

Research of Harmonics in Power System Signal using Gaussian's Distribution Overlapping by Receiver Operating Characteristics (Roc) Curve

S.Sangeetha, P.Venkatakrishnan, R.Shirisha

Abstract— Harmonic analysis of the power system signal is a proliferating research in the field of electronics technology. Whenever, we analyze odd and even harmonics are present in the signal, imperative operation is needed to transform from the time domain to the frequency domain. Hence, all the researchers are utilizing the Fourier Transform technique is very effective for the analysis of odd and even harmonics in the frequency domain. In the past two decades, Wavelet Transform is a wonderful technique to analyze the harmonics both frequency and time domain as well. The analysis of harmonic and its probability distribution are most important for the purpose to predict the harmonic effects in the present situation. We treated all the harmonics and its corresponding frequency distribution are considered as a zero mean unit variance. The overlapping these distributions (small, medium, large) are analyzed with help of statistical data processing technique. It is one of the most important basic plots in the decision theory and it provides the constructive decision about the overlapping of a frequency distribution in power system signal. The curvature as a plot of sensitivity and specificity underlying the harmonics are present and not present distributive (Gaussian). The above determined values are lying in the interval probability [0, 1] and it is depends only the nature of the dataset. In this paper, we explained with help of MATLAB and level of understand the basic concept of ROC is demonstrated. The dataset is drawn from the example of odd and even harmonics are generated and the probability distribution as input to our MATLAB program.

Keywords— Harmonics, Power system signal, Sensitivity, Specificity, ROC, Gaussian distribution

I. INTRODUCTION

In power system signal processing the serious effect of harmonics to make the system is impaired and very difficult process to handle is issued. Recently, few researchers are focused to analyze the harmonics in the power system signal still proliferating field in agriculture to IT sector. Actually, the imperative development of ROC analysis is during World War II for the main purpose to analyze the radar images. The theory of ROC is further very much implemented in the medical field to diagnosis of medical related prediction is correct or not. And the application about this theory is occurring during the late 1960s. Today the ROC analysis is a more often wide-spread method with all the field example, in

Revised Version Manuscript Received on August 19, 2019.

S.Sangeetha, Assistant Professor, EEE Department, CMR College of Engineering & Technology, Hyderabad, Telangana, India.(email: sangeee12681@cmrcet.org)

P.Venkatakrishnan, Professor, ECE Department, CMR Technical Campus, Hyderabad, Telangana, India.(email: pvkmephd@gmail.com)

R.Shirisha, Associate Professor, EEE Department, CMR College of Engineering & Technology, Hyderabad, Telangana, India.(email: rshirisha@cmrcet.org)

the signal and image processing, machine learning area in field [1]-[6]. Further, now emerging trends in the ROC analysis in healthcare related biomedical field [7], particularly in the radiology frequency analysis [8], where it is used to quantify amount of the accuracy of prediction tests. Mostly, it is a 2D plot of visual interpretation between the performance and classification of the prediction model [9] – [12]. This may lead the trade-off between the accuracy of prediction and the model order. On linear discrimination theory, this type of curves is devoted in communication theory to quantify the amount of discrimination between signal and noise [13]. In different computer processing perspective, from agriculture to information technology this type of ROC analysis is a powerful tool to determine the discriminated effects in very actively.

In theory, this type of characteristic curve is either parametric test or non-parametric test. That is, the probability distributions of samples are known very well in advance. Here in the parametric analysis, the underlying population distributions of the signal with odd harmonic distribution (disease) and the signal with even harmonic distribution (healthy) are often assumed to be normal(Gaussian). The normal distribution as a fundamental assumption about our signal samples worldwide. Because the basis of Fourier Transform and Wavelet Transform is always obeyed the central limit theorem. In the other parametric type of test like other types of distributions, such as gamma and negative exponential, are sometimes used. In different perspective, contrast, non-parametric analysis does not possible to make any assumptions about the form of the underlying harmonic population distributions in the power system signal.

In power system signal, the prediction of harmonics with underlying the principle curve as a ROC curve is the plot that displays the information about the separation prediction of the of trade-off between the sensitivity and (1-specificity) across a series of threshold cutoff points from [0 to 1] with each iteration is increased by .01 and totaled 100 iterations. This curve is very useful in (i) finding the very optimal cutoff points between the prediction accuracy to least classify incorrectly odd harmonics or even harmonics subjects in the power system signal, (ii) calculating the classification ability of a test to correctly pick the odd or even harmonics subjects from the two overlapping probability distributions, (iii) possible to comparing the efficiency of two or more tests for

RESEARCH OF HARMONICS IN POWER SYSTEM SIGNAL USING GAUSSIAN'S DISTRIBUTION OVERLAPPING BY RECEIVER OPERATING CHARACTERISTICS (ROC) CURVE

assessing the same and different harmonics, and (iv) the same possible to comparing two or more harmonics related measuring the same test in power system signal processing (inter-observer variability) of typical time-spatial operation.

The sustainable development of AI, Machine learning and deep learning concepts, utilized the ROC curve as a major part of their research findings to manipulate the hypothesis accordingly. The understanding concept in the plot of ROC graphs is easier to understand by everyone, but non-obvious complexities still exist. Because, the entire focus depends only on the odd and even harmonics overlapping and its distribution variation with respect to temporal – spatial distribution. When we used to analyze these concepts using MATLAB, due to common misconceptions and pitfalls the performance of the curve is not plotted perfectly. This paper attempts our goal to serve as a basic introduction to ROC graphs, characteristics of sensitivity and specificity, and as a guide for others using them in research. The main and pivotal goal of this article is to advance the general knowledge about ROC graphs, and area under the curve (AROC) so as to promote better evaluation practices in the field of percentage of harmonics exists in power system signal. This paper is organized and structured as follows for the help of readers to understand the concept. The brief introduction about the construction of the curve is clearly manifested with real time example. The example experiments with odd and even harmonics and its overlapping distribution is presented in section III. The overlapping of odd and even harmonics and its three categories are properly presented in section IV. At the end, the section V gives concluding remarks about the ROC curve for future use of our research related prediction.

II.ROC CURVES

For the purpose of understanding the situation where there are two types of events in medical diagnosis or in terms of communication theory, input data with clean data and noise data, is represented in shown in Fig.1 for the purpose to discriminate the properties of clean and noisy data [14]. From the Fig.1, possible to distinguish between the types of signal and noise distribution, it is most important in signal processing type classification. This is the definition of ROC curve analysis. Table I shows the 2 X 2 contingency table of the actual type of a signal and noise, against its test classification according to the threshold to determine the sensitivity and specificity of the data points. When the actual type of properties about the clean and noise data, this table is very helpful to construct the ROC curve.

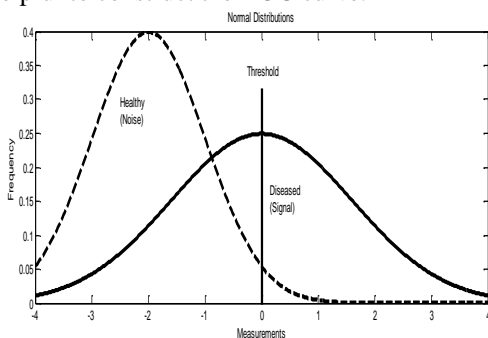


Fig.1 The basic construction of ROC model with clean and noise data with underlying model for ROC curves for signal and noise contamination

TABLE I

2 X 2 Table in with known variable threshold

		Observed data set (Tabulation)	
		Clean data	Noise data
By Matlab Predicted data set	Clean data	True Positive (TP)	False Positive (FP)
	Noise data	False Negative (FN)	True Negative (TN)

In practical, the above table is very much useful to take the possible decision about clean data and noise data contamination if existed. The observed data set and the predicted data set is act as a major role for the prediction of four categories of ROC curve construction. The four categories are true positive, false negative, false positive and true negative. The performance of the ROC curve depends only on these four categories values and the area under the ROC (AROC) curve is still active research is going on. Because, an AROC also one of the deciding parameter of classification accuracy [14]. These four parameters are most important to construct the ROC curve.

There are three equations are manipulated with respect to the Table I information. The first one is sensitivity, the second one is specificity and the final one is efficiency of the prediction. These three parameters are frequency used to construct the ROC curve by proper implementation of MATLAB code with recursive operation. The each operation totally 100 iterations are used from the threshold [0, 1]; for the purpose of ROC curve construction.

$$Se = \frac{TP}{(TP+FN)} \tag{1}$$

$$Sp = \frac{TN}{(TN+FP)} \tag{2}$$

$$EFF = TP + TN \tag{3}$$

III.EXPERIMENTAL PROCEDURE

In Fig. 2 there are three fundamental decision support system gives output in the interval [0, 1] of probability distribution with overlapping of odd and even harmonics. Based on the harmonics overlapping between the odd and even harmonics and its probability distribution, the ROC curve demonstrated by the following way, the sensitivity is arranged in the vertical scale direction and the opposite scale of the specificity is arranged in the horizontal axis. This type of construction is and the plot is very useful to analyze the overlapping between the odd and ever harmonics in the presence of power system signals. In other words, the scale of the horizontal axis with respect to each threshold from 0 to 1 is also called the false positive rate. The performance is based on the cutoff points of the ranges from 0 to 1. If several tests are performed in the three cases of overlapping, the ROC curves can be used to compare their performance of

overlapping in power system signal. Hence, another way to use ROC curves is to see the performance of a decision support system.

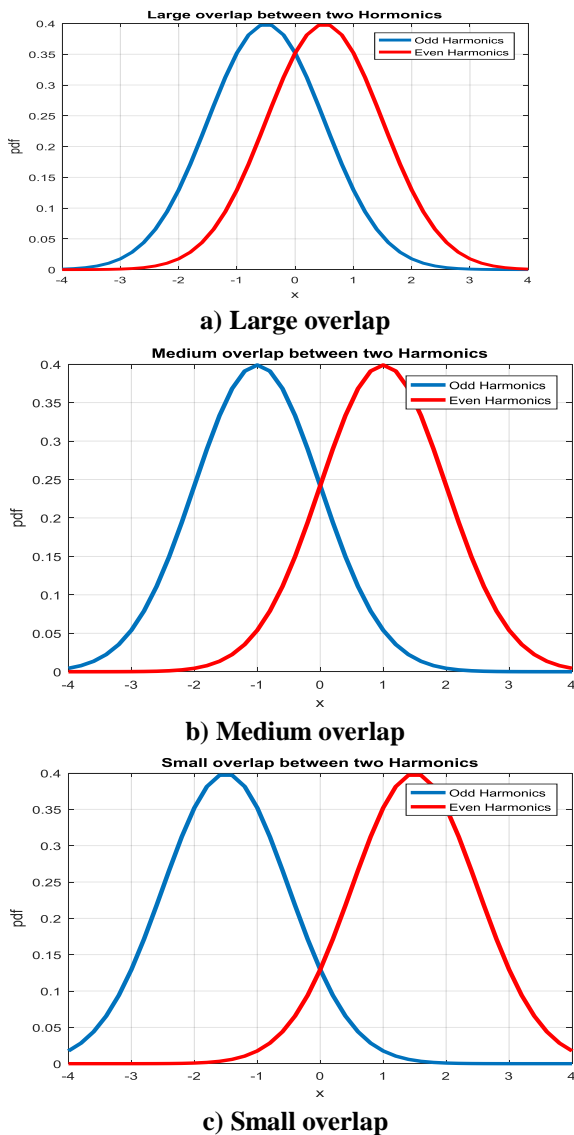


Fig.2 The odd and even harmonics with a) odd-even large overlapping, b) odd-even medium overlapping and c) odd-even small overlap

IV. RESULT ANALYSIS

The pure sinusoidal signal is generated with odd and even harmonics. Using statistical signal processing based on the probability Gaussian distribution to collect the each point on the curve corresponds to a specific pair of odd and even harmonics. In Fig. 3 a) and b) represents the clear understanding the concepts with odd harmonic presence in the pure sinusoidal waveform. The absence of few odd harmonics does not generate the pure square waveform for an exact signal generation with 60 or 50 Hz fundamental frequency. Based on the ROC curve concept with a sufficiently large number of thresholds changing in small discrete steps, a plot of 1- specificity (along the x axis) against sensitivity (along the y axis) for each threshold gives a ROC curve. A typically shaped curve in a multi threshold plot is given in Fig.4.

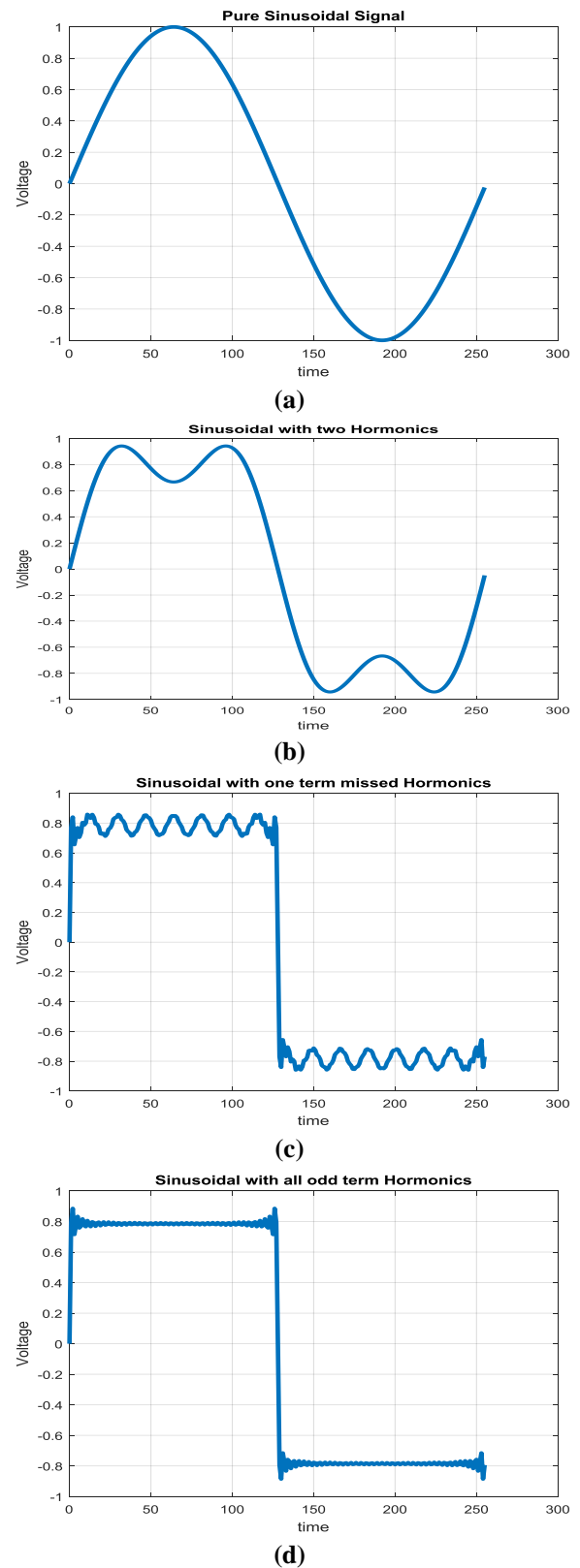


Fig. 3 The effects of harmonics presents in the pure sinusoidal signal. a) pure sinusoidal waveform, b) pure sinusoidal waveform with two odd harmonics presents, c) pure sinusoidal waveform with missed 15th odd harmonics component d) pure sinusoidal waveform with all odd harmonics (seems to be a square wave)

RESEARCH OF HARMONICS IN POWER SYSTEM SIGNAL USING GAUSSIAN'S DISTRIBUTION OVERLAPPING BY RECEIVER OPERATING CHARACTERISTICS (ROC) CURVE

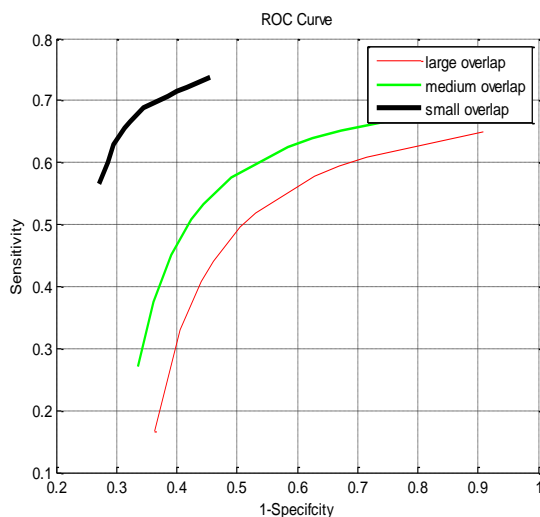
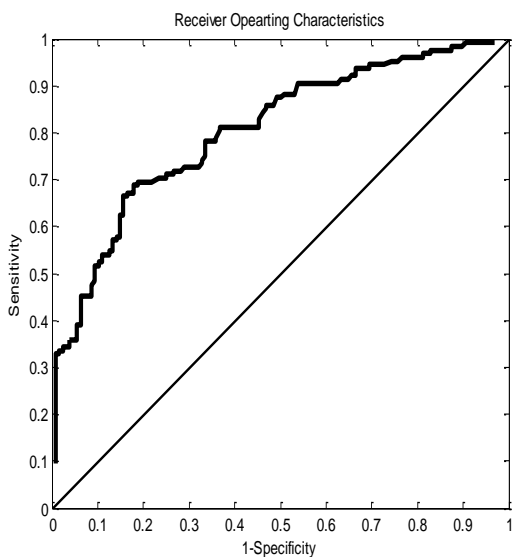
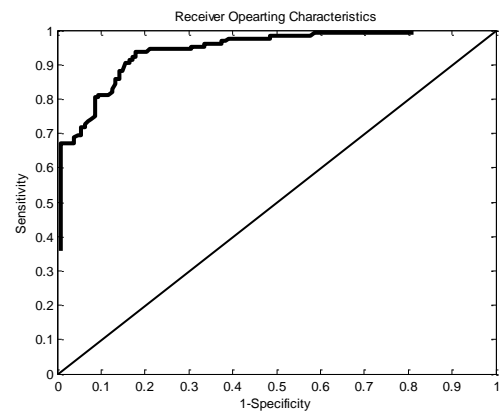


Fig.4 Example ROC curve with three examples cases of odd and even harmonics distribution overlapping (small, medium and large)

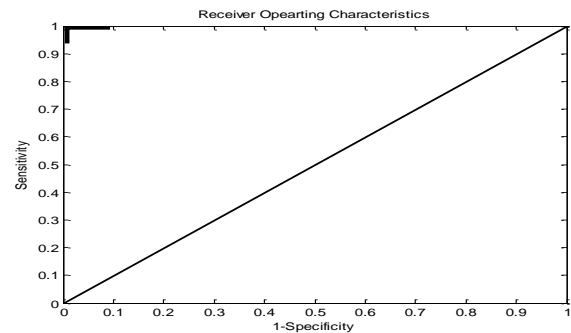
The above curve representation shows the tradeoff between the overlap between the odd and even harmonics probability distribution for a power system signal. If the odd and even harmonic and its probability distributions are clearly separated by well then the ROC curve is within few iterations the curve tends to top left corner and increasing towards horizontal axis. This is the pivotal point of our understanding of the ROC curve in the separation of odd and even harmonics. If the probability population distributions and tends to medium and more overlap, so that odd and harmonic components and its curve nearer to linear line show in Fig. Hence, the observations from the ROC curve prediction of overlapping between the harmonics very well. In Fig.5 shows the understanding of the x and y axis clearly from the tabulation details I and the values of sensitivity versus 1-specificity. The Fig. 5 a) marked by the arrow to represented the x and y axis values and the corresponding threshold is 0.325 is specified with some set of both odd-even harmonics are underlying with large distributions. Hence, from the observation point of view when the overlapping is more, then the curve towards to top left of the graph.



a) Large overlapping distributions



b) Medium overlapping distributions



c) Small overlapping distribution

Fig.5 ROC curves of odd-even harmonics population distribution for three pairs of distributions with a) large overlapping, b) medium overlapping and c) small overlap.

It is observed that the result of the decision support system about the contributions of odd and even harmonics present in the power system signals in the forms two normal distributions, one for the odd component and the other one for the even component density distributions. The threshold is placed in between the mean of odd and even harmonic distribution with different scales to approximate the positions distributions. The determined resulting points are plotted as a ROC curve of large, medium and small overlap show in Fig.5. The above figures information are excerpted the probability distribution in between +4 to -4. The overlap between the distributions is largest in the first example and the second is medium and the smallest in the third. All these examples are tested only in the form of parametric test only.

V.CONCLUSION

In this paper, we analyzed the harmonic presence and its contribution in power system signal with the help of ROC curve by three pairs of Gaussian distributions we presented. Depends on the data set generated randomly using MATLAB and the ROC analysis relies heavily based on the notations as sensitivity and specificity determination. Even in the values in theory, the practical also for computation purpose lie in the interval [0, 1]. Three data sets have been presented with the help of parametric approach and three different ways to create decision support systems of the identification of odd and even harmonics in real-time application to find the area



under the ROC curve.

VI. ACKNOWLEDGMENT

Behalf of authors, we express our ardent thanks to CMRCET management, principal, HOD for their wonderful support, laboratory utilization and valuable suggestions that helped to improve this research findings.

REFERENCES

1. Kraemer, H.C, Evaluating Medical Tests. Sage Publications, 1992.
2. Bamber, D. The area above the Ordinal Dominance Graph and the Area below the Receiver Operating Characteristics Graph. J.MathPsychol, 12: pp.387-415.1975
3. Metz, C.E, Basic Principles of ROC analysis. Seminars in Nuclear Medicine, VII (4): pp. 283-298, 1978.
4. DeLong, ER, DeLong D.M and Clarke-Person, D.L, Comparing the Areas Under Two or more correlated Receiver Operating Characteristics Curves: A Nonparametric Approach. Biometrics, 44: pp.837-845.1988.
5. Hanley, J.A, Receiver Operating Charactersitics (ROC) Methodology: The State of the Art. Critical Reviews in Diagnostic Imaging, 29 (3):pp 307-335, 1989.
6. Armitage, P and Berry, G. Statistical Methods in Medical Research. Blackwell Scientific Publications, 1994.
7. Adlassning K.P, and Scheithauer.W, "Performance evaluation of medical expert systems using ROC curves," Comput. Biomed. Res., vol.22, No.4, pp. 297-313, 1989.
8. Cheng H.D, Lui .Y.M., and Freimanis.R.I, "A novel approach to micro-calcification detection using fuzzy logic technique," IEEE Trans. Med, Imag., vol.17,no.3, pp.442-450, 1998.
9. Fawcett.T, ROC graphs: Notes and practical considerations for researchers. Machine Learning, 31, 2004.
10. Flach.P, Blockeel, H. Ferri.C., Hernandez - Orallo, J, and Struyf.J, Decision support for data mining: Introduction to ROC analysis and its applications. Data mining and decision support: Aspects of integration and collaboration, pp. 81-90, 2003.
11. Flach.P Tutorial at ICML: The many faces of ROC analysis in machine learning. Unpublished manuscript, 2004
12. Flach. P.A, The geometry of ROC space: Understanding machine learning metrics through ROC isometrics. Proceedings of the 20th international conference on Machine learning, 194-201, 2003.
13. Egan J.P, Signal Detection theory and ROC analysis. Academic press New York, 1975.
14. Julinan B. Tilbury, Peter W.J. Van Eetvelt, Jonathan M. Garibaldi, John S.H Curnow, and Emmanuel C.Ifeachor, "Receiver Operating Characteristics Analysis for Intelligent Medical Systems - A new approach for Finding Confidence Intervals", IEEE Trans. On Biomed. Engg, Vol.47, No.7,pp.952-963, 2000.