Statistical Health Reasoning System of Power Generator Stator Windings against Water Absorption

2014 2

Statistical Health Reasoning System of Power Generator Stator Windings against Water Absorption

2014 2

2013 12

( )

( )

( )

Abstract

Statistical Health Reasoning System of Power Generator Stator Windings against Water Absorption

Kyung Min Park
Department of Mechanical and Aerospace Engineering
The Graduate School
Seoul National University

The power generator, as one of the most critical components in a power plant, is typically maintained through use of a time- or usage-based strategy. Either strategy could result in a substantial waste of remaining useful life (RUL), high maintenance costs, and/or low plant availability due to excess, untimely, or missed maintenance. Recently, the field of prognostics and health management has offered new general diagnostic and prognostic techniques to precisely assess health conditions and robustly predict the RUL of engineered systems, with the aim of addressing the aforementioned deficiencies. This paper explores a smart health reasoning system that can be used to assess the health condition of power generator stator windings and their levels of water absorption. The system monitors health based on capacitance measurements of the winding insulations. In particular, a new relative health measure, namely the Directional Mahalanobis Distance (DMD), is proposed to quantify the health condition of stator windings. This paper also proposes an empirical health classification rule, based upon the DMD, which factors in maintenance history. The proposed smart health reasoning system is validated using eight years ôfield data from eight generators, each of which contains forty-two windings.

Keywords: Power Generator

Stator Winding

Statistical Correlation

Health Diagnostics

Directional Mahalanobis Distance

Water Absorption

Student Number: 2012-20664

Table of Contents

List of Tables

List of Figures

Nomenclature

$\lfloor \cdot \rfloor$	Round down function (= floor function)
A	Measurement area
C	Capacitance
Cov	Covariance
D	Distance
$\mathrm{E}[\cdot]$	Expected value
ε_0	Electric constant
ε_r	Relative static permittivity (= dielectric constant)
h	Threshold value/distance
k	The number of variables (= data size of group)
$\max[\cdot]$	Maximum value
$\min[\cdot]$	Minimum value
μ	Sample mean
n	Sample size
Q	Charge on conductor
ρ	Correlation coefficient
$\boldsymbol{\Sigma}$	Covariance matrix
σ	Standard deviation

V	Voltage
X	Random variable
$\tilde{X}$	Processed capacitance data

Abbreviations

CEB	Collector End ïBottom bar
CET	Collector End ïTop bar
CM	Condition Monitoring
DMD	Directional Mahalanobis Distance
GE	General Electric
ID	Identification
KEPCO	Korea Electric Power Corporation
MD	Mahalanobis Distance
MTS	Mahalanobis-Taguchi System
PHM	Prognostics and Health Management
RUL	Remaining Useful Life
SDMD	Scaled Directional Mahalanobis Distance
SMD	Scaled Mahalanobis Distance
TEB	Turbine End ïBottom bar
TET	Turbine End ïTop bar
USC	Ultra-Supercritical

Chapter 1. Introduction

1.1 Motivation

Power generators are critical elements of power plants. An unexpected breakdown of a generator can lead to plant shut-down and can result in substantial economic and societal loss. Recently, tremendous technological advancements have been achieved in the development and deployment of an ultra-supercritical (USC) steam generator, shown in Figure 1-1. The USC steam generator operates at an advanced steam temperature of 593°C or above, enabling it to achieve higher energy conversion efficiency, while at the same time reducing fuel consumption and waste emission. However, the large gap between the operation temperature and pressures of the advanced USC generator and those found in conventional subcritical generators leads to far harsher operating conditions in the USC. Thus, the USC has a much higher risk of catastrophic failure. To minimize the losses resulting from potential failures, the reliability of the USC-type power generator must be ensured throughout its life-cycle amidst uncertain operating conditions and manufacturing variability.

Recently, prognostics and health management (PHM) has emerged as a key technology to evaluate the current health condition (health diagnostics) and predict the future degradation behavior (health prognostics) of an engineered system throughout its lifecycle. In general, PHM consists of four basic functions: a health sensing function, a health reasoning function, a health prognostics function, and a health management function. PHM has shown success in lowering system maintenance costs of various engineered systems. Comprehensive exploration of

PHM techniques for power generator windings can enable early anticipation of failure. PHM can be used to develop cost-effective maintenance strategies and to seek opportunities for extending equipment life. Effective health reasoning systems are a crucial step towards a comprehensive exploration of PHM techniques.

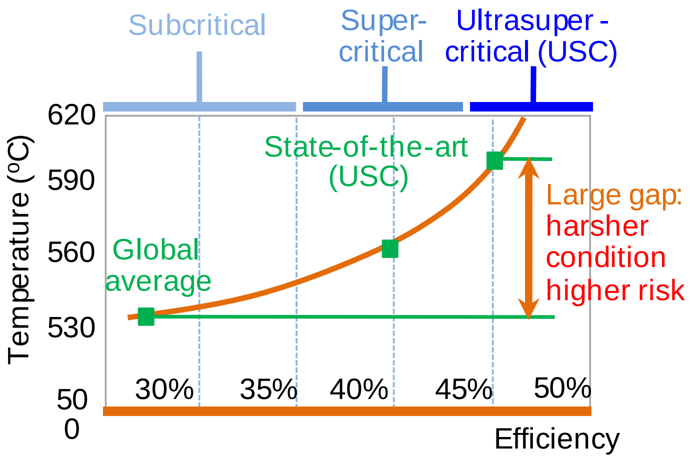


Figure 1-1.Comparison of performance between subcritical, supercritical and ultra-supercritical steam generators

1.2 Overview

This research aims to develop a health reasoning system for power generator stator windings through both physical and statistical analysis. A health reasoning system, also known as the integration of condition monitoring (CM) and health classification, is an algorithm-based system used to diagnose health conditions based on sensory signals and related health measures. Two steps are typically involved: (1) CM to extract relevant system health information through feature extraction techniques, (2) health classification to classify a system ô s health state into diverse health classes using health classification techniques such as artificial intelligence, a support vector machine, decision trees, and mahalanobis distance (MD).

This research proposes a new classification technique that can be applied to stator windings in power generators. The new technique eliminates the limitations found in existing methods and in MD, which is widely used in the PHM field. The proposed definition for health classification is carried out with data from the maintenance history that has been obtained from the field.

1.3 Thesis Layout

This thesis is organized as follows: Chapter 2 reviews existing methods for detecting leaks and water absorption in the insulation of stator windings found in power generators. Chapter 3 presents the sensing function for health monitoring and analysis of health data which can be measured regarding the insulation of stator windings. Chapter 4 discusses feature extraction techniques and introduces a health index for a smart health reasoning system. Chapter 5 presents a new health classification rule which includes consideration of the data size. The validation study for the proposed health index and grade system is given in this chapter as well. Finally, conclusions of the paper are presented in Chapter 6.

Chapter 2. Literature Review

This chapter reviews the existing state of knowledge related to health assessment of power generators ï the topic of this thesis. Topics reviewed include: (1) PHM techniques for engineered systems, including power generators, and (2) existing methods used in the field to guard against leak and water absorption.

2.1 Prognostics and Health Management Techniques used to Support the Health Reasoning Function

PHM is a method that permits assessment of the reliability of a system under its actual application conditions. Extensive research has been conducted in the application of PHM to various engineered systems.

Popular tools used for the health reasoning function include statistical methods (A. H. Christer et *al*. [1], and Y. Zhan et *al*. [2]), artificial intelligence (R. B. Chinnam et *al*. [3] and C. C. Lin et *al*. [4]), support vector machines (J. Yang et *al*. [5] and C. Cortes et *al*. [6]), kernel estimation (N. S. Altman [7]), decision trees (L. Rokach [8]), Mahalanobis distance (R. De Maesschalck et *al*. [9] and G. Taguchi et *al*. [10]), Kalman filter (J. D. Wu et *al*. [11] and S. K. Yang [12]), among others.

These algorithms can be applied to various engineered applications, including bearings (I. E. Alguindigue et *al*. [13]), gearboxes (S. Ebersbach et *al*. [14]), machine tools (D. E. Dimla Sr. [15] and K. F. Martin [16]), transformers (C. Bartoletti et *al*. [17], C. Bengtsson [18], C. Hu et *al*. [19] and C. Booth et *al*. [20]), generators (C. W. Park et *al*. [21], J. Finn et *al*. [22], A. Kheirmand et *al*. [23] and

G. C. Stone et *al*. [24]) and stator insulation used for winding in generators (G. A. Jayantha et *al*. [25] and Z. Jia et *al*. [26]).

2.2 Existing Tests to Detect Leaks or Water Absorption

General Electric (GE), one of the largest manufacturers of generators in the world, provides guidelines for a standard outage test program for periodic overhaul of generators. The schedule typically includes minor overhaul every 30 months and major overhaul every 48-60 months. Figure 2-1 summarizes the GE maintenance program plan. The program consists of four tests; (1) the Vacuum Decay Test, (2) the Pressure Decay Test, (3) the Helium Tracer Gas Test and (4) the Stator Bar Capacitance Mapping Test (J. A. Worden et *al*. [27]).

The first three tests are designed to detect leaks in the stator winding. The first test, the Vacuum Decay Test is a useful tool for determining the integrity of the entire water-cooled stator hydraulic system. The primary advantage of this test is its sensitivity. Decay measurements are made in units of microns. A typical pressure gage cannot detect one micron, which is equivalent to .00002 psi. Because of the high sensitivity of this test, ironically, extremely small leaks at flanges and connections can result in poor test results. The second test, the Pressure Decay Test, has two advantages over the Vacuum Decay Test. It provides a greater pressure differential and applies pressure in the normal direction of the leak flow. These factors may make it easier to find leaks undetectable in the Vacuum Decay Test. During this test, exposed potential leak sites can be tested using a bubble. Drawbacks to pressure testing are its insensitivity to small leaks, and relatively high sensitivity to changes in the environment. The third test, the Helium Tracer Gas Test, is a method of leak detection where the generator is pressurized with a helium gas so that possible leak points can be detected using a helium gas detector. In many cases, leaks that were missed by the Vacuum Decay Test and the Pressure Decay Test are found with the Tracer Gas Test.

Finally, the Stator Bar Capacitance Mapping Test is used to determine the extent of water absorption. This test assumes that good capacitance data provides a normal distribution when plotted; nearly all of the data should fall between -2 and +2 standard deviations from the average. This test uses +3 standard deviations of the capacitance data as a failure threshold.

Korea Electric Power Corporation ï Research Institute (KEPRI) has also developed methods using a capacitance reader (or wet bar detector) for detecting water absorption using statistical tools, including (1) a Normal Probability Plot and (2) a Box Plot (H. S. Kim et *al*. [28]).

The Normal Probability Plot method determines the health classes of winding insulation based on a normal probability plot. An operator can visually identify an outlier or anomaly case by examining the plot. It is assumed that the capacitance data of a healthy winding follows a normal distribution. The Box Plot method is another graphical method, which graphically depicts the health classes of the capacitance data through their 1st and 3rd quartiles. The drawback of both aforementioned methods is that the sensitivity of the winding health classification is relatively low because of improper statistical modeling of the capacitance data and a lack of consideration of data heterogeneity.

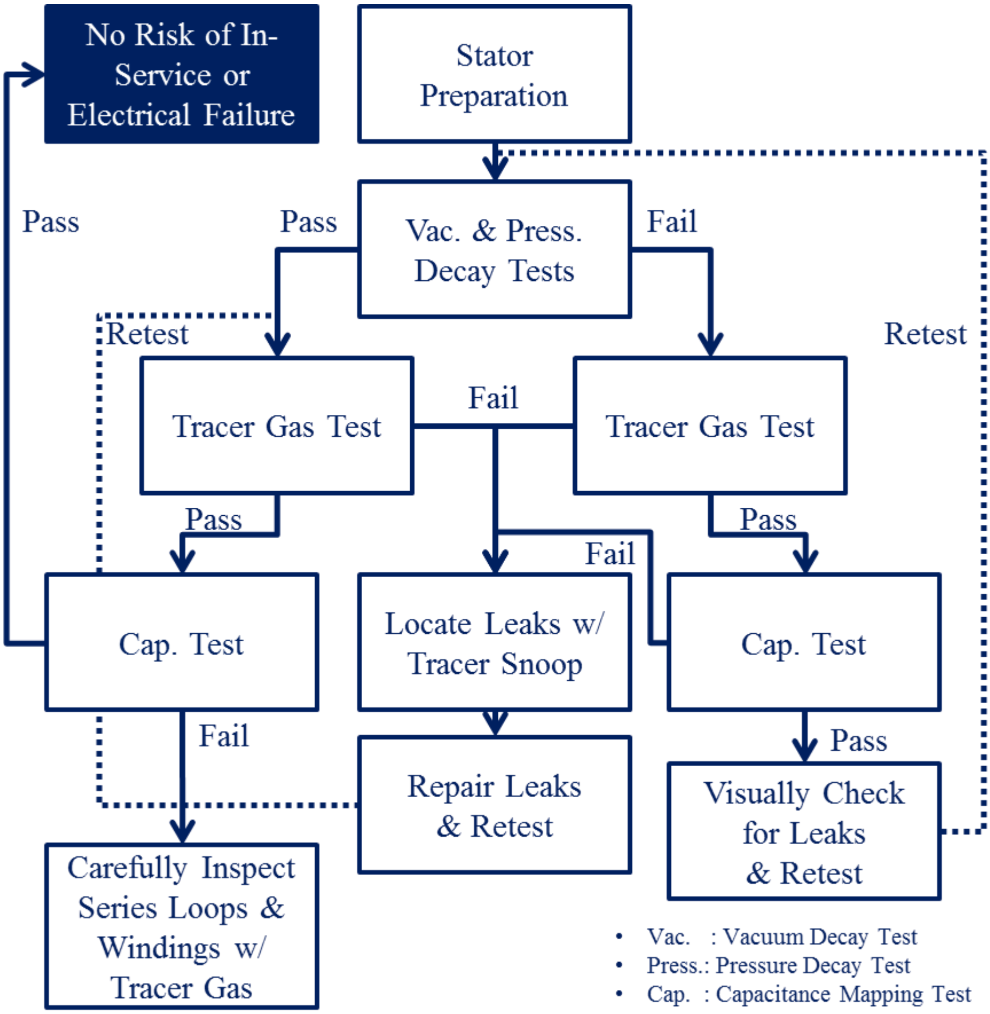


Figure 2-1.Major output leak test plan (GE) [27]

2.3 Summary and Discussion

The aforementioned PHM techniques can be applied to various engineered systems. Most monitoring systems for power generators use electrical signals to detect the faults of the generator. In the case of monitoring the system for water absorption, basic statistical ideas have been used to detect leaks and to detect the water absorbed in insulation based upon capacitance readings. Note that the direct use of capacitance measurements as the health index by the existing methods reviewed previously makes it difficult to easily and precisely infer the health status, especially when the measurements are of high dimensionality, high correlation, and/or high non-linearity. To improve upon the status quo, the work we propose develops a new health index through statistical analysis of multi-dimensional capacitance measurements for effectively determining the health of power generator windings. The ultimate goal of this work is to better prevent sudden failure and to enable a self-sustained generator, as shown in Figure 2-2.

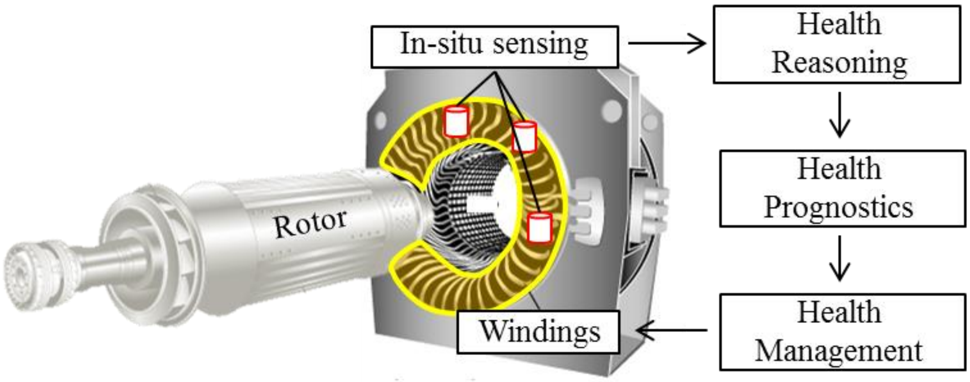


Figure 2-2.Processes involved in a self-sustained power generator

Chapter 3. Description of the Sensing Function and Data Analysis

The objective of the sensing function is to ensure high damage detectability and efficient data management by designing data acquisition logistics. In addition, the health condition of a power generator can be monitored by properly analyzing the capacitance of the winding insulation. This section discusses the fundamentals of capacitance measurements, locations of capacitance measurements, and characteristics of measurement data. This study examines eight power generators (nineteen datasets over eight years) which have the same specifications: (1) a 500 MW output, (2) a 2-path cooling system, and (3) a 60 Hz frequency.

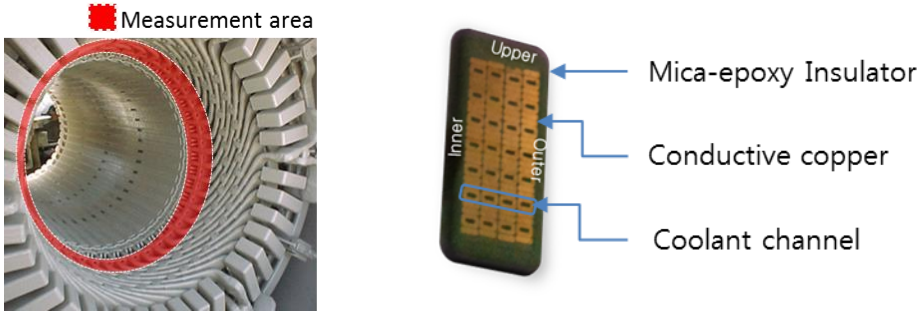
3.1 Fundamentals of Capacitance Measurements

When a power generator is water-cooled, coolant water flows into the water channels of the winding, as shown in Figure 3-1. Leakage into the surrounding insulation can occur due to various operational stresses, such as mechanical vibration, thermal shock, and crevice corrosion (see Figure 3-2). When leakage occurs, the water or moisture remains in the winding insulation. The remaining water degrades the winding insulation, which can cause insulation breakdown and power generator failure, as shown in Figure 3-3. For this reason, electric companies and manufacturing companies, such as KEPRI, GE and Toshiba, assess the health status of the winding insulation in their generators using a water absorption detector

[27-31]. The water absorption detector infers the level of water in the insulation by measuring the capacitance of the insulation. Because the relative static permittivity (or the dielectric constant) of water is higher than that of mica (which is generally used as the insulation material), wet insulation has a higher capacitance, *C*, based upon the following equation (see Figure 3-4 for a schematic representation):

$$C = \frac{Q}{V} = \varepsilon_r \varepsilon_0 \frac{A}{t} \tag{3.1}$$

where *Q* is the charge on each conductor, *V* is the voltage between the plates, *A* and *t* are, respectively, the measurement area and the thickness of the detector, ε_0 is the electric constant ($\varepsilon_0 \approx 8.854$pF-m^{-1}) and ε_r is the relative static permittivity of the material between the plates. Measures of capacitance as health data provide valuable information that can be used to infer the amount of moisture absorption of a stator winding. Health-relevant information about the winding can be extracted from this measured moisture level. It should be noted that various uncertainty factors, such as the measurement location, the ambient humidity, and the winding surface condition propagate uncertainties into the capacitance measurements. These uncertainties must be taken into account in the health reasoning process.



(a) (b)

Figure 3-1.Power generator stator (a) and cross-section view of a winding (b)

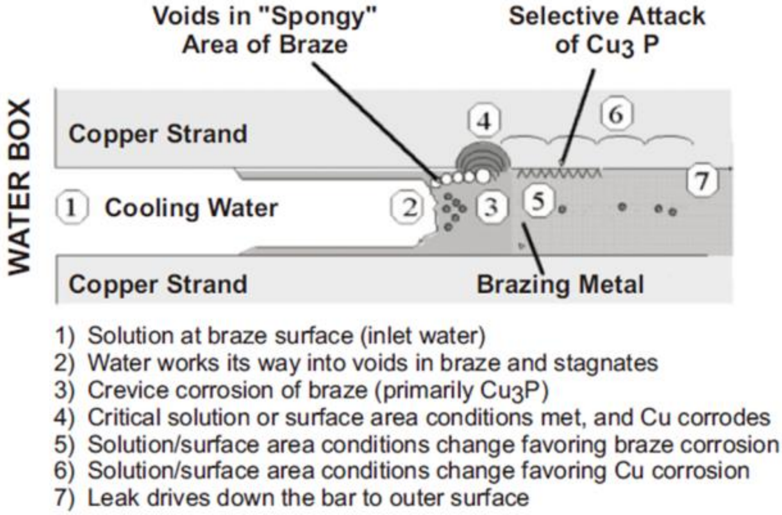
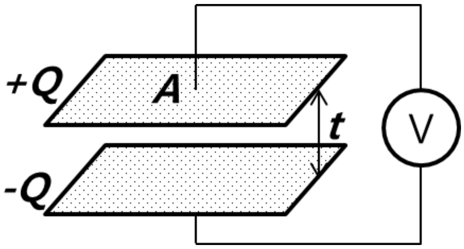


Figure 3-2.Diagram of a crevice corrosion mechanism [27]

Figure 3-3.Failure of a power generator stator

(a)



(b)

Figure 3-4.Capacitance reading using a detector (Model: GEN-SWAD I) (a) and the basic principle of the capacitance detector (b)

3.2 Capacitance Data Acquisition

As mentioned previously, each of the power generators employed in this study has forty-two stator windings and slots used for a water-cooled cooling system. As shown in Figure 3-5, the cooling water flows from the top bar inlet at the turbine end, through the top and bottom bars at the collector end, then back to the bottom bar outlet at the turbine end. At the turbine or collector end, an assembly slot in both the top and bottom bars contains ten measurements points. The ten measurement points are summarized in Table 3-1 and graphically illustrated together with the generator structure diagram in Figure 3-5. Note that since only an extremely small gap exists between the top and bottom bars, the capacitance on the top side of the bottom bar cannot be measured, resulting in only two measurement points for the bottom bar. As shown in Table 3-1, a unique identification (ID) code is assigned to each measurement point based on the location of the probing point.

For example, the ID code ìCET-TOP ò indicates that the measurement point is located on the **Top** side of the **C**ollector **E**nd **T**op bar. The capacitance data were acquired from the ten measurement points for each of the forty-two slots found on each power generator. The capacitance data measured at each measurement point can be modeled as a random variable (X).

Table 3-1.Summary of ten measurement points on a single stator winding

Stator side	**Winding location**	**Measurement point**	**Identification Code**
Collector End (CE)	Top Winding	TOP	CET-TOP (X_1)
		OUT	CET-OUT (X_2)
		IN	CET-IN (X_3)
	Bottom Winding	OUT	CEB-OUT (X_4)
		IN	CEB-IN (X_5)
Turbine End (TE)	Top Winding	TOP	TET-TOP (X_6)
		OUT	TET-OUT (X_7)
		IN	TET-IN (X_8)
	Bottom Winding	OUT	TEB-OUT (X_9)
		IN	TEB-IN (X_{10})

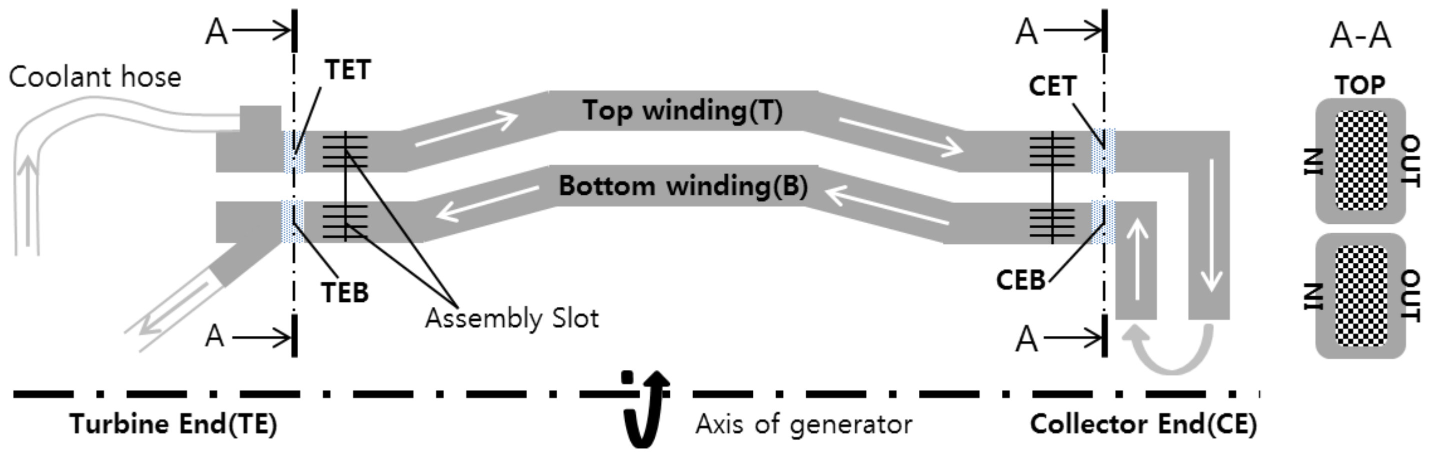


Figure 3-5.Structure diagram of a water-cooled power generator with a 2-path cooling system

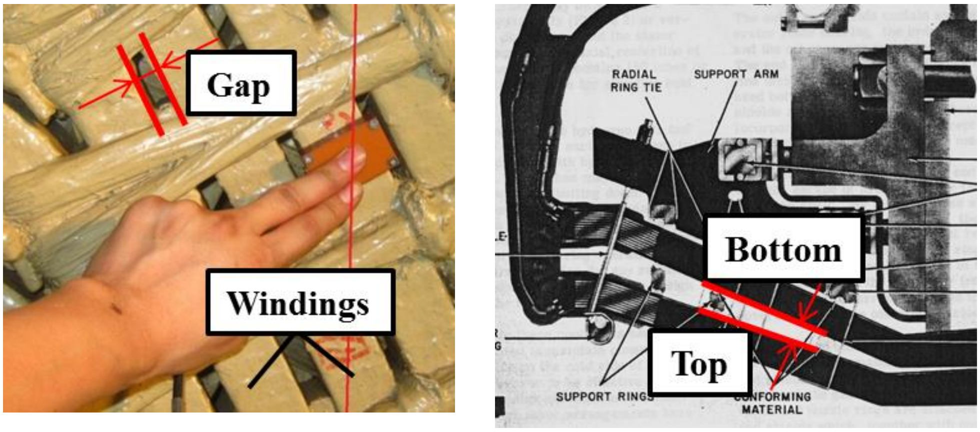
3.3 Statistical Characterization of the Capacitance Data

The capacitance data acquired at physically isolated measurement points can be modeled as statistically independent random variables (i.e. X_1 and X_{10}). Alternatively, the data can be modeled as statistically correlated random variables. For example, a physical gap between two different windings (as shown in Figure 3-6) is one reason why the related random variables might be statistically independent. Moreover, different winding locations (CET, CEB, TET, and TEB) in one winding are also physically distant. This also implies that the related random variables could be statistically independent. On the other hand, water absorption occurs concurrently at adjacent measurement points in the same group, such as CET-TOP (X_1), CET-OUT (X_2), and CET-IN (X_3). Checking statistical dependence between two random variables could confirm whether our intuitive observation is true or not. Before checking statistical correlations, a mean shift was applied to all datasets to take into account the inherent difference in the nominal states of water absorption between generators, as shown in Figure 3-7. After the mean shift, the correlation coefficients later become useful to develop the health reasoning process for a stator winding in a power generator.

In general, the correlation coefficient is used as a measure to imply statistical correlation. The most famous measure of correlation is the Pearson product-moment correlation coefficient. It is a quantitative measure of a linear dependence between two variables. Mathematically, a correlation coefficient can be calculated from the following form:

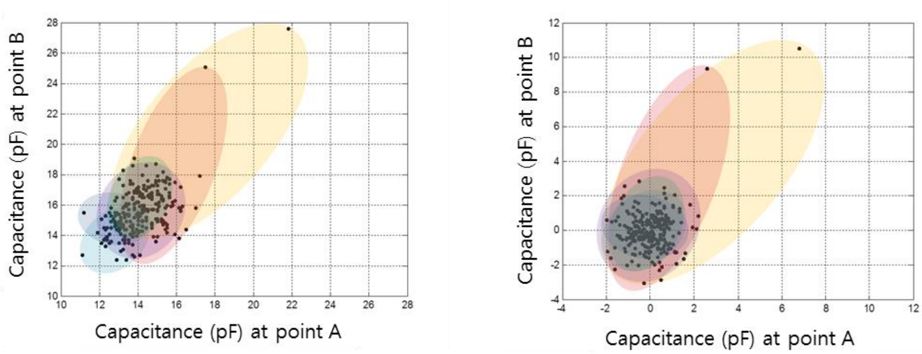
$$\rho_{X_i,X_j} = \frac{\mathrm{Cov}\left(X_i, X_j\right)}{\sigma_{X_i}\sigma_{X_j}} = \frac{\mathrm{E}\left[\left(X_i - \mu_{X_i}\right)\left(X_j - \mu_{X_j}\right)\right]}{\sigma_{X_i}\sigma_{X_j}} \tag{3.2}$$

where X_i and X_j are random variables, Cov($X_{i,}$ X_j) is the covariance between X_i and X_j, ε and ů are the mean and standard deviation of a random variable, respectively, and E[¶] is the expectation of a random variable.



(a) (b)

Figure 3-6.Gaps between two different windings (a) and between top and bottom stator bars (b)



(a) (b)

Figure 3-7.Scatter plots of the data between measurement points before (a) and after (b) mean shift

Table 3-2 summarizes the correlation coefficients for ten random variables, r_{X_i,X_j} for i, j = 1 to 10, in matrix form. The highlighted values in Table 3-2 are the coefficients between the correlated random variables in the same group. One can observe two features from the highlighted values: (1) a statistically positive correlation, and (2) a higher degree of correlation within the same group. These features indicate that the two or three capacitance data from the same group tend to behave (i.e. remain unchanged or grow) with linear dependence. This confirms the intuitive observations about the aforementioned statistical correlation and independence. Appendix A provides pairwise scatter plots between the ten measurement points (or ten random variables), from which two scatter plots are extracted to show the within-group and between-group correlations.

Table 3-2.Correlation coefficient matrix (symmetric) for ten random variables in matrix form

Correlation Matrix		CET TOP (X_1)	CET OUT (X_2)	CET IN (X_3)	CEB OUT (X_4)	CEB IN (X_5)	TET TOP (X_6)	TET OUT (X_7)	TET IN (X_8)	TEB OUT (X_9)	TEB IN (X_{10})
CET	**TOP (X_1)**	1									
	OUT (X_2)	**0.4761**	1								
	IN (X_3)	**0.4194**	**0.5503**	1							
CEB	**OUT (X_4)**	0.0849	0.1572	0.1354	1						
	IN (X_5)	-0.039	0.1686	0.0765	**0.3445**	1					
TET	**TOP (X_6)**	0.3341	0.1553	0.1868	0.0343	-0.052	1				
	OUT (X_7)	0.1972	0.2506	0.2729	0.0879	0.0171	**0.4377**	1			
	IN (X_8)	0.2295	0.1423	0.3296	0.0082	0.0457	**0.4269**	**0.4900**	1		
TEB	**OUT (X_9)**	0.0438	-0.128	-0.097	0.0186	-0.114	0.0887	-0.010	-0.003	1	
	IN (X_{10})	0.0354	-0.040	-0.004	0.0457	0.0870	-0.048	0.1084	0.0215	**0.3385**	1

3.4 Data Grouping

It is important to define a group of capacitance data with homogeneity prior to the data modeling and health reasoning process. Based upon the measurement location and correlation characteristic obtained in Section 3.3, the measurement points with high correlation can be conceived as individual data groups, such as CET, CEB, TET, and TEB. This implies that the entire dataset for ten random variables (or from ten-dimensional measurement points) would be split into four groups with two or three random variables. This data grouping will be used for the health reasoning process in the subsequent section, which defines a health index and models it in a statistical form, as shown in Figure 3-8 and Table 3-3. The data grouping makes the health reasoning process easier through dimensional reduction of the capacitance data.

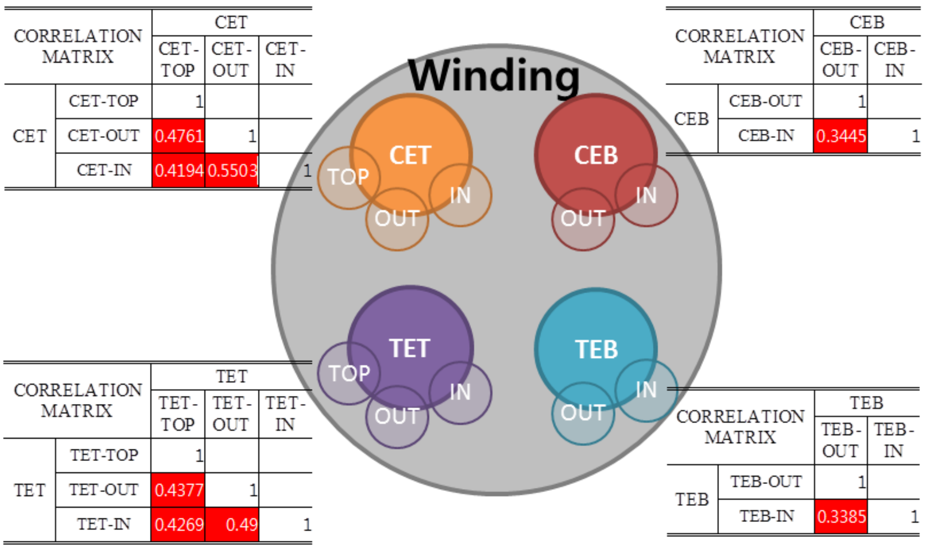


Figure 3-8.Data grouping using a statistical correlation

Table 3-3.List of measured datasets and the information employed in this study

Power Plant	**Generator Number**	**Year for Measurement**	**Winding Number**	**Measured dataset**			
				CE group		**TE group**	
A	1	'05, '07, '09	1	CET (X_1, X_2, X_3)	CEB (X_4, X_5)	TET (X_6, X_7, X_8)	TEB (X_9, X_{10})
			42	CET	CEB	TET	TEB
	2	'06, '09	1	CET	CEB	TET	TEB
			42	CET	CEB	TET	TEB
	3	'06, '10	1	CET	CEB	TET	TEB
			42	CET	CEB	TET	TEB
	4	'06, '07, '08	1	CET	CEB	TET	TEB
			42	CET	CEB	TET	TEB
B	1	'06, '10, '12	1	CET	CEB	TET	TEB
			42	CET	CEB	TET	TEB
	2	'09, '12	1	CET	CEB	TET	TEB
			42	CET	CEB	TET	TEB
C	4	'06, '10	1	CET	CEB	TET	TEB
			42	CET	CEB	TET	TEB
D	6	'09, '11	1	CET	CEB	TET	TEB
			42	CET	CEB	TET	TEB

Chapter 4. Statistical Health Reasoning System[1]

Although the capacitance data are relevant to the health status of the stator winding, its high dimensionality and non-linearity make it difficult to easily and precisely infer the health status. This section proposes a new health index, referred to as the Directional Mahalanobis Distance (DMD).

4.1 Review of Mahalanobis Distance

The Mahalanobis Distance (MD) is a relative health measure that quantifies the deviation of a measured data point from a clustered data center, which is generally a populated mean ($\boldsymbol{\mu}$) of a dataset. The MD degenerates multi-dimension data (**X**) to a one-dimension distance measure while taking into account the statistical correlation between random variables. Mathematically, the MD measure can be expressed as:

$$\mathrm{MD}\left(\mathbf{X}_i\right)=\sqrt{\left(\mathbf{X}_i-\boldsymbol{\mu}\right)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\left(\mathbf{X}_i-\boldsymbol{\mu}\right)} \tag{4.1}$$

where $\mathbf{X}_i=\left(X_{1,i},X_{2,i},\cdots,X_{N,i}\right)^{\mathrm{T}}$ is an N-dimensional capacitance data vector of the i^{th} winding unit which belongs to a group having the mean $\boldsymbol{\mu}=\left(\mu_1,\mu_2,\cdots,\mu_N\right)^{\mathrm{T}}$ and the covariance matrix $\boldsymbol{\Sigma}$. Figure 4-1 plots two-dimensional samples randomly drawn from two random variables with a positive correlation. Essentially, the MD

[1] Sections of this chapter have been submitted as the following journal article: Byeng D. Youn, Kyung Min Park, Hu Chao, Joung Taek Yoon, and Hee Soo Kim, "Statistical Health Reasoning of Water-Cooled Power Generator Stator Windings against Moisture Absorption," *Reliability Engineering and System Safety*, Submitted, 2014.

transforms an ellipsoid in the original random space to a circular shape in the standard Gaussian space, as shown in Figure 4-1. Since the distance (D_1) of the faulty point to the clustered data center is much shorter than that (D_2) of the healthy point, one could have concluded based upon the Euclidean distance that the faulty point is more likely to belong to the cluster. This misleading conclusion is mainly caused by not taking into account the correlation coefficient of the two random variables. Indeed, if we simply divide the distances D_1 and D_2 by the widths of the ellipsoid in the corresponding directions, respectively, we can easily come to the conclusion that the faulty point is much farther away from the clustered center than the healthy point. This can be clearly observed in Figure 4-1.

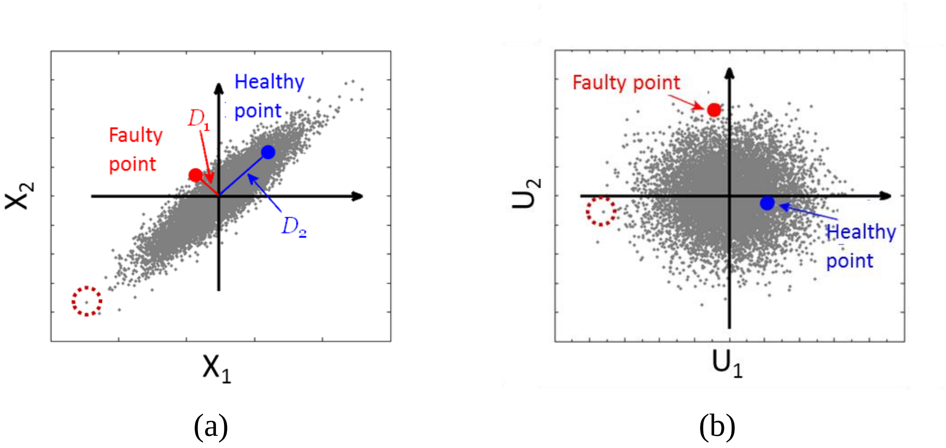


Figure 4-1.Healthy and faulty points located in the original space (a) and the normalized space (b)

As compared to the Euclidean distance, the MD measure possesses a few unique advantages, listed as follows: (1) The MD transforms a high-dimensional dataset that is complicated to handle into a one-dimensional measure capable of easy comprehension and quick computation. (2) The MD is robust to differing scales of the measurements, as the MD values are calculated after normalizing the data. (3) By taking into account the correlation of the dataset, the MD is sensitive to inter-variable changes in multivariate measurements.

4.2 A New Concept of Statistical Distance: Directional Mahalanobis Distance

This subsection introduces a new MD-based distance measure that can be used to imply the health condition of a stator winding in a power generator.

4.2.1 Data Projection

The MD, as a relative health measure, provides very useful information to characterize the health condition of a stator winding in a power generator. According to Equation (3.1), the capacitance values measured from a dry stator winding with a negligible amount of water on the insulation should be smaller than the mean value of the measurement population. Previous studies [27, 28] also reported that measured values smaller than the population mean should be treated as if they have no relation to the insulation water absorption. However, the MD, as a scalar distance measure, is a direction-independent health measure in the random capacitance space, as shown in Figure 4-2. In other words, two capacitance measurements with the same MD value but in two opposite directions are treated equally, although they most likely imply the different levels of water absorption.

Let us take the dashed circle datum in Figure 4-1 as an example. In this case, the dashed circle datum is a healthy point since this datum falls into the lower tails of the marginal distributions of the random capacitance variables. However, the MD declares this datum to be in the failure category simply because it is out of the data cluster. For this very reason it is necessary to refine the measure so that is better suited for this application. In order to rebuild the measure, this study employs a projection process which first identifies absolutely healthy variables(s) (e.g. a capacitance value less than its populated mean, say $X_i < \mu_i$) and then projects it onto the corresponding mean value(s), e.g. (X_i, μ_j) and (μ_i, X_j) shown in Figure 4-2. Through this projection process, the absolutely healthy data will be ignored in the subsequent transformation. The data projection underscores the consideration of the direction in the health reasoning process of the measurement data. This leads to the unique capability of the proposed health index to makes use of the distance and degradation direction as a health measure.

After the data projection, the capacitance data $\tilde{X}_{n,i}\left(n=1,\cdots,N\right)$, can be processed as:

$$\tilde{X}_{n,i} = \begin{cases} X_{n,i}, & \text{if } X_{n,i} > \mu_n \\ \mu_n, & \text{otherwise} \end{cases} \tag{4.2}$$

where $X_{n,i}$ denotes the raw capacitance data at the n^{th} measurement location of the i^{th} winding unit, μ_n is the mean of the capacitance data at the n^{th} measurement location, and $\tilde{X}_{n,i}$ denotes the processed capacitance data. The mean and variance of the dataset must be obtained before the data projection because they are physically meaningful in the original space.

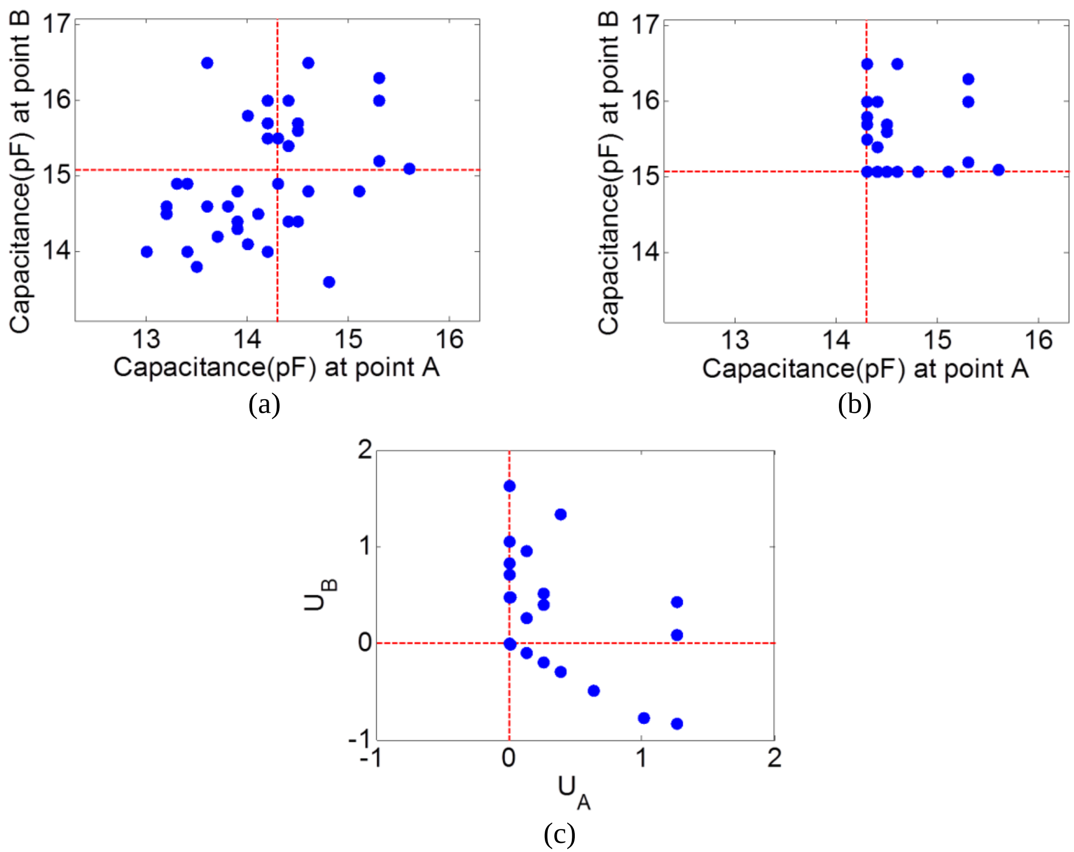


Figure 4-2.Scatter plots before data projection (a), after projection (b), and after transformation (c)

4.2.2 Transformation

The proposed index, namely the Directional Mahalanobis Distance (DMD), assesses the MD along the degradation direction of a stator winding insulation after data projection. Mathematically, the DMD shares a similar formula with the MD, except for consideration of the data projection. It is expressed as:

$$\mathrm{DMD}\left(\tilde{\mathbf{X}}_i\right) = \sqrt{\left(\tilde{\mathbf{X}}_i - \boldsymbol{\varepsilon}\right)^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \left(\tilde{\mathbf{X}}_i - \boldsymbol{\varepsilon}\right)} \tag{4.3}$$

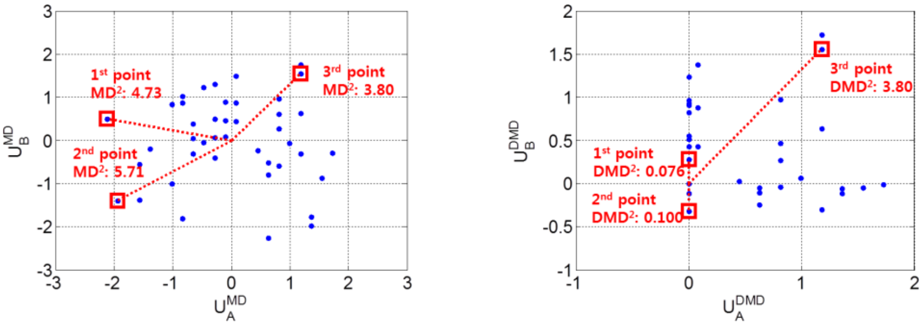
where $\tilde{\mathbf{X}}_i = \left(\tilde{X}_{1,i}, \tilde{X}_{2,i}, \cdots, \tilde{X}_{N,i}\right)^{\mathrm{T}}$ is an N-dimensional vector of the capacitance data from the i^{th} winding unit after the data projection, $\boldsymbol{\varepsilon} = \left(\mu_1, \mu_2, \cdots, \mu_N\right)^{\mathrm{T}}$ and $\boldsymbol{\Sigma}$ is the mean vector and covariance matrix of the reference dataset before the projection. Figure 4-2(c) shows the scatter plot of the DMD dataset after projection and transformation; the difference between the MD and DMD can be clearly observed in this figure.

One question remains: what is the proper sequence of data projection and transformation? We found that the former should be done prior to the latter. The justification for this sequence is the fact that the comparison between the absolutely healthy data and the mean value of each random variable is physically meaningful and valid only in the original space, not in the transformed space.

4.3 Comparison of Performance of Mahalanobis Distance (MD) and Directional Mahalanobis Distance (DMD)

The distances in this paper are squared in order to place progressively greater weight on objects that are farther apart. In general, squared distance is frequently used in instances where only distances need to be compared.

Figure 4-3 shows the scatter plot of MD and DMD with three highlighted data points, two of which represent a healthy state and the other a faulty state. In the case of MD, the 1st and 2nd data points (the two "healthy" points) are located in the 2nd (top left) and 3rd (bottom left) quadrants of the two-dimensional space composed of two Euclidean distances. The 3rd data point, the "faulty" point, is located in the 1st (top right) quadrant. The MD values of the 1st, 2nd and 3rd points are, respectively, 4.74, 5.71, and 3.80. In the case of DMD, the squared distances of the three points from the origin are, respectively, 0.076, 0.100, and 3.80. Without the projection before transformation, MD incorrectly treats the "healthy" data points as "faulty" while, with the projection before the transformation, the proposed index, DMD, correctly identifies these two points as "healthy" (with relatively small distance values). Therefore, the proposed DMD achieves better performance than the MD by accounting for the degradation direction in the health reasoning process of the capacitance data.



(a) (b)

Figure 4-3.Three cases to compare the performance of MD (a) and that of DMD (b)

Performance evaluation of these two indices requires an evaluation metric that assesses the effectiveness of a health index in quantifying the health condition of a generator winding. The evaluation metric considered here employs a score function with the health index value and true health condition of a generator winding as the inputs and a normalized score metric (ranging between 0 to 100) as the output. Mathematically, the proposed score function can be expressed as:

$$\mathrm{SF} = 100 \times \frac{\sum_{j=1}^{N} x_j y_j - \underbrace{\left(\sum_{j=1}^{l} x_j - \sum_{j=l+1}^{N} x_j\right)}_{W}}{\underbrace{\left(\sum_{j=N-l+1}^{N} x_j - \sum_{j=1}^{N-l} x_j\right)}_{B} - \underbrace{\left(\sum_{j=1}^{l} x_j - \sum_{j=l+1}^{N} x_j\right)}_{W}} \tag{4.4}$$

where x_j denotes the health index value of the j^{th} winding unit, y_j denotes the maintenance index of the j^{th} winding based on the actual repair history ($y_j = 1$ if the unit was maintained and $y_j = -1$ otherwise), N denotes the number of winding units, l denotes the number of winding units with maintenance histories, and B and W denote the best and worst score metric values, respectively. Figure 4-4 illustrates various combinations of the health index ranking, the maintenance history, and the score metric ranking of these combinations. The best and worst scenarios (represented respectively by B and W in Equation(4.4)) are depicted by the leftmost and rightmost plots, respectively.

Table 4-1 summarizes the distances and scores from the assessment results for health condition of the winding using MD and DMD, respectively. As one can see in Table 4-1, both MD and DMD can easily find windings that have been maintained; the capacitance of maintained insulation is relatively larger than the capacitance seen in the unmaintained samples. However, MD cannot find all the maintained windings in the top ten health indices in the CET group. In addition, some unmaintained windings ôhealth indices have higher values than those of maintained windings (e.g. 8th, 9th, and 10th health indices in TET group).

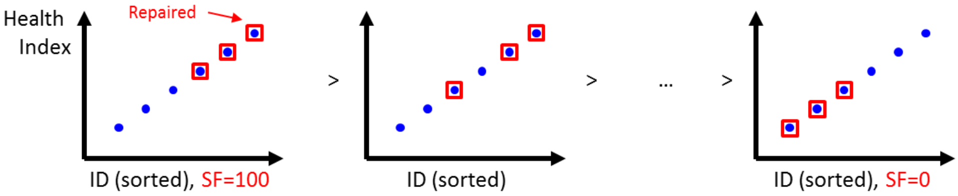


Figure 4-4.Combination of health index ranking and maintenance history and corresponding score metric ranking

Table 4-1.Distances and scores from the results examining health condition of windings using MD and DMD

Rank	**CET group**		**CEB group**		**TET group**		**TEB group**	
	MD^2	**DMD^2**	**MD^2**	**DMD^2**	**MD^2**	**DMD^2**	**MD^2**	**DMD^2**
1	33.64	33.64	21.14	21.14	20.37	20.37	11.65	11.65
2	28.41	28.41	20.63	20.63	18.93	18.93	9.97	9.01
3	21.96	21.96	17.96	14.36	17.93	17.93	9.64	8.88
4	18.60	18.60	15.87	14.00	17.67	17.67	9.18	8.08
5	18.06	18.06	14.36	13.61	15.88	15.88	9.01	7.87
6	17.58	17.58	14.27	12.68	15.75	14.65	8.88	7.34
7	16.42	16.42	12.68	8.18	14.65	13.57	8.69	7.04
8	16.39	16.39	11.05	7.86	14.11	13.15	8.08	6.86
9	13.36	12.25	10.85	6.75	13.98	9.33	7.87	6.45
10	12.76	9.90	10.78	6.70	13.55	8.91	7.74	6.23
11	12.55	9.51	9.67	6.27	13.15	8.39	7.68	6.18
12	12.52	9.41	9.51	6.08	11.79	8.09	7.66	5.96
13	11.47	9.36	9.37	5.30	11.56	6.80	7.50	5.74
14	11.46	8.57	8.98	5.27	11.11	6.68	7.48	5.49
15	11.43	8.03	8.76	5.19	10.63	6.62	7.34	5.48
21	10.56	6.67	7.82	4.38	9.33		6.86	4.60
27	9.51	6.25	6.77	4.06	8.52		6.22	4.39
29	9.36	5.60	6.52	3.94	8.39		6.17	4.32
798	0.01	0.00	0.00	0.00	0.02	0.00	0.00	0.00
Score	**96.31**	**98.30**	**98.47**	**100**	**94.52**	**96.80**	**91.26**	**95.49**

Chapter 5. Health Classification

This section is designed to construct a health grade system based upon the proposed health index, DMD, and field maintenance history. The historical maintenance data obtained from the operators of the power generators are presented in Section 5.1. An empirical health grade system is suggested in Section 5.3 with consideration of the scale that is discussed in Section 5.2.

5.1 Maintenance History Related to Water Absorption

In this section, the historical maintenance data related to water absorption is presented, including data from faulty windings. Field experts collected the maintenance records of generator windings and identified their health conditions using the Stator Bar Capacitance Mapping test method developed by GE. The maintenance records can be classified into four groups (CET, CEB, TET, and TEB). Table 5-1 summarizes the maintenance history related to water absorption over eight years. In this table, two data points (28.413 and 33.645) were obtained from the faulty winding and the others were measured from water-absorbed windings. These maintenance records provide the physical basis for determining an appropriate failure threshold, especially the data obtained from the faulty windings.

Table 5-1.Maintenance history related to water absorption

Group	Power Plant	Generator Number	Winding Number	Year	DMD^2	Condition
CET	A	1	20	05	18.598	Absorbed
				07	16.415	
				09	16.394	
		2	03	06	18.060	Absorbed
				09	21.963	
			40	09	9.511	Absorbed
		4	23	07	28.413	**Faulty**
				08	33.645	
	B	1	18	10	9.355	Absorbed
				12	17.580	
CEB	A	2	31	06	20.625	Absorbed
				09	21.141	
		4	20	07	14.359	Absorbed
				08	12.677	
	C	4	11	06	13.613	Absorbed
				10	14.001	
TET	A	1	11	07	9.334	Absorbed
				09	13.145	
			20	05	17.665	Absorbed
				07	17.929	
				09	18.932	
		2	40	06	20.368	Absorbed
				09	14.653	
	B	1	18	10	15.876	Absorbed
TEB	A	3	23	10	8.080	Absorbed

5.2 Review of Scaled Mahalanobis Distance

The aim of this section is to introduce an improvement to the concept of MD: Scaled Mahalanobis Distance (SMD) in the Mahalanobis-Taguchi System (MTS). As aforementioned, this study deals with the two groups which have different data size (i.e. Top group and Bottom group). Since MD and DMD do not consider the data size of each group, it is not suitable to compare two groups directly without any conversion. SMD is a way to solve this problem.

It has been shown that MD follows a χ^2-distribution with k degrees of freedom, when the sample size, n, is large and all characteristics follow the normal distribution (R. A. Johnson et *al*. [32]). A χ^2-distribution with k degrees of freedom has a mean equal to k. Hence, G. Taguchi et *al*. [10] proposed a new idea for MD known as SMD. This group suggested that MD should be scaled by dividing by the number of variables, k. Thus, the equation for calculating SMD in the MTS becomes:

$$\mathrm{SMD}^2 = \frac{1}{k}\mathrm{MD}^2 = \frac{1}{k}\left(\mathbf{X}_i - \boldsymbol{\mu}\right)^{\mathrm{T}} \boldsymbol{\Sigma}^{-1}\left(\mathbf{X}_i - \boldsymbol{\mu}\right) \tag{5.1}$$

where k denotes the number of variables or the data size of each group, and $\mathbf{X}_i$ is a capacitance data vector of the i^{th} winding unit, which belongs to a group having the mean vector, $\boldsymbol{\mu}$, and the covariance matrix, $\boldsymbol{\Sigma}$.

Since the expected value of MD^2 is equal to k, the expected value of SMD^2 becomes:

$$\mathrm{E}\left[\mathrm{SMD}^2\right] = \mathrm{E}\left[\frac{1}{k}\mathrm{MD}^2\right] = \frac{1}{k}\mathrm{E}\left[\mathrm{MD}^2\right] = \frac{1}{k}(k) = 1 \tag{5.2}$$

where E[·] is a function of the expectation. This scaling process thus allows direct comparison between the top and bottom groups. The scaled distance offers an advantage that it can be applicable to any number of variables.

For the purposes of this paper, it is important to make sure whether or not the proposed idea of the SMD works for the capacitance data measured from the stator windings found in generators. Table 5-2 summarizes the properties of each group's data; top and bottom.

Table 5-2.The properties of top and bottom groups' data

Properties	Top Group	Bottom Group
The number of variables (or, the data size of group) (k)	3	2
The number of samples(n)	1,680	1,680
Expected value (μ)	2.88	1.96
Expected value of SMD^2 (μ/k)	0.96	0.98

As shown in Table 5-2, the expected value of the top group (CET and TET) is close to the number of variables in the top groups. Likewise, the expected value of the bottom group (CEB and TEB) approximately equals to the number of variables in the bottom groups. Thus, it can be concluded that the scaled distance measures can be uniformly used regardless of the group. The scaling idea can also be applied to Scaled Directional Mahalanobis Distance (SDMD), just like SMD.

5.3 Health Grade System

This section aims to define a health grade system in which the windings' health states are classified into diverse classes according to the health-relevant distance measure, SDMD.

Based upon the maintenance strategies for the stator windings and the opinions of field experts, three health classes are proposed: (1) a faulty condition (or, water-absorbed), (2) a warning condition (or, close to water absorption), and (3) a healthy condition (or, not water-absorbed). The failure listed in Table 5-1 which was caused by water absorption resulted in two meaningful data. These data can define a failure threshold for the distance measure, *h*, expressed as:

$$h = \left\lfloor k \cdot \mathrm{E}\left[\min\left[\mathrm{SDMD}^2\left(\tilde{\mathbf{X}}_{\mathbf{faulty}}\right)\right], \max\left[\mathrm{SDMD}^2\left(\tilde{\mathbf{X}}_{\mathbf{warning}}\right)\right]\right]\right\rfloor_{5} \quad (5.3)$$

where $\lfloor \cdot \rfloor$ denotes a round down function, E[·] denotes the expected value, k is the data size of each group (e.g. In case of top group, k = 3) and $\tilde{\mathbf{X}}_{\mathbf{faulty}}$ and $\tilde{\mathbf{X}}_{\mathbf{warning}}$ are capacitance data vectors obtained from the faulty or water-absorbed windings on each group, respectively. Sixty percent of the threshold value (0.6 h) is defined as a boundary line between the warning condition and the healthy condition, based upon field experts' experience and historic information on inspection and maintenance for stator windings.

Since the bottom group does not have any faulty winding in the maintenance history, the failure threshold of bottom group cannot be calculated from the equation 5.3. Thus the failure threshold of top group, h_{top}, is applied to define the failure threshold for the bottom bar, h_{bottom}, which can be defined by the following expression:

$$h_{top} : h_{bottom} = 3 : 2 \quad (5.4)$$

Finally, Table 5-3 summarizes the definition of the three health grades and suggested maintenance actions.

Table 5-3.Definition of health grades and related maintenance actions

Health Grade	**Range**		**Suggested Maintenance Actions**
	Top	**Bottom**	
Faulty	$\mathrm{DMD}^2 \geq 25$	$\mathrm{DMD}^2 \geq 16.7$	Immediate replacement
Warning	$15 < \mathrm{DMD}^2 < 25$	$10 < \mathrm{DMD}^2 < 16.7$	Frequent inspection
Healthy	$\mathrm{DMD}^2 < 15$	$\mathrm{DMD}^2 < 10$	No immediate maintenance required

5.4 Validation Study

In this section, the feasibility of the proposed SDMD-based health grade system is verified by comparison with the maintenance history of generator stator windings. Figure 5-1 shows the scatter plot of DMD with operating time. Let us examine the points highlighted with circles in the figure. These circles label the data obtained from the faulty winding or water absorbed windings. Each class contains several data points highlighted with circles, as shown in Table 5-4.

Table 5-4.Summary of the number of the data and the circled data in each grade

Health grade	The number of the data (A)	The number of the circled data (B)	B/A (%)
Faulty	4	4	100
Warning	15	15	100
Healthy	3,173	6	0.19

Most of the circled data points belong to either “faulty” or “warning” class. This indicates that the proposed health grade system properly defines the health condition of the generator stator windings against water absorption.

In the faulty class, there are two cases (Case 1 and 2 in Figure 5-1) which can be split into four data points, although only one failure case was actually recorded in the maintenance history. In order to make sure if the failure threshold was correct, we looked at the maintenance history in detail.

The data points in the first case are obtained from the failed winding which was burned in 2008. According to the proposed index, DMD^2, a substantial increase of the index was found from 2006 to 2007. The index of this winding is equal to 3.05 in 2006 and 28.41 in 2007, respectively. This increase would have suggested by

DMD^2 that a preventative maintenance action be taken the year (2007) before the failure. This is a critically important observation in order to build the health grade system correctly. Thus, there is no problem that the data points observed from this winding belong to the faulty class.

On the other hand, the second case contains two data points which are measured from non-faulty winding. It seems that the winding studied here should be contained within the warning or healthy class. This is because of stator windings ô electrical characteristics. A voltage is induced in the stator winding when the rotor is rotated. Typically, large synchronous generators are designed for a terminal voltage of several thousand volts. According to the opinion of experts in the power generation field, several windings per stator continuously retain zero-volts. This implies that these windings, like the second case ôs winding, may not have failed even though the water is absorbed enough to otherwise indicate a problem. In summary, it can be concluded based upon the above observation that replacement should be carried out on this winding in spite of its good external appearance. Actually, this winding was replaced in 2011.

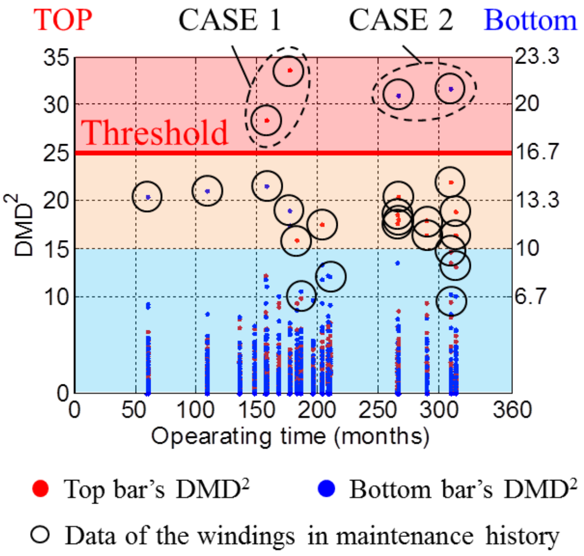


Figure 5-1.Scatter plot of DMD with operating time

Chapter 6. Conclusion

This paper explores a new health reasoning system to assess the health condition of stator windings in power generators. The proposed system extracts health-relevant features from the capacitance data in a statistical manner, to assess the health condition of power generator stator winding, to classify the health classes into three groups (healthy, warning, and faulty states), and to develop health grade system for condition-based maintenance. Correlation analysis of measured capacitance data is used to help understand the statistical features of the data and to divide the variables into four groups (CET, CEB, TET, and TEB) per winding. This paper proposes a statistical health measure Directional Mahalanobis Distance (DMD). DMD incorporates the correlation between variables and provides the degree of health condition considering the health degradation direction. Due to the unique capability of DMD to make use of the distance and degradation direction as a health measure, it can also be applicable to a health grade system designed to monitor a building. The health grade system outlined in this paper was developed with guidance from field maintenance records. Moreover, it takes into account the data size of each group with a scaling factor. This study employed the datasets from eight generators over eight years to validate the proposed health reasoning system. The proposed system can be generally applicable to health degradation trend analysis of engineered systems. In order to accomplish the condition-based maintenance system, health prognostics using machine learning techniques must be further studied.

Bibliography

[1] CHRISTER, A. H.; WALLER, W. M. Delay time models of industrial inspection maintenance problems. Journal of the Operational Research Society, 1984, 401-406.

[2] ZHAN, Yimin; MAKIS, Viliam; JARDINE, Andrew KS. Adaptive state detection of gearboxes under varying load conditions based on parametric modelling.Mechanical Systems and Signal Processing, 2006, 20.1: 188-221.

[3] CHINNAM, Ratna Babu; BARUAH, Pundarikaksha. A neuro-fuzzy approach for estimating mean residual life in condition-based maintenance systems.International Journal of Materials and Product Technology, 2004, 20.1: 166-179.

[4] LIN, Chang-Ching; TSENG, Hsien-Yu. A neural network application for reliability modelling and condition-based predictive maintenance. The International Journal of Advanced Manufacturing Technology, 2005, 25.1-2: 174-179.

[5] YANG, Junyan; ZHANG, Youyun; ZHU, Yongsheng. Intelligent fault diagnosis of rolling element bearing based on SVMs and fractal dimension. Mechanical Systems and Signal Processing, 2007, 21.5: 2012-

2024.

[6] CORTES, Corinna; VAPNIK, Vladimir. Support-vector networks. Machine learning, 1995, 20.3: 273-297.

[7] ALTMAN, N. S. An introduction to kernel and nearest-neighbor nonparametric regression. The American Statistician, 1992, 46.3: 175-185.

[8] ROKACH, Lior. Data mining with decision trees: theory and applications. World Scientific, 2007.

[9] DE MAESSCHALCK, Roy; JOUAN-RIMBAUD, Delphine; MASSART, Désiré L. The mahalanobis distance. Chemometrics and Intelligent Laboratory Systems, 2000, 50.1: 1-18.

[10] TAGUCHI, Genichi; WU, Yuin; CHODHURY, Subir. Mahalanobis-Taguchi System. McGraw-Hill Professional, 2000.

[11] WU, Jian-Da; HUANG, Chin-Wei; HUANG, Rongwen. An application of a recursive kalman filtering algorithm in rotating machinery fault diagnosis. NDT & E International, 2004, 37.5: 411-419.

[12] YANG, S. K. An experiment of state estimation for predictive maintenance using Kalman filter on a DC motor. Reliability engineering & system safety, 2002, 75.1: 103-111.

[13] ALGUINDIGUE, Isreal E.; LOSKIEWICZ-BUCZAK, Anna; UHRIG, Robert E. Monitoring and diagnosis of rolling element bearings using artificial neural networks. Industrial Electronics, IEEE Transactions on, 1993, 40.2: 209-217.

[14] EBERSBACH, S.; PENG, Z.; KESSISSOGLOU, N. J. The investigation of the condition and faults of a spur gearbox using vibration and wear debris analysis techniques. Wear, 2006, 260.1: 16-24.

[15] SNR, Dimla; DIMLA, E. Sensor signals for tool-wear monitoring in metal cutting operations ð a review of methods. International Journal of Machine Tools and Manufacture, 2000, 40.8: 1073-1098.

[16] MARTIN, K. F. A review by discussion of condition monitoring and fault diagnosis in machine tools. International Journal of Machine Tools and Manufacture, 1994, 34.4: 527-551.

[17] BARTOLETTI, Cipriano, et al. Vibro-acoustic techniques to diagnose power transformers. Power Delivery, IEEE Transactions on, 2004, 19.1: 221-229.

[18] BENGTSSON, C. Status and trends in transformer monitoring. Power Delivery, IEEE Transactions on, 1996, 11.3: 1379-1384.

[19] HU, Chao, et al. Copula-based statistical health grade system against mechanical faults of power transformers. Power Delivery, IEEE Transactions on, 2012, 27.4: 1809-1819.

[20] BOOTH, Campbell; MCDONALD, Jim R. The use of artificial neural networks for condition monitoring of electrical power transformers. Neurocomputing, 1998, 23.1: 97-109.

[21] PARK, Chul-Won, et al. Generator fault detection technique using detailed coefficients ratio by daubechies wavelet transform. In: Power & Energy Society General Meeting, 2009. PES'09. IEEE. IEEE, 2009. p. 1-7.

[22] FINN, Joshua; WAGNER, John; BASSILY, Hany. Monitoring strategies for a combined cycle electric power generator. Applied Energy, 2010, 87.8: 2621-2627.

[23] KHEIRMAND, Akbar; LEIJON, Mats; GUBANSKI, S. M. Advances in online monitoring and localization of partial discharges in large rotating machines.Energy Conversion, IEEE Transactions on, 2004, 19.1: 53-59.

[24] STONE, G. C., et al. Impact of slot discharges and vibration sparking on stator winding life in large generators. Electrical Insulation Magazine, IEEE, 2008, 24.5: 14-21.

[25] JAYANTHA, G. A.; FERNANDO, M. A. R. M.; EKANAYAKE, C. M.

B. Monitoring of moisture on stator winding insulation during drying process. In:Industrial and Information Systems (ICIIS), 2009 International Conference on. IEEE, 2009. p. 483-488.

[26] JIA, Zhidong, et al. Evaluation of the degradation of generator stator ground wall insulation under multistresses aging. Energy Conversion, IEEE Transactions on, 2008, 23.2: 474-483.

[27] WORDEN, Joseph A.; MUNDULAS, Jorge M. Understanding, Diagnosing, and Repairing Leaks in Water-Cooled Generator Stator Windings. GE Reference Library, GER-3751A, 2003.

[28] KIM, Hee Soo; BAE, Yong Chae; KEE, Chang Doo. Wet bar detection by using water absorption detector. Journal of Mechanical Science and Technology, 2008, 22.6: 1163-1173.

[29] KIM, Hee-Dong. Analysis of Leak and Water Absorption Test Results for Water-Cooled Generator Stator Windings. Journal of Electrical Engineering & Technology, 2012, 7.2: 230-235.

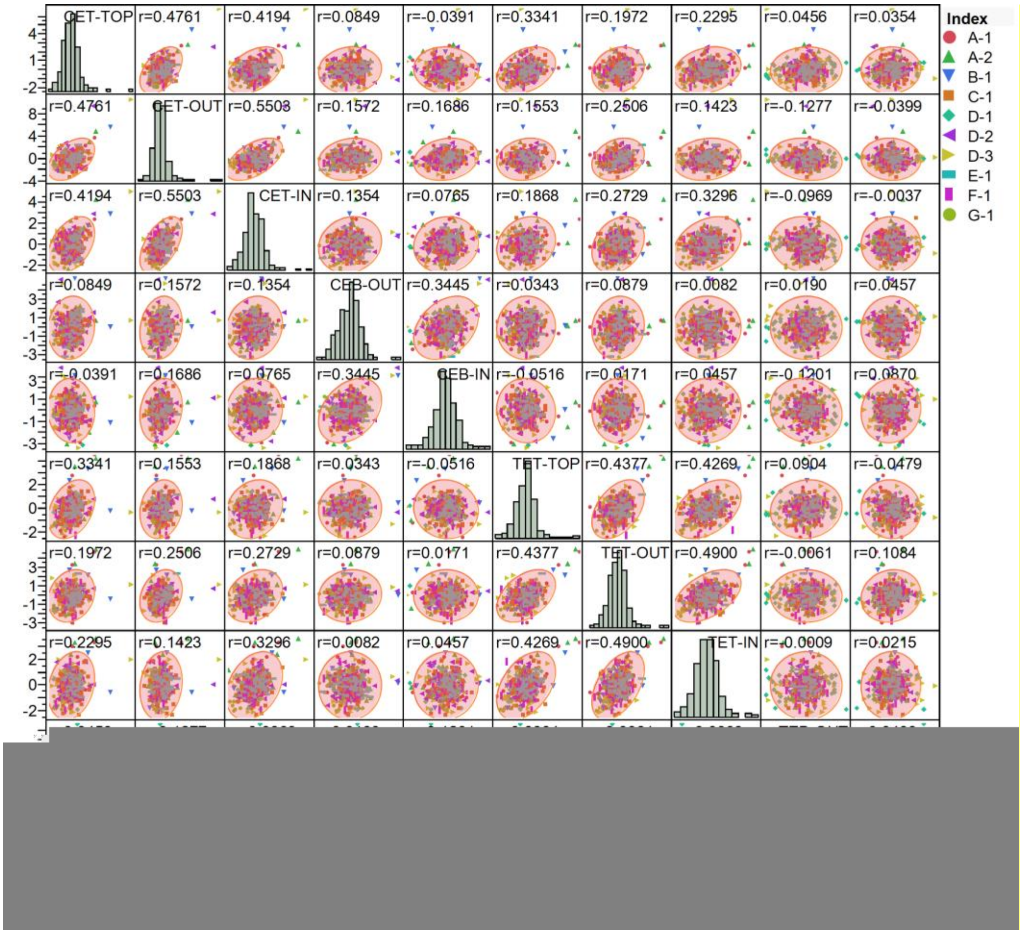
[30] KIM, Hee-Dong. Study on Reliability of Water Absorption Diagnosis through Precise Water Absorption Test. Journal of Electrical Engineering & Technology, 2012, 7.5: 772-777.

[31] INOUE, Y., et al. Technology for detecting wet bars in water-cooled

stator windings of turbine generators. In: Electric Machines and Drives Conference, 2003. IEMDC'03. IEEE International. IEEE, 2003. p. 1337-1343.

[32] HRDLE, Wolfgang; SIMAR, Léopold. Applied multivariate statistical analysis. Springer, 2007.

APPENDIX A.



Appendix A. Pairwise scatter plots between the ten measurements

.

.

/

.

.

Directional Mahalanob is Distance (DMD)

.

DMD

.

/ 8 8

.

:

: 2012 - 20664

^ **(--)**; [ ]

.

. (,

)

09 ,

. (ñ)

? ò

. 2 ,

. ó . ô

2

.

. ó ô

.

, .

.

12 2

.

.

2

. ó ô

.

.

. .

, ,

.

.

.

.

. , , ,

,

/

. .

.

(University of Florida),

(University of Maryland)

.

.

(The best is yet to come.).

.

. **See your tomorrow!**

.

201 3 ( )

(KETEP)

. (: 2012101010001C)

공학석사학위논문

발전기 고정자 권선 흡습에 대한 통계적 추론 시스템 개발

Statistical Health Reasoning System of Power Generator Stator Windings against Water Absorption

2014 년 2 월

서울대학교 대학원

기계항공공학부

박 경 민

발전기 고정자 권선 흡습에 대한 통계적 추론 시스템 개발

Statistical Health Reasoning System of Power Generator Stator Windings against Water Absorption

지도교수 윤 병 동

이 논문을 공학석사 학위논문으로 제출함

2014 년 2 월

서울대학교 대학원

기계항공공학부

박 경 민

박경민의 공학석사 학위논문을 인준함

2013 년 12 월

위 원 장	**조 맹 효**	**(인)**
부위원장	**윤 병 동**	**(인)**
위 원	**김 남 호**	**(인)**

Abstract

Statistical Health Reasoning System of Power Generator Stator Windings against Water Absorption

Kyung Min Park
Department of Mechanical and Aerospace Engineering
The Graduate School
Seoul National University

The power generator, as one of the most critical components in a power plant, is typically maintained through use of a time- or usage-based strategy. Either strategy could result in a substantial waste of remaining useful life (RUL), high maintenance costs, and/or low plant availability due to excess, untimely, or missed maintenance. Recently, the field of prognostics and health management has offered new general diagnostic and prognostic techniques to precisely assess health conditions and robustly predict the RUL of engineered systems, with the aim of addressing the aforementioned deficiencies. This paper explores a smart health reasoning system that can be used to assess the health condition of power generator stator windings and their levels of water absorption. The system monitors health based on capacitance measurements of the winding insulations. In particular, a new relative health measure, namely the Directional Mahalanobis Distance (DMD), is proposed to quantify the health condition of stator windings. This paper also proposes an empirical health classification rule, based upon the DMD, which factors in maintenance history. The proposed smart health reasoning system is validated using eight years' field data from eight generators, each of which contains forty-two windings.

Keywords: Power Generator

Stator Winding

Statistical Correlation

Health Diagnostics

Directional Mahalanobis Distance

Water Absorption

Student Number: 2012-20664

Table of Contents

List of Tables

List of Figures

Nomenclature

$\lfloor \cdot \rfloor$	Round down function (= floor function)
A	Measurement area
C	Capacitance
Cov	Covariance
D	Distance
$\mathrm{E}[\cdot]$	Expected value
ε_0	Electric constant
ε_r	Relative static permittivity (= dielectric constant)
h	Threshold value/distance
k	The number of variables (= data size of group)
$\max[\cdot]$	Maximum value
$\min[\cdot]$	Minimum value
μ	Sample mean
n	Sample size
Q	Charge on conductor
ρ	Correlation coefficient
$\mathbf{\Sigma}$	Covariance matrix
σ	Standard deviation

V	Voltage
X	Random variable
$\tilde{X}$	Processed capacitance data

Abbreviations

CEB	Collector End – Bottom bar
CET	Collector End – Top bar
CM	Condition Monitoring
DMD	Directional Mahalanobis Distance
GE	General Electric
ID	Identification
KEPCO	Korea Electric Power Corporation
MD	Mahalanobis Distance
MTS	Mahalanobis-Taguchi System
PHM	Prognostics and Health Management
RUL	Remaining Useful Life
SDMD	Scaled Directional Mahalanobis Distance
SMD	Scaled Mahalanobis Distance
TEB	Turbine End – Bottom bar
TET	Turbine End – Top bar
USC	Ultra-Supercritical

Chapter 1. Introduction

1.1 Motivation

Power generators are critical elements of power plants. An unexpected breakdown of a generator can lead to plant shut-down and can result in substantial economic and societal loss. Recently, tremendous technological advancements have been achieved in the development and deployment of an ultra-supercritical (USC) steam generator, shown in Figure 1-1. The USC steam generator operates at an advanced steam temperature of 593℃ or above, enabling it to achieve higher energy conversion efficiency, while at the same time reducing fuel consumption and waste emission. However, the large gap between the operation temperature and pressures of the advanced USC generator and those found in conventional subcritical generators leads to far harsher operating conditions in the USC. Thus, the USC has a much higher risk of catastrophic failure. To minimize the losses resulting from potential failures, the reliability of the USC-type power generator must be ensured throughout its life-cycle amidst uncertain operating conditions and manufacturing variability.

Recently, prognostics and health management (PHM) has emerged as a key technology to evaluate the current health condition (health diagnostics) and predict the future degradation behavior (health prognostics) of an engineered system throughout its lifecycle. In general, PHM consists of four basic functions: a health sensing function, a health reasoning function, a health prognostics function, and a health management function. PHM has shown success in lowering system maintenance costs of various engineered systems. Comprehensive exploration of

PHM techniques for power generator windings can enable early anticipation of failure. PHM can be used to develop cost-effective maintenance strategies and to seek opportunities for extending equipment life. Effective health reasoning systems are a crucial step towards a comprehensive exploration of PHM techniques.

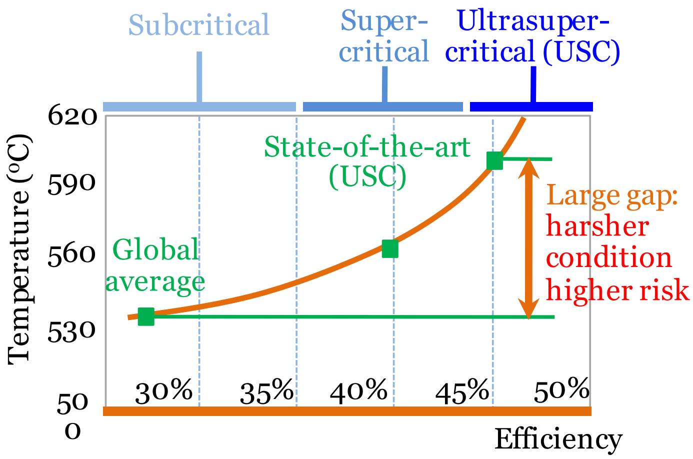


Figure 1-1.Comparison of performance between subcritical, supercritical and ultra-supercritical steam generators

1.2 Overview

This research aims to develop a health reasoning system for power generator stator windings through both physical and statistical analysis. A health reasoning system, also known as the integration of condition monitoring (CM) and health classification, is an algorithm-based system used to diagnose health conditions based on sensory signals and related health measures. Two steps are typically involved: (1) CM to extract relevant system health information through feature extraction techniques, (2) health classification to classify a system's health state into diverse health classes using health classification techniques such as artificial intelligence, a support vector machine, decision trees, and mahalanobis distance (MD).

This research proposes a new classification technique that can be applied to stator windings in power generators. The new technique eliminates the limitations found in existing methods and in MD, which is widely used in the PHM field. The proposed definition for health classification is carried out with data from the maintenance history that has been obtained from the field.

1.3 Thesis Layout

This thesis is organized as follows: Chapter 2 reviews existing methods for detecting leaks and water absorption in the insulation of stator windings found in power generators. Chapter 3 presents the sensing function for health monitoring and analysis of health data which can be measured regarding the insulation of stator windings. Chapter 4 discusses feature extraction techniques and introduces a health index for a smart health reasoning system. Chapter 5 presents a new health classification rule which includes consideration of the data size. The validation study for the proposed health index and grade system is given in this chapter as well. Finally, conclusions of the paper are presented in Chapter 6.

Chapter 2. Literature Review

This chapter reviews the existing state of knowledge related to health assessment of power generators – the topic of this thesis. Topics reviewed include: (1) PHM techniques for engineered systems, including power generators, and (2) existing methods used in the field to guard against leak and water absorption.

2.1 Prognostics and Health Management Techniques used to Support the Health Reasoning Function

PHM is a method that permits assessment of the reliability of a system under its actual application conditions. Extensive research has been conducted in the application of PHM to various engineered systems.

Popular tools used for the health reasoning function include statistical methods (A. H. Christer et *al*. [1], and Y. Zhan et *al*. [2]), artificial intelligence (R. B. Chinnam et *al*. [3] and C. C. Lin et *al*. [4]), support vector machines (J. Yang et *al*. [5] and C. Cortes et *al*. [6]), kernel estimation (N. S. Altman [7]), decision trees (L. Rokach [8]), Mahalanobis distance (R. De Maesschalck et *al*. [9] and G. Taguchi et *al*. [10]), Kalman filter (J. D. Wu et *al*. [11] and S. K. Yang [12]), among others.

These algorithms can be applied to various engineered applications, including bearings (I. E. Alguindigue et *al*. [13]), gearboxes (S. Ebersbach et *al*. [14]), machine tools (D. E. Dimla Sr. [15] and K. F. Martin [16]), transformers (C. Bartoletti et *al*. [17], C. Bengtsson [18], C. Hu et *al*. [19] and C. Booth et *al*. [20]), generators (C. W. Park et *al*. [21], J. Finn et *al*. [22], A. Kheirmand et *al*. [23] and

G. C. Stone et *al*. [24]) and stator insulation used for winding in generators (G. A. Jayantha et *al*. [25] and Z. Jia et *al*. [26]).

2.2 Existing Tests to Detect Leaks or Water Absorption

General Electric (GE), one of the largest manufacturers of generators in the world, provides guidelines for a standard outage test program for periodic overhaul of generators. The schedule typically includes minor overhaul every 30 months and major overhaul every 48-60 months. Figure 2-1 summarizes the GE maintenance program plan. The program consists of four tests; (1) the Vacuum Decay Test, (2) the Pressure Decay Test, (3) the Helium Tracer Gas Test and (4) the Stator Bar Capacitance Mapping Test (J. A. Worden et *al*. [27]).

The first three tests are designed to detect leaks in the stator winding. The first test, the Vacuum Decay Test is a useful tool for determining the integrity of the entire water-cooled stator hydraulic system. The primary advantage of this test is its sensitivity. Decay measurements are made in units of microns. A typical pressure gage cannot detect one micron, which is equivalent to .00002 psi. Because of the high sensitivity of this test, ironically, extremely small leaks at flanges and connections can result in poor test results. The second test, the Pressure Decay Test, has two advantages over the Vacuum Decay Test. It provides a greater pressure differential and applies pressure in the normal direction of the leak flow. These factors may make it easier to find leaks undetectable in the Vacuum Decay Test. During this test, exposed potential leak sites can be tested using a bubble. Drawbacks to pressure testing are its insensitivity to small leaks, and relatively high sensitivity to changes in the environment. The third test, the Helium Tracer Gas Test, is a method of leak detection where the generator is pressurized with a helium gas so that possible leak points can be detected using a helium gas detector. In many cases, leaks that were missed by the Vacuum Decay Test and the Pressure Decay Test are found with the Tracer Gas Test.

Finally, the Stator Bar Capacitance Mapping Test is used to determine the extent of water absorption. This test assumes that good capacitance data provides a normal distribution when plotted; nearly all of the data should fall between -2 and +2 standard deviations from the average. This test uses +3 standard deviations of the capacitance data as a failure threshold.

Korea Electric Power Corporation – Research Institute (KEPRI) has also developed methods using a capacitance reader (or wet bar detector) for detecting water absorption using statistical tools, including (1) a Normal Probability Plot and (2) a Box Plot (H. S. Kim et *al*. [28]).

The Normal Probability Plot method determines the health classes of winding insulation based on a normal probability plot. An operator can visually identify an outlier or anomaly case by examining the plot. It is assumed that the capacitance data of a healthy winding follows a normal distribution. The Box Plot method is another graphical method, which graphically depicts the health classes of the capacitance data through their 1st and 3rd quartiles. The drawback of both aforementioned methods is that the sensitivity of the winding health classification is relatively low because of improper statistical modeling of the capacitance data and a lack of consideration of data heterogeneity.

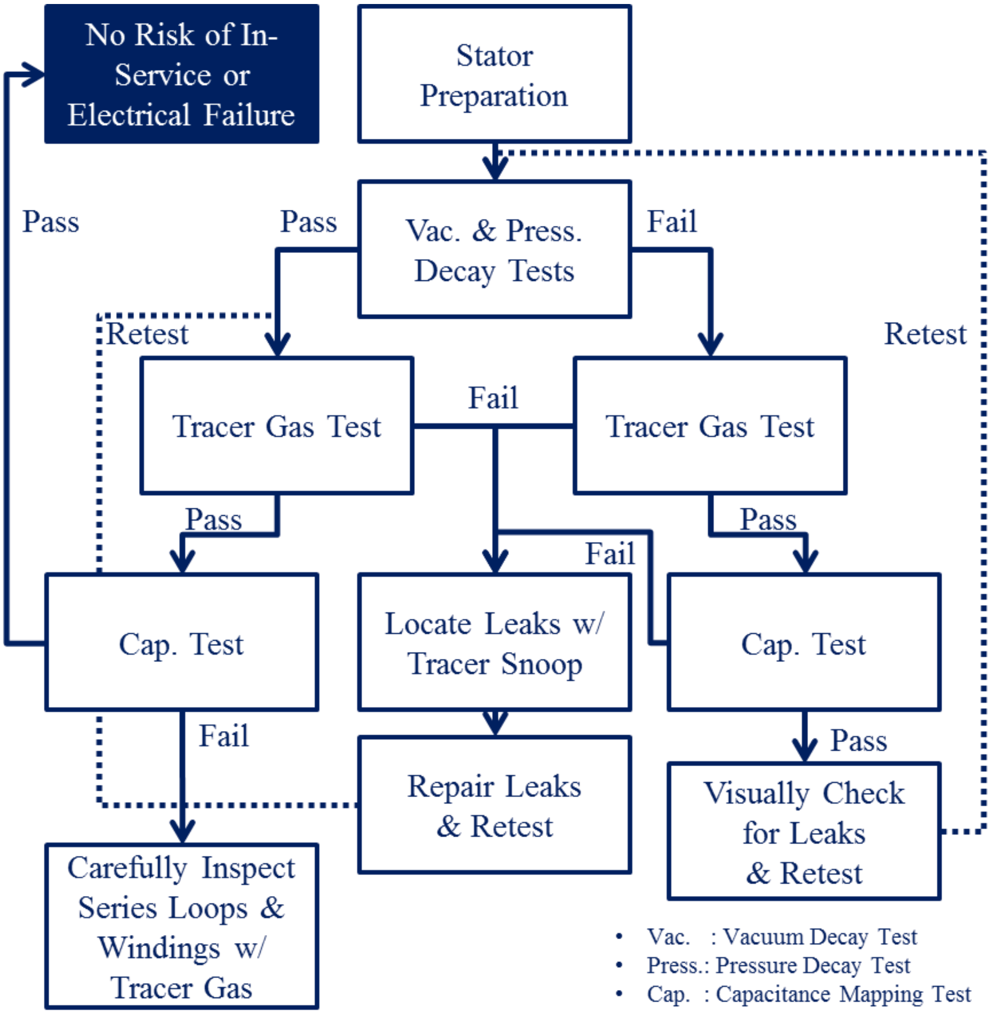


Figure 2-1.Major output leak test plan (GE) [27]

2.3 Summary and Discussion

The aforementioned PHM techniques can be applied to various engineered systems. Most monitoring systems for power generators use electrical signals to detect the faults of the generator. In the case of monitoring the system for water absorption, basic statistical ideas have been used to detect leaks and to detect the water absorbed in insulation based upon capacitance readings. Note that the direct use of capacitance measurements as the health index by the existing methods reviewed previously makes it difficult to easily and precisely infer the health status, especially when the measurements are of high dimensionality, high correlation, and/or high non-linearity. To improve upon the status quo, the work we propose develops a new health index through statistical analysis of multi-dimensional capacitance measurements for effectively determining the health of power generator windings. The ultimate goal of this work is to better prevent sudden failure and to enable a self-sustained generator, as shown in Figure 2-2.

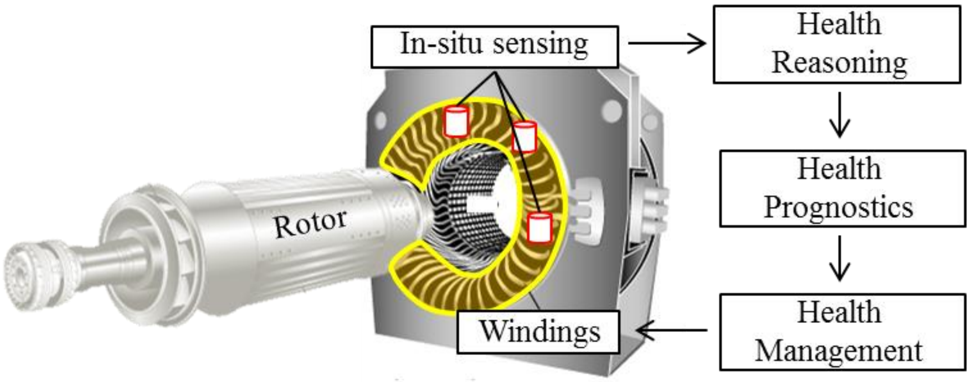


Figure 2-2.Processes involved in a self-sustained power generator

Chapter 3. Description of the Sensing Function and Data Analysis

The objective of the sensing function is to ensure high damage detectability and efficient data management by designing data acquisition logistics. In addition, the health condition of a power generator can be monitored by properly analyzing the capacitance of the winding insulation. This section discusses the fundamentals of capacitance measurements, locations of capacitance measurements, and characteristics of measurement data. This study examines eight power generators (nineteen datasets over eight years) which have the same specifications: (1) a 500 MW output, (2) a 2-path cooling system, and (3) a 60 Hz frequency.

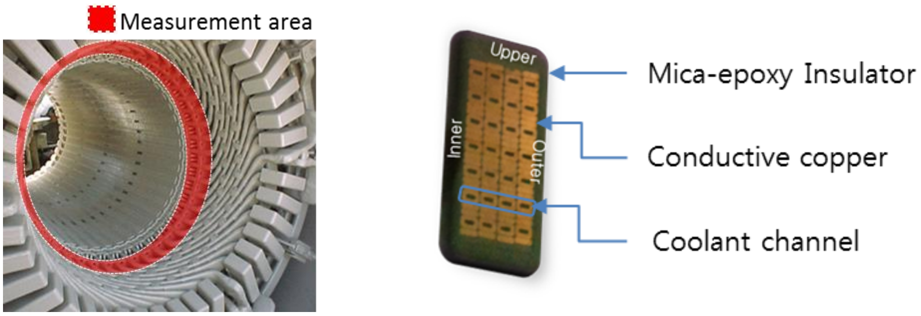
3.1 Fundamentals of Capacitance Measurements

When a power generator is water-cooled, coolant water flows into the water channels of the winding, as shown in Figure 3-1. Leakage into the surrounding insulation can occur due to various operational stresses, such as mechanical vibration, thermal shock, and crevice corrosion (see Figure 3-2). When leakage occurs, the water or moisture remains in the winding insulation. The remaining water degrades the winding insulation, which can cause insulation breakdown and power generator failure, as shown in Figure 3-3. For this reason, electric companies and manufacturing companies, such as KEPRI, GE and Toshiba, assess the health status of the winding insulation in their generators using a water absorption detector

[27-31]. The water absorption detector infers the level of water in the insulation by measuring the capacitance of the insulation. Because the relative static permittivity (or the dielectric constant) of water is higher than that of mica (which is generally used as the insulation material), wet insulation has a higher capacitance, C, based upon the following equation (see Figure 3-4 for a schematic representation):

$$C = \frac{Q}{V} = \varepsilon_r \varepsilon_0 \frac{A}{t} \tag{3.1}$$

where Q is the charge on each conductor, V is the voltage between the plates, A and t are, respectively, the measurement area and the thickness of the detector, ε_0 is the electric constant ($\varepsilon_0 \approx 8.854\text{pF-m}^{-1}$) and ε_r is the relative static permittivity of the material between the plates. Measures of capacitance as health data provide valuable information that can be used to infer the amount of moisture absorption of a stator winding. Health-relevant information about the winding can be extracted from this measured moisture level. It should be noted that various uncertainty factors, such as the measurement location, the ambient humidity, and the winding surface condition propagate uncertainties into the capacitance measurements. These uncertainties must be taken into account in the health reasoning process.



(a) (b)

Figure 3-1.Power generator stator (a) and cross-section view of a winding (b)

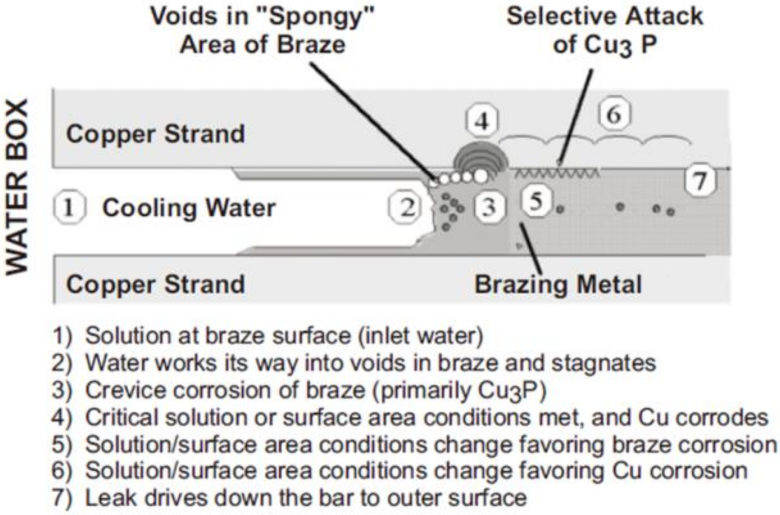
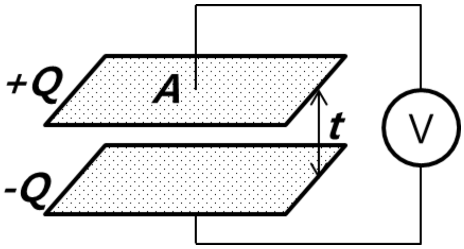


Figure 3-2.Diagram of a crevice corrosion mechanism [27]

Figure 3-3.Failure of a power generator stator



(a) (b)

Figure 3-4.Capacitance reading using a detector (Model: GEN-SWAD I) (a) and the basic principle of the capacitance detector (b)

3.2 Capacitance Data Acquisition

As mentioned previously, each of the power generators employed in this study has forty-two stator windings and slots used for a water-cooled cooling system. As shown in Figure 3-5, the cooling water flows from the top bar inlet at the turbine end, through the top and bottom bars at the collector end, then back to the bottom bar outlet at the turbine end. At the turbine or collector end, an assembly slot in both the top and bottom bars contains ten measurements points. The ten measurement points are summarized in Table 3-1 and graphically illustrated together with the generator structure diagram in Figure 3-5. Note that since only an extremely small gap exists between the top and bottom bars, the capacitance on the top side of the bottom bar cannot be measured, resulting in only two measurement points for the bottom bar. As shown in Table 3-1, a unique identification (ID) code is assigned to each measurement point based on the location of the probing point.

For example, the ID code “CET-TOP” indicates that the measurement point is located on the **Top** side of the **C**ollector **E**nd **T**op bar. The capacitance data were acquired from the ten measurement points for each of the forty-two slots found on each power generator. The capacitance data measured at each measurement point can be modeled as a random variable (X).

Table 3-1.Summary of ten measurement points on a single stator winding

Stator side	**Winding location**	**Measurement point**	**Identification Code**
Collector End (CE)	Top Winding	TOP	CET-TOP (X_1)
		OUT	CET-OUT (X_2)
		IN	CET-IN (X_3)
	Bottom Winding	OUT	CEB-OUT (X_4)
		IN	CEB-IN (X_5)
Turbine End (TE)	Top Winding	TOP	TET-TOP (X_6)
		OUT	TET-OUT (X_7)
		IN	TET-IN (X_8)
	Bottom Winding	OUT	TEB-OUT (X_9)
		IN	TEB-IN (X_{10})

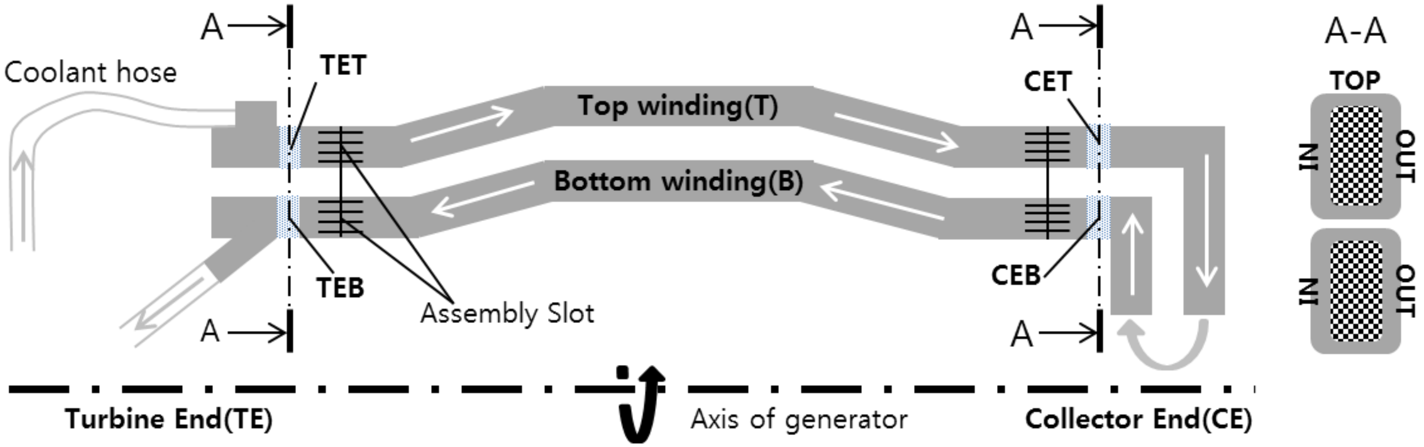


Figure 3-5.Structure diagram of a water-cooled power generator with a 2-path cooling system

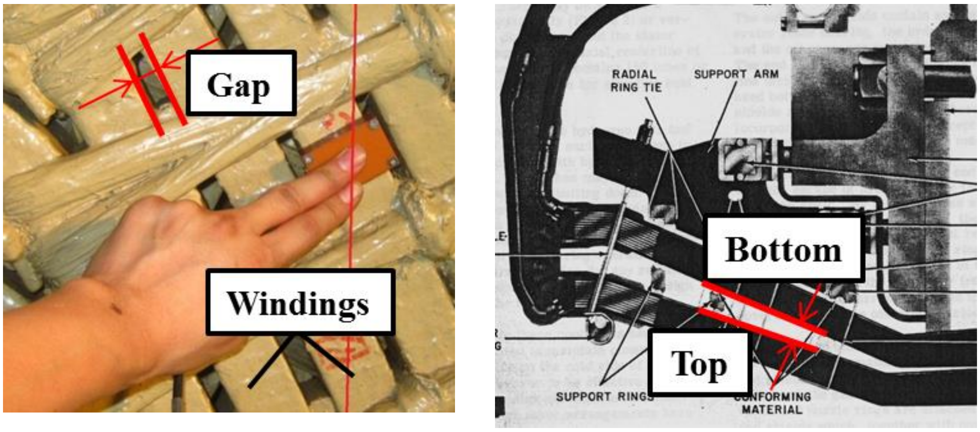
3.3 Statistical Characterization of the Capacitance Data

The capacitance data acquired at physically isolated measurement points can be modeled as statistically independent random variables (i.e. X_1 and X_{10}). Alternatively, the data can be modeled as statistically correlated random variables. For example, a physical gap between two different windings (as shown in Figure 3-6) is one reason why the related random variables might be statistically independent. Moreover, different winding locations (CET, CEB, TET, and TEB) in one winding are also physically distant. This also implies that the related random variables could be statistically independent. On the other hand, water absorption occurs concurrently at adjacent measurement points in the same group, such as CET-TOP (X_1), CET-OUT (X_2), and CET-IN (X_3). Checking statistical dependence between two random variables could confirm whether our intuitive observation is true or not. Before checking statistical correlations, a mean shift was applied to all datasets to take into account the inherent difference in the nominal states of water absorption between generators, as shown in Figure 3-7. After the mean shift, the correlation coefficients later become useful to develop the health reasoning process for a stator winding in a power generator.

In general, the correlation coefficient is used as a measure to imply statistical correlation. The most famous measure of correlation is the Pearson product-moment correlation coefficient. It is a quantitative measure of a linear dependence between two variables. Mathematically, a correlation coefficient can be calculated from the following form:

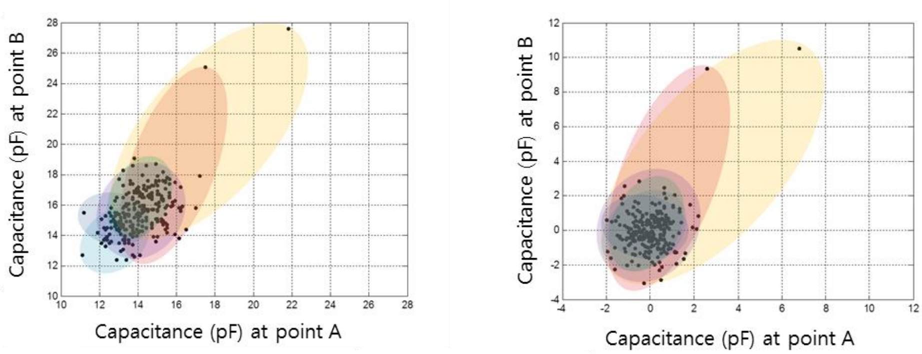
$$\rho_{X_i,X_j} = \frac{\mathrm{Cov}\left(X_i, X_j\right)}{X_i X_j} = \frac{\mathrm{E}\left[\left(X_i - \mu_{X_i}\right)\left(X_j - \mu_{X_j}\right)\right]}{\sigma_{X_i}\sigma_{X_j}} \tag{3.2}$$

where X_i and X_j are random variables, Cov($X_{i,}$ X_j) is the covariance between X_i and X_j, μ and σ are the mean and standard deviation of a random variable, respectively, and E[•] is the expectation of a random variable.



(a) (b)

Figure 3-6.Gaps between two different windings (a) and between top and bottom stator bars (b)



(a) (b)

Figure 3-7.Scatter plots of the data between measurement points before (a) and after (b) mean shift

Table 3-2 summarizes the correlation coefficients for ten random variables, ρ_{X_i,X_j} for i, j = 1 to 10, in matrix form. The highlighted values in Table 3-2 are the coefficients between the correlated random variables in the same group. One can observe two features from the highlighted values: (1) a statistically positive correlation, and (2) a higher degree of correlation within the same group. These features indicate that the two or three capacitance data from the same group tend to behave (i.e. remain unchanged or grow) with linear dependence. This confirms the intuitive observations about the aforementioned statistical correlation and independence. Appendix A provides pairwise scatter plots between the ten measurement points (or ten random variables), from which two scatter plots are extracted to show the within-group and between-group correlations.

Table 3-2.Correlation coefficient matrix (symmetric) for ten random variables in matrix form

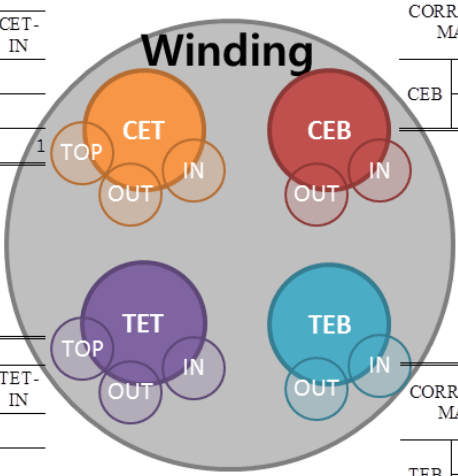
Correlation Matrix		**TOP (X_1)**	**CET OUT (X_2)**	**IN (X_3)**	**CEB OUT (X_4)**	**IN (X_5)**	**TOP (X_6)**	**TET OUT (X_7)**	**IN (X_8)**	**TEB OUT (X_9)**	**IN (X_{10})**
CET	**TOP (X_1)**	1									
	OUT (X_2)	**0.4761**	1								
	IN (X_3)	**0.4194**	**0.5503**	1							
CEB	**OUT (X_4)**	0.0849	0.1572	0.1354	1						
	IN (X_5)	-0.039	0.1686	0.0765	**0.3445**	1					
TET	**TOP (X_6)**	0.3341	0.1553	0.1868	0.0343	-0.052	1				
	OUT (X_7)	0.1972	0.2506	0.2729	0.0879	0.0171	**0.4377**	1			
	IN (X_8)	0.2295	0.1423	0.3296	0.0082	0.0457	**0.4269**	**0.4900**	1		
TEB	**OUT (X_9)**	0.0438	-0.128	-0.097	0.0186	-0.114	0.0887	-0.010	-0.003	1	
	IN (X_{10})	0.0354	-0.040	-0.004	0.0457	0.0870	-0.048	0.1084	0.0215	**0.3385**	1

3.4 Data Grouping

It is important to define a group of capacitance data with homogeneity prior to the data modeling and health reasoning process. Based upon the measurement location and correlation characteristic obtained in Section 3.3, the measurement points with high correlation can be conceived as individual data groups, such as CET, CEB, TET, and TEB. This implies that the entire dataset for ten random variables (or from ten-dimensional measurement points) would be split into four groups with two or three random variables. This data grouping will be used for the health reasoning process in the subsequent section, which defines a health index and models it in a statistical form, as shown in Figure 3-8 and Table 3-3. The data grouping makes the health reasoning process easier through dimensional reduction of the capacitance data.

CORRELATION MATRIX		CET		
		CET-TOP	CET-OUT	CET-IN
CET	CET-TOP	1		
	CET-OUT	0.4761	1	
	CET-IN	0.4194	0.5503	1

CORRELATION MATRIX		CEB	
		CEB-OUT	CEB-IN
CEB	CEB-OUT	1	
	CEB-IN	0.3445	1



CORRELATION MATRIX		TET		
		TET-TOP	TET-OUT	TET-IN
TET	TET-TOP	1		
	TET-OUT	0.4377	1	
	TET-IN	0.4269	0.49	1

CORRELATION MATRIX		TEB	
		TEB-OUT	TEB-IN
TEB	TEB-OUT	1	
	TEB-IN	0.3385	1

Figure 3-8.Data grouping using a statistical correlation

Table 3-3.List of measured datasets and the information employed in this study

Power Plant	**Generator Number**	**Year for Measurement**	**Winding Number**	**Measured dataset**			
				CE group		**TE group**	
A	1	’05, ’07, ‘09	1	CET (X_1, X_2, X_3)	CEB (X_4, X_5)	TET (X_6, X_7, X_8)	TEB (X_9, X_{10})
			⋮	⋮	⋮	⋮	⋮
			42	CET	CEB	TET	TEB
	2	’06, ‘09	1	CET	CEB	TET	TEB
			⋮	⋮	⋮	⋮	⋮
			42	CET	CEB	TET	TEB
	3	’06, ‘10	1	CET	CEB	TET	TEB
			⋮	⋮	⋮	⋮	⋮
			42	CET	CEB	TET	TEB
	4	’06, ’07, ‘08	1	CET	CEB	TET	TEB
			⋮	⋮	⋮	⋮	⋮
			42	CET	CEB	TET	TEB
B	1	’06, ’10, ‘12	1	CET	CEB	TET	TEB
			⋮	⋮	⋮	⋮	⋮
			42	CET	CEB	TET	TEB
	2	’09, ‘12	1	CET	CEB	TET	TEB
			⋮	⋮	⋮	⋮	⋮
			42	CET	CEB	TET	TEB
C	4	’06, ‘10	1	CET	CEB	TET	TEB
			⋮	⋮	⋮	⋮	⋮
			42	CET	CEB	TET	TEB
D	6	’09, ‘11	1	CET	CEB	TET	TEB
			⋮	⋮	⋮	⋮	⋮
			42	CET	CEB	TET	TEB

Chapter 4. Statistical Health Reasoning System[1]

Although the capacitance data are relevant to the health status of the stator winding, its high dimensionality and non-linearity make it difficult to easily and precisely infer the health status. This section proposes a new health index, referred to as the Directional Mahalanobis Distance (DMD).

4.1 Review of Mahalanobis Distance

The Mahalanobis Distance (MD) is a relative health measure that quantifies the deviation of a measured data point from a clustered data center, which is generally a populated mean (μ) of a dataset. The MD degenerates multi-dimension data (**X**) to a one-dimension distance measure while taking into account the statistical correlation between random variables. Mathematically, the MD measure can be expressed as:

$$\mathrm{MD}(\mathbf{X}_i) = \sqrt{(\mathbf{X}_i - \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{X}_i - \boldsymbol{\mu})} \quad (4.1)$$

where $\mathbf{X}_i = (X_{1,i}, X_{2,i}, \cdots, X_{N,i})^{\mathrm{T}}$ is an N-dimensional capacitance data vector of the i^{th} winding unit which belongs to a group having the mean $\boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_N)^{\mathrm{T}}$ and the covariance matrix $\boldsymbol{\Sigma}$. Figure 4-1 plots two-dimensional samples randomly drawn from two random variables with a positive correlation. Essentially, the MD

1 Sections of this chapter have been submitted as the following journal article: Byeng D. Youn, Kyung Min Park, Hu Chao, Joung Taek Yoon, and Hee Soo Kim, "Statistical Health Reasoning of Water-Cooled Power Generator Stator Windings against Moisture Absorption," *Reliability Engineering and System Safety*, Submitted, 2014.

transforms an ellipsoid in the original random space to a circular shape in the standard Gaussian space, as shown in Figure 4-1. Since the distance (D_1) of the faulty point to the clustered data center is much shorter than that (D_2) of the healthy point, one could have concluded based upon the Euclidean distance that the faulty point is more likely to belong to the cluster. This misleading conclusion is mainly caused by not taking into account the correlation coefficient of the two random variables. Indeed, if we simply divide the distances D_1 and D_2 by the widths of the ellipsoid in the corresponding directions, respectively, we can easily come to the conclusion that the faulty point is much farther away from the clustered center than the healthy point. This can be clearly observed in Figure 4-1.

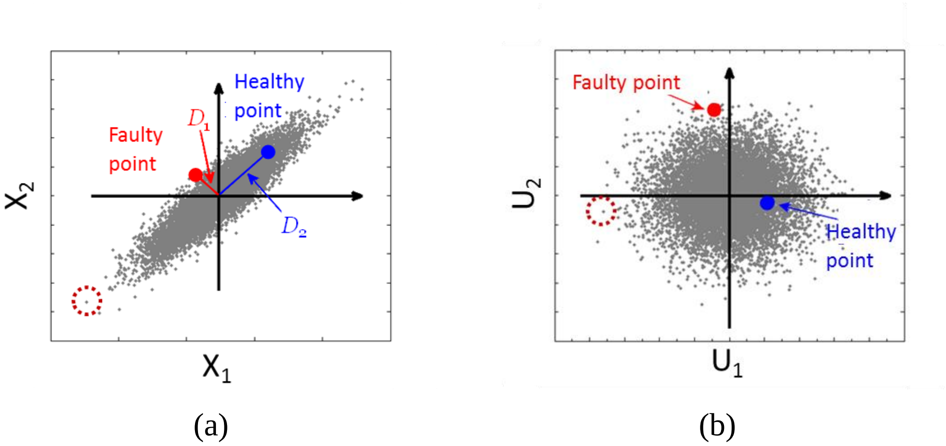


Figure 4-1.Healthy and faulty points located in the original space (a) and the normalized space (b)

As compared to the Euclidean distance, the MD measure possesses a few unique advantages, listed as follows: (1) The MD transforms a high-dimensional dataset that is complicated to handle into a one-dimensional measure capable of easy comprehension and quick computation. (2) The MD is robust to differing scales of the measurements, as the MD values are calculated after normalizing the data. (3) By taking into account the correlation of the dataset, the MD is sensitive to inter-variable changes in multivariate measurements.

4.2 A New Concept of Statistical Distance: Directional Mahalanobis Distance

This subsection introduces a new MD-based distance measure that can be used to imply the health condition of a stator winding in a power generator.

4.2.1 Data Projection

The MD, as a relative health measure, provides very useful information to characterize the health condition of a stator winding in a power generator. According to Equation (3.1), the capacitance values measured from a dry stator winding with a negligible amount of water on the insulation should be smaller than the mean value of the measurement population. Previous studies [27, 28] also reported that measured values smaller than the population mean should be treated as if they have no relation to the insulation's water absorption. However, the MD, as a scalar distance measure, is a direction-independent health measure in the random capacitance space, as shown in Figure 4-2. In other words, two capacitance measurements with the same MD value but in two opposite directions are treated equally, although they most likely imply the different levels of water absorption.

Let us take the dashed circle datum in Figure 4-1 as an example. In this case, the dashed circle datum is a healthy point since this datum falls into the lower tails of the marginal distributions of the random capacitance variables. However, the MD declares this datum to be in the failure category simply because it is out of the data cluster. For this very reason it is necessary to refine the measure so that is better suited for this application. In order to rebuild the measure, this study employs a projection process which first identifies absolutely healthy variables(s) (e.g. a capacitance value less than its populated mean, say $X_i < \mu_i$) and then projects it onto the corresponding mean value(s), e.g. (X_i, μ_j) and (μ_i, X_j) shown in Figure 4-2. Through this projection process, the absolutely healthy data will be ignored in the subsequent transformation. The data projection underscores the consideration of the direction in the health reasoning process of the measurement data. This leads to the unique capability of the proposed health index to makes use of the distance and degradation direction as a health measure.

After the data projection, the capacitance data $\tilde{X}_{n,i}\left(n=1,\cdots,N\right)$, can be processed as:

$$\tilde{X}_{n,i} = \begin{cases} X_{n,i}, & \text{if } X_{n,i} > \mu_n \\ \mu_n, & \text{otherwise} \end{cases} \tag{4.2}$$

where $X_{n,i}$ denotes the raw capacitance data at the n^{th} measurement location of the i^{th} winding unit, μ_{n} is the mean of the capacitance data at the n^{th} measurement location, and $\tilde{X}_{n,i}$ denotes the processed capacitance data. The mean and variance of the dataset must be obtained before the data projection because they are physically meaningful in the original space.

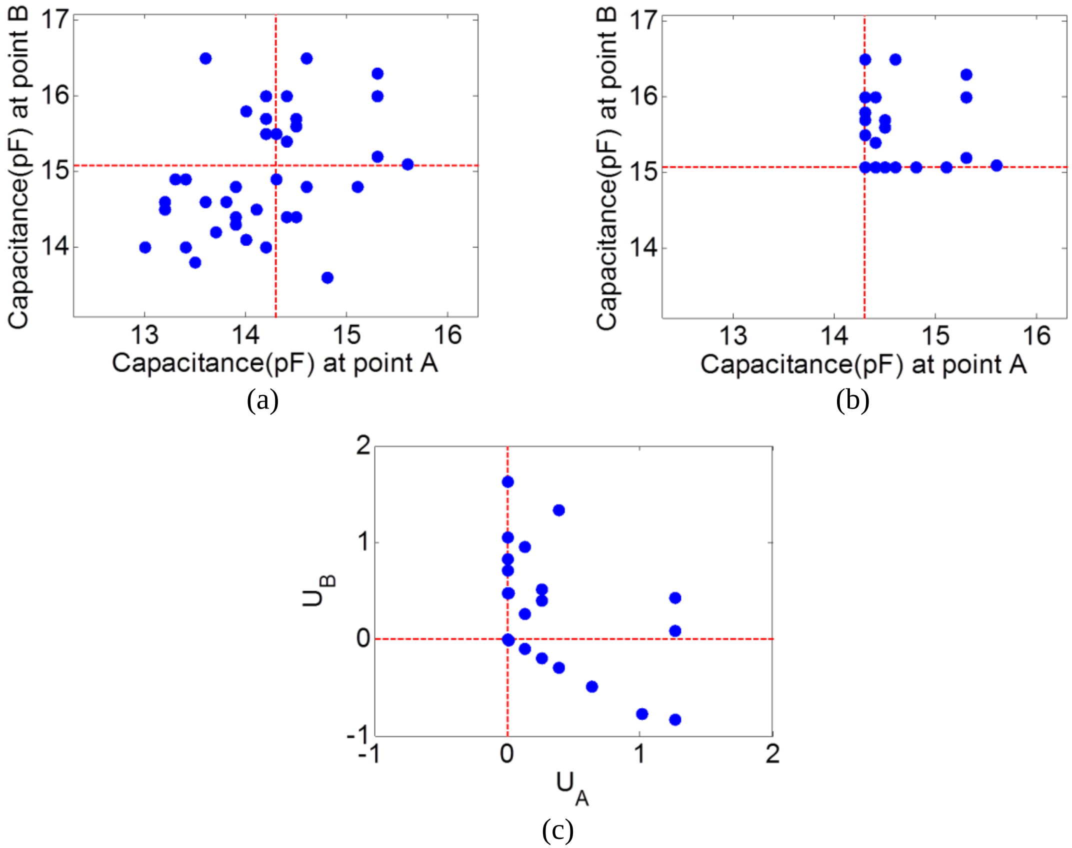


Figure 4-2.Scatter plots before data projection (a), after projection (b), and after transformation (c)

4.2.2 Transformation

The proposed index, namely the Directional Mahalanobis Distance (DMD), assesses the MD along the degradation direction of a stator winding insulation after data projection. Mathematically, the DMD shares a similar formula with the MD, except for consideration of the data projection. It is expressed as:

$$\mathrm{DMD}\left(\tilde{\mathbf{X}}_i\right)=\sqrt{\left(\tilde{\mathbf{X}}_i-\boldsymbol{\mu}\right)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\left(\tilde{\mathbf{X}}_i-\boldsymbol{\mu}\right)} \tag{4.3}$$

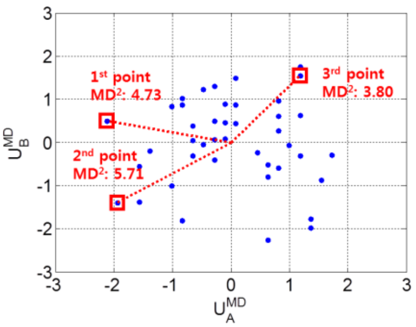
where $\tilde{\mathbf{X}}_i=\left(\tilde{X}_{1,i},\tilde{X}_{2,i},\cdots,\tilde{X}_{N,i}\right)^{\mathrm{T}}$ is an N-dimensional vector of the capacitance data from the i^{th} winding unit after the data projection, $\boldsymbol{\mu}=\left(\mu_1,\mu_2,\cdots,\mu_N\right)^{\mathrm{T}}$ and $\boldsymbol{\Sigma}$ is the mean vector and covariance matrix of the reference dataset before the projection. Figure 4-2(c) shows the scatter plot of the DMD dataset after projection and transformation; the difference between the MD and DMD can be clearly observed in this figure.

One question remains: what is the proper sequence of data projection and transformation? We found that the former should be done prior to the latter. The justification for this sequence is the fact that the comparison between the absolutely healthy data and the mean value of each random variable is physically meaningful and valid only in the original space, not in the transformed space.

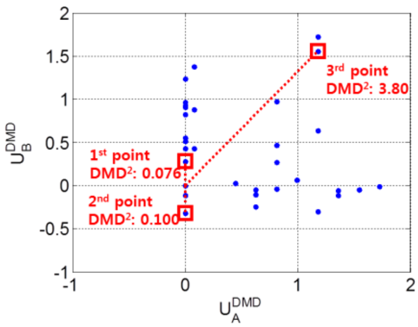
4.3 Comparison of Performance of Mahalanobis Distance (MD) and Directional Mahalanobis Distance (DMD)

The distances in this paper are squared in order to place progressively greater weight on objects that are farther apart. In general, squared distance is frequently used in instances where only distances need to be compared.

Figure 4-3 shows the scatter plot of MD and DMD with three highlighted data points, two of which represent a healthy state and the other a faulty state. In the case of MD, the 1st and 2nd data points (the two "healthy" points) are located in the 2nd (top left) and 3rd (bottom left) quadrants of the two-dimensional space composed of two Euclidean distances. The 3rd data point, the "faulty" point, is located in the 1st (top right) quadrant. The MD values of the 1st, 2nd and 3rd points are, respectively, 4.74, 5.71, and 3.80. In the case of DMD, the squared distances of the three points from the origin are, respectively, 0.076, 0.100, and 3.80. Without the projection before transformation, MD incorrectly treats the "healthy" data points as "faulty" while, with the projection before the transformation, the proposed index, DMD, correctly identifies these two points as "healthy" (with relatively small distance values). Therefore, the proposed DMD achieves better performance than the MD by accounting for the degradation direction in the health reasoning process of the capacitance data.



(a)



(b)

Figure 4-3.Three cases to compare the performance of MD (a) and that of DMD (b)

Performance evaluation of these two indices requires an evaluation metric that assesses the effectiveness of a health index in quantifying the health condition of a generator winding. The evaluation metric considered here employs a score function with the health index value and true health condition of a generator winding as the inputs and a normalized score metric (ranging between 0 to 100) as the output. Mathematically, the proposed score function can be expressed as:

$$\mathrm{SF} = 100 \times \frac{\sum_{j=1}^{N} x_j y_j - \underbrace{\left(\sum_{j=1}^{l} x_j - \sum_{j=l+1}^{N} x_j \right)}_{W}}{\underbrace{\left(\sum_{j=N-l+1}^{N} x_j - \sum_{j=1}^{N-l} x_j \right)}_{B} - \underbrace{\left(\sum_{j=1}^{l} x_j - \sum_{j=l+1}^{N} x_j \right)}_{W}} \tag{4.4}$$

where x_j denotes the health index value of the j^{th} winding unit, y_j denotes the maintenance index of the j^{th} winding based on the actual repair history ($y_j = 1$ if the unit was maintained and $y_j = -1$ otherwise), N denotes the number of winding units, l denotes the number of winding units with maintenance histories, and B and W denote the best and worst score metric values, respectively. Figure 4-4 illustrates various combinations of the health index ranking, the maintenance history, and the score metric ranking of these combinations. The best and worst scenarios (represented respectively by B and W in Equation(4.4)) are depicted by the leftmost and rightmost plots, respectively.

Table 4-1 summarizes the distances and scores from the assessment results for health condition of the winding using MD and DMD, respectively. As one can see in Table 4-1, both MD and DMD can easily find windings that have been maintained; the capacitance of maintained insulation is relatively larger than the capacitance seen in the unmaintained samples. However, MD cannot find all the maintained windings in the top ten health indices in the CET group. In addition, some unmaintained windings' health indices have higher values than those of maintained windings (e.g. 8^{th}, 9^{th}, and 10^{th} health indices in TET group).

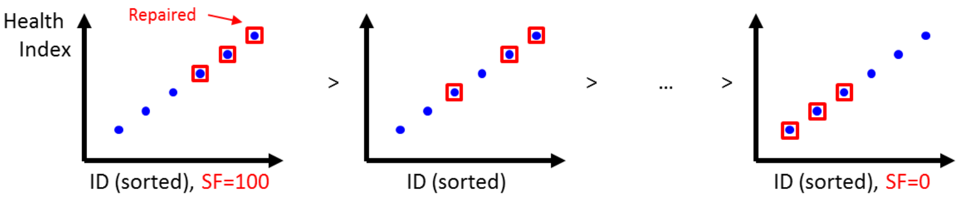


Figure 4-4.Combination of health index ranking and maintenance history and corresponding score metric ranking

Table 4-1.Distances and scores from the results examining health condition of windings using MD and DMD

Rank	**CET group**		**CEB group**		**TET group**		**TEB group**	
	MD^2	**DMD^2**	**MD^2**	**DMD^2**	**MD^2**	**DMD^2**	**MD^2**	**DMD^2**
1	33.64	33.64	21.14	21.14	20.37	20.37	11.65	11.65
2	28.41	28.41	20.63	20.63	18.93	18.93	9.97	9.01
3	21.96	21.96	17.96	14.36	17.93	17.93	9.64	8.88
4	18.60	18.60	15.87	14.00	17.67	17.67	9.18	8.08
5	18.06	18.06	14.36	13.61	15.88	15.88	9.01	7.87
6	17.58	17.58	14.27	12.68	15.75	14.65	8.88	7.34
7	16.42	16.42	12.68	8.18	14.65	13.57	8.69	7.04
8	16.39	16.39	11.05	7.86	14.11	13.15	8.08	6.86
9	13.36	12.25	10.85	6.75	13.98	9.33	7.87	6.45
10	12.76	9.90	10.78	6.70	13.55	8.91	7.74	6.23
11	12.55	9.51	9.67	6.27	13.15	8.39	7.68	6.18
12	12.52	9.41	9.51	6.08	11.79	8.09	7.66	5.96
13	11.47	9.36	9.37	5.30	11.56	6.80	7.50	5.74
14	11.46	8.57	8.98	5.27	11.11	6.68	7.48	5.49
15	11.43	8.03	8.76	5.19	10.63	6.62	7.34	5.48
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
21	10.56	6.67	7.82	4.38	9.33		6.86	4.60
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
27	9.51	6.25	6.77	4.06	8.52		6.22	4.39
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
29	9.36	5.60	6.52	3.94	8.39		6.17	4.32
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
798	0.01	0.00	0.00	0.00	0.02	0.00	0.00	0.00
Score	**96.31**	**98.30**	**98.47**	**100**	**94.52**	**96.80**	**91.26**	**95.49**

Chapter 5. Health Classification

This section is designed to construct a health grade system based upon the proposed health index, DMD, and field maintenance history. The historical maintenance data obtained from the operators of the power generators are presented in Section 5.1. An empirical health grade system is suggested in Section 5.3 with consideration of the scale that is discussed in Section 5.2.

5.1 Maintenance History Related to Water Absorption

In this section, the historical maintenance data related to water absorption is presented, including data from faulty windings. Field experts collected the maintenance records of generator windings and identified their health conditions using the Stator Bar Capacitance Mapping test method developed by GE. The maintenance records can be classified into four groups (CET, CEB, TET, and TEB). Table 5-1 summarizes the maintenance history related to water absorption over eight years. In this table, two data points (28.413 and 33.645) were obtained from the faulty winding and the others were measured from water-absorbed windings. These maintenance records provide the physical basis for determining an appropriate failure threshold, especially the data obtained from the faulty windings.

Table 5-1.Maintenance history related to water absorption

Group	Power Plant	Generator Number	Winding Number	Year	DMD^2	Condition
CET	A	1	20	‘05	18.598	Absorbed
				‘07	16.415	
				‘09	16.394	
		2	03	‘06	18.060	Absorbed
				‘09	21.963	
			40	‘09	9.511	Absorbed
		4	23	‘07	28.413	**Faulty**
				‘08	33.645	
	B	1	18	‘10	9.355	Absorbed
				‘12	17.580	
CEB	A	2	31	‘06	20.625	Absorbed
				‘09	21.141	
		4	20	‘07	14.359	Absorbed
				‘08	12.677	
	C	4	11	‘06	13.613	Absorbed
				‘10	14.001	
TET	A	1	11	‘07	9.334	Absorbed
				‘09	13.145	
			20	‘05	17.665	Absorbed
				‘07	17.929	
				‘09	18.932	
		2	40	‘06	20.368	Absorbed
				‘09	14.653	
	B	1	18	‘10	15.876	Absorbed
TEB	A	3	23	‘10	8.080	Absorbed

5.2 Review of Scaled Mahalanobis Distance

The aim of this section is to introduce an improvement to the concept of MD: Scaled Mahalanobis Distance (SMD) in the Mahalanobis-Taguchi System (MTS). As aforementioned, this study deals with the two groups which have different data size (i.e. Top group and Bottom group). Since MD and DMD do not consider the data size of each group, it is not suitable to compare two groups directly without any conversion. SMD is a way to solve this problem.

It has been shown that MD follows a χ^2-distribution with k degrees of freedom, when the sample size, n, is large and all characteristics follow the normal distribution (R. A. Johnson et *al*. [32]). A χ^2-distribution with k degrees of freedom has a mean equal to k. Hence, G. Taguchi et *al*. [10] proposed a new idea for MD known as SMD. This group suggested that MD should be scaled by dividing by the number of variables, k. Thus, the equation for calculating SMD in the MTS becomes:

$$\mathrm{SMD}^2 = \frac{1}{k}\mathrm{MD}^2 = \frac{1}{k}\left(\mathbf{X}_i - \boldsymbol{\mu}\right)^{\mathrm{T}} \boldsymbol{\Sigma}^{-1}\left(\mathbf{X}_i - \boldsymbol{\mu}\right) \tag{5.1}$$

where k denotes the number of variables or the data size of each group, and $\mathbf{X}_i$ is a capacitance data vector of the i^{th} winding unit, which belongs to a group having the mean vector, $\boldsymbol{\mu}$, and the covariance matrix, $\boldsymbol{\Sigma}$.

Since the expected value of MD^2 is equal to k, the expected value of SMD^2 becomes:

$$\mathrm{E}\left[\mathrm{SMD}^2\right] = \mathrm{E}\left[\frac{1}{k}\mathrm{MD}^2\right] = \frac{1}{k}\mathrm{E}\left[\mathrm{MD}^2\right] = \frac{1}{k}\cdot k = 1 \tag{5.2}$$

where E[•] is a function of the expectation. This scaling process thus allows direct comparison between the top and bottom groups. The scaled distance offers an advantage that it can be applicable to any number of variables.

For the purposes of this paper, it is important to make sure whether or not the proposed idea of the SMD works for the capacitance data measured from the stator windings found in generators. Table 5-2 summarizes the properties of each group's data; top and bottom.

Table 5-2.The properties of top and bottom groups' data

Properties	Top Group	Bottom Group
The number of variables (or, the data size of group) (k)	3	2
The number of samples(n)	1,680	1,680
Expected value (μ)	2.88	1.96
Expected value of SMD^2 (μ/k)	0.96	0.98

As shown in Table 5-2, the expected value of the top group (CET and TET) is close to the number of variables in the top groups. Likewise, the expected value of the bottom group (CEB and TEB) approximately equals to the number of variables in the bottom groups. Thus, it can be concluded that the scaled distance measures can be uniformly used regardless of the group. The scaling idea can also be applied to Scaled Directional Mahalanobis Distance (SDMD), just like SMD.

5.3 Health Grade System

This section aims to define a health grade system in which the windings' health states are classified into diverse classes according to the health-relevant distance measure, SDMD.

Based upon the maintenance strategies for the stator windings and the opinions of field experts, three health classes are proposed: (1) a faulty condition (or, water-absorbed), (2) a warning condition (or, close to water absorption), and (3) a healthy condition (or, not water-absorbed). The failure listed in Table 5-1 which was caused by water absorption resulted in two meaningful data. These data can define a failure threshold for the distance measure, h, expressed as:

$$h = \left\lfloor k \cdot \mathrm{E}\left[\min\left[\mathrm{SDMD}^2\left(\tilde{\mathbf{X}}_{\mathbf{faulty}}\right)\right], \max\left[\mathrm{SDMD}^2\left(\tilde{\mathbf{X}}_{\mathbf{warning}}\right)\right]\right] + 0.5 \right\rfloor \quad (5.3)$$

where $\lfloor \bullet \rfloor$ denotes a round down function, E[•] denotes the expected value, k is the data size of each group (e.g. In case of top group, k = 3) and $\tilde{\mathbf{X}}_{\mathbf{faulty}}$ and $\tilde{\mathbf{X}}_{\mathbf{warning}}$ are capacitance data vectors obtained from the faulty or water-absorbed windings on each group, respectively. Sixty percent of the threshold value (0.6 h) is defined as a boundary line between the warning condition and the healthy condition, based upon field experts' experience and historic information on inspection and maintenance for stator windings.

Since the bottom group does not have any faulty winding in the maintenance history, the failure threshold of bottom group cannot be calculated from the equation 5.3. Thus the failure threshold of top group, h_{top}, is applied to define the failure threshold for the bottom bar, h_{bottom}, which can be defined by the following expression:

$$h_{\mathrm{top}} : h_{\mathrm{bottom}} = 3:2 \quad (5.4)$$

Finally, Table 5-3 summarizes the definition of the three health grades and suggested maintenance actions.

Table 5-3.Definition of health grades and related maintenance actions

Health Grade	**Range**		**Suggested Maintenance Actions**
	Top	**Bottom**	
Faulty	$\mathrm{DMD}^2 \geq 25$	$\mathrm{DMD}^2 \geq 16.7$	Immediate replacement
Warning	$15 < \mathrm{DMD}^2 < 25$	$10 < \mathrm{DMD}^2 < 16.7$	Frequent inspection
Healthy	$\mathrm{DMD}^2 < 15$	$\mathrm{DMD}^2 < 10$	No immediate maintenance required

5.4 Validation Study

In this section, the feasibility of the proposed SDMD-based health grade system is verified by comparison with the maintenance history of generator stator windings. Figure 5-1 shows the scatter plot of DMD with operating time. Let us examine the points highlighted with circles in the figure. These circles label the data obtained from the faulty winding or water absorbed windings. Each class contains several data points highlighted with circles, as shown in Table 5-4.

Table 5-4.Summary of the number of the data and the circled data in each grade

Health grade	The number of the data (A)	The number of the circled data (B)	B/A (%)
Faulty	4	4	100
Warning	15	15	100
Healthy	3,173	6	0.19

Most of the circled data points belong to either the “faulty” or “warning” class. This indicates that the proposed health grade system properly defines the health condition of the generator stator windings against water absorption.

In the faulty class, there are two cases (Case 1 and 2 in Figure 5-1) which can be split into four data points, although only one failure case was actually recorded in the maintenance history. In order to make sure if the failure threshold was correct, we looked at the maintenance history in detail.

The data points in the first case are obtained from the failed winding which was burned in 2008. According to the proposed index, DMD^2, a substantial increase of the index was found from 2006 to 2007. The index of this winding is equal to 3.05 in 2006 and 28.41 in 2007, respectively. This increase would have suggested by

DMD^2 that a preventative maintenance action be taken the year (2007) before the failure. This is a critically important observation in order to build the health grade system correctly. Thus, there is no problem that the data points observed from this winding belong to the faulty class.

On the other hand, the second case contains two data points which are measured from non-faulty winding. It seems that the winding studied here should be contained within the warning or healthy class. This is because of stator windings' electrical characteristics. A voltage is induced in the stator winding when the rotor is rotated. Typically, large synchronous generators are designed for a terminal voltage of several thousand volts. According to the opinion of experts in the power generation field, several windings per stator continuously retain zero-volts. This implies that these windings, like the second case's winding, may not have failed even though the water is absorbed enough to otherwise indicate a problem. In summary, it can be concluded based upon the above observation that replacement should be carried out on this winding in spite of its good external appearance. Actually, this winding was replaced in 2011.

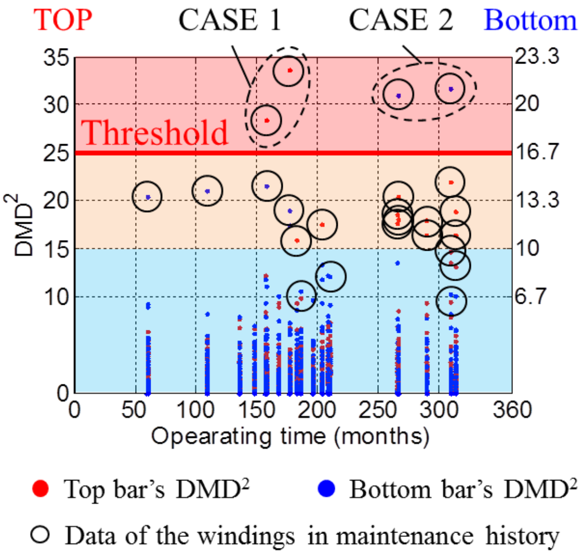


Figure 5-1.Scatter plot of DMD with operating time

Chapter 6. Conclusion

This paper explores a new health reasoning system to assess the health condition of stator windings in power generators. The proposed system extracts health-relevant features from the capacitance data in a statistical manner, to assess the health condition of power generator stator winding, to classify the health classes into three groups (healthy, warning, and faulty states), and to develop health grade system for condition-based maintenance. Correlation analysis of measured capacitance data is used to help understand the statistical features of the data and to divide the variables into four groups (CET, CEB, TET, and TEB) per winding. This paper proposes a statistical health measure Directional Mahalanobis Distance (DMD). DMD incorporates the correlation between variables and provides the degree of health condition considering the health degradation direction. Due to the unique capability of DMD to make use of the distance and degradation direction as a health measure, it can also be applicable to a health grade system designed to monitor a building. The health grade system outlined in this paper was developed with guidance from field maintenance records. Moreover, it takes into account the data size of each group with a scaling factor. This study employed the datasets from eight generators over eight years to validate the proposed health reasoning system. The proposed system can be generally applicable to health degradation trend analysis of engineered systems. In order to accomplish the condition-based maintenance system, health prognostics using machine learning techniques must be further studied.

Bibliography

[1] CHRISTER, A. H.; WALLER, W. M. Delay time models of industrial inspection maintenance problems. Journal of the Operational Research Society, 1984, 401-406.

[2] ZHAN, Yimin; MAKIS, Viliam; JARDINE, Andrew KS. Adaptive state detection of gearboxes under varying load conditions based on parametric modelling.Mechanical Systems and Signal Processing, 2006, 20.1: 188-221.

[3] CHINNAM, Ratna Babu; BARUAH, Pundarikaksha. A neuro-fuzzy approach for estimating mean residual life in condition-based maintenance systems.International Journal of Materials and Product Technology, 2004, 20.1: 166-179.

[4] LIN, Chang-Ching; TSENG, Hsien-Yu. A neural network application for reliability modelling and condition-based predictive maintenance. The International Journal of Advanced Manufacturing Technology, 2005, 25.1-2: 174-179.

[5] YANG, Junyan; ZHANG, Youyun; ZHU, Yongsheng. Intelligent fault diagnosis of rolling element bearing based on SVMs and fractal dimension. Mechanical Systems and Signal Processing, 2007, 21.5: 2012-

2024.

[6] CORTES, Corinna; VAPNIK, Vladimir. Support-vector networks. Machine learning, 1995, 20.3: 273-297.

[7] ALTMAN, N. S. An introduction to kernel and nearest-neighbor nonparametric regression. The American Statistician, 1992, 46.3: 175-185.

[8] ROKACH, Lior. Data mining with decision trees: theory and applications. World Scientific, 2007.

[9] DE MAESSCHALCK, Roy; JOUAN-RIMBAUD, Delphine; MASSART, Désiré L. The mahalanobis distance. Chemometrics and Intelligent Laboratory Systems, 2000, 50.1: 1-18.

[10] TAGUCHI, Genichi; WU, Yuin; CHODHURY, Subir. Mahalanobis-Taguchi System. McGraw-Hill Professional, 2000.

[11] WU, Jian-Da; HUANG, Chin-Wei; HUANG, Rongwen. An application of a recursive kalman filtering algorithm in rotating machinery fault diagnosis. NDT & E International, 2004, 37.5: 411-419.

[12] YANG, S. K. An experiment of state estimation for predictive maintenance using Kalman filter on a DC motor. Reliability engineering & system safety, 2002, 75.1: 103-111.

[13] ALGUINDIGUE, Isreal E.; LOSKIEWICZ-BUCZAK, Anna; UHRIG, Robert E. Monitoring and diagnosis of rolling element bearings using artificial neural networks. Industrial Electronics, IEEE Transactions on, 1993, 40.2: 209-217.

[14] EBERSBACH, S.; PENG, Z.; KESSISSOGLOU, N. J. The investigation of the condition and faults of a spur gearbox using vibration and wear debris analysis techniques. Wear, 2006, 260.1: 16-24.

[15] SNR, Dimla; DIMLA, E. Sensor signals for tool-wear monitoring in metal cutting operations—a review of methods. International Journal of Machine Tools and Manufacture, 2000, 40.8: 1073-1098.

[16] MARTIN, K. F. A review by discussion of condition monitoring and fault diagnosis in machine tools. International Journal of Machine Tools and Manufacture, 1994, 34.4: 527-551.

[17] BARTOLETTI, Cipriano, et al. Vibro-acoustic techniques to diagnose power transformers. Power Delivery, IEEE Transactions on, 2004, 19.1: 221-229.

[18] BENGTSSON, C. Status and trends in transformer monitoring. Power Delivery, IEEE Transactions on, 1996, 11.3: 1379-1384.

[19] HU, Chao, et al. Copula-based statistical health grade system against mechanical faults of power transformers. Power Delivery, IEEE Transactions on, 2012, 27.4: 1809-1819.

[20] BOOTH, Campbell; MCDONALD, Jim R. The use of artificial neural networks for condition monitoring of electrical power transformers. Neurocomputing, 1998, 23.1: 97-109.

[21] PARK, Chul-Won, et al. Generator fault detection technique using detailed coefficients ratio by daubechies wavelet transform. In: Power & Energy Society General Meeting, 2009. PES'09. IEEE. IEEE, 2009. p. 1-7.

[22] FINN, Joshua; WAGNER, John; BASSILY, Hany. Monitoring strategies for a combined cycle electric power generator. Applied Energy, 2010, 87.8: 2621-2627.

[23] KHEIRMAND, Akbar; LEIJON, Mats; GUBANSKI, S. M. Advances in online monitoring and localization of partial discharges in large rotating machines.Energy Conversion, IEEE Transactions on, 2004, 19.1: 53-59.

[24] STONE, G. C., et al. Impact of slot discharges and vibration sparking on stator winding life in large generators. Electrical Insulation Magazine, IEEE, 2008, 24.5: 14-21.

[25] JAYANTHA, G. A.; FERNANDO, M. A. R. M.; EKANAYAKE, C. M.

B. Monitoring of moisture on stator winding insulation during drying process. In:Industrial and Information Systems (ICIIS), 2009 International Conference on. IEEE, 2009. p. 483-488.

[26] JIA, Zhidong, et al. Evaluation of the degradation of generator stator ground wall insulation under multistresses aging. Energy Conversion, IEEE Transactions on, 2008, 23.2: 474-483.

[27] WORDEN, Joseph A.; MUNDULAS, Jorge M. Understanding, Diagnosing, and Repairing Leaks in Water-Cooled Generator Stator Windings. GE Reference Library, GER-3751A, 2003.

[28] KIM, Hee Soo; BAE, Yong Chae; KEE, Chang Doo. Wet bar detection by using water absorption detector. Journal of Mechanical Science and Technology, 2008, 22.6: 1163-1173.

[29] KIM, Hee-Dong. Analysis of Leak and Water Absorption Test Results for Water-Cooled Generator Stator Windings. Journal of Electrical Engineering & Technology, 2012, 7.2: 230-235.

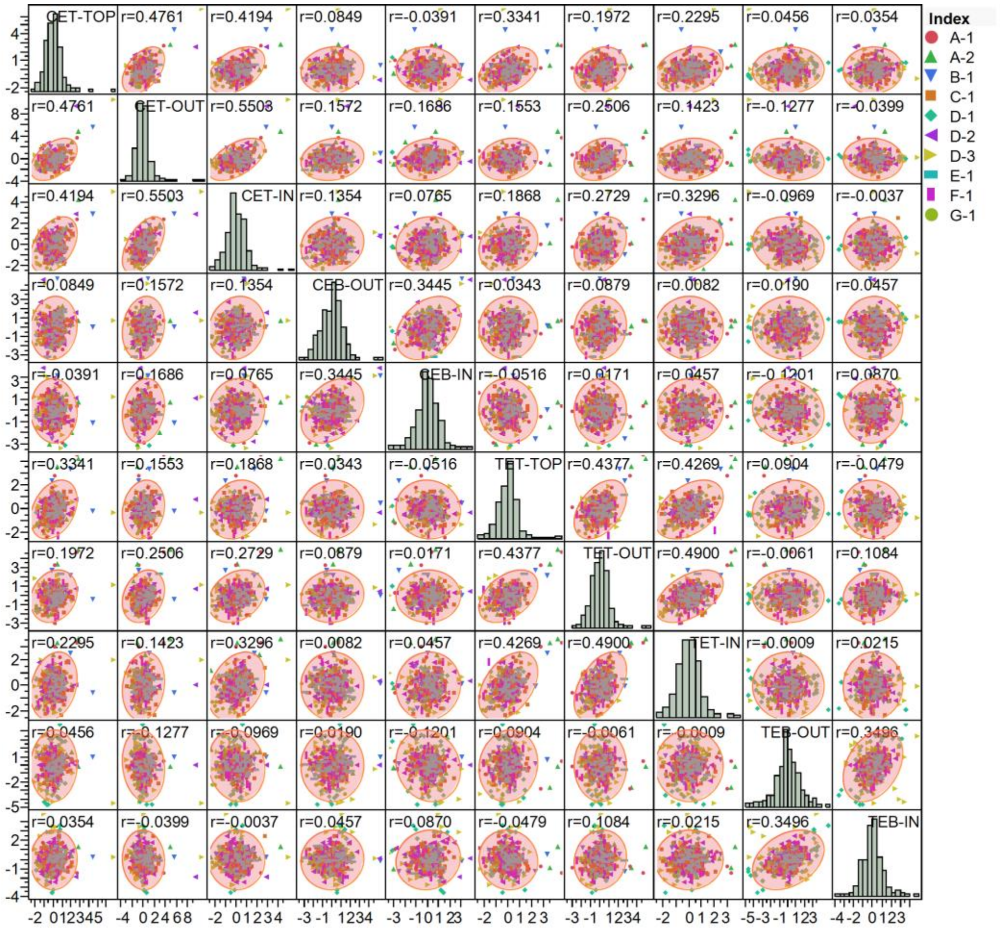
[30] KIM, Hee-Dong. Study on Reliability of Water Absorption Diagnosis through Precise Water Absorption Test. Journal of Electrical Engineering & Technology, 2012, 7.5: 772-777.

[31] INOUE, Y., et al. Technology for detecting wet bars in water-cooled

stator windings of turbine generators. In: Electric Machines and Drives Conference, 2003. IEMDC'03. IEEE International. IEEE, 2003. p. 1337-1343.

[32] HRDLE, Wolfgang; SIMAR, Léopold. Applied multivariate statistical analysis. Springer, 2007.

APPENDIX A.



Appendix A. Pairwise scatter plots between the ten measurements

국문 초록

발전소 내 가장 중요한 설비들 중 하나인 발전기는 일반적으로 운영 시간 혹은 사용 정도에 따라 유지 보수를 진행한다. 이런 종류의 유지 보수 형태는 분명 정확한 일정에 따라 움직일 수 있다는 장점이 있으나 아직 사용이 가능한 부품을 보수 한다거나 보수 하기 이전에 고장이 날 수 있는 단점을 항시 내재하고 있다. 최근 들어 건전성 예지 및 관리 분야에서는 앞서 언급한 단점들을 해결하기 위해 진단/예지 기술들을 활용하여 다양한 공학적 설비들의 현재 상태를 진단하고 잔여 수명을 예측하는 연구가 활발히 진행되고 있다. 본 논문에서는 발전기 고정자의 권선 절연체에서 발생하는 흡습에 따른 사고를 사전에 방지하기 위하여 정전용량을 측정하고 그 측정치로 하여금 현재 권선 절연체의 흡습 건전성을 평가하고자 한다. 본 논문에서 제안된 건전성 지수인 Directional Mahalanobis Distance (DMD)는 권선 절연체의 흡습 건전성을 확인하기에 최적화된 값으로써 정량적인 분석이 가능하게 해줄 뿐만 아니라 정전용량이 가지고 있는 특수성을 해결해줄 수 있다. 제안된 지수 DMD 는 전문가 측으로부터 받은 실제 발전기 고정자 권선의 흡습 의심 이력들을 바탕으로 하여 실증적 건전성 등급제를 구축하는 데에 사용된다. 마지막으로 본 연구에서 제안된 지능형 건전성 추론/진단 시스템은 8 개의 발전기로부터 8 년 동안 축적된 데이터를 이용하여 그 유효성을 입증하였다.

주요어: 발전기
고정자 권선

통계적 상관성

건전성 진단

디렉셔널 마하라노비스 거리

흡습

학 번: 2012-20664

감사의 글

나비^효과(--效果); [명사] 어느 한 곳에서 일어난 작은 나비의 날갯짓이 뉴욕에 태풍을 일으킬 수 있다는 이론. 초기 조건의 사소한 변화가 전체에 막대한 영향을 미칠 수 있음을 이르는 말. (국립 국어원, 표준국어대사전 발췌)

09 년 어느 날 저녁, 동아리 방에 처박혀 고장 난 로봇을 바라보며 친구와 나누던 대화가 있었습니다. “애(로봇)가 고장이 났는지 아닌지 미리 알 수 있는 방법이 없나?” 친구와 열띤 토론을 벌이며 각자 나름대로 방법도 생각해봤지만 실현되진 않았습니다. 그로부터 2 년 뒤, 대학원 진학을 앞두고 여러 연구실들을 알아보던 중 저의 시선을 멈추게 만든 것이 있었습니다. ‘고장 진단 및 예지.’ 인생에 있어 나비의 날갯짓 같았던 몇십 분의 친구와의 대화가 2 년 뒤의 저에게 와서 저의 진로를 바꾸는 태풍이 되는 순간이었습니다.

사실 많은 사람들이 본인이 원하는 연구 분야를 공부하기 위해 준비하는 기간이 짧지 않다고 합니다. 하지만 저는 ‘고장 진단 및 예지’ 분야에 대해 전혀 아는 것이 없었기에 그 기간이 짧았고 그만큼 부족했습니다. 그럼에도 불구하고 면담 이후에 이 분야에 대한 저의 관심과 제가 가지고 있는 가능성에 대해 믿어주신 윤병동 교수님 덕분에 저의 연구가 시작될 수 있었고, 그 내용들을 모아 이 논문을 완성시킬 수 있었습니다. 연구가 방황하지 않고 방향을 찾아갈 수 있도록 해주시고 그 과정에서 연구뿐만 아니라 인생 전반에 있어 필요한 말씀을 해주신 교수님께 이 자리를 빌려 깊은 감사의 말씀을 전합니다.

12 년도부터 2 년 간 동고동락을 함께 한 시스템 건전성 및 리스크 관리 연구실 식구들에게도 깊은 감사를 드립니다. 여러분들이 없었다면 이

논문은 완성되지 않았을 것입니다. 제가 궁금한 것이 있을 때 다가가도 본인 일처럼 알아봐주었고 제가 혼자 고민하고 있으면 비록 자신의 고민이 아니더라도 같이 고민해주는 사람들이 있었기에 저의 2 년 간의 연구실 생활은 고독하지 않았습니다. ‘멀리 가려면 함께 가라’ 라는 말이 있듯 저와 함께 해준 인연들이 있었기에 이렇게 멀리 올 수 있었습니다. 앞으로는 제가 떠남으로 인해 연구는 함께 하지 못하지만 인생에 있어서 함께 걸어나가는 인연이 되길 바랍니다.

저의 인생에 있어 가장 소중한 인연이라면 당연히 부모님과 저의 동생입니다. 부모님은 제가 하고자 하는 일이라면 항상 믿어주셨습니다. 고등학생 때에도, 대학생 때에도, 그리고 지금까지도 늘 저를 믿어주시는 부모님께서 계셨기에 자신감을 가지고 달릴 수 있었습니다. 동생 역시 형답지 않은 형을 믿어주고 항상 응원해주었기에 제게 큰 힘이 되었습니다. 아직 부족함이 많지만 그 성원에 보답할 수 있도록 언제나 좋은 모습 보여드리겠습니다. 표현을 잘 하지 않는 무뚝뚝한 아들이자 형이지만 모두 사랑한다는 말을 이 지면을 통해 전하고 싶습니다.

만약 저에게 또 다른 중요한 인연을 찾으라 한다면 친구들과 선후배들이 있습니다. 고등학교 친구들, 학부 친구들, 동아리 동기 및 선후배님들, 함께 봉사했던 친구들, 소프트웨어 멤버십 사람들 그리고 동문회 선후배님들 모두 저에게 연구 외적으로 물리적/정신적으로 도움을 주었던 사람들입니다. 모두 고맙습니다. 더불어 연구 내용과는 별개로 저의 졸업을 도와주셨던 모든 분들께 감사의 인사를 전합니다.

마지막으로 논문을 심사해주시기 위해 귀한 시간을 내어주신 조맹효 교수님과 김남호 교수님 (University of Florida), 그리고 마치 본인의 제자인 마냥 조언을 아끼지 않아주셨던 한봉태 교수님(University of Maryland)과 과제를 진행하면서 나이 어린 카운터파트에게 많은 조언을 해주셨던

한국 전력 전력연구원 김희수 부장님께도 감사 드립니다.

이제 학교를 떠나 사회에 첫 발을 내딛고자 합니다. 오바마 미 대통령이 말했듯 아직 최고의 순간은 오지 않았습니다(The best is yet to come.). 앞으로 더 좋은 날들이 지금부터 저에게 다가오겠지요. 그런 의미에서 오늘은 각자의 밝은 미래를 보고 나아가자는 의미로 다음 말을 끝으로 감사의 글 맺도록 하겠습니다. **See your tomorrow!**

감사합니다.

본 연구는 2013년도 정부(산업통상자원부)의 재원으로 한국에너지기술평가원(KETEP)의 지원을 받아 수행한 연구 과제입니다. (과제번호: 20121010100001C)