

## Detection of power transmission lines faults based on voltages and currents values using K-nearest neighbors

Nisreen Khalil Abed, Faisal Theyab Abed, Hamdalla F. Al-Yasriy, Haider TH. Salim ALRikabi

Department of Electrical Engineering, College of Engineering, Wasit University, Al Kut, Iraq

### Article Info

#### Article history:

Received Dec 26, 2022

Revised Feb 12, 2023

Accepted Feb 29, 2023

#### Keywords:

Faults detection

Input features

K-nearest neighbor algorithm

Output features

Transmission line

### ABSTRACT

The critical factors to consider when implementing a maintenance plan for energy transmission lines are, accuracy, speed, and time, because of the increased global demand for electricity power caused by rapid development, and overuse of electric power transmission lines (both underground cables and overhead transmission lines), which in turn reduces the efficiency of the lines. Consequently, the efficiency of the lines may be reduced as a result of overuse or other activities like excavation that may have tampered with the cables. Thus, it becomes important to investigate the faults to which the lines are exposed. To this end, this article focuses on the detection of fault in transmission lines through the use of k-nearest neighbor algorithm. Using this algorithm, the characteristics were obtained (voltage, current), and these characteristics enable the identification of faults in the transmission lines, and in the specific location (the entire system, phase B, and phase A). The benefits that can be derived from the use of this algorithm include time, accuracy, speed, which are the requirements for the maintenance of transmission lines. Euclidean distance used in the application of the k-nearest neighbor technique for weights, and  $K = 3$  for number of neighbors. The dataset was split into two parts, 70% training set and 30% testing set.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



### Corresponding Author:

Haider TH. Salim ALRikabi

Department of Electrical Engineering, College of Engineering, Wasit University

Al Kut, Iraq

Email: hdhiyab@uowasit.edu.iq

## 1. INTRODUCTION

Presently, the cost of developing electric power transmission lines is higher than ever before, because of the higher demand for electricity. This high cost covers for power generation, transmission, and distribution as contained in Figure 1 [1]. The performance of these transmission lines is affected by heavy and continuous use, as well as other external factors [2]. Undetected faults can be a major obstacle in the functioning of any power system, as they can stop the operation of the entire electrical system [3]–[5]. There are different kinds of faults that can be found in transmission lines, and these different faults can be categorized as either asymmetric or symmetrical. An example of such faults that can arise in transmission lines include phase fault such as phase-to-ground fault, phase-to-phase fault, phase-to-earth fault, and three-phase fault. Nevertheless, their presence does not affect the functionality of the power system. More so, other faults like overlapping faults, circuit fault, and other faults are faults that are also regarded as unimportant faults compared to the aforementioned faults. Traditionally, these lines are maintained through the use of megger device which facilitates the detection of faults. Alternatively, the faults are also detected through physical inspection of lines [6]–[12]. These two methods are time-consuming, and as such, countries around the world are exploring new ways through which the faults can be detected and addressed within a short period of time. In this work the ability of the k-nearest neighbor (KNN) algorithm to detect the faults is explored, and the specific kind of fault

is identified through the use of MATLAB software, whether phase-ground fault or phase-to-phase fault. The elements that should be considered when a fault is detected in any type of power system include voltage, current, resistance, power factor, and frequency. Several techniques of fault detection show the presence of a fault by comparing the post-fault values with the pre-fault values of the systems.

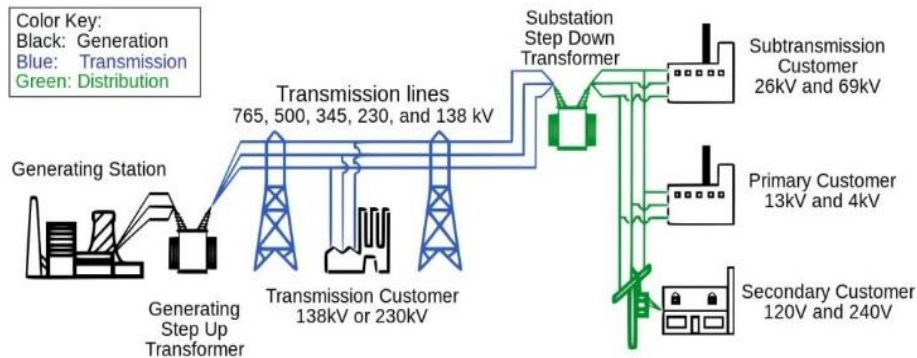


Figure 1. Power system structure [1]

## 2. A LITERATURE SURVEYS

A wide range of approaches and techniques have been used in the detection of faults in electric power transmission lines. In the work done by Muir and Lopatto a new method was introduced; the method helps in detecting fault in digital relays-based power system through the use of Petri nets [1]. The authors made use of Petri nets for modelling and detection of location, and with the proposed technique, the power system is monitored in a hierarchical manner. Their experimental results revealed that the use of Petri nets reduced the time required to process information, and the precision of fault detection increased. As early as in 1994, the use of microprocessors was employed by Barros and Drake to detect faults in real time [13] based on the estimation of the three phase voltage phasors by mean of a set of Kalman filters, and on the calculation of the fault probability. Subsequently, in 2004, wavelet transform was proposed by [3] for the detection of fault in a transformer by measuring neural currents. The analysis of the wavelet transform was carried out based on the Morlet wavelet (mother wavelet). It was concluded that significant improvement was achieved in terms of the fault detection sensitivity by the use of wavelet analysis approaches for the assessment of impulse tests on transformer. Similar efforts geared towards fault detection were made by Bracho and Martinez [14], who used dynamic power supply current test in 1997. Subsequently, in 1998, Chowdhury and Aravena [15] introduced a new technique through which faults can be detected by the use of a modular methodology. The method which is relatively flexible also allows classification in power system. Upon detection of the fault, the fault indicator is processed by a Kohonen network for the classification of faults. Abed and AIRikabi [8] who presented a conference paper in 2021 focused on the detection of faults in underground cables as transmission lines, used IoT applications to monitor and detect underground cable faults. In the work done by Majd *et al.* [16], the protection and control of power systems were investigated. In their work, a technique for the detection of transmission line faults was presented. In their proposed approach, the use of KNN based fault detection and classification approaches was employed. Similar efforts made by Samet *et al.* [17] led to the production of a technique for the detection and classification fault for transmission lines through; the authors used an improved alienation coefficients method. In the research carried out by Gafoor and Rao, a wavelet-based fault detection technique was proposed. The proposed technique is able to detect, classify faults as well as the location of the fault in the transmission lines [18], [19].

## 3. THE PROPOSED METHOD

The dataset used in this work was used in modeling a power system in MATLAB was simulated for fault analysis. As seen in Figure 2, the power system is made up of four power generators of  $11 \times 10^3$  V, with each pair sitting at the end of the transmission line. Transformers are present in between for the simulation and investigation of different faults ta the midpoint of the transmission line.

The authors of the database carried out the simulation of the circuit under both normal conditions and abnormal conditions (with faults). Afterwards, they collected and saved the measured line voltages and line currents at the power system's output side. Here, about 12000 data points were collected, and then the data was

labeled. The dataset can be accessed through Kaggle [20]–[22]. The dataset is made up of input features including voltages ( $V_a$ ,  $V_b$ ,  $V_c$ ) and currents ( $I_a$ ,  $I_b$ ,  $I_c$ ) of the three phases, the description of statistics of the input features as seen in Table 1, as well as their histogram distribution as presented in Figure 3 (see Appendix).

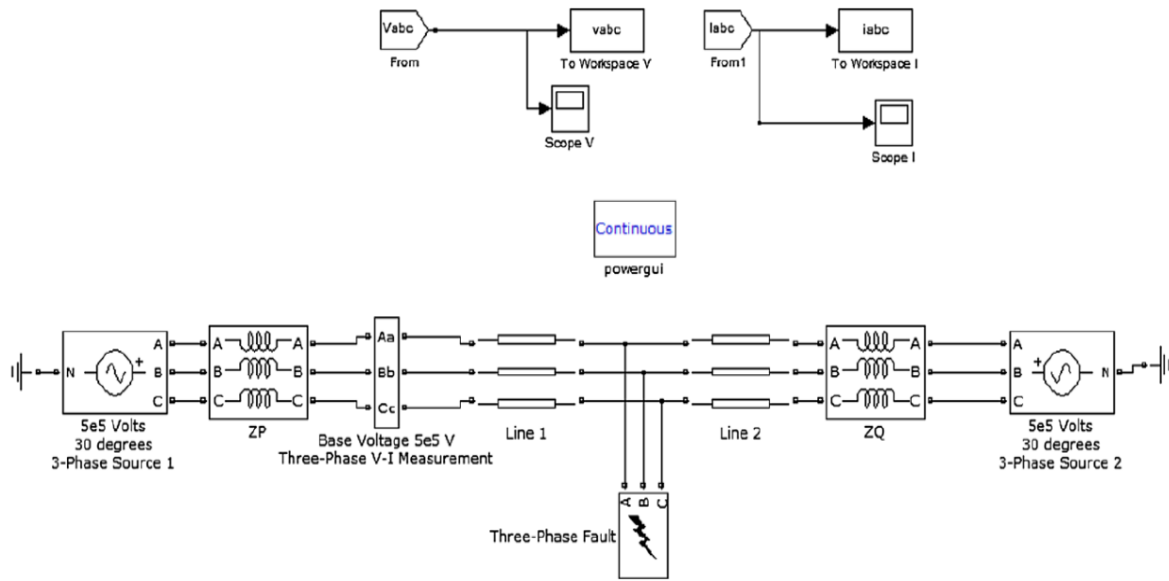


Figure 2. Power system diagram [20], [21]

Table 1. Statistics description of input features

Statistics	Ia	Ib	Ic	Va	Vb	Vc
Mean	25.476	-26.633	-0.607	0.002	-0.002	-0.001
Median	8.195	-0.118	-8.573	-0.002	-0.001	0.005
Mode	-9.677	-93.940	85.800	0.000	-0.115	0.136
Minimum	-883.542	-900.527	-900.527	-0.621	-0.608	-0.613
Maximum	885.739	889.869	901.274	0.606	0.628	0.600
Count	14395	14395	14395	14395	14395	14395

Also contained in the database are the values for the outputs (G, A, B, C), which possess just two values; the value of 0 denotes no fault, while the value of 1 denotes the presence of faults. In this work, additional output parameter (S) has been added to the entire system. Figure 4 (see Appendix) shows the distribution of the output features. A summary of the dataset is presented using the correlation matrix in Table 2, which presents the correlation between all features, whereby, the value-100% represents a perfectly negative linear relationship between all feature, while the value 0% means there is no linear relationship between two features, and the value 100% denotes a perfectly positive linear relationship between two features.

Table 2. Correlation matrix of features

	Ia	Ib	Ic	Va	Vb	Vc	FG	FC	FB	FA	FS
Ia	100%										
Ib	-31.31%	100%									
Ic	-30.72%	-33.58%	100%								
Va	10.93%	-1.39%	12.32%	100%							
Vb	-14.41%	2.84%	10.37%	-53.31%	100%						
Vc	1.93%	-1.22%	3.48%	58.01%	-37.99%	100%					
FG	-1.16%	2.56%	-1.87%	2.11%	-1.60%	-0.78%	100%				
FC	-0.06%	-3.07%	2.53%	1.26%	-3.87%	2.34%	9.20%	100%			
FB	-0.06%	-8.19%	7.71%	1.26%	-4.23%	2.69%	9.20%	8.88%	100%		
FA	6.05%	-7.54%	0.88%	0.99%	-3.23%	2.03%	9.20%	8.88%	8.88%	100%	
FS	3.38%	-6.19%	2.97%	2.99%	-4.86%	1.42%	49.05%	47.11%	47.11%	47.11%	100%

## 4. METHOD

### 4.1. K-nearest neighbor

A KNN algorithm can be described as simple and efficient supervised machine learning method employed in regression and classification operations [23]–[25]. Given that, the algorithm carries out classification directly and based on the training examples, it is categorized as case-based classification or example-based classification it classified as example-based classification, or case-based classification [26]. This algorithm performs the classification operation based on similarity criteria, giving consideration to the distance measure. Here, "K" denotes the integer value that ranges from 3 to +10. Compared to even values, the odd values are mostly preferred when seeking to get a good prediction. A given class is selected based on majority votes given by neighboring points that correspond to the nearest class. The neighbors are assigned weights so that the nearer neighbor adds more weight to the average than that of the farther one. Weights are assigned to the assigned to the neighbors based on their Euclidean distance [27]. A flowchart for KNN algorithm modeling is shown in Figure 5.

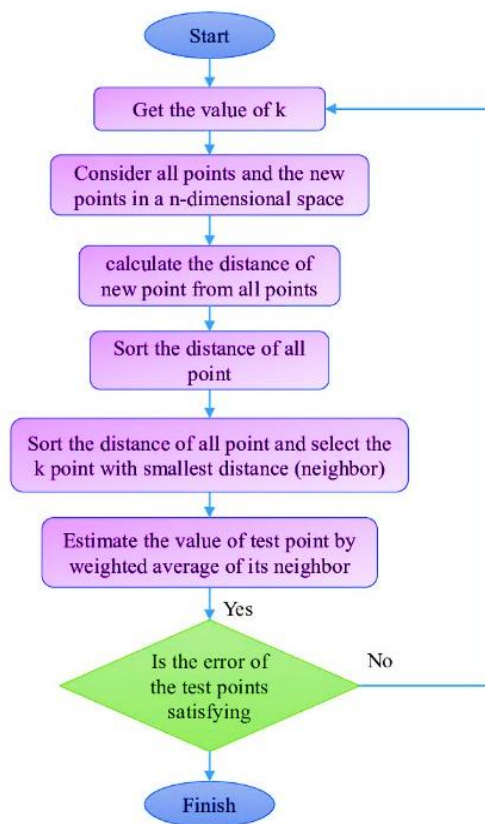


Figure 5. A simple flowchart for the k-nearest neighbor modeling [28]

The detection of faults in transmission lines is carried in five stages. In the first stage, the faults in phase A will be detected, followed by the second stage which involves the detection of faults in phase B. In the third stage, the fault in phase C is detected, and followed by the detection of fault in the ground, and lastly, the overall faults in the entire system are detected. The operations are performed according to the values of currents and voltages. The application of these values is done in the following manner: phase A only features, phase B only features, phase C only features, phase A and phase B features, phase A and phase C features, phase B and phase C features, voltages only features, currents only features, and all features. The application of the KNN technique involved the use of Euclidean distance for weights, while  $K = 3$  for number of neighbors. The dataset was divided into two for training and testing, with 70% of the dataset designated for training the algorithm, and 30% for testing it. Table 3 shows the kind of features that were used in this paper.

Table 3. Description of features

Types of features	Features	Number of features
Phase A only	Ia, Va	2
Phase B only	Ib, Vb	3
Phase C only	Ic, Vc	3
Phase A and phase B	Ia, Va, Ib, Vb	6
Phase A and phase C	Ia, Va, Ic, Vc	6
Phase B and phase C	Ib, Vb, Ic, Vc	6
Voltages only	Va, Vb, Vc	3
Currents only	Ia, Ib, Ic	3
All features	Va, Vb, Vc, Ia, Ib, Ic	9

## 5. RESULTS AND DISCUSSION

Majority of the parameters used in measuring the performance of the algorithm are based on the confusion matrix, are classified as ‘True’ prediction/reality matches (TP and TN) and ‘False’ non-matches or errors (FP and FN) [29].

- True positive, means that the actual and predicted outcomes both fall under “faults” class.
- False positive, means that the predicted is in faults class whereas the actual is classified as “no fault” class.
- False negative, means the predicted fault is classified as “no fault”, while the actual is classified under “fault class”.
- True negative, means that both the predicted and actual faults are classified as “no fault”. the values of true positives, false positives, true negatives and false negatives is shown in the Tables 4 and 5.

Table 4. Confusion matrix of fault detection phases A, B, and C

Features	Fault detection in phase A				Fault detection in phase B				Fault detection in phase C			
	TN	TP	FP	FN	TN	TP	FP	FN	TN	TP	FP	FN
Phase A only	2,197	2,053	36	33	1,198	958	1,086	1,077	1,188	945	1,121	1,065
Phase B only	1,539	1,342	694	744	2,183	1,983	101	52	1,681	1,408	628	602
Phase C only	1,349	1,182	884	904	1,332	1,194	952	841	2,165	1,955	144	55
Phase A and phase B	2,225	2,067	8	19	2,221	2,005	63	30	2,140	1,807	169	203
Phase A and phase C	2,223	2,066	10	20	2,150	1,831	134	204	2,240	1,983	69	27
Phase B and phase C	2,131	1,886	102	200	2,240	2,018	44	17	2,216	1,982	93	28
Voltages only	2,158	2,029	75	57	2,180	1,985	104	50	2,222	1,924	87	86
Currents only	2,233	2,078	0	8	2,242	2,027	42	8	2,260	2,007	49	3
All features	2,233	2,078	0	8	2,242	2,027	42	8	2,260	2,007	49	3

Table 5. It shows the faults detection of phases in the ground and system

Features	Fault detection in ground				Fault detection in system			
	TN	TP	FP	FN	TN	TP	FP	FN
Phase A only	2,089	1,416	291	523	3,440	475	152	252
Phase B only	1,928	1,444	452	495	3,442	536	150	191
Phase C only	2,131	1,415	249	524	3,442	558	150	169
Phase A and phase B	2,114	1,699	266	240	3,572	717	20	10
Phase A and phase C	2,078	1,687	302	252	3,576	716	16	11
Phase B and phase C	2,130	1,699	250	240	3,562	710	30	17
Voltages only	2,145	1,708	235	231	3,563	710	29	17
Currents only	2,235	1,789	145	150	3,582	727	10	0
All features	2,235	1,789	145	150	3,582	727	10	0

The proposed models were evaluated based on parameters in the confusion matrix including accuracy, sensitivity, specificity, and precision. Accuracy refers to the ratio of total number of correct faults and no faults predictions to sample size [29]. Sensitivity (recall) is the measure of faults points that correctly detected [30]. Specificity is the measure of no-fault points that are detected correctly [30]. Precision or confidence is the measure of predicted faults that are actual faults [31]. These metrics were calculated for the results of the methods that were used in this work. The results are presented in Tables 6 and 7.

For the detection of faults phase, A, only currents were used as input features, and optimal results were obtained, which will be the same even if all features are used as inputs. Very good results were obtained when the features of phase A were used (phase A only, phase A and phase B, phase A and phase C). Also, the result obtained from the use of only voltages is better than the results of the features used in phase B and phase C. For the detection of faults phase B, only currents were used as input features, and optimal results were obtained, which will be the same even if all features are used as inputs. Very good results were obtained when

the features of phase B were used (phase B only, phase A and phase C, phase B and phase C). Also, the result obtained from the use of only voltages is better than the results of the features used in phase B and phase C.

Table 6. Performance metrics of the accuracy, precision, and sensitivity of the phases

Features	Fault detection in phase A				Fault detection in phase B				Fault detection in phase C			
	Acc	Pr	Se	Sp	Acc	Pr	Se	Sp	Acc	Pr	Se	Sp
Phase A only	0.984	0.983	0.984	0.984	0.499	0.469	0.471	0.525	0.494	0.457	0.470	0.515
Phase B only	0.667	0.659	0.643	0.689	0.965	0.952	0.974	0.956	0.715	0.692	0.700	0.728
Phase C only	0.586	0.572	0.567	0.604	0.585	0.556	0.587	0.583	0.954	0.931	0.973	0.938
Phase A and phase B	0.994	0.996	0.991	0.996	0.978	0.970	0.985	0.972	0.914	0.914	0.899	0.927
Phase A and phase C	0.993	0.995	0.990	0.996	0.922	0.932	0.900	0.941	0.978	0.966	0.987	0.970
Phase B and phase C	0.930	0.949	0.904	0.954	0.986	0.979	0.992	0.981	0.972	0.955	0.986	0.960
Voltages only	0.969	0.964	0.973	0.966	0.964	0.950	0.975	0.954	0.960	0.957	0.957	0.962
Currents only	0.998	1.000	0.996	1.000	0.988	0.980	0.996	0.982	0.988	0.976	0.999	0.979
All features	0.998	1.000	0.996	1.000	0.988	0.980	0.996	0.982	0.988	0.976	0.999	0.979

Table 7. It shows that the Performance metrics of accuracy, precision, and sensitivity of the phases with ground and the system

Features	Fault detection in ground				Fault detection in system			
	Acc	Pr	Se	Sp	Acc	Pr	Se	Sp
Phase A only	0.812	0.830	0.730	0.878	0.906	0.758	0.653	0.958
Phase B only	0.781	0.762	0.745	0.810	0.921	0.781	0.737	0.958
Phase C only	0.821	0.850	0.730	0.895	0.926	0.788	0.768	0.958
Phase A and phase B	0.883	0.865	0.876	0.888	0.993	0.973	0.986	0.994
Phase A and phase C	0.872	0.848	0.870	0.873	0.994	0.978	0.985	0.996
Phase B and phase C	0.887	0.872	0.876	0.895	0.989	0.959	0.977	0.992
Voltages only	0.892	0.879	0.881	0.901	0.989	0.961	0.977	0.992
Currents only	0.932	0.925	0.923	0.939	0.998	0.986	1.000	0.997
All features	0.932	0.925	0.923	0.939	0.998	0.986	1.000	0.997

For the detection of faults phase C, only currents were used as input features, and optimal results were obtained, which will be the same even if all features are used as inputs. Very good results were obtained when the features of phase C were used (phase C only, phase A and phase C, Phase B and phase C). Also, the result obtained from the use of only voltages is better than the results of the features used in phase A and phase B. For ground fault detection, the best results were obtained by using only currents as inputs, which will be the same if all features are used as inputs. Higher results were obtained for voltages only features as compared to the results of features used in phase A and phase B and phase C. The use of the features in the two phases at the same time yielded optimal results in comparison to when a single phase is used. For the detection of faults in the entire system, optimal results were obtained using only current as input features, which will be the same if all the features were used as inputs. Optimal results were obtained by using the features of two phases at the same time as inputs. The results were better than using only single phase and voltages only. Generally, it was found that better results were achieved in the detection of ground faults and those in the entire system. More so, the least performance was recorded in the detection of ground faults. High values of sensitivity and specificity were achieved in the case ground faults detection. This reveals that the algorithm is able to accurately differentiate faults points from no fault points, indicating that the algorithm can be used reliably for faults detection based on the values of voltages and currents of the transmission lines.

## 6. CONCLUSION

In this study, the process of faults detection in transmission lines was performed in five phases. The algorithm proposed in this work successfully detected faults in phase A, phase B, phase C, ground, and whole system. The detection of faults in the transmission lines was done through the use of K-nearest neighbor model on a simulated power system that is made up of 11 KV generators. Also, the detection involved the use of values of voltages and currents of the transmission lines and in different combinations. The algorithm's performance was evaluated using different parameters from the confusion matrix, including accuracy, sensitivity, precision, and specificity. Analysis and discussion of the findings have been presented, showing the best feature combinations for the detection of faults in electrical transmission lines, as well as the worst combination.

APPENDIX

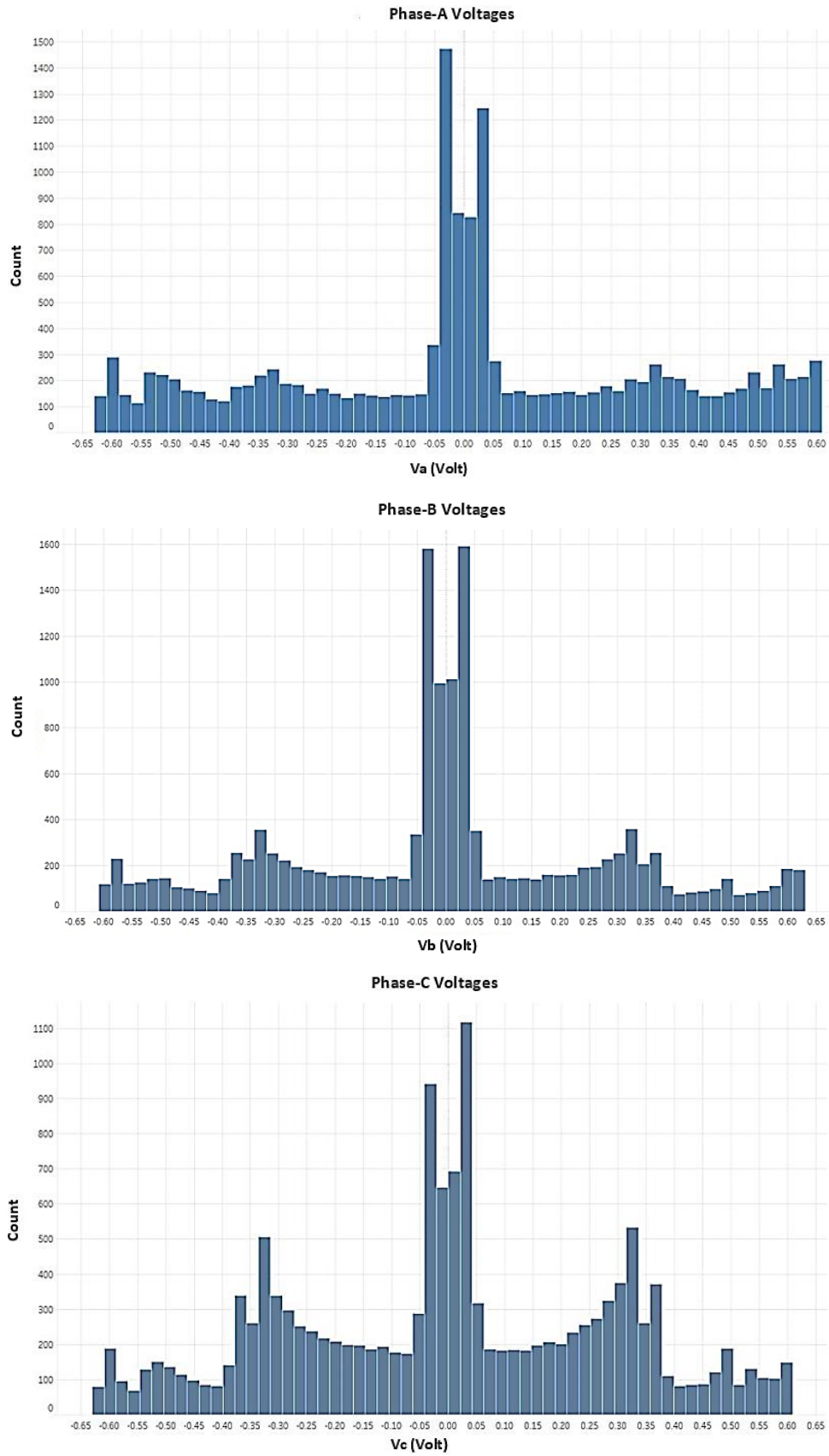


Figure 3. Histogram distribution of input features

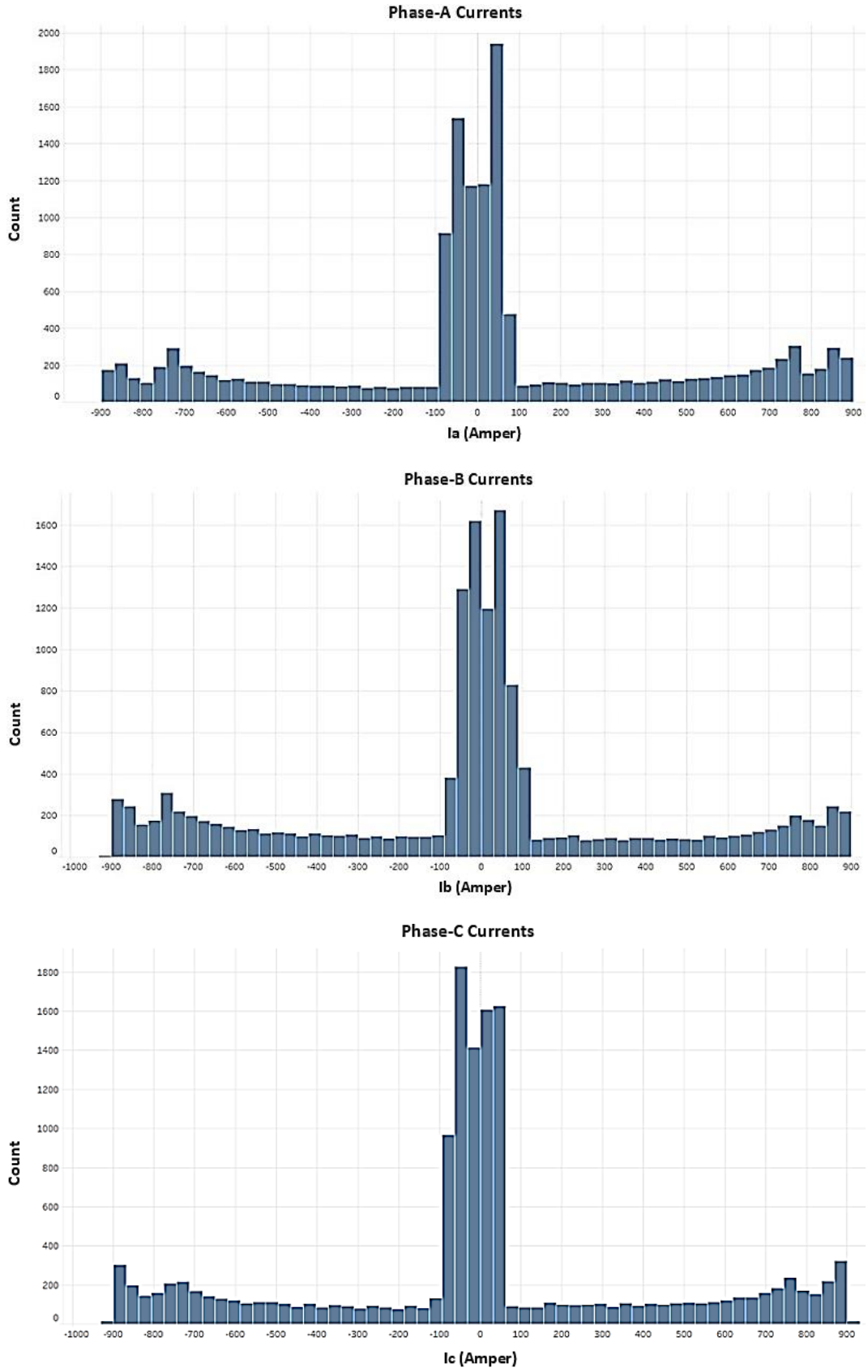


Figure 3. Histogram distribution of input features (continue)



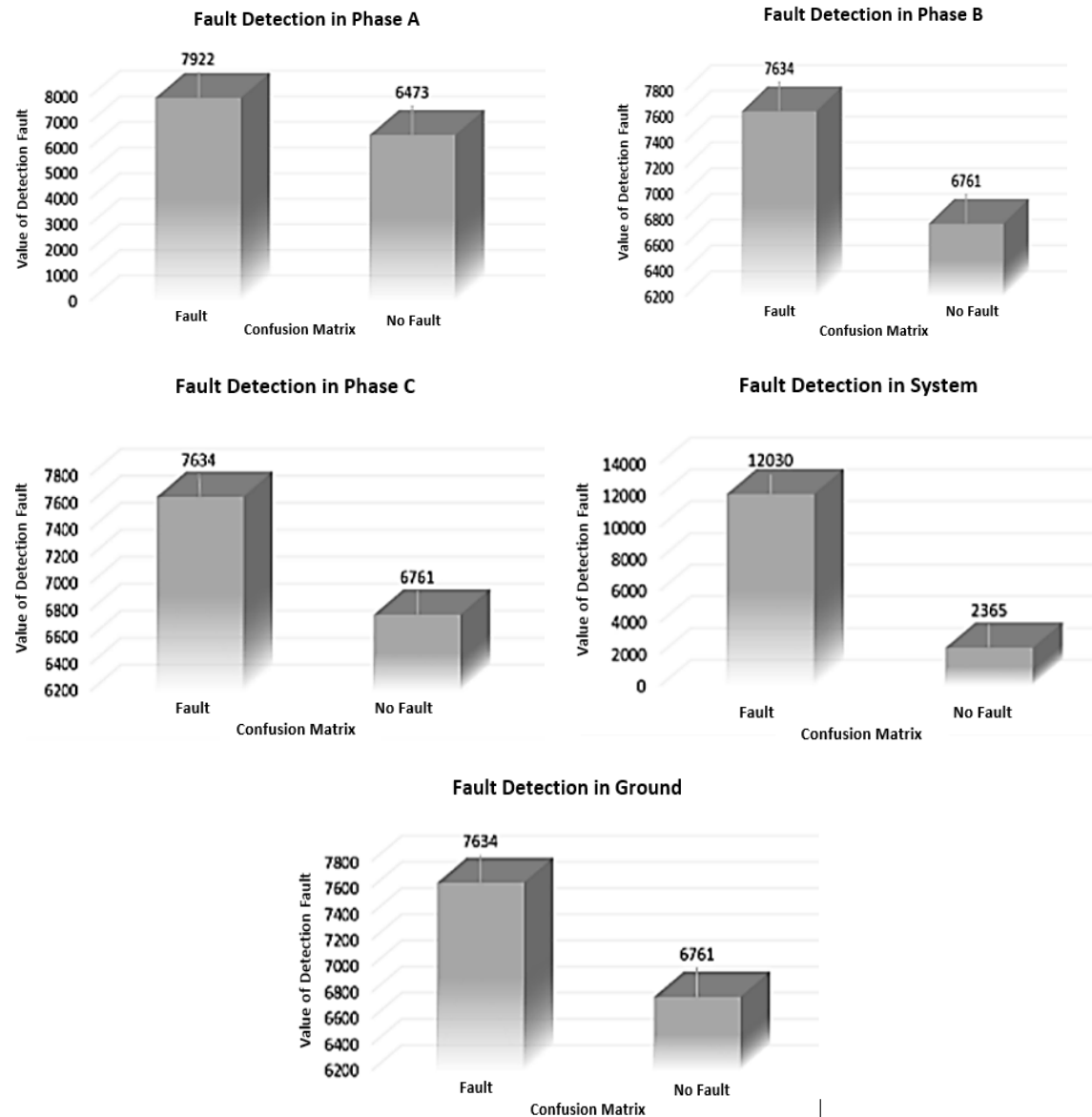


Figure 4. Output features distribution





## REFERENCES

- [1] A. Muir and J. Lopatto, "Final report on the August 14, 2003 blackout in the United States and Canada: causes and recommendations," 2004. [Online]. Available: <https://www.energy.gov/oe/articles/blackout-2003-final-report-august-14-2003-blackout-united-states-and-canada-causes-and-recommendations>.
- [2] N. M. Zainuddin *et al.*, "Review of thermal stress and condition monitoring technologies for overhead transmission lines: issues and challenges," *IEEE Access*, vol. 8, pp. 120053–120081, 2020, doi: 10.1109/ACCESS.2020.3004578.
- [3] N. Yu, S. Shah, R. Johnson, R. Sherick, M. Hong, and K. Loparo, "Big data analytics in power distribution systems," in *2015 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, Feb. 2015, pp. 1–5, doi: 10.1109/ISGT.2015.7131868.
- [4] H. Akhavan-Hejazi and H. Mohsenian-Rad, "Power systems big data analytics: an assessment of paradigm shift barriers and prospects," *Energy Reports*, vol. 4, pp. 91–100, Nov. 2018, doi: 10.1016/j.egy.2017.11.002.
- [5] L. Abd Ali, F. Al-Rufae, V. Kuvshinov, B. Krit, A. Al-Antaki, and N. Morozova, "Study of hybrid wind–solar systems for the Iraq energy complex," *Applied Solar Energy*, vol. 56, pp. 284–290, 2020, <https://doi.org/10.3103/S0003701X20040027>.
- [6] J. Antonino-Daviu, P. Jover Rodriguez, M. Riera-Guasp, M. Pineda-Sánchez, and A. Arkkio, "Detection of combined faults in induction machines with stator parallel branches through the DWT of the startup current," *Mechanical Systems and Signal Processing*, vol. 23, no. 7, pp. 2336–2351, Oct. 2009, doi: 10.1016/j.ymsp.2009.02.007.
- [7] N. A. Hussien, A. A. D. Al-Magsoosi, H. T. S. AlRikabi, and F. T. Abed, "Monitoring the consumption of electrical energy based on the internet of things applications," *International Journal of Interactive Mobile Technologies (IJIM)*, vol. 15, no. 07, p. 17, Apr. 2021, doi: 10.3991/ijim.v15i07.20183.




- [8] F. T. Abed and H. T. AlRikabi, "Using IoT applications for detection and monitoring of underground cable fault," *IOP Conference Series: Materials Science and Engineering*, vol. 1184, no. 1, p. 012003, Sep. 2021, doi: 10.1088/1757-899X/1184/1/012003.
- [9] G. Dileep, "A survey on smart grid technologies and applications," *Renewable energy*, vol. 146, pp. 2589-2625, 2020, doi.org/10.1016/j.renene.2019.08.092
- [10] Y. Du, S. Zhou, X. Jing, Y. Peng, H. Wu, and N. Kwok, "Damage detection techniques for wind turbine blades: a review," *Mechanical Systems and Signal Processing*, vol. 141, p. 106445, Jul. 2020, doi: 10.1016/j.ymssp.2019.106445.
- [11] S. Al-Hamadani, "Solar energy as a potential contributor to help bridge the gap between electricity supply and growing demand in Iraq: A review," *Int J Adv Appl Sci ISSN*, vol. 2252, no. 8814, p. 8814, 2020, doi: 10.11591/ijaas.v9.i4.pp302-312.
- [12] H. D. S. Altai, F. T. Abed, M. H. Lazim, and H. T. S. AlRikabi, "Analysis of the problems of electricity in Iraq and recommendations of methods of overcoming them," *Periodicals of Engineering and Natural Sciences (PEN)*, vol. 10, no. 1, pp. 607-614, Feb. 2022, doi: 10.21533/pen.v10i1.2722.
- [13] J. Barros and J. M. Drake, "Realtime fault detection and classification in power systems using microprocessors," *IEEE Proceedings: Generation, Transmission and Distribution*, vol. 141, no. 4, pp. 315-322, 1994, doi: 10.1049/ip-gtd:19949968.
- [14] S. Bracho and M. Martinez, "Catastrophic and parametric fault detection by dynamic power supply current test," in *IEEE Colloquium on Testing Mixed Signal Circuits and Systems*, 1997, vol. 1997, p. 10, doi: 10.1049/ic:19971201.
- [15] F. N. Chowdhury and J. L. Aravena, "A modular methodology for fast fault detection and classification in power systems," *IEEE Transactions on Control Systems Technology*, vol. 6, no. 5, pp. 623-634, 1998, doi: 10.1109/87.709497.
- [16] A. A. Majd, H. Samet, and T. Ghanbari, "K-NN based fault detection and classification methods for power transmission systems," *Protection and Control of Modern Power Systems*, vol. 2, no. 1, p. 32, Dec. 2017, doi: 10.1186/s41601-017-0063-z.
- [17] H. Samet, A. Shabanpour-Haghighi, and T. Ghanbari, "A fault classification technique for transmission lines using an improved alienation coefficients technique," *International Transactions on Electrical Energy Systems*, vol. 27, no. 1, p. e2235, Jan. 2017, doi: 10.1002/etep.2235.
- [18] S. A. Gafoor and P. V. R. Rao, "Wavelet based fault detection, classification and location in transmission lines," in *2006 IEEE International Power and Energy Conference*, Nov. 2006, pp. 114-118, doi: 10.1109/PECON.2006.346630.
- [19] H. T. S. AlRikabi and N. Ali Jasim, "Design and implementation of smart city applications based on the internet of things," *International Journal of Interactive Mobile Technologies (IJIM)*, vol. 15, no. 13, pp. 4-15, Jul. 2021, doi: 10.3991/ijim.v15i13.22331.
- [20] E. S. Prakash, "Electrical fault detection and classification dataset," *kaggle*. <https://www.kaggle.com/datasets/esathyaparakash/electrical-fault-detection-and-classification?q=Electrical+Fault+in%3Adatasets>.
- [21] M. Jamil, S. K. Sharma, and R. Singh, "Fault detection and classification in electrical power transmission system using artificial neural network," *SpringerPlus*, vol. 4, no. 1, p. 334, Dec. 2015, doi: 10.1186/s40064-015-1080-x.
- [22] C. S. Bojer and J. P. Meldgaard, "Kaggle forecasting competitions: an overlooked learning opportunity," *International Journal of Forecasting*, vol. 37, no. 2, pp. 587-603, Apr. 2021, doi: 10.1016/j.ijforecast.2020.07.007.
- [23] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, "KNN model-based approach in classification," in *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE*, Berlin, Germany: Springer, 2003, pp. 986-996.
- [24] O. Adepoju, J. Woswei, S. Lawte, and H. Jaiman, "Comparative evaluation of credit card Fraud detection using machine learning techniques," in *2019 Global Conference for Advancement in Technology (GCAT)*, Oct. 2019, pp. 1-6, doi: 10.1109/GCAT47503.2019.8978372.
- [25] B. Charbuty and A. Abdulazeez, "Classification based on decision tree algorithm for machine learning," *Journal of Applied Science and Technology Trends*, vol. 2, no. 01, pp. 20-28, Mar. 2021, doi: 10.38094/jastt20165.
- [26] P. Cunningham and S. J. Delany, "K-nearest neighbour classifiers - a tutorial," *ACM Computing Surveys*, vol. 54, no. 6, pp. 1-25, Jul. 2022, doi: 10.1145/3459665.
- [27] S. K. Shukla and E. Koley, "Detection and classification of open conductor faults in six-phase transmission system using k-nearest neighbour algorithm," in *2017 7th International Conference on Power Systems (ICPS)*, Dec. 2017, pp. 157-161, doi: 10.1109/ICPS.2017.8387285.
- [28] M. R. Mahdiani, E. Khamchi, S. Hajirezaie, and A. Hemmati-Sarapardeh, "Modeling viscosity of crude oil using k-nearest neighbor algorithm," *Advances in Geo-Energy Research*, vol. 4, no. 4, pp. 435-447, Dec. 2020, doi: 10.46690/ager.2020.04.08.
- [29] G. Canbek, S. Sagioglu, T. T. Temizel, and N. Baykal, "Binary classification performance measures/metrics: a comprehensive visualized roadmap to gain new insights," in *2017 International Conference on Computer Science and Engineering (UBMK)*, Oct. 2017, pp. 821-826, doi: 10.1109/UBMK.2017.8093539.
- [30] D. G. Altman and J. M. Bland, "Statistics notes: diagnostic tests 1: sensitivity and specificity," *British Medical Journal*, vol. 308, no. 6943, pp. 1552-1552, Jun. 1994, doi: 10.1136/bmj.308.6943.1552.
- [31] D. M. W. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," *arXiv:2010.16061*, 2020, doi: 10.48550/arXiv.2010.16061.

## BIOGRAPHIES OF AUTHORS






**Nisreen Khalil Abed**     is a lecturer in the Engineering College, at the Wasit University, Iraq. His area of research focuses on power systems and their applications. Year of Graduation 2006 University of Technology/Electromechanical Engineering Department – Electrical branch Bachelor degree. M.Sb. Year of Graduation 2020 University of Wasit/Electrical engineering department–general branch Master degree. She can be contacted at email: nabed@uowasit.edu.iq.






**Faisal Theyab Abed**    is a lecturer in the Engineering College, at the Wasit University, Iraq. His area of research focuses on power systems and their applications. He received his B.Sc. degree in Electrical Engineering in 2006 from the Al Mustansiriya University in Baghdad, Iraq. His M.Sc. degree in Electrical Engineering focusing on power systems from Russia in 2017. Al Kut city – Hay ALRabee, Wasit, Iraq. The number of articles in national databases – 4, and the number of articles in international databases – 10. He can be contacted at email: ftheyab@uowasit.edu.iq.



**Hamdalla F. Al-Yasriy**    is an Assistant Lecturer in the Engineering College, at the Wasit University, Iraq. At 2016 receives his BSc degree in electrical engineering from Wasit University, and at 2020, he receives his M.Sc. in the College of Engineering in Wasit University. The number of articles in national databases – 2. He can be contacted at email: hamdallak@uowasit.edu.iq.



**Haider TH. Salim ALRikabi**    is presently Asst. Prof and one of the faculty College of Engineering, Electrical Engineering Department, Wasit University in Al Kut, Wasit, Iraq. He received his B.Sc. degree in Electrical Engineering in 2006 from the Al Mustansiriya University in Baghdad, Iraq. His M.Sc. degree in Electrical Engineering focusing on communications systems from California State University/Fullerton/USA in 2014. He is author, coauthor, and editor of some national and international journals and conferences. His current research interests include communications systems with the mobile generation, control systems, intelligent technologies, smart cities, renewable energies, and the internet of things (IoT). Al Kut city – Hay ALRabee, Wasit, Iraq. The number of articles in national databases – 10, and the number of articles in international databases – 70. He can be contacted at email: hdhiyab@uowasit.edu.iq.