Neural Network Based Classification of Partial Discharge in HV Motors

Yahya Asiri *Alfred Vouk*
Electrical Systems Division.
Saudi Aramco
Dhahran, Saudi Arabia
yahya.asiri@aramco.com *alfred.vouk@aramco.com*

Lee Renforth
HVPD Ltd
Manchester, UK
lee@hvpd.co.uk

David Clark *Roger Shuttleworth*
The University of Manchester
Manchester, UK
dave@hvpd.co.uk *roger.shuttleworth@manchester.ac.uk*

Jack Copper
NeuralWare
Pittsburgh PA, USA
jack.copper@neuralware.com

***Abstract*—This paper discusses the general application of using Neural Networks (NN) to classify six different types of Partial Discharge (PD). Stator winding failures contribute about 30-40% of the total motor failures according to IEEE and EPRI. Ninety percent (90%) of electrical failures on High-Voltage (HV) equipment are related to insulation deterioration. Large datasets were collected for motors with PD defects as well as PD-free machines. The datasets of PD were pre-processed and prepared for use with a NN using statistical means. It was possible to utilise the advantages offered by multiple NN models to classify the PD defects with a maximum recognition rate of 94.5% achieved, whereas previous research work did not exceed a classification accuracy of 79%.**

Keywords- Partial Discharge; Neural Networks; Pre-processing; Multiple Defects

I. INTRODUCTION

Electric motors play a pivotal role in various industrial applications for electrical to mechanical energy conversion. As a result, their reliability and availability are of the utmost importance to industry [1]. In order to prevent productivity losses and achieve minimum machinery downtime, it is extremely important that such critical machines are constantly monitored and diagnosed for potential faults. Sources of failures may extend from mechanical faults such as broken rotor bars, damaged motor bearings and air-gap eccentricities to electrical faults such as stator winding shorts and supply voltage imbalance. A majority of these faults is often based on the physical degradation of parts, and hence require condition monitoring and preventative maintenance actions.

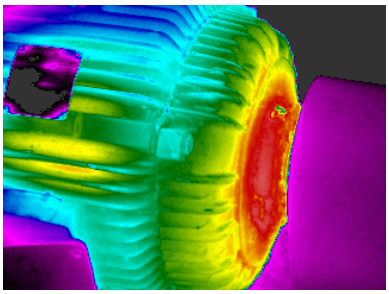


Figure 1. Overheated inboard bearing detected by infrared imaging

II. MOTOR FAILURES STATISTICS

One of the major causes of failures in high voltage motors is stator winding failure which contributes to around 37% of motor failures, according to a studies carried out by both the IEEE and EPRI (Electrical Power Research Institute) as shown below in Figures 2 and 3.

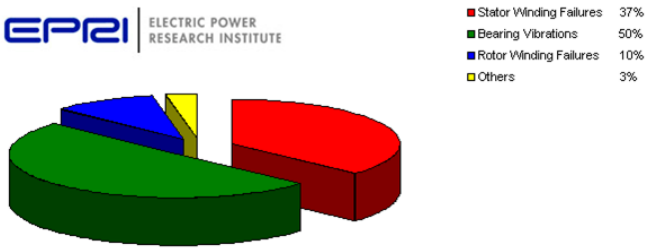


Figure 2. EPRI HV Motor Failure Causes Distribution [2]

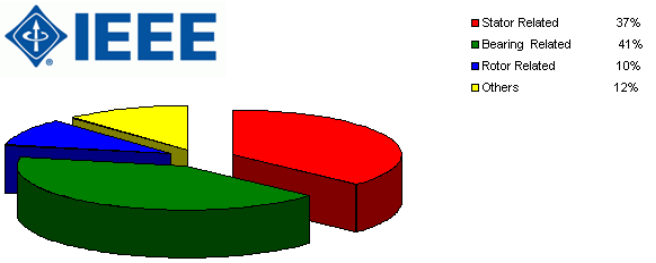


Figure 3. IEEE HV Motor Failure Causes Distribution [3]

III. PD DATA COLLECTION

In this project, a limited data-set of PD test and monitoring data was provided by an oil and gas producer with data from a single gas plant where 16 large HV motors have been in service since the 1970's [4]. Ten of these motors were equipped with a PD monitoring system which utilized HV Coupling Capacitor PDA sensors. These 16x HV Motors are shown below in TABLE 1.

TABLE I. 13.2 kV HV MOTORS MONITORING

Location	HP	Application	Type
KM002A	21000	Fuel & Sales Gas	Synchronous
KM002B	21000	Fuel & Sales Gas	Synchronous
KM002C	21000	Fuel & Sales Gas	Synchronous
KM001A	16000	Fuel & Sales Gas	Induction
KM001B	16000	Fuel & Sales Gas	Induction
GM070A	7000	Product Surge	Induction
GM070B	7000	Product Surge	Induction
GM070C	7000	Product Surge	Induction
KM161A	21000	Chilldown	Synchronous
KM161B	21000	Chilldown	Synchronous
KM261A	21000	Chilldown	Synchronous
KM261B	21000	Chilldown	Synchronous
KM361A	21000	Chilldown	Synchronous
KM361B	21000	Chilldown	Synchronous

At the project start, a review was made on the datasets available for the training, validation and testing of the proposed NN. While the online PD monitoring systems' performance was satisfactory for operation and maintenance purposes, it was found that the collected PD datasets were inadequate to train a NN to classify the different PD types. Another shortcoming was the limited number of stator winding failures that could be related to PD activity.

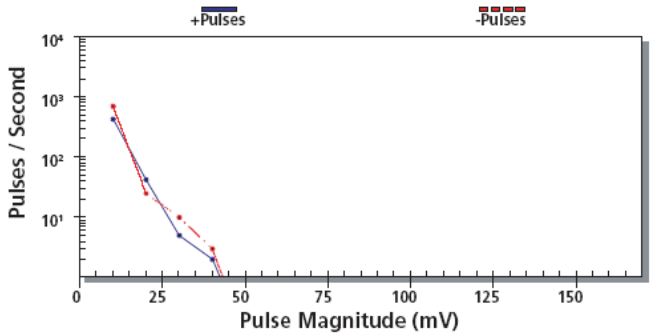


Figure 4. Measurement Taken at the Oil and Gas End-user

Work was thus undertaken to convince a number of PD test and measurement manufacturers to contribute in the collection of more detailed PD measurements. Whilst the PD test manufacturers were initially reluctant to disclose data from their customers PD measurements, access to the data was allowed after an anonymity agreement was reached. The PD datasets have thus been identified as manufacturer A, B, and C. Phase Resolved Partial Discharge (PRPD) patterns over the complete 360º power cycle were needed to apply the power of an NN pattern recognition function to classify *six different types of PD, namely* corona, end-winding, internal, slot, surface and no PD (for healthy machines).

Manufacturer A: The R&D department of this manufacturer was contacted to request PD measurement datasets, which have been well documented in technical papers prepared by experiment's at a HV Lab. PD trend patterns were supplied as an image taken from the PD meter without any tabulated details. A trial was done to digitize these PRPD trends to convert the image to pC-deg without success.

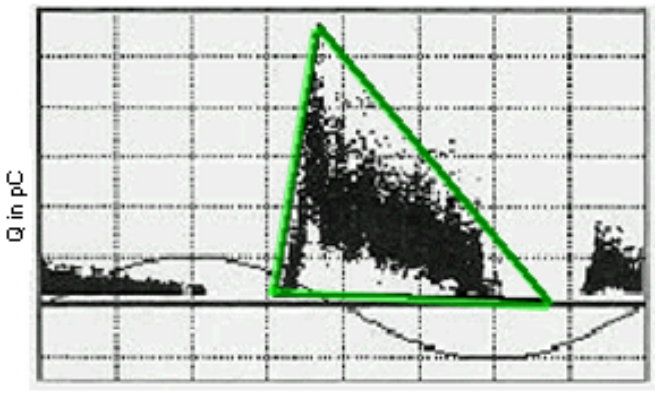


Figure 5. PD Data provided by Manufacturer A

Manufacturer B: This supplier offered an excellent set of PRPD colour-coded PD measurements, and the corresponding data text was tabulated and shown in Fig. 6. Although high quality PD trends were given, it was not feasible to generate satisfactory datasets to cover all targeted PD types since the trends were limited and inconsistent.

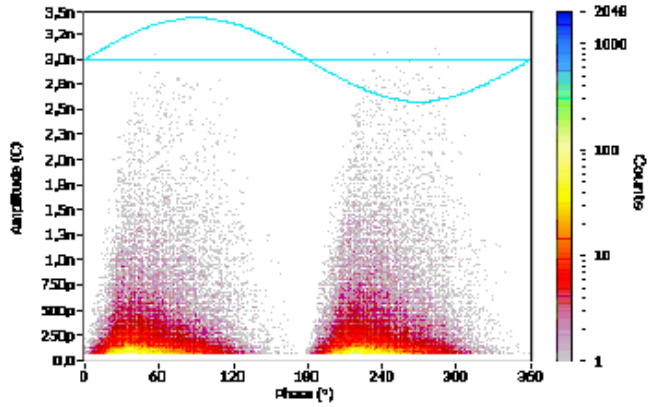


Figure 6. PD Data provided by Manufacturer B

Manufacturer C: This manufacturer was introduced through some customers in the region who were using PD instruments from the supplier. The manufacturer was able to collect, classify and communicate suitable PRPD datasets with both PRPD and pulse waveshape details from the 360º power cycle. The text files were made available in tabulated format so that they could be employed to calculate the PD pulses charges at all positions across the 60Hz power cycle. Figure 7 shows the PRPD mV-deg.

IV. High Resolution Data Analysis of PD Signals

The idea of making PD measurements on the multi-channel data from high voltage motors and generators using various types of PD sensors is not new and has been around in one form or another for at least 40 years. The analysis of PD data obtained from these measurements over this period has largely consisted of looking at the PRPD patterns and 'clusters' across the 50/60Hz power cycle. Diagnostic 'knowledge rules' based

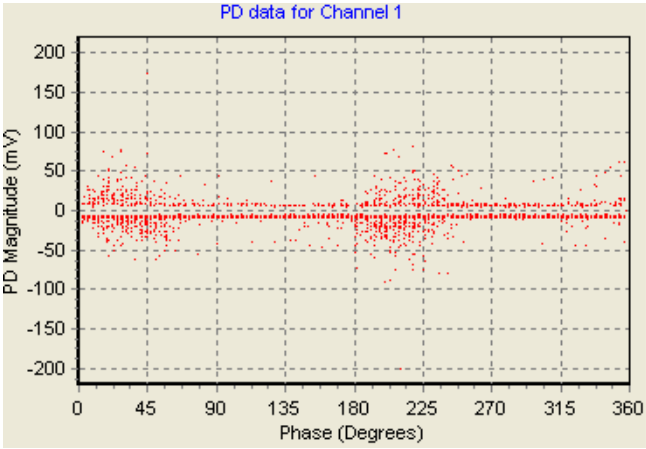


Figure 7. PD Data provided by Manufacturer C

on the magnitude, intensity and grouping of pulses across the power cycle using this PRPD-type analysis has been published many times by a number of authors from around the world. Some of these 'knowledge rules' are also included in the two main guides published by IEEE and CIGRE [5, 6].

Over the past 5 years the authors have developed new measurement methods for high-resolution data capture (in the data acquisition range of 100MS/s up to 250MS/s) which now make it possible to consider a new type of analysis of PD activity. This is based on *individual pulse shape analysis* in which the analysis of high resolution 'waveforms' of individual PD pulses is made. It is proposed that this method allows for a much more detailed analysis of the PD pulses emanating from the machines' windings to be made. This detailed recording of the waveshapes of individual PD pulses allows for different PD pulses, from different defects in the machine insulation, to be separated and analysed.

Manufacturer C have developed robust algorithms for diagnosing PD in power cables and switchgear over the past 10 years which uses the wave shapes and pulse timing between phases to get useful diagnostic data in the form of 'Event Recognisers' for different types of PD which was not available before. It is seen as a natural extension to allow the same type of data analysis to be made for rotating HV plant, thereby greatly improving the potential for on-line diagnostic testing, particularly when diagnosing multiple discharging sites in stator windings.

V. Example of a PD Waveshape From a 100MVA Hydrogenerator

A. *Multi Channel Comparison Methods*

Multi-channel data across the 60Hz Power Cycle is shown below in Figure 1 from a single power cycle of PD data from an on-line PD test on a large hydro generator. The measurements here were made using permanently installed 80pF high voltage coupling capacitors (PDA's).

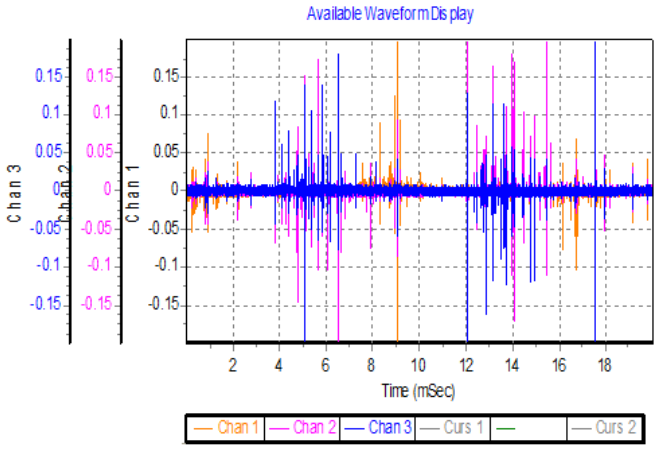


Figure 8. PRPD data from three phases of a hydro generator

With reference to Figure 8, it is clear that some pulses exist singularly on one channel and others exist on several or all channels simultaneously. If the pulse propagation path through the generator windings could be established, then simple locations (or at least categories) of PD can be established based on pulse shapes. These will be mainly separated into phase/phase PD events and phase/ground PD events.

In addition to the PRPD analysis methods, with high resolution signal acquisition it is possible to consider using the *waveshapes* of the PDs to give clues to the nature of the discharge type and/or the location of the PD. Two examples of different PD pulses are shown below in Fig. 9 and Fig 10.

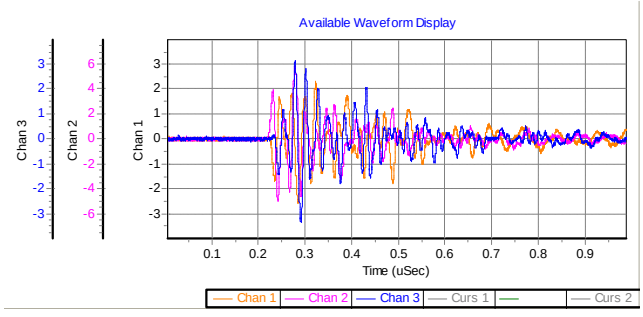


Figure 9. Type 1 PD: 'Surface' type PD event

The Type 1 event shown in Fig. 9 is a *phase-to-phase* event as the signal is seen on all three PDA couplers at once. Notice the *oscillatory nature* of this Type 1 PD event in the above case. This is typical of a surface discharge/tracking PD event in the end winding of the machine.

Figure 10 shows a Type 2 PD event detected on Ch 1 from the same test data on the same hydro generator. This type of PD is a *phase-to-earth* discharge as it is largely monopolar in shape

with predominance on Ch1 (with a small amount of 'cross-talk' on the other two phases) and is more typical of internal PD originating within the winding insulation of defined impedance.

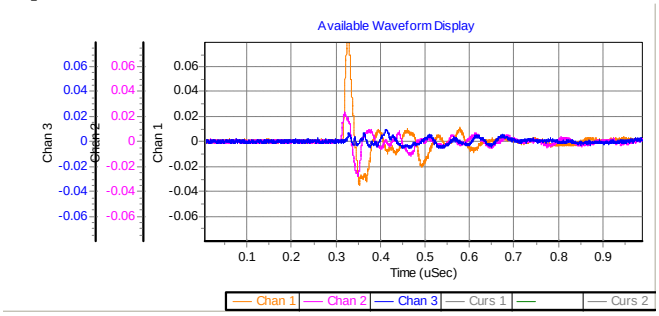


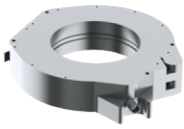
Figure 10. Type 2 PD: Internal PD type event

The two graphs shown in Figures 9 and 10 illustrate the different waveshapes which can be captured in the detection method. It can also be noted that there is a very large disparity in the magnitudes of the 2x types of PD activity (surface discharges up to 4.0V+, internal discharges = 80mV).

VI. PD Sensor Options for Rotating HV Machines

There are a number of sensor options available for measuring PD activity in rotating machines. These include HV Coupling Capacitors (PDA's), ferrite-based High Frequency Current Transformers (HFCT's) and Rogoswki Coil (RC) Sensors. These sensors are shown below in TABLE 2 along with their relative sensitivity (relative sensor output in mV) at 10 MHz. The best sensor solution for any application will depend on the power rating of the machine under test and practical issue of the space available in the cable box of the machine to install the sensors.

TABLE II. PD Sensor Options For Rotating Machines

Sensor	**PD Sensor Options for Rotating Machines**		
		Coupling Method	***Relative Sensitivity at 10MHz***
High Voltage Coupling Capacitor		Capacitive	80
Ferrite-cored High Frequency Current Transformer		Inductive	20
Rogowski Coil		Inductive	1

It is important also to know frequency response of the sensor across the wideband frequency range of 100 kHz to 50 MHz. With a known sensitivity it is possible to make on-line measurements of PD in Pico Coulombs (pC's) with a suitable, high-resolution wideband PD measurement system. Fig. 11 shows the output of the three types of sensor across the frequency range of interest. It can be noted that the HV Coupling Capacitor is most sensitive followed by the ferrite HFCT sensor and then the Rogoswki Coil. All of these sensors were evaluated in the project.

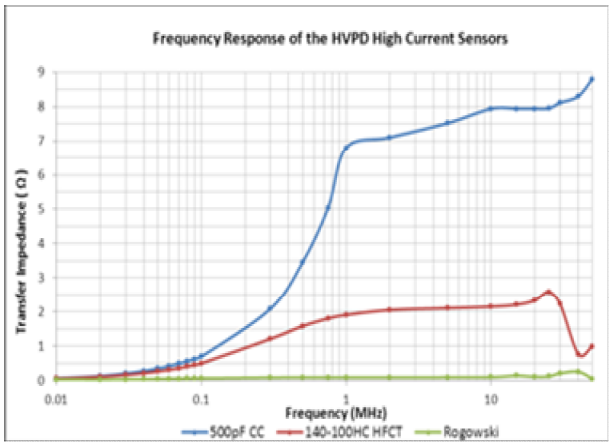


Figure 11. Frequency Response (10 kHz to 50MHz) for the 3 main types of Rotating Machine PD sensor

VII. Neural Network Overview

The fundamental objective of this project was to develop neural network classification models capable of differentiating between the following six types of PD:

a. Corona
b. Endwinding
c. Internal
d. Slot
e. Surface
f. No PD

Neural networks are machine-learning algorithms which in essence perform a nonlinear mathematical mapping from a set of input values to a set of output values. Neural networks "learn" the mapping by iterating through available data and making small adjustments in the coefficients of the mapping function until some stopping criterion is met. As a consequence, neural networks are considered empirical models – and they are particularly well suited to learning patterns in data acquired from sensors attached to physical operating equipment.

While a neural network can employ one of two basic forms of learning – unsupervised or supervised, given the availability of labeled datasets the supervised learning form was selected for this project. Supervised learning entails minimizing the error between the model output and the target output (i.e.,

label) while iterating through the training data. The stopping criterion is typically an error threshold, with an iteration threshold as a secondary criterion to ensure that training ultimately terminates. The initial values of coefficients are set randomly.

To actually construct the neural networks, NeuralWare's NeuralSight model generator and the NeuralWorks Predict neural network engine were used to create feed-forward neural networks trained with back-propagation supervised learning. The platform was chosen not only for its ability to create robust neural networks, but also for its ability to automate other critical tasks associated with any empirical model development effort.

In traditional neural network development environments there are several critical decisions that a model builder must make. One is to specify the basic network architecture (number of hidden layers and number of hidden units); another is to specify internal transfer and output functions; another is to specify the internal evaluation function. NeuralSight makes these decisions automatically based on the type of model required (prediction, classification, or clustering). For prediction and classification models (supervised learning), the neural network is dynamically constructed using cascade correlation [7].

With cascade correlation, a network initially has no hidden units; as training progresses, hidden units are added until network performance on the test set ceases to improve. For classification models (as developed in this work) Predict uses a softmax output transfer function, which enforces the mutual dependencies of outputs; the internal evaluation function is relative entropy – which maximizes the probability that a training record belongs to its appropriate class [8]. This approach provides an elegant interpretation of raw neural network output, and is the basis for identifying multiple PD faults. Since network output values are probabilities of class membership, the natural engineering interpretation is that the class with the highest probability is the dominant fault, and the class with the next highest probability is a contributing fault. If a single probability dominates (e.g., > .90) then likely there is only one fault. Furthermore, the use of numeric probabilities means that the results of multiple models may be averaged to provide a better estimate of the type of PD fault(s) detected.

In addition to choosing the appropriate internal evaluation function and output transfer function, NeuralSight also automatically analyzes all training data and applies mathematical transformations to generate (transformed) input data sets whose distributions of values more closely match the distribution of output values. The set of transformed inputs then are candidate inputs for the neural network model, but before the model is trained a genetic algorithm (GA) optimizer is run in order to find the best set of inputs for the neural network. The fitness function for the GA is either a linear regression model or a "small" neural network; the inputs which most frequently produce the best model over a 50 GA populations are the inputs which are used to train the neural network.

NeuralSight minimizes the risk of over-fitting models (memorizing patterns rather than learning patterns) by internally partitioning data into training and test sets, and using performance of models when run with the test set as the primary criterion for stopping training.

The figure below illustrates results achieved on the selected reduced matrix dataset. Models were ranked by average accuracy, and the best model run using the entire dataset.

VIII. PD Data Preproccesing for NN Input

The pre-processing phase commenced by converting the PRPD data shown previously in Figure 7 to ASCI file's with numeric values. The numeric values were used to reconstruct the PRPD to an Excel spreadsheet chart showing PD magnitude (in mV) vs. the phase angle.

The PRPD plot measured is cumulative and is built-up by synchronously acquiring data over a number of power cycles to produce the final PRPD plot. This consists of the PD magnitude (in mV or pC) at various phase angles of the power cycle AC sine wave. Based on the above, the PRPD was reconstructed into the numerical chart shown in Figure 12 which was used to prepare the PRPD vs. Phase Angle matrix to be ready as an input for the NN. Statistical values of minimized PD vs. phase angle were the foundation of preprocessing phase.

The following reduced matrices are as follows:

2. Max and Min chart as in Figure 13
3. Per (Max and Min) chart as in Figure 14
4. Max and AbsMin chart as in Figure 15

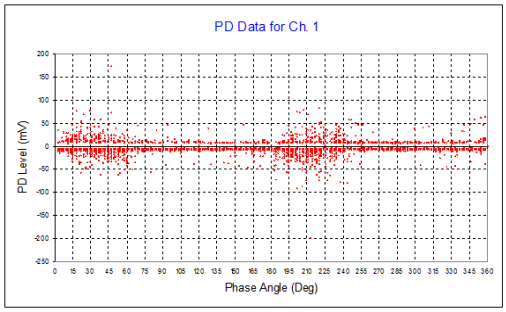


Figure 12. Reconstructed PRPD

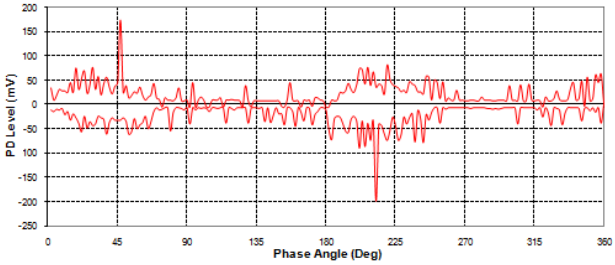


Figure 13. Max and Min PRPD Reduced Matrix

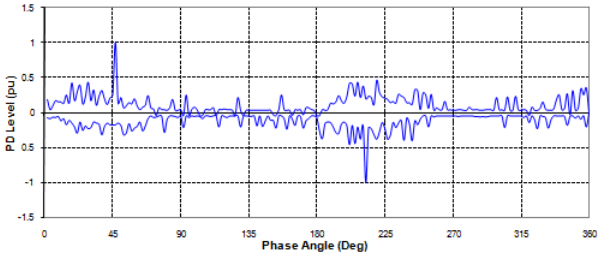


Figure 14. Per (Max and Min) PRPD Reduced Matrix

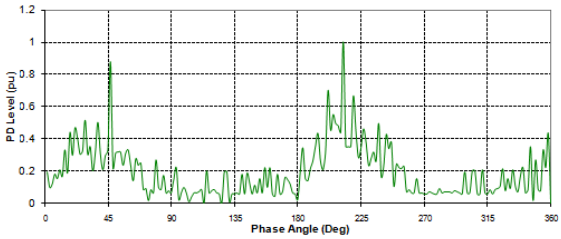


Figure 15. Max and AbsMin Reduced Matrix

The validation phase was carried out through field measurements on 13.2kV, 9MW motor which was examined for PD activity. Figure 16 shows a photo of Manufacture C's PD test engineer and the Oil and Gas Company's field engineer conducting on-line PD measurements which were used to test and validate the developed NN.

The *pulse shape analysis* of the high resolution 'waveforms' made by the subject matter experts revealed that two distinct types of PD were present in one of the motors tested.

1. *Internal PD*
2. *Surface PD*

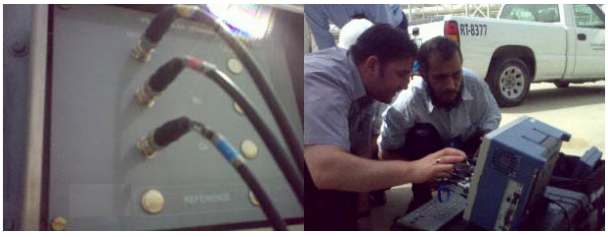


Figure 16. PD Field Measurement

A feature of the NN used was the use of multiple models that enabled ranking the PD activities according to the type as a percentage of all of the PD pulse activity recorded. Table III shows the percentage weight of each type.

A. Envelop (Max_Min) reduced matrix

As shown in Fig. 14, the PRPD matrix was pre-processed using the maximum and minimum values of the PRPD. The reduced matrix was fed to the NN and the results for 9MW motor are presented in Fig.s 17, 18 and 19.

TABLE III. MOTOR I - PD CLASSIFICATION USING MAX_MIN REDUCED MATRIX

PD Type	Percentage
Corona	0.33
End-winding	4.02
Internal	45.10
Slot	49.33
Surface	0.244
No PD	0.96

The types of PD characterized by the NN for Motor I are shown graphically as shown in Fig. 17.

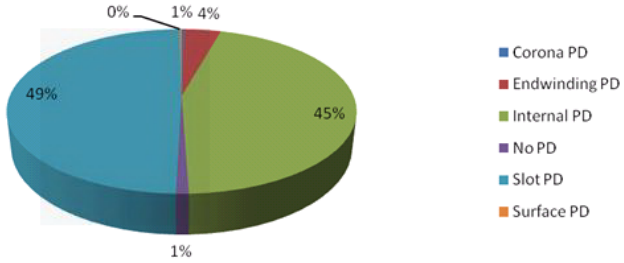


Figure 17: PD classification using Max_Min reduced matrix

The pie chart shows 45.10% of the PD activity was classified as Internal PD and 49.33% was classified as Slot PD. The dissimilarity between the human classification and the NN performance was discussed with Manufacturer C and the manufacturer accepted that the NN results stating surface PD may appear as slot PD. Thus the two types of PD were classified by the human operator and the NN.

Error calculation:
Error = suberrors
Suberror = 0.33 + 4.02 + 0.24 + 0.96 = 5.55 (1)
Accuracy = 100 - error (2)
Accuracy = 100 - 5.55 = 94.45 %

Calculations above showed that the NN achieved an accuracy of 94.5% in classifying multiple types of PD defects. Such accuracy exceeded previous results of 79% given in the literature [5].

TABLE IV. PD CLASSIFICATION USING THE MAX_MIN REDUCED MATRIX

PD Type	Percentage
Corona PD	0.90
End-winding PD	4.02
Internal PD	0.01
Slot PD	7.95
Surface PD	43.02
No PD	44.10

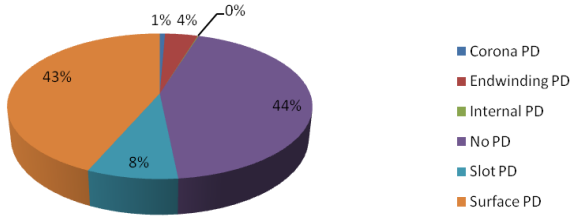


Figure 18: PD Classification using the Max_Min reduced matrix

Suberror = 0.9 + 4.02 + 43.02 + 44.1 = 92.04 (3)

Accuracy = 100 – 92.04 = 7.96 % (4)

B. The Max AbsMin reduced matrix

As shown in Figure 14, the PRPD matrix was pre-processed using the maximum and absolute values of the minimum curve. The reduced matrix was fed to the NN and the results were presented in Fig.19.

TABLE V. PD CLASSIFICATION USING THE MAX ABS MIN REDUCED MATRIX

PD Type	Percentage
Corona	93.07
End-winding	1.18
Internal	2.36
Slot	0.88
Surface	1.64
No PD	0.87

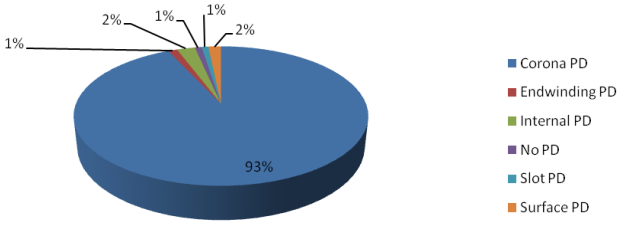


Figure 19: PD classification using Max AbsMin reduced matrix

Suberror = 93.07 + 1.18 + 1.64 + 0.87 = 96.76 (5)

Accuracy = 100 – 96.76 = 4.24 % (6)

IX. FUTURE WORK

In an attempt to understand the concepts of pulse propagation and reflections of PD signals in HV rotating machinery better, the authors intend to carry out pulse injection experiments on HV machine stator windings. Carrying out such tests will allow a better understanding of how the machines' geometry and construction affects pulse attenuation and distortion from source to sensor. The accurate location of PD sites within a rotating machine insulation will be dependent upon the combined knowledge of the propagation paths throughout the machine, and the correct application of a suitable differential equations to describe the pulse propagation process.

Developing a relationship between the sites of PD activity within a machine, and the pulse shapes of the signal at different locations, will be dependent upon the frequency response of the PD sensor, and understanding certain types of differential equations which permit so-called traveling wave solutions. Such equations help in describing the PD pulse propagation process as they exhibit decaying modes of solution that model how a wave in a dispersive medium travels. Further work will be carried out to examine the validity of using differential equations to predict PD pulse attenuation, dispersion and reflections as a function of distance travelled through the windings.

X. CONCLUSION

The NN was examined to classify different types of PD according to the location of PD activity. Main challenging phase in the process was the PD data collection. Only the tabulated PD measurements over the 360º were the base of generating the Phase Resolved partial Discharge. The statistical reduction techniques revealed that the envelop chart was the most accurate preprocessed feed to the NN. NN classification was benchmarked against human being classification with a maximum recognition rate of 94.5% achieved, whereas previous research work did not exceed a classification accuracy of 79%.

ACKNOWLEDGMENTS

The authors would like to thank the following entities for their assistance to develop this paper: Saudi Arabian Oil Company (Saudi Aramco), King Fahd University of Petroleum & Minerals, Siemens Saudi Arabia, ABB Saudi Arabia, HVPD UK and NeuralWare USA.

REFERENCES

[1] Lourenco Mafika Thusi, "Determining the Optimal Technique for EarlyDetection of Broken Rotor Bars in Medium Voltage Squirrel Cage Induction Motors during Operation," Thesis of Master of Science in Engineering at University of Witwatersrand, Johannesburg, South Africa, 2009.

[2] EPRI, "Improved motors for utility applications and improved motors for utility applications industry assessment study," Vol. 1, EPRI EL-2678, 1763-1, final report, and Vol. 2, 1763-1 final report, October 1982.

[3] Motor Reliability Working Group, IEEE Transactions on Industry Applications "Report of Large Motor Reliability Survey of Industrial and Commercial Installations," Part I and II, Vol. IA-21, No. 4, pp.863-872.

[4] A.Vouk, Y. Asiri, Z. Al Hamouz " Online PD monitoring insures reliable operation of 24000 hp motors in the Saudi Arabian deserts" Insucon 2009, Birmingham, UK, May 2009

[5] IEEE Std 1434-2000 - IEEE Trial-Use Guide to the Measurement of Partial Discharges in Rotating Machinery, IEEE Power Engineering Society, 26 April 2000

[6] CIGRE TF 15.11/33.03.02 – Knowledge Rules for Partial Discharge Diagnosis In Service, CIGRE 226, April 2003

[7] S. E. Fahlman, C. Lebiere, "The Cascade Correlation Architecture," Advances in Neural Information Processing Systems Volume 2, San Mateo, CA: Morgan Kaufmann Publishers, 1990.

[8] Bishop, C.M. (1995). Neural Networks for Pattern Recognition. New York NY: Oxford University Press Inc.