Statistical Health Reasoning of Water-Cooled Power Generator Stator Bars Against Moisture Absorption

Byeng D. Youn, Kyung Min Park, Chao Hu, Joung Taek Yoon, Hee Soo Kim, Beom Chan Jang, and Yong Chae Bae

***Abstract*—The power generator is typically maintained with a time- or usage-based strategy, which could result in a substantial waste of remaining useful life, high maintenance cost, and low plant availability. Recently, the field of prognostics and health management offers diagnostic and prognostic techniques to precisely assess the health condition and robustly predict the remaining useful life (RUL) of an engineered system, with an aim to address the aforementioned deficiencies. This paper explores a smart health reasoning system to assess the health condition of power generator stator bars against moisture absorption based on the statistical analysis of the capacitance measurements on bar insulators. In particular, a relative health measure, namely the directional Mahalanobis distance, is proposed to quantify the health condition of a stator bar. The smart health reasoning system is validated using five years' field data from seven generators, each of which contains 42 turns.**

***Index Terms*—Directional Mahalanobis distance (DMD), health diagnostics, moisture absorption, power generator stator bar, statistical correlation.**

I. Introduction

THE POWER generator is a critical power element in a power plant, since an unexpected breakdown of the generator leads to plant shut down and substantial economic and societal loss. Recently, tremendous technological advancements have been achieved in the development and deployment of an ultrasupercritical (USC) steam generator, which operates at an advanced steam temperature of 593 °C or above to achieve a higher energy conversion efficiency, and reduced fuel consumption and waste emission. However, the harsher operating condition of high temperatures (and pressures) of the advanced USC generator leads to 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 operational condition 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, health sensing function, health reasoning function, health prognostics function, and health management function, which are detailed as follows.

1) *Health sensing function:* The objective of this function is to ensure high damage detectability and efficient data management. Designing an optimal sensor network determines sensor types, sensor nodes, and their locations to maximize the degree of damage detectability, while designing data acquisition (DAQ) logistics decides the period and the time duration of DAQ for an optimal data management.
2) *Health reasoning function:* This function is the process of diagnosing health conditions based on sensory signals and related health measures. Two steps are typically involved: a) condition monitoring (CM) to extract system health relevant information with feature extraction techniques; b) health classification to classify a system health state into diverse health classes using health classification techniques. Popular tools used for this function include statistical methods [1], artificial intelligence [2], support vector machines [3], kernel estimation [4], and Mahalanobis distance (MD) [5]–[7]. CM has been applied to various engineering applications: machine components [8], [9], machine tools [10], and power transformers [11]–[13] or generators [14]–[20].
3) *Health prognostics function:* This function aims at predicting the remaining time before an engineered system no longer performs the required function(s) or the remaining useful life (RUL) in real time. In general, prognostics approaches can be categorized into model-based approaches [21], [22], data-driven approaches [23], [24] and hybrid approaches [25], [26]. PHM research has been conducted with various systems [27]–[30].
4) *Health management function:* This function enables optimal decision making on maintenance based on RUL predictions and/or field expert opinions. The condition-based maintenance is mainly involved in this function.

Comprehensive exploration of PHM techniques for power generator bars enables early anticipation of failure to develop cost-effective maintenance strategies and to seek opportunities for life extensions, and effective health reasoning is a crucial step toward the comprehensive exploration of PHM. Park *et al*.

Manuscript received August 18, 2014; revised December 7, 2014 and February 17, 2015; accepted June 8, 2015. Date of publication July 5, 2015; date of current version November 20, 2015. This work was supported in part by the Seoul National University-Institute of Advanced Machinery and Design (SNU-IAMD); in part by the Mid-Career Researcher Program under Grant NRF-2013R1A2A2A01068627, through the National Research Foundation (NRF) of Korea, funded by the Ministry of Science ICT and Future Planning; and in part by the Power Generation and Electricity Delivery Core Technology Program of the Korea Institute of Energy Technology Evaluation and Planning (KETEP) granted financial resource from the Ministry of Trade, Industry, and Energy, Republic of Korea under Grant 2012101010001C. Paper no. TEC-00587-2014.

B. D. Youn, K. M. Park, J. T. Yoon, and B. C. Jang are with the School of Mechanical and Aerospace Engineering, Seoul National University, Seoul 151-742, Korea (e-mail: bdyoun@snu.ac.kr; kyungmin@snu.ac.kr; kaekol@snu.ac.kr; bob1333@naver.com).

C. Hu is with the Department of Mechanical Engineering, Iowa State University, IA 50011, USA (e-mail: chaohu@iastate.edu).

H. S. Kim and Y. C. Bae are with the Korea Electric Power Research Institute, Daejeon 305-380, Korea (e-mail: saeromy1@kepco.co.kr; ycbae@kepri.re.kr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TEC.2015.2444873

applied the Daubechies wavelet transform to overcome the limits of a conventional discrete Fourier transform approach in winding-fault detection [15]. Kheirmand *et al*. developed the enhanced partial discharge (PD) technique by decoupling PD pulses using inductive sensors [16]. Stone *et al.* [17] reviewed the mechanisms of slot discharge and vibration sparking, and the methods to detect and repair them. Yue *et al*. proposed two health measure of the skewness of PD phase resolved distribution and the high-frequency crest of PD impulse based upon a multistress accelerating aging test of stator windings [18]. Fernando *et al.* found that the frequency-domain spectroscopy and estimated parameters are informative to assess the generator insulation condition with a laboratory generator bar experiment [19]. Li *et al.* used a partial least-square approach and a statistical parameter skewness to diagnose the stator bar insulation [20]. General electric (GE) suggests the standard outage test program [31] to indicate the presence of a water leakage through a water-cooled generator stator bar insulation and possible avert stator bar insulation failure. This failure can cause an extensive collateral damage to the generator stator core and field. This can extend a forced outage several months, depending upon the extent of the damage. This program consists of vacuum decay test, pressure decay test, Helium tracer gas test, and stator bar capacitance mapping. The stator bar capacitance mapping test developed by GE, measures the capacitance level on a bar surface and possibly detects the insulation failure due to moisture absorption. Alternative methods to the capacitance mapping include the normal probability plot method and the box plot method, developed by Korea Electric Power Corporation (KEPCO) [32]. The capacitance-based methods use the nondestructive capacitance measurements from a generator stator bar as sensory data. The first technique, known as the stator bar capacitance mapping test, is a quantitative health reasoning method that uses the $+3$ standard deviations of the capacitance data as a failure threshold. The second one is a graphical health reasoning method, which determines the health classes of a bar based on a normal probability plot. It is assumed in both methods that the capacitance data of healthy bar follow a normal distribution. The last method is another graphical method, so called the box plot, which graphically depicts the health classes of the capacitance data through their 1st and 3rd quartiles. The drawback of aforementioned methods is that the sensitivity of the bar health classification is relatively low because of the improper statistical modeling of the capacitance data and no consideration of data heterogeneity.

It is noted that the direct use of capacitance measurements as the health index by the existing methods makes it difficult to infer the health status easily and precisely, especially when the measurements are of high dimensionality, high correlation, and/or high nonlinearity. To this end, this work develops a new health index through statistical analysis of multidimensional capacitance measurements for an effective health reasoning of power generator bars.

The paper is organized as follows. Section II describes the sensing function for a power generator health monitoring and the analysis of bar capacitance data. Section III discusses the feature extraction and health diagnostics for smart health reasoning. Section IV presents the validation study using five years' maintenance history from seven generators. The paper is concluded in Section V.

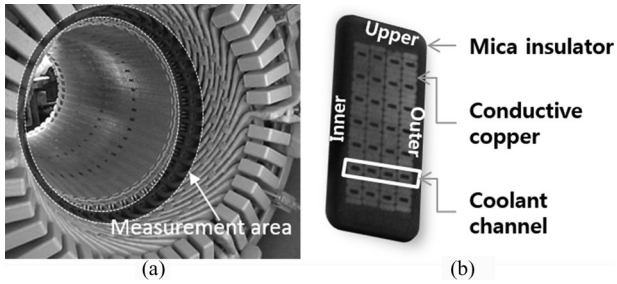


Fig. 1. (a) Power generator stator and (b) cross-sectional view of a bar.

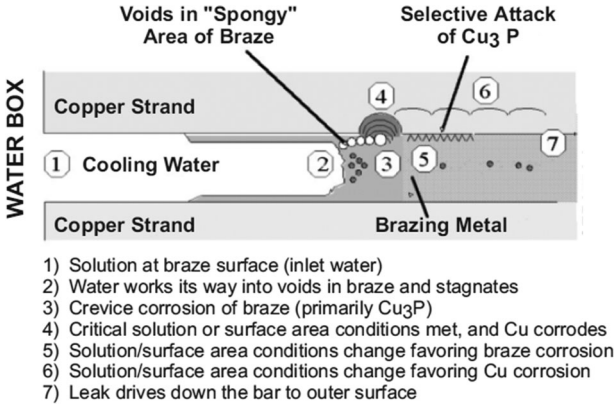


Fig. 2. Diagram of the crevice corrosion mechanism [31].

II. Description of Sensing Function and Data Analysis

The health condition of a power generator can be monitored by analyzing the capacitance of a bar insulator properly. This section discusses the fundamentals of capacitance measurements, locations of capacitance measurements, and characteristics of measurement data.

A. Fundamentals of Capacitance Measurements

When a power generator is water-cooled, coolant water flows into the water channels of a bar (see Fig. 1). When leakage occurs in insulators caused by failure mechanisms, i.e., crevice corrosion (see Fig. 2), the water or moisture remains in bar insulators. The remaining moisture degrades the bar insulation, which can cause an insulation breakdown and the power generator failure. For this reason, electric companies or manufacturing companies, such as KEPCO, GE, and Toshiba, assess the health status of the bar insulators using the moisture absorption detector [see Fig. 3(a)] [31]–[33]. The moisture absorption detector infers the moisture of the insulators by measuring an insulator capacitance. Because the relative static permittivity (or the dielectric constant) of the water ($\varepsilon_{\mathrm{water}(20^{\circ}\mathrm{C})} = 80.4$) is higher than that of the mica ($\varepsilon_{\mathrm{mica}} = 5.6 - 6.0$) which is used as a material for insulator generally, wet bar insulator has higher

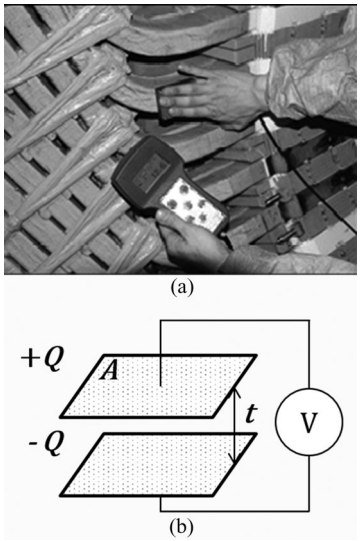


Fig. 3. (a) Capacitance measurement using detector and (b) the basic principle of detector.

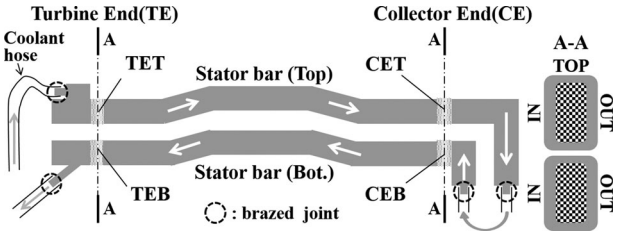


Fig. 4. Structure diagram of a water-cooled power generator with two-path cooling system and four brazed locations.

capacitance C based upon the following (see Fig. 3(b) for a schematic representation):

$$C = \frac{Q}{V} = \varepsilon_r \varepsilon_0 \frac{A}{t} \tag{1}$$

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 insulation, 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. Capacitance as health data provide valuable information that can be used to infer the amount of the moisture absorption of a stator bar and from which the health-relevant information of the bar can be extracted. It is noted that various uncertainty factors, such as measurement location, ambient humidity, and bar surface condition contribute to uncertainties in the capacitance measurements, which should be taken into account in the health reasoning process.

B. Capacitance Data Acquisition

This study employs seven power generators (ten datasets over five years) with the same specifications: 500-MW output, two-path cooling system, and 60-Hz frequency. The employed power generators have 42 turns and slots with a water-cooled cooling system. As shown in Fig. 4, the cooling water flows from the top bar inlet at the turbine end, through the top and bottom bars at the collector end, and back to the bottom bar outlet at the turbine end. At the turbine or collector end, an assembly slot with both the top and bottom bars contains totally ten measurement points graphically illustrated together with the generator structure diagram in Fig. 4. For the top bars, there are three measurement points (TOP, IN, and OUT), whereas 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 two measurement points for the bottom bar. In Table I, a unique identification (ID) code is assigned to ten measurement points based on the stator side (collector end/turbine end), bar location (top bar/bottom bar), and measurement point (top/out/in). For example, the ID code "CET-TOP" indicates that the measurement point is located on the top side of the collector end top bar. The capacitance data were acquired from the ten measurement points of the entire 42 slots of the power generator. The capacitance data measured at each measurement point can be modeled as a random variable (X).

C. 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}). Otherwise, the data can be modeled as statistically correlated random variables. For example, a physical gap between two different bars is one reason which makes the related random variables statistically independent. Moreover, different bar locations (CET, CEB, TET, and TEB) in one bar are also physically distant. This implies that the related random variables could be statistically independent as well. On the other hand, the moisture absorption occurs concurrently at adjacent measurement points, such as CET-TOP (X_1), CET-OUT (X_2), and CET-IN (X_3) in the same bar. Checking statistical dependence between two random variables could confirm whether our intuitive observation is true or not.

In general, the correlation coefficient is used as a measure to imply statistical correlation. One of the most famous measures 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:

$$\rho_{X_i,X_j} = \frac{\mathrm{Cov}\left(X_i, X_j\right)}{X_i X_j} = \frac{E\left[\left(X_i - \mu_{X_i}\right)\left(X_j - \mu_{X_j}\right)\right]}{\sigma_{X_i}\sigma_{X_j}} \tag{2}$$

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

Table I summarizes the correlation coefficients for ten random variables, ρ_{X_i,X_j} for i, $j = 1$ to 10, in a matrix form. The highlighted values in Table I are the coefficients between the correlated random variables in the same group. One can observe two features from the highlighted values: 1) statistically positive correlation, and 2) 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 (remain unchanged

TABLE I
CORRELATION COEFFICIENT MATRIX (SYMMETRIC) FOR TEN RANDOM VARIABLES IN A MATRIX FORM

	CET			CEB		TET			TEB	
Correlation Matrix	TOP (X_1)	OUT (X_2)	IN (X_3)	OUT (X_4)	IN (X_5)	TOP (X_6)	OUT (X_7)	IN (X_8)	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

or grow) with linear dependence. This confirmed the intuitive observations about the statistical correlation and independence aforementioned.

We note that several correlation coefficients between the measurement points in the CET group and those in the TET group are positive and of significant values (e.g., 0.3341 between CET-TOP and TET-TOP). It appears to be counterintuitive to have a significant positive correlation between two measurements points at two opposite ends (i.e., CE and TE) of a generator. A possible explanation of the phenomenon is a leakage of coolant water into the mica insulator (see Fig. 1) due to initial brazing defects at both ends (i.e., CE and TE) of the generator. The brazing defects could be caused by a misalignment (at TE) between the coolant hose and bar end of the generator (see Fig. 4). As the bar is highly stiff, the misalignment at one end (i.e., TE) of the generator would lead to a similar misalignment at the other end (i.e., CE), which would then cause an initial brazing defect at CE. This common-cause scenario, if it occurred, would produce relatively high-capacitance values on both ends of the generator, even though the two ends are physically distant from each other. It is also noted that compared to the top bars (i.e., CET and TET), the bottom bars (CEB and TEB) exhibit in general smaller correlation coefficients. This is possibly due to the fact that as the bottom bars are farther away from the coolant inlet than the top bars, the former carry lower pressure water coolant than the latter. As a result, the bottom bars are less likely to have coolant leakage into the mica-epoxy insulator, in the presence of the brazing defects, than the top bars.

D. Data Grouping

It is important to define a group of capacitance data with homogeneity prior to data modeling and health reasoning process. Based upon the measurement location and correlation characteristic obtained in Section II-C, the measurement points with high correlation can be conceived as individual data groups, such as CET, CEB, TET, and TEB. It implies that one entire dataset for ten random variables (or from ten-dimensional measurement points) would be split into four data groups (CET, CEB, TET, and TEB) 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. The data grouping makes the health reasoning process easier through the dimensional reduction of the capacitance data.

III. STATISTICAL HEALTH REASONING SYSTEM

Although the capacitance data are relevant to the health status of the stator bar, its high dimensionality and nonlinearity make it difficult to infer the health status easily and precisely. This section introduces a new health index, referred to as the directional Mahalanobis distance (DMD).

A. Review of MD

The 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 multidimension data (X) to 1-D distance measure, while taking into account the statistical correlation between random variables. Mathematically, the MD measure can be expressed as

$$\mathrm{MD}\left(\boldsymbol{X}_i\right) = \sqrt{\left(\boldsymbol{X}_i - \boldsymbol{\mu}\right)^T \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{X}_i - \boldsymbol{\mu}\right)} \tag{3}$$

where $\boldsymbol{X}_i = \left(X_{1,i}, X_{2,i}, \ldots, X_{N,i}\right)^T$ is an N-dimensional capacitance data vector of the ith bar unit, which belongs to a group having the mean $\boldsymbol{\mu} = \left(\mu_1, \mu_2, \ldots, \mu_N\right)^T$ and the covariance matrix $\boldsymbol{\Sigma}$. Fig. 5(a) plots 2-D samples randomly drawn from two random variables with a positive correlation. Essentially, the MD transforms an ellipsoid in the original random space [see Fig. 5(a)] to a circular shape in the standard Gaussian space [see Fig. 5(b)]. 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 [see Fig. 5(a)], one could have concluded based on 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 of the correlation coefficient of the two random variables.

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 1-D measure capable of easy comprehension and quick computation.

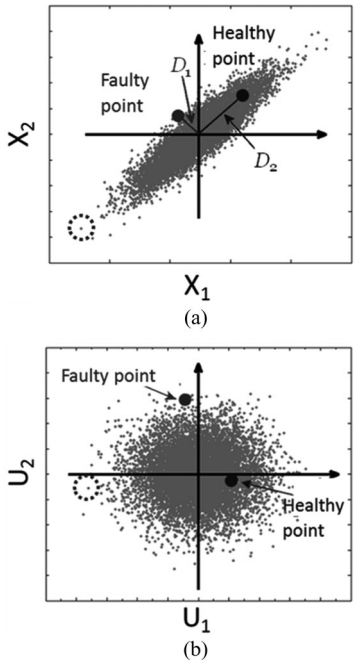


Fig. 5. Healthy and faulty points located in (a) the original space and (b) the normalized space.

2) The MD is robust to the 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 intervariable changes in multivariate measurements.

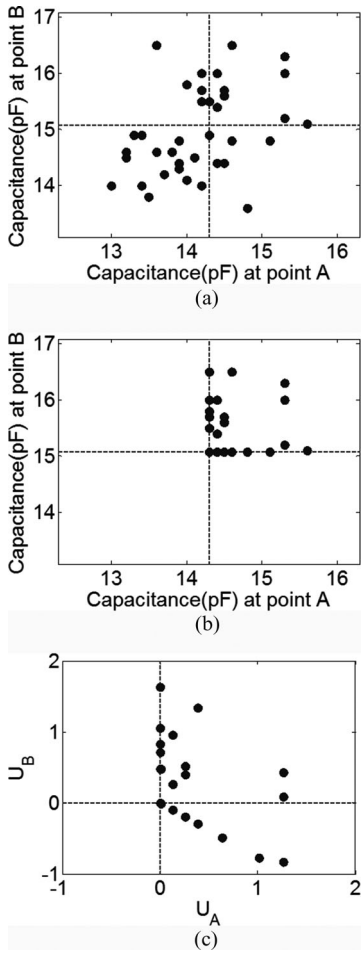


Fig. 6. Scatter plots (a) before projection, (b) after projection, and (c) after transformation.

B. New Concept of Statistical Distance: DMD

This section introduces a new MD-based distance measure that implies the health condition of a power generator stator bar.

1) Data Projection: The MD, as a relative health measure, provides very useful information to characterize the health condition of a power generator stator bar. According to (1), the measured capacitance values from a dry stator bar with a negligible amount of moisture on the insulation should be smaller than the mean value of the measurement population. It means that the measured values smaller than the population mean should be treated as of no relation to the moisture absorption by an insulator. However, the MD, as a scalar distance measure, is a direction-independent health measure in the random capacitance space [see Fig. 6(a)]. 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 the moisture absorption.

Let us take the data point marked with a dashed circle in Fig. 5 as an example. In this example, the data point is a healthy one since this point falls into the lower tails of the marginal distributions of the random capacitance variables. However, the MD declares this point to be in the failure because it is simply out of the data cluster. For the very reason, it becomes necessary to redefine the measure that is better suited for this application. In order to rebuild the measure, this study employs a projection process, which first identifies absolutely healthy variable(s) (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, X_j) and (μ_i, X_j) shown in Fig. 6(a) and (b). Through this projection process, the absolutely healthy data would be ignored in the subsequent MD transformation. The data projection underscores the consideration of the direction in the health reasoning process of the measurement data and leads to the unique capability of a new approach that makes use of the distance and degradation direction as a health measure.

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

$$\tilde{X}_{n,i} = \begin{cases} X_{n,i}, & \text{if } X_{n,i} > \mu_n \\ \mu_n, & \text{otherwise} \end{cases} \tag{4}$$

where $X_{n,i}$ denotes the raw capacitance data at the *n*th measurement location of the ith bar unit, μ_n is the mean of the capacitance data at the nth measurement location, and $\tilde{X}_{n,i}$ denotes the processed capacitance data.

2) MD Transformation: A proposed health index, namely DMD, assesses MD along the degradation direction of a stator

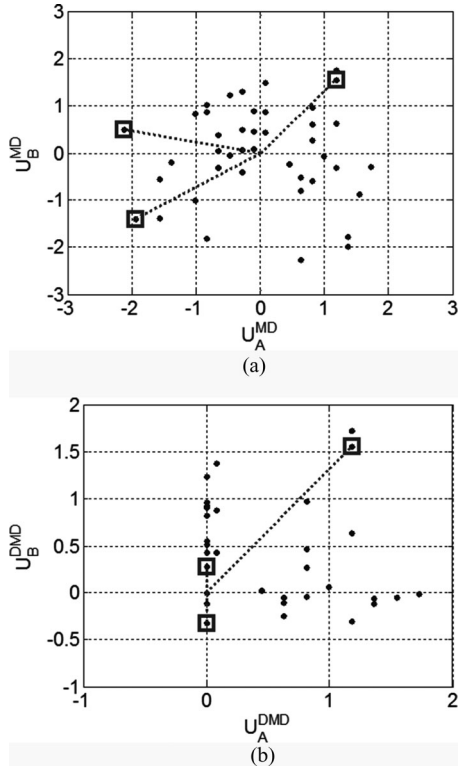


Fig. 7. Data points located in (a) the original and (b) the normalized space.

bar insulator after the data projection. Mathematically, DMD shares a similar formula with MD except the data projection, expressed as

$$\mathrm{DMD}\left(\tilde{\boldsymbol{X}}_i\right) = \sqrt{\left(\tilde{\boldsymbol{X}}_i - \boldsymbol{\mu}\right)^T \boldsymbol{\Sigma}^{-1} \left(\tilde{\boldsymbol{X}}_i - \boldsymbol{\mu}\right)} \qquad (5)$$

where $\tilde{\boldsymbol{X}}_i = \left(\tilde{X}_{1,i}, \tilde{X}_{2,i}, \ldots, \tilde{X}_{N,i}\right)^T$ is an *N*-dimensional vector of the capacitance data from the ith bar unit after the data projection $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_N)^T$ and $\boldsymbol{\Sigma}$ are the mean vector and covariance matrix of the reference dataset before the projection. Fig. 6(c) shows the scatter plots of DMD dataset after the projection and MD transformation where the difference between MD and DMD can be clearly observed.

One question still remains in the sequence of the data projection and MD transformation. We found that the former should be done prior to the latter. This is because the projection maps healthy data located lower than a data mean onto the mean is physically meaningful and valid only in the original space not in the transformed space.

C. Comparison of MD and DMD

Fig. 7 shows the scatter plots 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 2-D space composed of by two Euclidean distances [see Fig. 7(a)], while

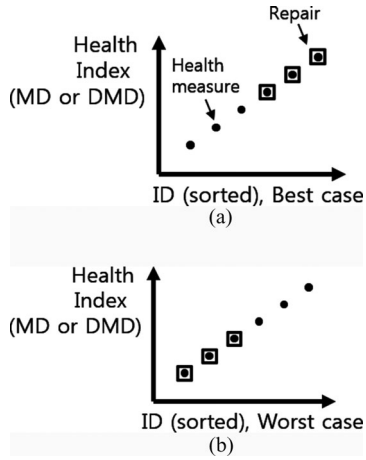


Fig. 8. (a) Best case and (b) the worst case of health index ranking and maintenance history and corresponding score metric ranking.

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, 4.73, 5.71, and 3.80. In the case of DMD, the squared distances of the three points from the origin are 0.076, 0.100, and 3.80. Without the projection before transformation, the MD incorrectly treats the "healthy" data points as "faulty," while with the projection before transformation, the proposed 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.

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

$$\mathrm{SF} = 100 \times \frac{\sum_{j=1}^{N} x_j y_j - \underbrace{\left(\sum_{j=1}^{k} x_j - \sum_{j=k+1}^{N} x_j\right)}_{W}}{\underbrace{\left(\sum_{j=N-k+1}^{N} x_j - \sum_{j=1}^{N-k} x_j\right)}_{B} - \underbrace{\left(\sum_{j=1}^{k} x_j - \sum_{j=k+1}^{N} x_j\right)}_{W}} \qquad (6)$$

where x_j denotes the health index value of the jth bar unit, y_j denotes the maintenance index of the jth bar based on the true repar history ($y_j = 1$ if the unit was maintained and $y_j = -1$ otherwise), *N* denotes the number of bar units, *k* denotes the number of bar units with maintenance histories, and *B* and *W* denote the best and the worst score metric values, respectively. Fig. 8 illustrates various combinations of the health index

TABLE II
TOP TEN SCORES OF MD AND DMD IN EACH GROUP

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	18.93	18.93	11.65	11.92
2	28.41	28.41	14.36	14.36	17.93	17.93	9.97	8.08
3	21.96	21.96	14.27	12.68	15.88	15.88	9.64	7.37
4	16.42	16.42	13.24	11.16	14.65	14.65	8.67	6.86
5	16.39	16.39	12.68	6.75	14.11	13.57	8.08	5.74
6	12.55	9.51	11.05	6.70	13.98	13.15	7.66	5.49
7	11.47	9.41	10.85	6.08	13.55	9.33	7.48	5.48
8	11.46	9.36	9.67	5.19	13.15	8.91	7.37	4.74
9	11.43	7.97	9.51	5.03	11.56	8.09	7.19	4.64
10	10.85	7.31	3.97	4.87	11.11	6.80	6.86	4.53
Score	96.32	**99.96**	97.47	**100**	94.10	**95.50**	91.26	**100**

ranking and the maintenance history and the score metric ranking of these combinations. The best and worst scenarios (represented respectively by B and W in (6)) are depicted by the left-most and right-most plots, respectively.

Table II summarizes the top ten distances and scores from the assessment results for health condition of bars using MD and DMD, respectively. As one can see in Table II, both MD and DMD can find maintained bars easily when the capacitance of maintained bar's insulator is relatively larger than others. However, the MD cannot find all maintained bars in top ten health indices in CET group, as well as some unrepaired bars' health indices have higher value than those of maintained (e.g., 5th, 6th and 7th health indices in TET group).

IV. VALIDATION STUDY WITH MAINTENANCE HISTORY

This section is designed to demonstrate the effectiveness of the proposed health index by assessing the health condition of bars and compare the DMD health reasoning results with historical maintenance data in Section IV-A, and, then, the MD results in Section IV-B.

To verify the effectiveness of the proposed health index, it is important to validate the DMD health reasoning results with maintenance history for the generators of interest. The field experts collected the health information of generator bars with maintenance records, and identified their health conditions using the stator bar capacitance mapping test method that was developed by GE [31].

As aforementioned, the capacitance data were grouped into CET, CEB, TET, and TEB. Based upon this grouping, the maintenance histories can also be divided. Table III summarizes the histories of the generators over five years, which are related to the moisture absorption. The maintained bar history is named with its own ID code ("a" to "n"). For example, the CET group in one bar of the first generator in the power plant A was maintained in 2009 for preventive maintenance, and this history is indexed with ID code "a." The TET in the same bar was maintained twice in 2009 and two indices ("b" and "c") appear in Table III.

As shown in Table III, 14 maintenance histories (a–n) are found in total over five years—five (a, d, e, i, l) in the CET group, three (f, j, n) in the CEB group, four (b, c, g, m) in the TET group, and two (h, k) in the TEB group. Fig. 9 shows the scatter plots of the health index (DMD^2), where a black-square labels the health index of a bar unit right before the maintenance, a gray circle labels the health index of that unit one year before the maintenance, and a black triangle labels the health index of a bar unit with no maintenance history but with a relatively high ranking. As shown in Table III, three maintained bar units (with ID codes a, b, and c) of the 1st generator in the power plant A and three units (with ID codes i, j, and k) of the 4th generator in the same power plant possess the health index data one year prior to the year of maintenance and are labeled with both red and orange circles.

TABLE III
MAINTENANCE HISTORY RELATED TO MOISTURE ABSORPTION

Power plant	Generator	Measurement year	Measurement group	ID code (number of maintained bar)
A	1st	2007	–	–
		2009	CET	a (20)
			TET	b (11), c (20)
	2nd	2009	CET	d (3), e (40)
			CEB	f (31)
			TET	g (40)
	3rd	2010	TEB	h (23)
	4th	2006	–	–
		2007	–	–
		2008	CET	i (23)
			CEB	j (20)
			TEB	k (31)
B	1st	2010	CET	l (18)
			TET	m (18)
	2nd	2009	–	–
C	4th	2010	CEB	n (11)

Overall speaking, most of the high-ranked values are labeled with the red squares in all groups. This indicates that the proposed health index in this paper assesses the health condition of the generator bars consistently, as compared to the maintenance history. An interesting example worth further elaboration here is the 1st ranked bar unit (i) in the CET group. This bar unit was burned in 2008 because none of the conventional inspection and maintenance techniques (leak test, capacitance mapping test, normal probability plot method, and box plot) were able to detect the moisture absorption. This bar unit has a 98th ranked health index ($= 3.05$) in 2006, a 2nd ranked one ($= 28.41$) in 2007, and a 1st ranked one ($= 33.64$) in 2008 among all the bar units in the CET group. According to the proposed health index (DMD^2), the substantial increase of the health index was found from 2006 to 2007. 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 since the proposed health index is capable of detecting abnormal health degradation and enabling health prognostics for failure anticipation and prevention.

It can be observed that some health indices with no maintenance history hold relatively high rankings (see the black triangles in Fig. 9). One case is the 7th ranked health index ($= 9.41$) in the CET group, and the other is ranked at the 5th ($= 13.57$) in

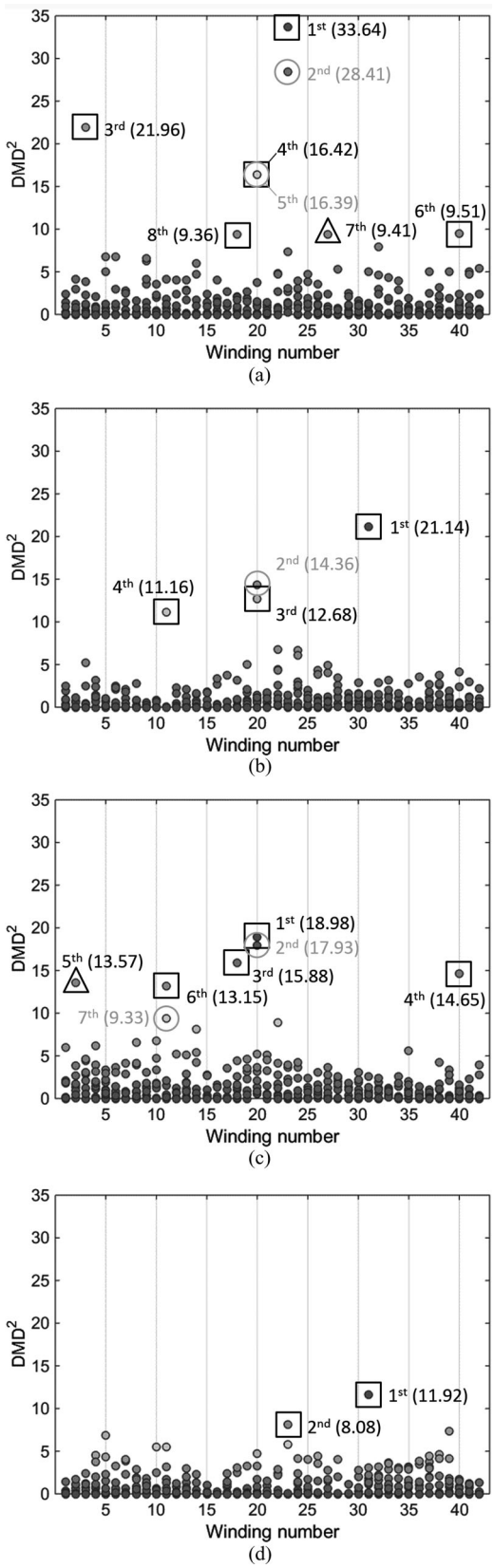


Fig. 9. Scatter plots of the health index DMD^2. (a) CET group. (b) CEB group. (c) TET group. (d) TEB group.

the TET group. Let us take a close look at these two cases carefully. The capacitance of the 7th ranked bar at the measurement point CET-TOP (X_1) reads 16.5 pF. Given the sample mean and standard deviation being 14.9 and 0.6 pF, respectively, 16.5 pF indicates that the winding is highly moisture absorbed and prone to failure. The same analysis can be reached for the 5th ranked bar in the TET group of which the capacitance reads 15.5 pF with the sampled mean and standard deviation being 12.8 and 0.9 pF, respectively. It can be concluded based on the proposed health index (DMD^2) that maintenance should be carried out on these bars.

V. Conclusion

This paper explores the health reasoning function to assess the health condition of power generator stator bars by extracting the health-relevant features from the capacitance measurements on the bars in a statistical manner.

This assessment is based on the fact that it is intuitive to use the capacitance value to infer the moisture absorption in a stator bar, and the change in the capacitance value and the rate of change in this value are good predictors for anticipating the stator bar failure due to the moisture absorption. First, the correlation analysis of the capacitance measurements helps to understand the statistical features of the data and divide the variables into four groups. Second, a statistical health index, the DMD, is proposed to quantify the health condition of the bar unit, while taking into consideration both the correlation between multivariate measurements, and the degradation direction. The effectiveness of the proposed health index is verified using the five years' maintenance records from seven generators. A good agreement between the health index ranking and the maintenance history suggests that the proposed health index is capable of precisely assessing the health condition of power generator stator bars. Future research will investigate the integration of the proposed health index with advance machine learning techniques in order to realize failure prognostics and prevention for power generator stator bars.

References

[1] Y. Zi, Y. Zhan, and A. K. S. Jardine, "Adaptive state detection of gearboxes under varying load conditions based on parametric modeling," *Mech. Syst. Signal Process.*, vol. 20, no. 1, pp. 188–221, 2006.

[2] C. C. Lin and H. Y. Tseng, "A neural network application for reliability modeling and condition-based predictive maintenance," *Int. J. Adv. Manuf. Technol.*, vol. 25, nos. 1/2, pp. 174–179, 2005.

[3] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[4] N. S. Altman, "An Introduction to kernel and nearest-neighbor nonparametric regression," *Amer. Statist.*, vol. 46, no. 3, pp. 175–185, 1992.

[5] H. Oh, M. H. Azarian, and M. Pecht, "Estimation of fan bearing degradation using acoustic emission analysis and mahalanobis distance," presented at the The Applied System Health Management Conference, Virginia Beach, VA, USA, 2011.

[6] Y. Wang, Q. Miao, and M. Pecht, "Health monitoring of hard disk based on Mahalanobis distance," in *Proc. Progn Syst. Health Manage. Conf.*, Shenzhen, China, 2011, pp. 1–8.

[7] S. Kumar, T. W. S. Chow, and M. Pecht, "Approach to fault identification for electronic products using Mahalanobis distance," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 8, pp. 2055–2064, Aug. 2010.

[8] I. E. Alguindigue, A. Loskiewicz-Buczak, and R. E. Uhrig, "Monitoring and diagnosis of rolling element bearings using artificial neural networks," *IEEE Trans. Ind. Electron.*, vol. 40, no. 2, pp. 209–217, Apr. 1993.
[9] S. Ebersbach, Z. Peng, and N. J. Kessissoglou, "The investigation of the condition and faults of a spur gearbox using vibration and wear debris analysis techniques," *Wear*, vol. 260, nos. 1/2, pp. 16–24, 2006.
[10] K. F. Martin, "Review by discussion of condition monitoring and fault diagnosis in machine tools," *Int. J. Mach. Tools Manuf.*, vol. 34, no. 4, pp. 527–551, 1994.
[11] C. Bartoletti, M. Desiderio, D. Di Carlo, G. Fazio, F. Muzi, G. Sacerdoti, and F. Salvatori, "Vibroacoustic techniques to diagnose power transformers," *IEEE Trans. Power Del.*, vol. 19, no. 1, pp. 221–229, Jan. 2004.
[12] C. Hu, P. Wang, B. D. Youn, W.-R. Lee, and J. T. Yoon, "Copula-based statistical health grade system against mechanical faults of power transformers," *IEEE Trans. Power Del.*, vol. 27, no. 4, pp. 1809–1819, Oct. 2012.
[13] C. Booth and J. R. McDonald, "The use of artificial neural networks for condition monitoring of electrical power transformers," *Neurocomputing*, vol. 23, pp. 97–109, 1998.
[14] J. Finn, J. Wagner, and H. Bassily, "Monitoring strategies for a combined cycle electric power generator," *Appl. Energy*, vol. 87, no. 8, pp. 2621–2627, 2010.
[15] C.-W. Park, K.-C. Shin, S.-S. Lee, J.-K. Park, and M.-C. Shin, "Generator fault detection technique using detailed coefficients ratio by daubechies wavelet transform," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, 2009, pp. 1–7.
[16] A. Kheirmand, M. Leijon, and S. M. Gubanski, "Advances in online monitoring and localization of partial discharges in large rotating machines," *IEEE Trans. Energy Convers.*, vol. 19, no. 1, pp. 53–59, Mar. 2004.
[17] G. C. Stone, C. V. Maughan, D. Nelson, and R. P. Schultz, "Impact of slot discharges and vibration sparking on stator winding life," *IEEE Elect. Insul. Mag.*, vol. 24, no. 5, pp. 14–21, Sep./Oct. 2008.
[18] B. Yue, X. Chen, Y. Cheng, J. Song, and H. Xie, "Diagnosis of stator winding insulation of large generator based on partial discharge measurement," *IEEE Trans. Energy Convers.*, vol. 21, no. 2, pp. 387–395, Jun. 2006.
[19] M. A. R. M. Fernando, W. M. L. B. Naranpanawa, R. M. H. M. Rathnayake, and G. A. Jayantha, "Condition assessment of stator insulation during drying, wetting and electrical ageing," *IEEE Trans. Dielect. Elect. Insul.*, vol. 20, no. 6, pp. 2081–2090, Dec. 2013.
[20] R. Li, L. Pan, C. Yan, H. Li, and B. Hu, "Condition evaluation of large generator stator insulation based on partial discharge measurement," *Adv. Mech. Eng.*, vol. 2013, pp. 765374-1–765374-10, 2013.
[21] J. Luo, K. R. Pattipati, L. Qiao, and S. Chigusa, "Model-based prognostic techniques applied to a suspension system," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 38, no. 5, pp. 1156–1168, Sep. 2008.
[22] N. Gebraeel and J. Pan, "Prognostic degradation models for computing and updating residual life distributions in a time-varying environment," *IEEE Trans. Rel.*, vol. 57, no. 4, pp. 539–550, Dec. 2008.
[23] M. Schwabacher, "A survey of data-driven prognostics," presented at the AIAA Infotech at Aerospace Conference, Arlington, VA, USA, 2005.
[24] E. Zio and F. Di Maio, "A data-driven fuzzy approach for predicting the remaining useful life in dynamic failure scenarios of a nuclear power plant," *Rel. Eng. Syst. Saf.*, vol. 95, no. 1, pp. 49–57, 2010.
[25] K. Goebel, N. Eklund, and P. Bonanni, "Fusing competing prediction algorithms for prognostics," presented at the IEEE Aerospace Conference, New York, NY, USA, 2005.
[26] B. Saha, K. Goebel, S. Poll, and J. Christophersen, "Prognostics methods for battery health monitoring using a Bayesian framework," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 2, pp. 291–296, Feb. 2009.
[27] A. Yanagisawa and T. Kojima, "Degradation of InGaN blue light-emitting diodes under continuous and low-speed pulse operations," *Microelectron. Rel.*, vol. 43, no. 6, pp. 977–980, 2006.
[28] N. G. Shrive, "Intelligent structural health monitoring: A civil engineering perspective," in *Proc. IEEE Syst., Man Cybern., Int. Conf.*, 2005, vol. 2, pp. 1973–1977.
[29] P. Wang, B. D. Youn, and C. Hu, "A generic probabilistic framework for structural health prognostics and uncertainty management," *Mech. Syst. Signal Process.*, vol. 28, pp. 622–637, 2013.
[30] C. Hu, B. D. Youn, P. Wang, and J. T. Yoon, "Ensemble of data-driven prognostic algorithms for robust prediction of remaining useful life," *Rel. Eng. Syst. Saf.*, vol. 103, pp. 120–135, 2012.
[31] J. A. Worden and J. M. Mundulas, "Understanding, diagnosing, and repairing leaks in water-cooled generator stator windings," *GE Power Syst.*, GER-3751A, 2001.
[32] H. S. Kim, Y. C. Bae, and C. D. Kee, "Wet bar detection by using water absorption detector," *J. Mech. Sci. Technol.*, vol. 22, no. 6, pp. 1163–1173, 2008.
[33] Y. Inoue, H. Hasegawa, S. Sekito, M. Sotodate, H. Shimada, and T. Okamoto, "Technology for detecting wet bars in water-cooled stator winding of turbine generators," in *Proc. IEEE Int. Elect. Mach. Drives Conf.*, 2003, vol. 2, pp. 1337–1343.

Byeng D. Youn received the Ph.D. degree from the Department of Mechanical Engineering, University of Iowa, Iowa City, IA, USA, in 2001.

He was a Research Associate with the University of Iowa, until 2004; an Assistant Professor with Michigan Technical University, Houghton, MI, USA, until 2007; and an Assistant Professor with the University of Maryland, College Park, MD, USA, until 2010. He is currently an Associate Professor with the School of Mechanical and Aerospace Engineering, Seoul National University, Seoul, Korea. His current research interests include system risk-based design, prognostics and health management, and energy harvester design.

Dr. Youn's research and educational dedication has led him to many notable awards, including the winner of the international PHM competitions in 2014 hosted by the IEEE PHM and PHM Society, and the ISSMO/Springer Prize for the Best Young Scientist in 2005.

Kyung Min Park received the B.S. degree in mechanical engineering from Korea University, Seoul, Korea, in 2012, and the Master's degree in mechanical engineering from Seoul National University, Seoul, in 2014.

His research interests include prognostics and health management.

Chao Hu received the B.E. degree in engineering physics from Tsinghua University, Beijing, China, in 2007, and the Ph.D. degree in mechanical engineering from the University of Maryland, College Park, MD, USA, in 2011.

He worked as a Principle Scientist at Medtronic, Inc., MN, USA, from 2011 to 2015. He is currently an Assistant Professor in the Department of Mechanical Engineering, Iowa State University, Ames, IA, USA. His research interests include reliability-based design, prognostics and health management, and design of Li-ion rechargeable battery. His research work has led to more than 50 publications in the above areas.

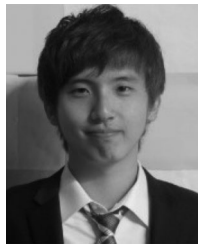
Dr. Hu has received several awards and recognitions, including the Star of Excellence Individual Award at Medtronic in 2014; the Best Paper Award at the American Society of Mechanical Engineering (ASME) Design Automation Conferences (DAC) in 2013; a nomination for the 2013 Eni Award Renewable Energy Prize, known as the Nobel Prize in Energy Sector; the Best Paper Award at the IEEE PHM Conference in 2012; and he is a two-time recipient of the Top 10 Best Paper Award at the ASME DACs in 2011 and 2012.

Joung Taek Yoon received the B.S. degree in mechanical engineering from Seoul National University, Seoul, Korea, in 2011, where he is currently working toward the combined Master's and Doctorate degrees in mechanical engineering.

His research interests include prognostics and health management, and resilient system design.

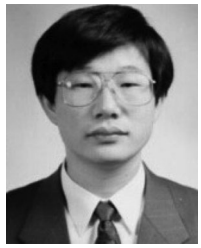
Hee Soo Kim received the Ph.D. degree from the Department of Mechanical Engineering, Chonnam National University, Gwangju, Korea, in 2009.

Since 1995, he has been working with the Korea Electric Power Research Institute, Daejeon, Korea. His research interests include vibration analysis/testing and mechanical integrity evaluation of generator stator windings.



Beom Chan Jang received the B.S. degree in mechanical engineering from Seoul National University, Seoul, Korea, in 2014, where he is currently working toward the combined Master's and Doctorate degrees in mechanical engineering.

His research interests include prognostics and health management.



Yong Chae Bae received the Ph.D. degree in mechanical engineering from Chonnam National University, Gwangju, Korea, in 1995.

Since 1986, he has been working with the Korea Electric Power Corporation Research Institute, Daejeon, Korea. He is currently a Chief Researcher with the Power Generation Lab., KEPCO Research Institute, Daejeon. His research interests include vibration analysis, mechanical integrity of generator stator windings, and monitoring and diagnostics of main equipment in power plant.