$$\mathbf{arg}\left[\mathbf{max}[P(C|L) = P(C)P(L|C)]\right]$$
(8)

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

$$\mathrm{P}(\mathrm{C}|\mathrm{L}_1, L_2, \ldots, \mathrm{L}_n) = \frac{\mathrm{P(C)P}(\mathrm{L}_1|\mathrm{C})\mathrm{P}(\mathrm{L}_2|\mathrm{C})\ldots\mathrm{P}(\mathrm{L}_n|\mathrm{C})}{\mathrm{P(L)}} \quad (9)$$

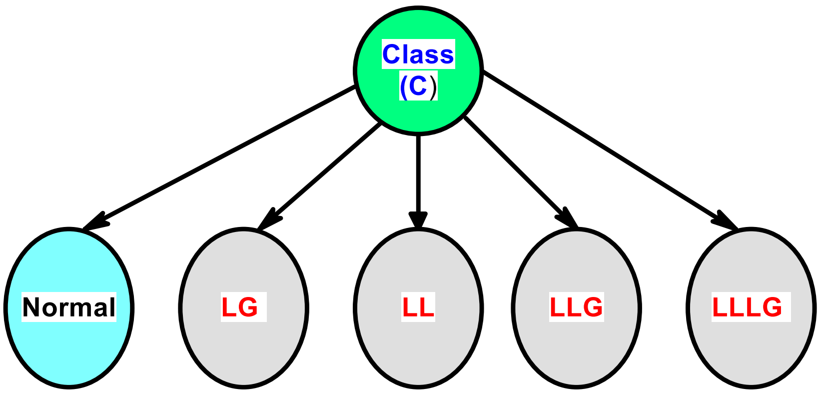


Figure 5. NB classifier of proposed work

4.2.1 Performance Indices of classifier

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

$$K = \frac{P(OF) - P(EF)}{(1 - P(EF))} \quad (10)$$

where P(OF) is the probability of observed fault, P(EF) is the probability of 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 error calculated between the predicted and observed value and is depicted as follows [21, 38, 39],

$$MAE = \frac{\left|\sum_{i=1}^{n}(E_P - E_O)\right|}{n} \tag{11}$$

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

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(E_P - E_O)^2}{n}} \tag{12}$$

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 proposed probabilistic NB based classifier to classify the fault and location of fault in transmission line. The effect of probabilistic classifier is studied for the transmission line with and without compensations. The simulation is carried out for power system model depicted in Figure 1 and various plausible faults such as LG, LL, LLG and LLLG fault in the system considering the variation in fault resistances. The simulation is carried out for time period of one cycle and the fault is applied during 0.1 to 0.12 s. Figure 6 and 7 depicts the three phase current waveform of the system without and with STATCOM respectively. The minimum and maximum of peak magnitude of three phase current signal are captured for the system with and without compensation that are illustrated in Table 2 and 3. It is seen the magnitude of current signal increases for the system with STATCOM device and the same is presented in the form waveform for case of LG fault in the system with and without STATCOM are portrayed in Figures 10 and 11 respectively. Then, the current signal obtained for various cases of fault are analyzed using 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 8 and 9 represent the DWT analysis of current waveform under normal operation of the system without and with STATCOM respectively. In general, the coefficients are high for the compensated system compare to the uncompensated system. Figures 12 and 13 portray the DWT analysis of LG fault current waveform considering without and with STATCOM respectively. Also, it is observed that the coefficients of detailed coefficient is low when fault occurs after the location of STATCOM (at 150 km) device. This effect is due to the STATCOM, the system fault current reduces as the distance of fault increase from, the fault location point. Table 4 and 5 represents 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 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.

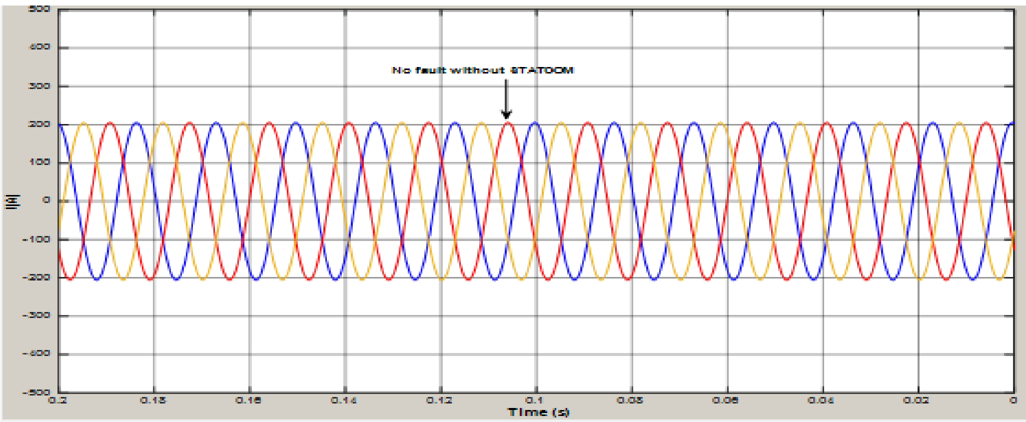


Figure 6. Three phase current waveform under normal condition without STATCOM compensation

Table 2. Normal and LG faults at Different Locations without STATCOM compensation

Fault Distance	Type of fault	Without STATCOM					
		Minimum current			Maximum current		
		Ia (10^3)	I b(10^3)	I c(10^3)	Ia(10^3)	I b(10^3)	I c(10^3)
100 km	No fault	-0.25	-0.25	-0.25	0.25	0.25	0.25
	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	-0.71	1.34	4.3	0.65
	LLLG	-3.88	-12	-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	-1560	-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 3. Normal and SLG faults at Different Locations with STATCOM compensation

Fault Distance	Type of fault	With STATCOM					
		Minimum current			Maximum current		
		Ia(10^3)	I b(10^3)	I c(10^3)	I a(10^3)	I b (10^3)	I c (10^3)
100 km	No f	1.11	-1.24	-1.32	1.4	**1.2**	**1.11**
	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	1.2.6	4.82	1.18
	LLLG	-4.57	-11.5	-1.1.9	1.4.3	7.02	4.91
200 km	LG	-2.2	-1.12	-1.23	3.97	1.23	1.08
	LL	-2.8	-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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 October 2019 doi:10.20944/preprints201910.0148.v1

Peer-reviewed version available at Energies 2020, 13, 243; doi:10.3390/en13010243

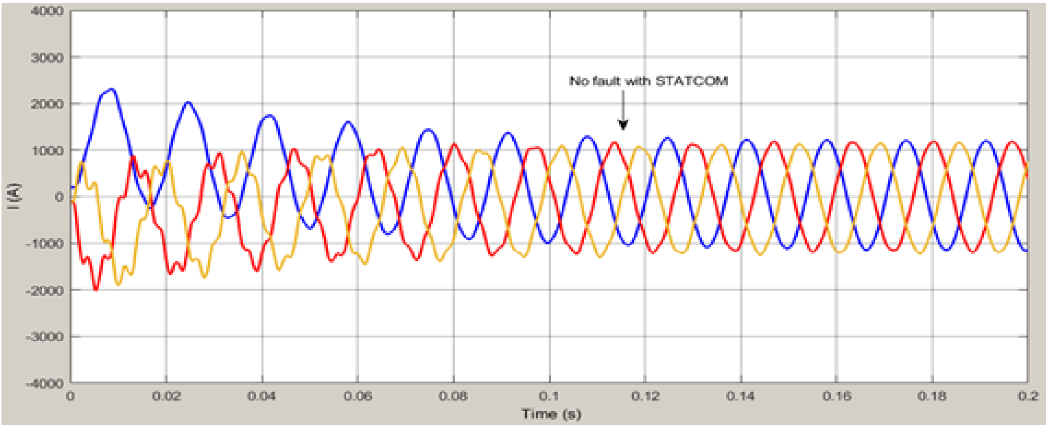


Figure 7. Three phase current waveform under normal condition with midpoint compensation

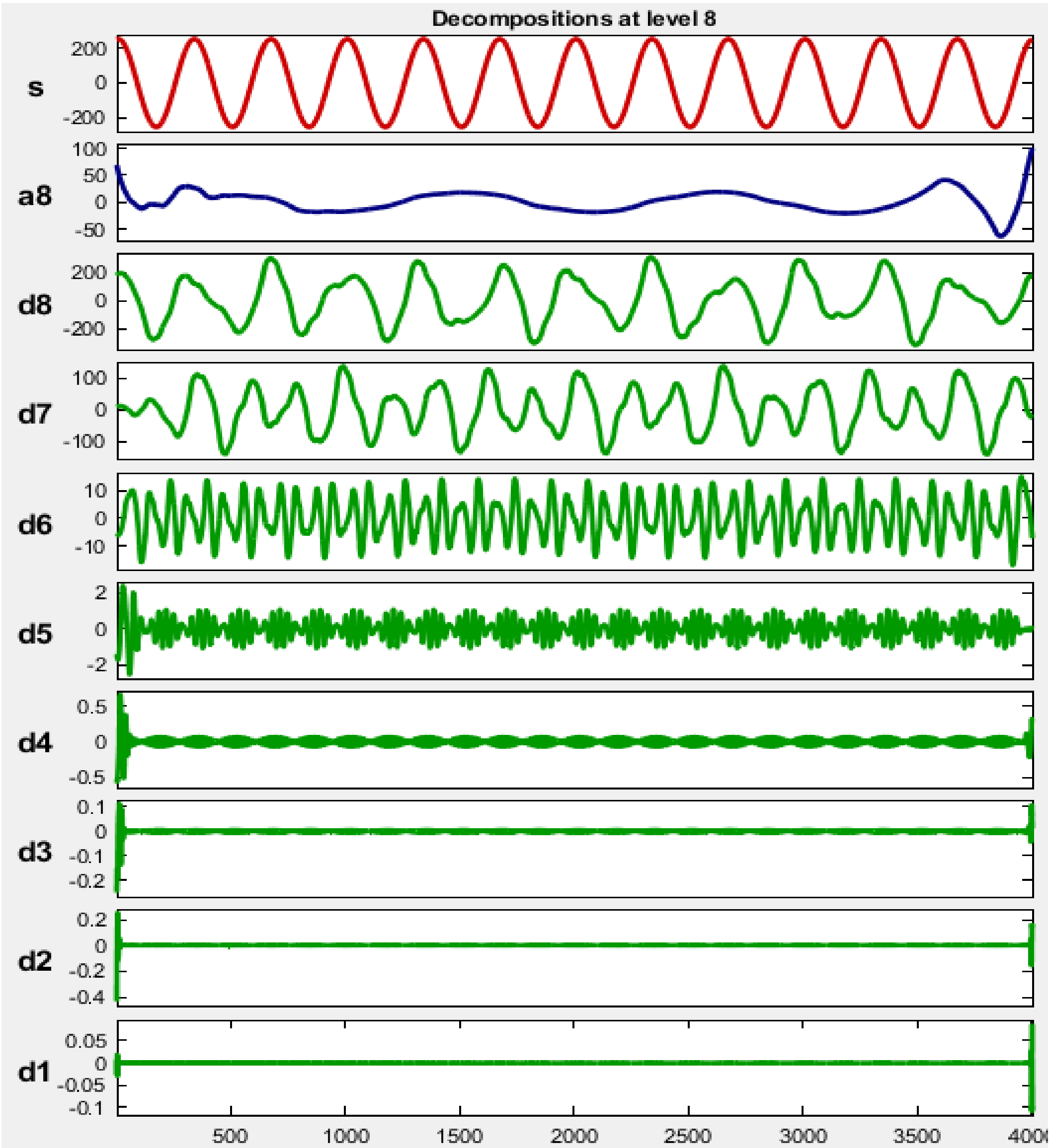


Figure 8. DWT analysis of Phase A under normal condition without compensation

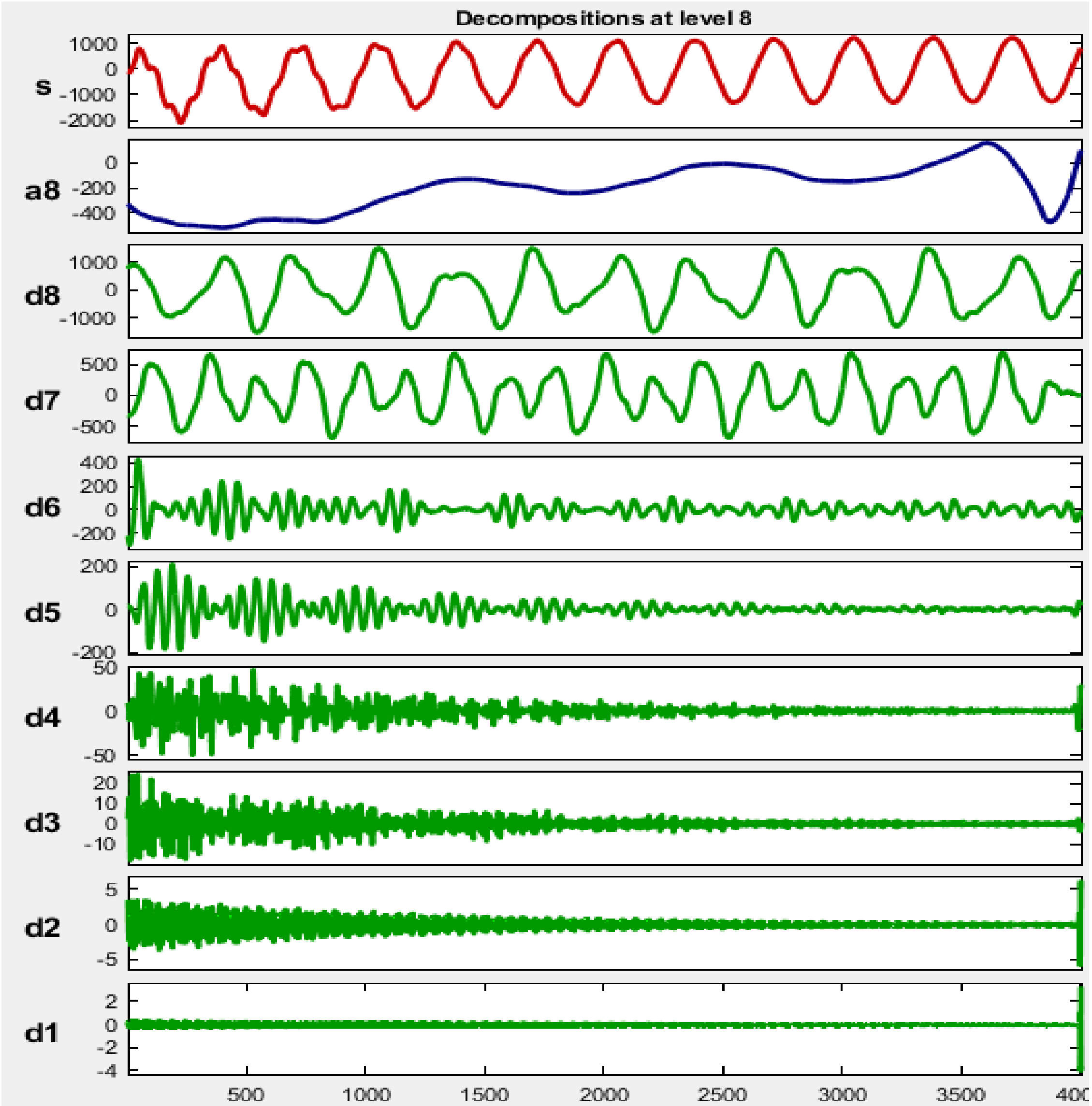
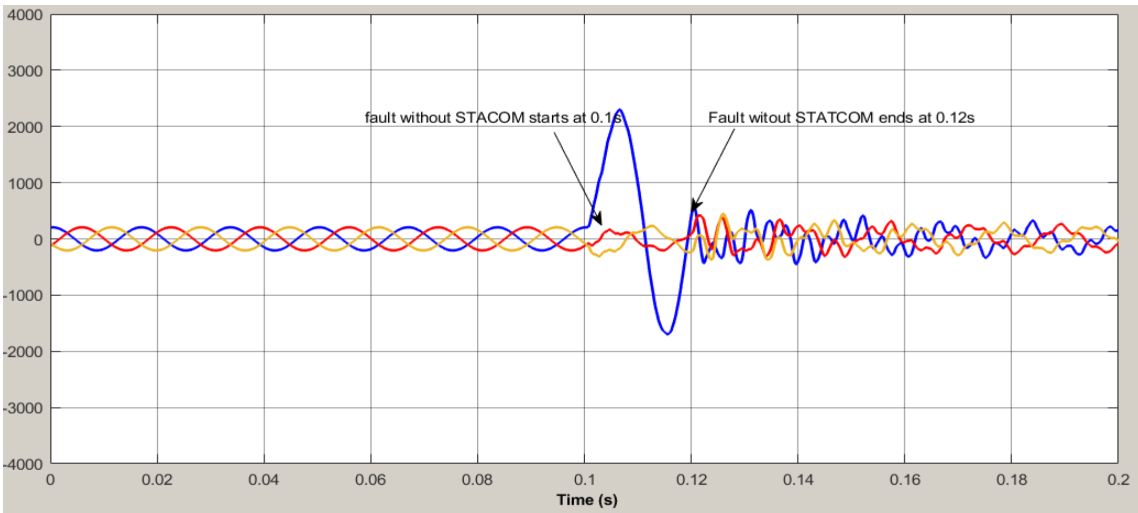


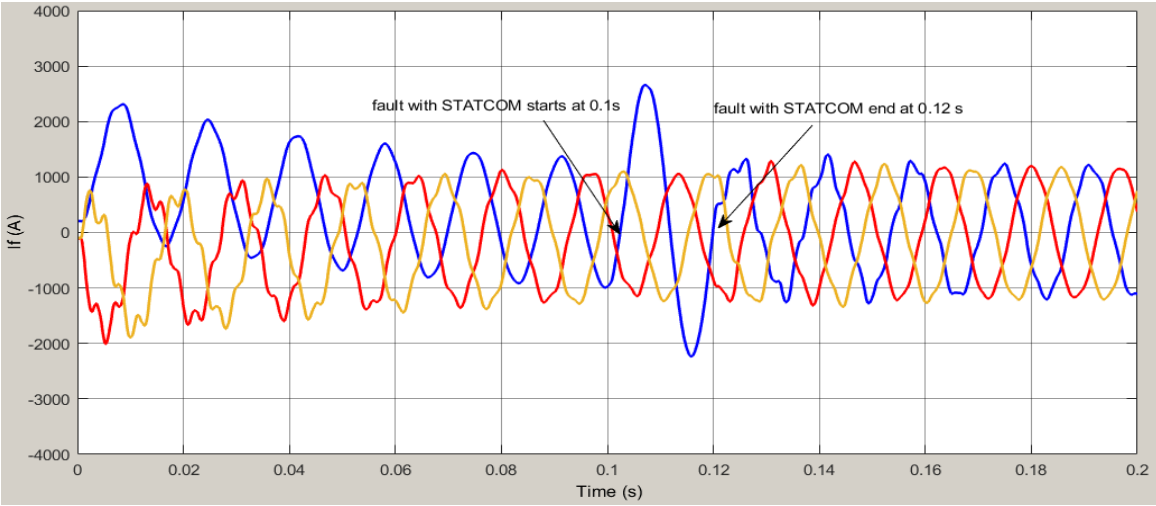
Figure 9. DWT analysis of Phase A under normal condition with STATCOM compensation



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 October 2019 doi:10.20944/preprints201910.0148.v1

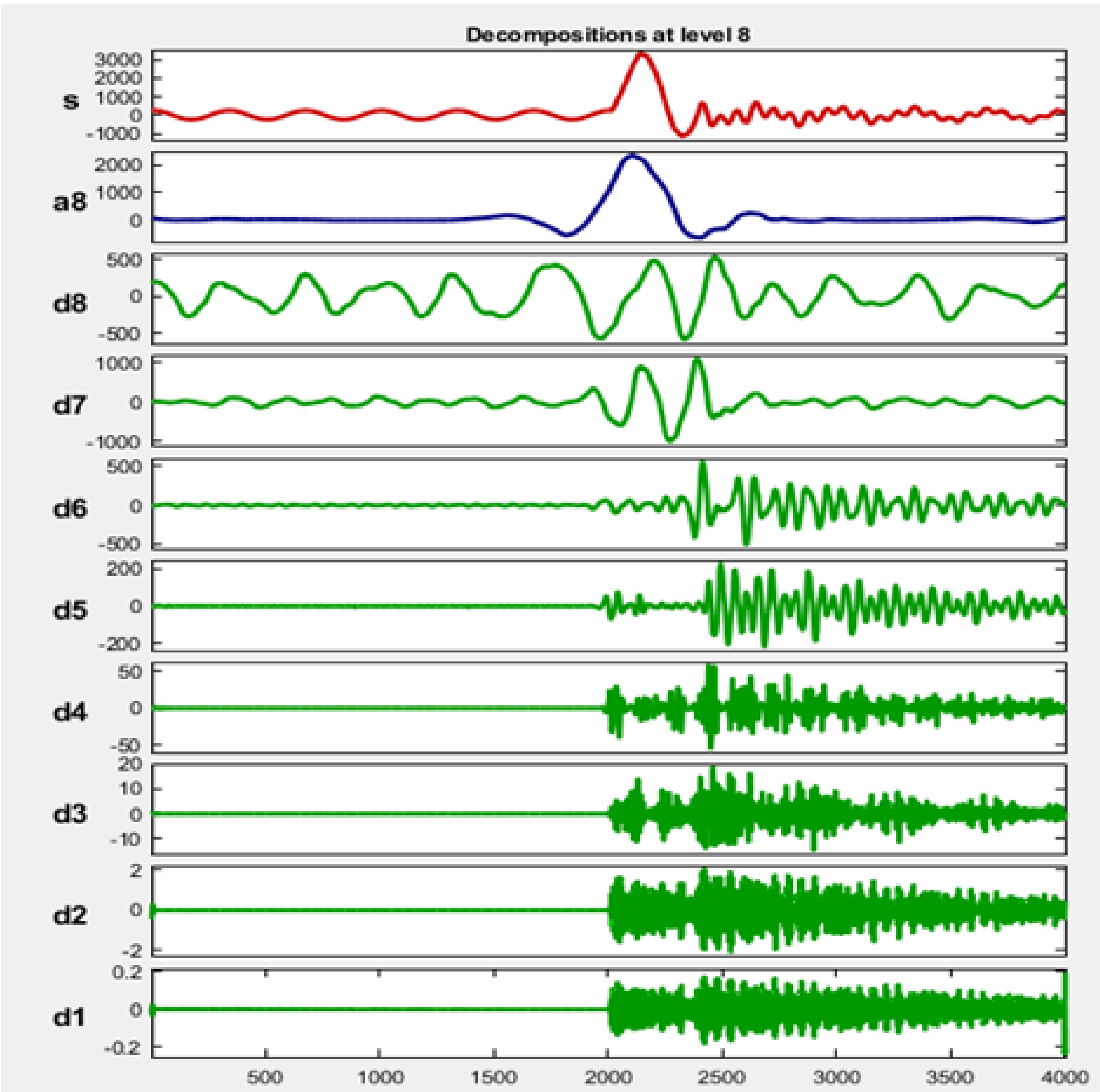
Peer-reviewed version available at *Energies* **2020**, *13*, 243; doi:10.3390/en13010243

282 **Figure 10.** Three phase current during LG fault in Phase A without compensation



283

284 **Figure 11.** Three phase current during LG fault in Phase A with compensation



285

286 **Figure 12.** DWT analysis of Phase A during LG fault without compensation

Peer-reviewed version available at Energies 2020, 13, 243; doi:10.3390/en13010243

Table 4. SD based feature values for classification

			Without STATCOM			With STATCOM		
Condition	Type of fault	Location km	SD- A (×10³)	SD-B (×10³)	SD- C (×10³)	SD- A (×10³)	SD-B (×10³)	SD- C (×10³)
Normal	No fault	100	0.177	0.177.1	0.177	0.875	0.877	00.866.1
		200	0.177	0.177.1	0.177	0.875	0.877	0866.1
		300	0.177	0.177.1	.0177	0.875.1	0.877	0.866.1
LG	AG	100	3.087	0.166	0.204	3.394	0.800	0.797.4
		200	1.582	0.154	0.196	2.046	0.825	0.817.5
		300	1.058	0.145	0.190	1.674.7	0.851	0.835.9
	BG	100	0.300	3.170	0.267	0.793.1	3.49	0.835
		200	0.245	1.630	0.198	0.821	2.11	0.836
		300	0.238	1.100	0.196	0.838	1.72	0.859
	CG	100	0.263	29.9	2.66	0.854	0.811	0.3057.4
		200	0.193	24.30	1.37	0.852	0.831	1.888.6
		300	0.193	23.80	92.1	0.874	0.849	1.569
LLG	ABG	100	5.865	5.65	28.2	5.810	5.690	76.600
		200	3.158	3.03	20.6	3.188	3.140	80.300
		300	2.140	2.05	20.5	2.357	2.330	83.200
	BCG	100	0.188	5.65	4.99	0.755	5.67	5.120
		200	0.170	3.06	2.71	0.799	3.14	2.870
		300	0.161	2.09	1.84	0.834	2.35	2.160
	CAG	100	5.108	28.70	5.15	5.247	0.759	5.210
		200	2.749	20.30	2.79	2.932	0.794	2.900
		300	1.842	20.20	1.87	2.202	0.833	2.170
LL	AB	100	5.723	5.67	17.7	5.633	5.67	83.800
		200	3097	3.04	17.7	3.085	3.11	84.200
		300	2105.6	2.05	17.70	2.279	2.30	84.900
	BC	100	177.3	5255.5	5.691	0.851	5.281	5245
		200	177.3	2868.9	2.832	0.856	2.944	2905
		300	177.3	1964	1.929	0.860	2.204	2164.5
	CA	100	4998	1.77	5.06	5.112	0.846	5.04
		200	2693.5	1.77	2.75	2.850	0.852	2.80
		300	1800.5	1.77	1.86	2.131	0.858	2.09
LLLG	ABCG	100	6254	6.48	5.69	6.224	6.430	5.75
		200	3368.6	3.51	3.10	3.397	3.370	3.19
		300	2263.6	2.39	2.10	2.493	2.580	2.36

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 October 2019 doi:10.20944/preprints201910.0148.v1

Peer-reviewed version available at *Energies* **2020**, *13*, 243; doi:10.3390/en13010243

Table 5. Energy based feature values for classification

Condition	Type of fault	Location Km	Without STATCOM E- A (×10⁸)	Without STATCOM E-B (×10⁸)	Without STATCOM E- C (×10⁸)	With STATCOM E-A (×10⁸)	With STATCOM E- B (×10⁸)	With STATCOM E-C (×10⁸)
		100	1.25	0.49	0.4	22.7	6.26	13.1
Normal	No fault	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
		100	96.4	0.56	0.51	128	5.36	11.4
	AG	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
LG	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
	ABG	100	3.01.0	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
LLG	BCG	100	1.36	184	179	20.8	185	200
		200	1.2	54.6	53.4	21.3	58.1	670
	CAG	300	1.18	25	22.7	22.1	32.80	44.4
		100	318	0.73	313	326	5.17	305
		200	94.6	0.51	93	106	5.09	94.9
	AB	300	41.6	0.52	41.2	66.9	5.53	56.5
		100	255	254	4.05	265	234	12.8
		200	74.7	73.9	0.4	92.6	68.3	12.9
LL	BC	300	35.6	35	0.4	56.8	35.8	12.9
		100	1.24	174	169	22.4	170	18.6
		200	1.24	53	50.2	22.4	49.5	62.4
	CA	300	1.24	23.5	22.3	22.4	30.2	40.8
		100	308	0.5	312	314	5.8	300
		200	91.5	0.49	93.4	103	5.87	91.7
LLLG	ABCG	300	40.50	0.49	41.3	65.5	5.98	54.3
		100	42.5	241	315	414	236	315
		200	125	70.5	94.7	1.30E+10	71.9	97
		300	57.5	33.3	40.7	76.6	38.6	5.91

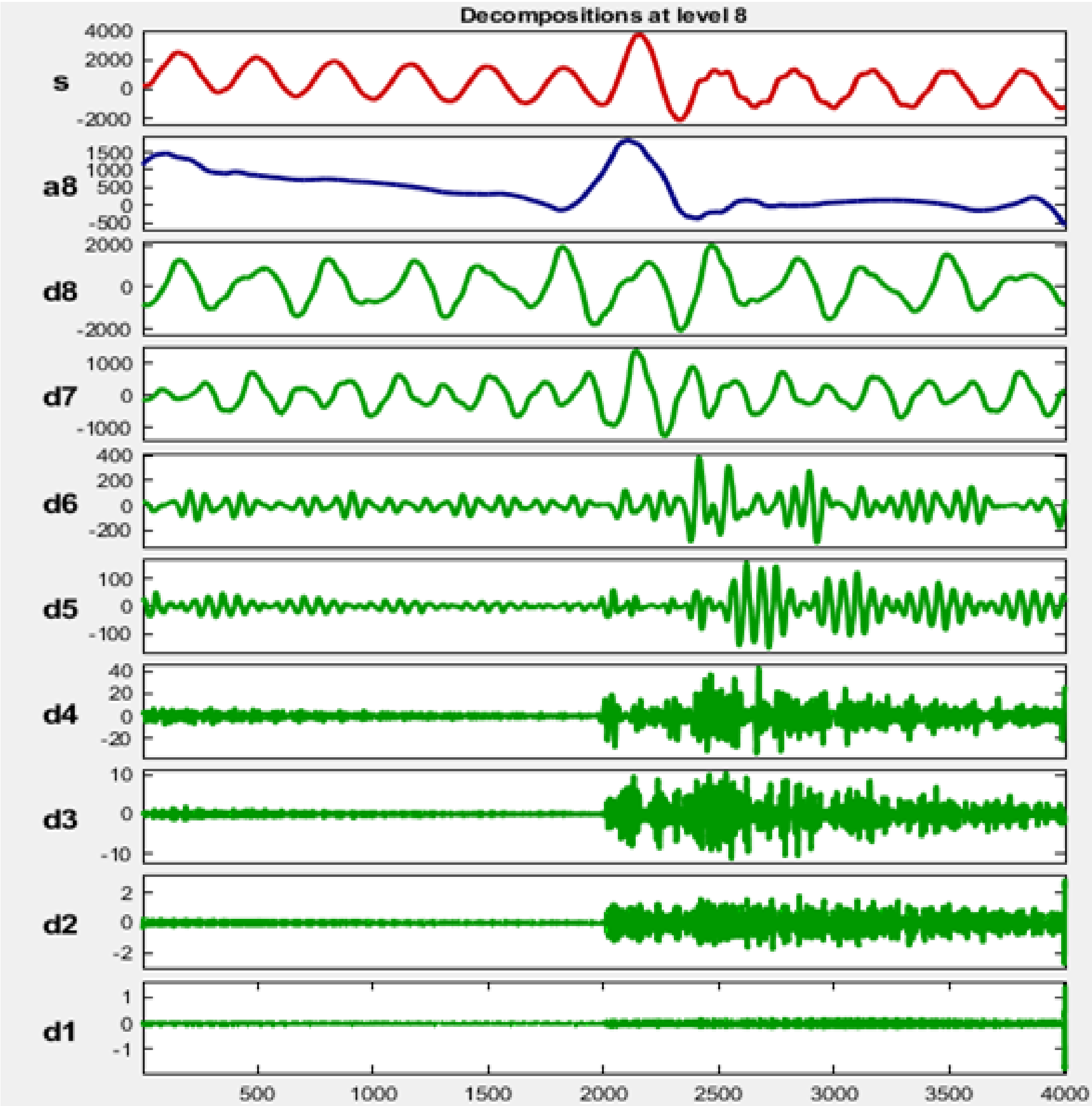


Figure 13. DWT analysis of Phase A during LG fault with compensation

Table 6. 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-1: In this study, the transmission fault classification and identification in a transmission network is done without STATCOM. Table 6 presents the confusion matrix for classification of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 October 2019 doi:10.20944/preprints201910.0148.v1

Peer-reviewed version available at *Energies* **2020**, *13*, 243; doi:10.3390/en13010243

different state 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 distance of 100 km, 200km and 300 km of an overhead transmission line is given in Table 4. Then these data's are used for training the neural network and the classification results obtained are presented in the Table 7. The result shows that the proposed Naive Bayes (NB) method of classifier is more accurate compared to the MLP and Bayes method of classification. Moreover, the % misclassification rate of the proposed method is 0%, whereas the rate is 20% and 80 % for MLP and Bayes approach 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 type of fault and whose performance is inferior compared to other methods. It is inferred from the Fig.. and Table.. that the NB classifier is the mostsignificant method, to classify the various type of fault in the system compared to all other methods.

Case-2: Here in this study, the classification and identification offault is done without STATCOM as like 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 the training the network and which is illustrated in the Table5. The results obtained reveals that 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 is 60 % and 20 % for MLP and Bayes network respectively. The MLP method of classification fails to detect LG and LLG faults whilst Bayes classifier unable to detect all type of faults. It is seen that the propounded NB has 0% misclassification rate, the MLP has 40% and Bayes method has 80% of misclassification rate as depicted in Table 7.

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 location such as 100 km, 200 km and 300 km are studied. The SD values obtained are used to train the network as like the case-1 and the results for classification are shown in Table 4. It is observed from the results that the proposed NB classifier performance is more predominant in terms of accuracy and % misclassification rate compared to the MLP and Bayes method of classification and is shown in Figure 14. The Bayes method fails to identify all type of fault expect when the system is operating in normal condition and MLP method fails to detect the LLG type of fault as like case-1. It is inferred from the results, both the MLP and Bayes classifier performance is same for transmission line involving with and without STATCOM and the proffered NB method classifier outperforms compared to these approaches.

Table7. ClassifiersAccuracy andMisclassification Rate

	Accuracy Rate			Misclassification Rate					
				MLP		Bayes		Naive Bayes	
Cases	MLP	Bayes	Naïve Bayes	% Rate	Type of faul	Rate	Type of Fault	Rate	Type of Fault
Case-1	80	20	100	2	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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 October 2019 doi:10.20944/preprints201910.0148.v1

Peer-reviewed version available at *Energies* **2020**, *13*, 243; doi:10.3390/en13010243

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 which is portrayed in Table 5. Figure 14 represents the proposed NB classifier is very efficient compared to the MLP and Bayes method. The % accuracy of NB and MLP are 100% MLP, butthe Bayes method is only 20 % accurate. On the flipside, the % misclassification rate is 0% for NB and MLP method and it is 80% for 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 7.

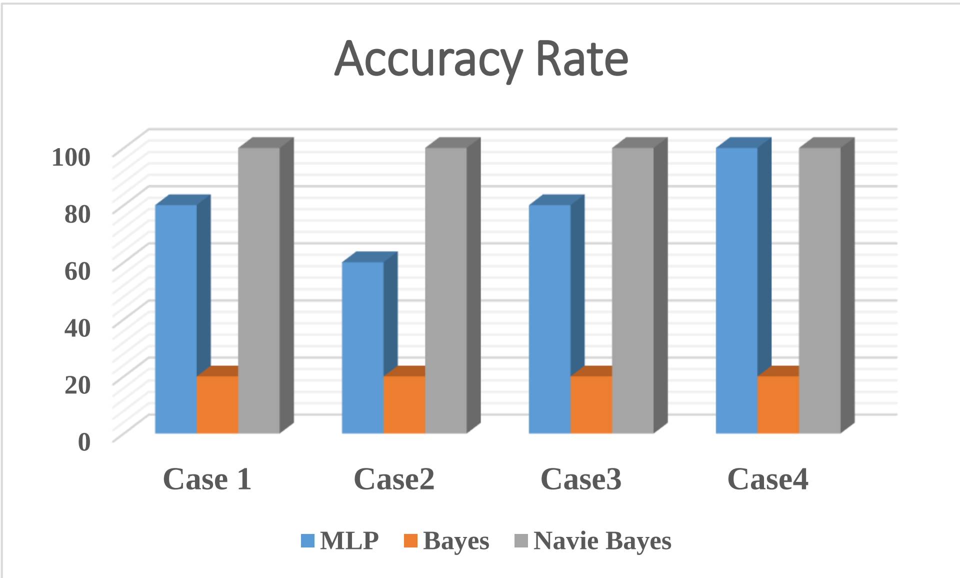


Figure 14. Comparision of Accuracy rate of classifiers

5.1 Performance Evaluation of Classifiers

The robustness of the classifier are evaluated by various performance indices such as Kappa Statistics (KS), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Percentage Relative Absolute Error (% MAE) and Percentage Root Relative Square Error (%RRSE) for classifiers namely Bayes, MLP and NB approach. Firstly, the KS index for various classifier is presented in Table 8 and Figure 15. The result shows that the indices is '1' for the proposed NB classifier for all the cases and the values lies in the range of 0.5-1 for MLP classifier (for various cases) and is almost '0' for 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 MLP method and it is greater than 0.3 for 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 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 Figsures 16 and 17 for the intended NB method of classifier than other approaches presented, proves 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 MLP and Bayes classifier as depicted in Table 9 and Figure 18. It is observed the results outperforms for all the cases by the NB approach rather than the MLP and Bayes classifier method.

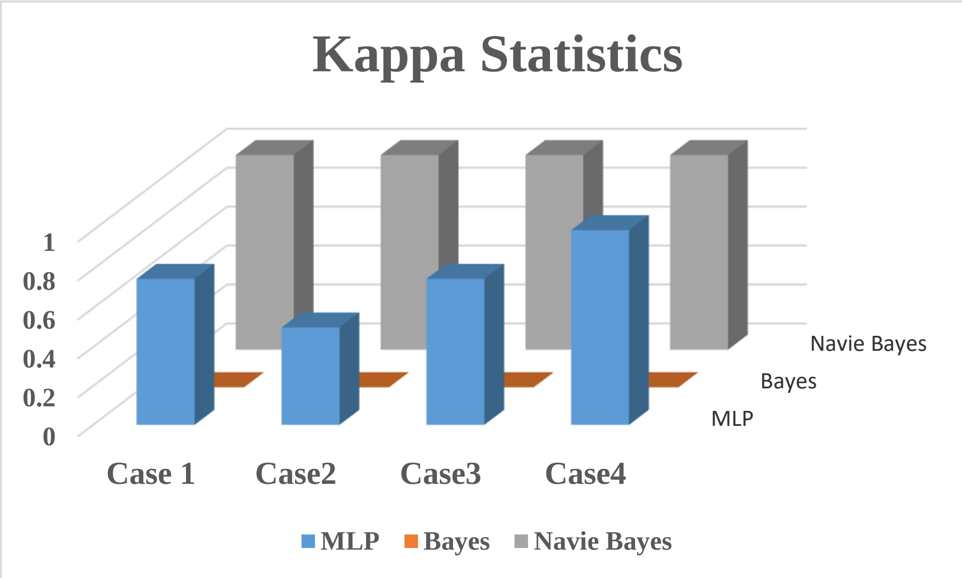
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 October 2019 doi:10.20944/preprints201910.0148.v1

Peer-reviewed version available at Energies 2020, 13, 243; doi:10.3390/en13010243

367 **Tabl 8.** Performance comparison of various Classifiers

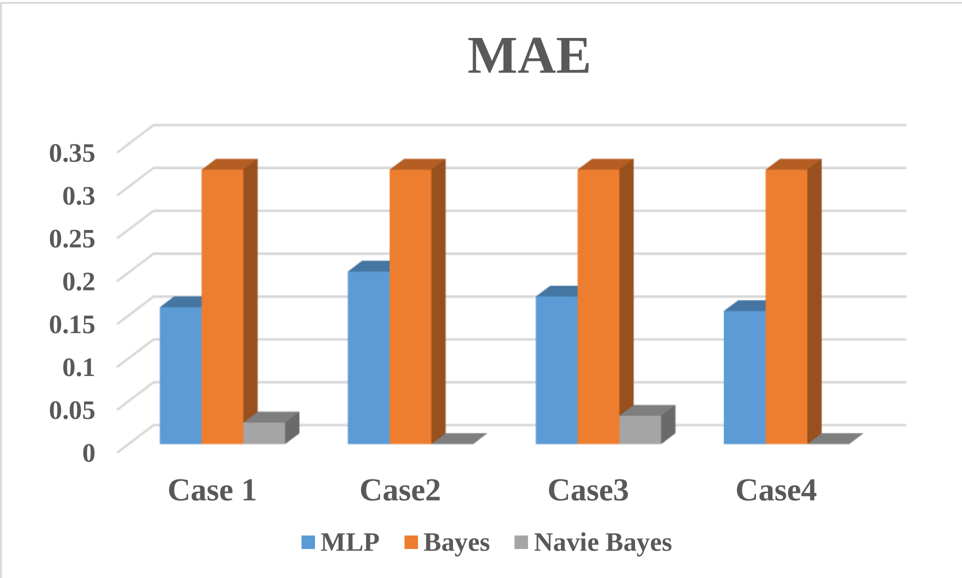
	Kappa Statistics			MAE			RMSE		
	MLP	Bayes	Naive Bayes	MLP	Bayes	Naive Bayes	MLP	Bayes	Naive Bayes
Case-1	0.75	0	1	0.1596	0.32	0.0251	0.2369	0.4	0.0888
Case-2	0.5	0	1	0.2012	0.32	0	0.2929	0.4	0
Case-3	0.75	0	1	0.172	0.32	0.033	0.248	0.4	0.0979
Case-4	1	0	1	0.1551	0.32	0	0.2276	0.4	0

368



369

370 **Figure 15.** Kappa Statics comparison of various classifiers



371

372 **Figure 16.** MAE comparison of various classifiers

 doi:10.20944/preprints201910.0148.v1

Peer-reviewed version available at Energies 2020, 13, 243; doi:10.3390/en13010243

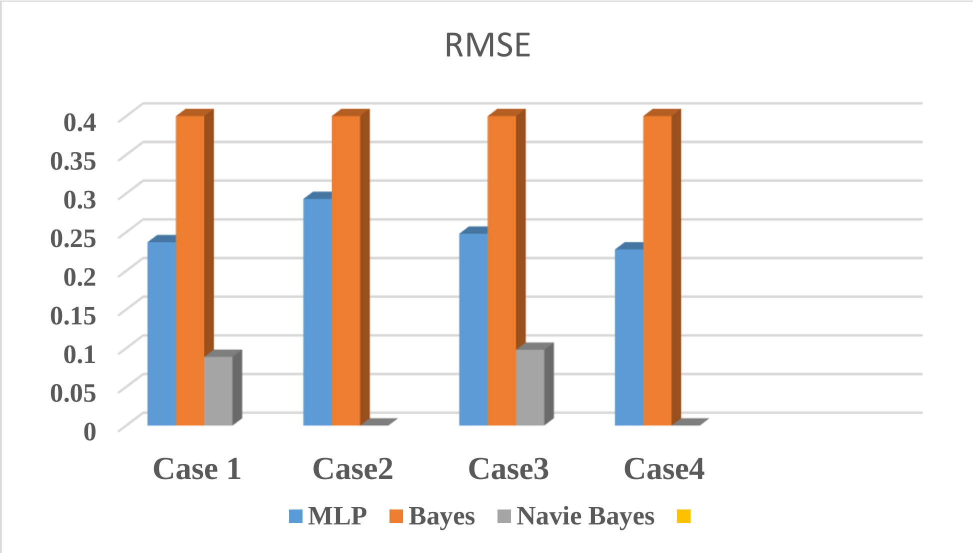


Figure 17. RMSE comparison of various classifiers

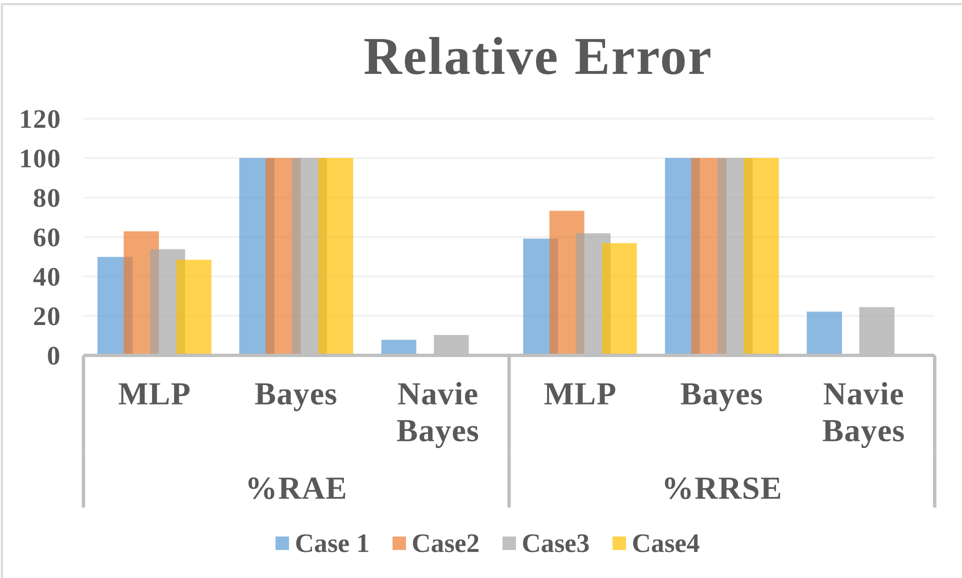


Figure 18. %RAE comparison of various classifiers

Table 9. %RAE and %RRSE comparison of various classifiers

	%RAE			%RRSE		
Cases	MLP	Bayes	**Naive Bayes**	MLP	Bayes	**Naive Bayes**
Case-1	49.89	100	7.85	59.23	100	22.21
Case-2	62.8627	100	0	73.2233	100	0
Case-3	53.7439	100	10.2998	62.0026	100	24.46
Case-4	48.4605	100	0	56.8981	100	0

6. Conclusion

This paper presents a novel probabilistic based Navie Bayes approach to locate the fault in 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 depicts 1st, 2nd, 3rd, 4th, 5th, 6th , 7th and 8th level of detailed coefficient and its approximation coefficient at 8th level. The SD and Energy values have been obtained for different faults with various fault resistance. The obtained features are used to train the classifiers to classify the type of fault. The results obtained showed that the proposed NB classifier outperforms with 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, the proffered NB approach gives the predominant result compared to the MLP and Bayes classifier method.

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 assisted to write paper; proof reading and final drafting was done by I.A and H.H.

Funding: This research received no external funding

Acknowledgment: 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. *Meas. J. Int. Meas. Confed.* **2018**, *116*, 96–105.
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.
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.
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.
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.
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**, *0123456789*.
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*.
8. Gupta, O.H.; Tripathy, M. An innovative pilot relaying scheme for shunt-compensated line. *IEEE Trans. Power Deliv.* **2015**, *30*, 1439–1448.
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.

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.
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.
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.
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.
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.
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.
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*, 1–21.
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.
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.
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.
21. Wilkinson, W.A.; Cox, M.D. Discrete wavelet analysis of power system transients. *IEEE Trans. Power Syst.* **1996**, *11*, 2038–2044.
22. Osman, A.H.; Malik, O.P. Protection of Parallel Transmission Lines Using Wavelet Transform. *IEEE Trans. Power Deliv.* **2004**, *19*, 49–55.
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*, 1–11.
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.
25. Zhang, N.; Wu, L.; Yang, J.; Guan, Y. Naive bayes bearing fault diagnosis based on enhanced independence of data. *Sensors (Switzerland)* **2018**, *18*, 1–17.
26. Hindarto; Muntasa, A. Wavelet sub-band energy for feature extraction of electro encephalo graph (EEG) signals. *J. Eng. Sci. Technol.* **2019**, *14*, 578–588.
27. Patil, K.; Jadhav, N. Multi-Layer Perceptron Classifier and Paillier Encryption Scheme for Friend Recommendation System. *2017 Int. Conf. Comput. Commun. Control Autom. ICCUBEA 2017* **2018**, 1–5.
28. Deep, K.; Nagar, A.; Pant, M.; Bansal, J.C. *Proceedings of the Second International Conference on Soft Computing for Problem Solving (SocProS 2012), December 28-30, 2012*; 2014; Vol. 236; ISBN 978-81-322-1601-8.
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.
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.

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.