Reliability

February 2015

Special Issue on Prognostics and Health Management

IEEE

Reliability Society

Reliability Society Administrative Committee (AdCom) Member

OFFICERS(EXCOM)

President
Christian Hansen

Sr. Past President
Jeffrey Voas

Jr. Past President
Dennis Hoffman

Vice Presidents

Technical Activities
Shiuhpyng Winston Shieh

Publications
W. Eric Wong

Meetings and Conferences
Alfred Stevens

Membership
Marsha Abramo

Secretary
Scott Abrams

Treasurer
Bob Loomis

ELECTED MEMBERS-AT-LARGE (WITH VOTE)

TERM EXPIRES 2015 (DEC 31)		*TERM EXPIRES 2016 (DEC 31)*		*TERM EXPIRES 2017 (DEC 31)*	
Marsha Abramo	Shiuhpyng Winston Shieh	Lou Gullo	Zhaojun (Steven) Li	Joseph A. Childs	Carole Graas
Loretta Arellano	Alfred Stevens	Christian Hansen	Bob Loomis	Pierre Dersin	Samuel J. Keene
Lon Chase	Rex Sallade	Pradeep Lall	Pradeep Ramuhalli	Lance Fiondella	W. Eric Wong

STANDING COMMITTEES AND ACTIVITIES/INITIATIVES

Web Master
Lon Chase

Standards
Lou Gullo

Chapters Coordinator
Loretta Arellano

Professional Development
Marsha Abramo

Constitution and Bylaws
Dennis Hoffman

Fellows
Sam Keene

Academic Education/Scholarship
Sam Keene

Meetings Organization
Alfred Stevens

Membership Development
Marsha Abramo

Transactions Editor-in-Chief
Way Kuo

Newsletter Editor-in-Chief
Joe Childs

Video Tutorials
Christian Hansen

Nominations and Awards
Jeffrey Voas

Life Science Initiative
Peter Ghavami

Transportation Electification
Sam Keene
Pradeep Lall
Michael Austin

IEEE Press Liaison
Dev Raheja

IEEE GOVERNMENT RELATIONS COMMITTEES

Energy Policy
Jeff Voas

Transportation and Aerospace Policy
Scott Abrams

Medical Technology Policy
Jeff Voas

Critical Infrastructure Protection
Sam Keene

Career and Workforce Policy
Christian Hansen
(corresponding only)

Intellectual Property Committee
Carole Graas
(corresponding only)

Research and Development Committee
Pradeep Lall

Combined TAB Ad Hoc Committee on Attracting Industrial Members
Dennis Hoffman

2013 TAB Awards and Recognition Committee (TABARC)
Dennis Hoffman

Technical Committees

Technical Committee on Internet of Things (IoT)

Chair: Jeffrey M. Voas, National Institute of Standards and Technology

Email: jeff.voas@nist.gov

Co–chair: Irena Bojanova, Program Director, Telecommunications Management, and Collegiate Professor, University of Maryland, University College

Email: irena.bojanova@umuc.edu

Committee Member:

1. George F. Hurlburt: CEO of Change Index

Technical Committee on System and Software Assurance

Chair: Eric Wong, University of Texas at Dallas

Email: ewong@utdallas.edu

Technical Committee on Prognostics and Health Management (PHM)

Chair: Rex Sallade, Sikorsky Aircraft Co.

Email: Rex.Sallade@SIKORSKY.COM

Co–chair: Pradeep Lall, Auburn University

Email: lall@eng.auburn.edu

Technical Committee on Trustworthy Computing and Cybersecurity

Chair: Wen-Guey Tzeng, National Chiao Tung University

Email: wgtzeng@cs.nctu.edu.tw

Committee Member:

1. Raul Santelices: Assistant Professor, Department of Computer Science and Engineering, University of Notre Dam
2. Brahim Hamid: Associate Professor, IRIT Research Laboratory, University of Toulouse, France

Technical Committee on Reliability Science for Advanced Materials & Devices

Chair: Carole Graas, Colorado School of Mine & IBM Systems and Technology Group

Email: cdgraas@mines.edu

Technical Committee on Systems of Systems

Chair: Pierre Dersin, Alstom Transport

Email: pierre.dersin@transport.alstom.com

BIG Data Committee

Chair: David Belanger, Stevens Institute of Technology

Email: david.belanger@stevens.edu

Resilient Cyber-Systems Committee

Chair: Pradeep Ramuhalli, Pacific Northwest National Laboratory

Email: pradeep.ramuhalli@pnnl.gov

Standards Committee

Chair: Louis J Gullo, IEEE RS Standards

Email: Lou.Gullo@RAYTHEON.COM

Committee Member:

1. Ann Marie Neufelder: Softrel, LLC – Owner; IEEE P1633 Standard Working Group Chair
2. Lance Fiondella: Assistant Professor, Dept. of Electrical and Computer Engineering, University of Massachusetts Dartmouth; IEEE P1633 Standard Working Group Vice Chair
3. Steven Li: Assistant Professor, Industrial Engineering and Engineering Management, Western New England University; IEEE P61014 Standard Working Group Chair
4. Diganta Das: Research Staff at the Center for Advanced Life Cycle Engineering, University of Maryland
5. Sony Mathews: Engineer at Halliburton, IEEE P1856 Standard Working Group Chair
   Mike Pecht: Chair Professor and Director of Center for Advanced Life Cycle Engineering, University of Maryland
6. Arvind Sai Sarathi Vasan: Research Assistant at Center for Advanced Life Cycle Engineering , University of Maryland; IEEE P1856 Standard Working Group Vice-Chair
7. Joe Childs: Staff Reliability/Testability Engineer at Lockheed Martin

Working Group on Education

Chair: Zhaojun (Steven) Li, Western New England University

Email: zhaojun.li@wne.edu

Committee Member:

1. Emmanuel Gonzalez, Jardine Schindler Elevator Corporation

Editorial Board

Editorial Staff

Message from the Editor

Commercial electronic components are increasingly being used in safety critical high reliability applications with significantly longer life time than consumer electronics. Examples include automotive electronics such as lane departure warning systems, collision avoidance systems, antilock braking systems, airbag actuation systems and cruise control. There is need for methods for identification of imminent failure in electronics with sufficient advanced warning to allow for intervention for repair or replacement of the affected system. While diagnostics has been used in electronics for a number of years, the development of solutions for prognostic health management have yet to be developed and implemented although a few implementations do exist on some applications. Examples of diagnostics in consumer products include the on-board diagnostic system or the OBD which provides an indication of system fault with a dashboard indicator, and the built-in self-test or BIST which tests the systems for faults during power-up of the system. Diagnostic methods are limited to the detection of faults that have already occurred in the electronic system and do not provide in an indication of imminent failure in the system. In this issue, several papers are presented examining various aspects of prognostic health management for electronic systems have been presented.

Pradeep Lall, Walter Professor and Director, Auburn University
Editor, IEEE Reliability Digest Special Issue on PHM

Reliability Digest

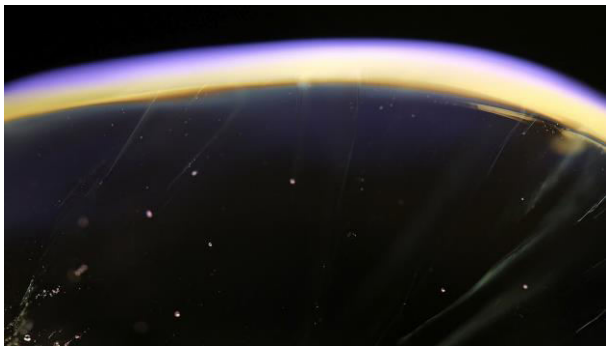
February Issue
Special Issue on Prognostics and Health Management

Special Issue on Prognostics and Health Management (PHM)

A Circuit Component Modeling Approach Based on Vehicle Control System Health Simulation

Zeng Qinghua, Zhao Wei
Science and Technology on Scramjet Laboratory
College of Aerospace Science and Engineering
National University of Defense Technology, China
z.qinghua@nudt.edu.cn, aaaa01234567890@qq.com

Jia Tao
China Aerodynamics R&D Center
tobejiatao@163.com



***Abstract*—The simulation of vehicle control system based on transfer function or nonlinear differential equation, which implicates the health state information of every component in the model parameters, can't directly describe physical character of component. With the emergence of a large number of powerful professional circuit simulation software and the mechanism-electric design & simulation software, a system-level health simulation framework is presented based on the component-level response approximation of the circuit simulating data. In this framework, the circuit schematic model based on the component health behavior is created first; then a lot of data under the various health degradation or fault is obtained using the professional circuit design and simulation software such as Multisim or Proteus; thus the response identification model of the BP neural network which approximates the health behavior of the component, the component-level model, is applied into the system-level simulation. In this case, the simulated results of system-level can be obtained under the condition of every component health behavior. The example shows that the response approximation model is capable of well imitating the various health behaviors of the component.**

***Keywords*—circuit fault simulation; response approximation method; health behavior model; system-level health simulation**

I. INTRODUCTION

Recently, a lot of attention and positive responses from the international community are obtained on the research on Integrated Vehicle Health Management (IVHM) system [1-4]. With the help of this technology, a lot of physical mechanism becomes a white box, which leads to corresponding professional simulation software becoming mature, such as the Adams for mechanical power simulation, the Fluent for computational fluid dynamics simulation. Meanwhile, the corresponding fault simulation technology is also booming. Especially in the areas of circuit design and analysis, the simulation software was improved greatly. The corresponding circuit fault simulation is becoming mature. So it is possible to provide a lot of health or fault information of electronic components in vehicle control systems [5]. However, with the development of the electronic technology, the electromechanical systems are becoming more and more complex. The traditional modelling methods, such as the transfer function, the state space model, are not able to build the relationship between the parameter variables and the faults or health status of the system [6]. Hence, the integrated health monitoring system is becoming more and more important [7], particularly for the vehicle control system. In order to establish an ideal health monitoring system, the computer simulation becomes the main technology which can provide various performance degradation information of the component. In this case, the research on an advanced health monitor and management technology can be carried out deeply, which is a critical basis of IVHM.

In the traditional vehicle control system, the research is often based on the transfer function, the nonlinear differential equation, etc. The health information of every component is implicated in the parameters of the model above. Therefore, the simulation of IVHM system can be carried out but indirectly based on these parameters of model. Currently, the simulation of the vehicle control system is generally based on the parametric model [7]. The corresponding fault modeling approach is also based on the parametric model, which implies that the coverage of faults is relatively limited. In the real case, the vehicle control system is composed of electric circuit and mechanism-electric components, such as the inertia measure unit, the controller, the servo motor, etc. The circuit fault simulation can be carried out by a lot of mature commercial simulation software currently [5]. This is a direct way to do research on the component health, but the modeling consistency between the system-level and the component-level must be considered.

Manuscript received April 28, 2014. This work was supported by the National Natural Science Foundation of China (Grant No：61174120).

The Author is with the College of Aerospace Science and Engineering, NUDT, 151 SanYi Ave. ChangSha City, Hunan Province, 410073, P.R.China (e-mail: z.qinghua@nudt.edu.cn).

Some fault simulation, whose model only contains the degeneration parameter, is easy to be carried out, because model structure need not to be changed. While most fault simulation, whose topology configuration of the model will change under different fault conditions, is difficult, since there are enormous challenges on how to build the complex fault model. Jean-Nicolas Paquino has built a multi-platform simulation system using SIMULINK and SimPower, which is a competent for parallel simulation of different circuit components for complex electronic equipment [8]. The parallel simulation platform can be extended as a scheduler to carry out simulation on different mechanistic models, such as electronic, mechanical and fluid systems. This method is expected to solve the problems of multi-mechanism and co-simulation. However, there is a lot of simulation software in this parallel simulation platform. The interface of different simulation software is very complicated. The work required by co-computation of different simulation software is numerous. Especially the simulation results may easily diverge sometimes. If a response approximation model can be used to replace the corresponding mechanism platform, various approximate health models of components can be transformed into the same environment for system-level simulation. Then a hierarchical and simplified health simulation system is constructed. The health behaviors model based on fault simulated database by professional simulation software can be concentrated into a unified environment for system-level simulation. The multi-interface requirements for different kinds of simulation software can be avoided.

II. Health Simulation Frame

Generally, a vehicle control system is composed of multiple components. For a specific circuit component, it is composed of many electrical elements, whose performance degradation will deteriorate the health behaviors of the component. As shown in Figure 1, the large amounts of health simulation data can be produced by component-level simulation. The simulation is related to various component health behaviors in the environment of professional circuit design and simulation software. Any complex component, with enough response test data, can be identified through in-depth study with various algorithms to get its mathematics model [9]. So research on the component health behavior modeling method is carried out in order to map the relationship between the health simulated data and the health behaviors. At last, the traditional simulation model of the vehicle control system is reconstructed whose component-level model is replaced by its corresponding component health behavior model (CHBM)[11]. The system-level simulation can be carried out under different health conditions of various components.

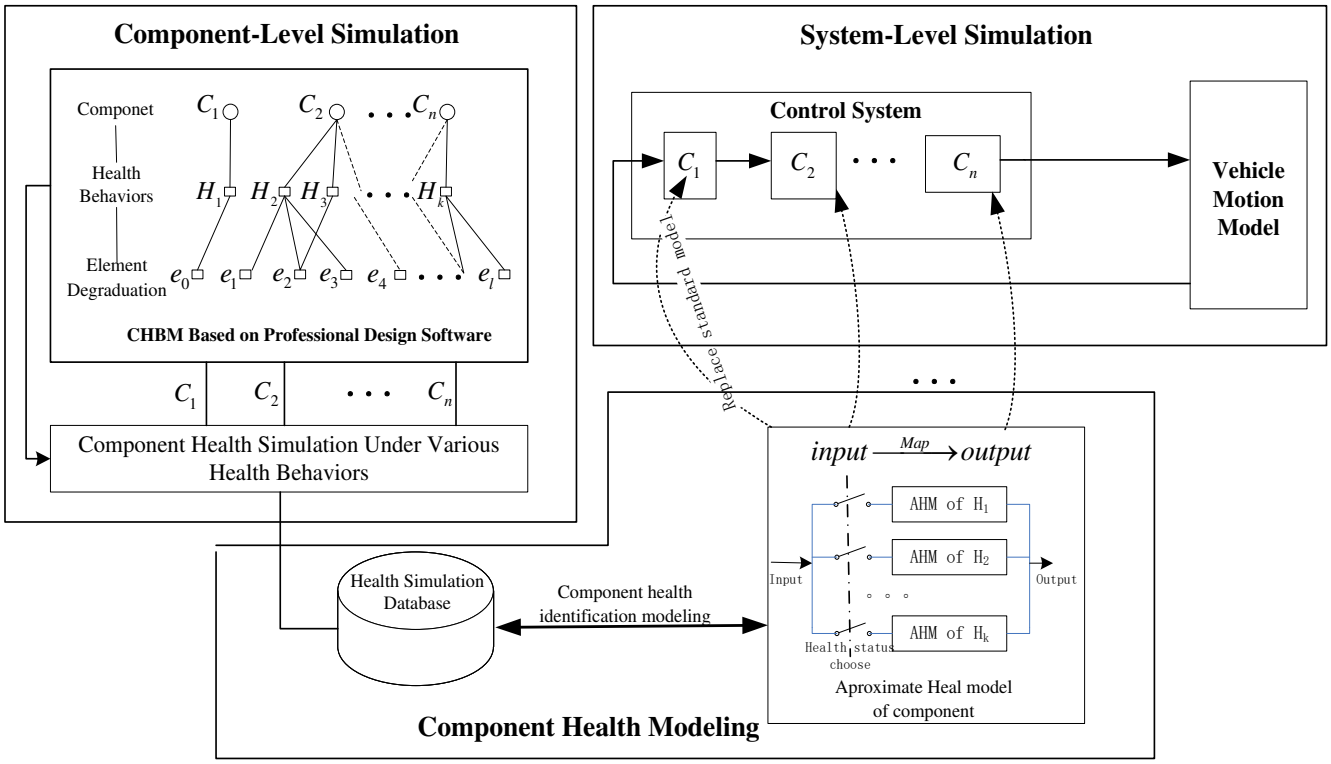


Figure 1 Health Simulation Frame

III. Component Health Behavior Modeling Method

In the health simulation frame, the exact circuit model of a component can be built by professional circuit design software. But the method to produce and utilize the component-level simulation database is the key to identifying CHBM. In theory, the essence of CHBM is the design of the component input signal in simulation test design and the design of identification algorithm.

Component input signal should be carefully designed to effectively stimulate the interior mode of the system, in order to get the information related to every health status of the component as much as possible [10]. Assuming the input of a component is $r(t)$ and its output is $y(t)$, a set of step functions can be applied to reconstruct $r(t)$. The time margin $[0,t]$ is divided into several pieces, the left end of

each interval is the beginning of a step function, denoted as $S_i(t)$, $(i=0,1,...,k,k+1,...,n)$, see Figure 2.

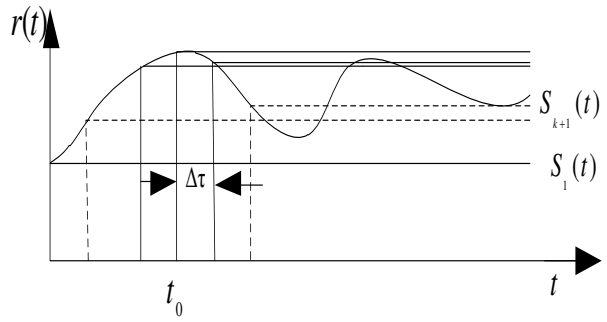


Figure 2 Input Signal Equivalent Principle

Assume the unit step function as $u(t-i\cdot\Delta\tau)=\begin{cases}1, & t\geq i\cdot\Delta\tau\\ 0, & t<i\cdot\Delta\tau\end{cases}$, then the range of the step function can be expressed as $S_i(t)=r(i\cdot\Delta\tau)\cdot u(t-i\cdot\Delta\tau)$. So,

$$r(t)\approx S_0(t)+\sum_{i=0}^{n}(S_{i+1}(t)-S_i(t)) \tag{1}$$

$$r(t)\approx r(0)+\sum_{i=0}^{n}\frac{r(i\cdot\Delta\tau+\Delta\tau)-r(i\cdot\Delta\tau)}{\Delta\tau}u(t-i\cdot\Delta\tau-\Delta\tau)\cdot\Delta\tau \tag{2}$$

For $\Delta\tau\to 0$, $r(t)=r(0)+\int_0^t r'(\tau)\cdot u(t-\tau)d\tau$, thus $r(t)$ can be reconstructed by the unit step function.

Assume $h(t-\tau)$ as the system response function of the unit step function $u(t-\tau)$.

In the interval of $[0,t]$, with the application of the principle of superposition , the system response $y(t)$ of arbitrary excitations $r(t)$ can be written as the summation of all the step responses of the system.

$$y(t)=r(0)\cdot h(t)+\sum_{0}^{n}[r(i\cdot\Delta\tau+\Delta\tau)-r(i\cdot\Delta\tau)]\cdot h[t-(i+1)\cdot\Delta\tau] \tag{3}$$

Consider $\Delta\tau\to 0$,

$$y(t)=\int_0^t r(\tau)\cdot h(t-\tau)d\tau \tag{4}$$

Assume that the step response $h(t-(i-1)\cdot\Delta\tau)$ will not cause too much impact on the step response $h(t-i\cdot\Delta\tau)$.

That is, when $t>t_0$, $\int_0^{t_0} r(\tau)h(t-\tau)d\tau=C$, C is a constant.

$$y(t)=\int_0^{t_0} r(\tau)\cdot h(t-\tau)d\tau+\int_{t_0}^{t} r(\tau)h(t-\tau)d\tau \tag{5}$$

So,

$$y(t)=\int_{t_0}^{t} r(\tau)h(t-\tau)d\tau+\mathrm{C} \tag{6}$$

In this case, the system dynamics is determined only by the step excitation at the time of t_0.

On the other hand, if there are some methods to establish the map relationship between the step signals and their responses of a system, then the response of any input can be obtained by these given system step response data.

In fact, the system dynamics is a set of solutions of the ordinary differential equations which can usually be expressed as follows,

$$\frac{d^n y}{dt^n}+q_{n-1}\frac{d^{n-1}y}{dt^{n-1}}+...+q_0 y=p_{n-1}\frac{d^{n-1}r}{dt^{n-1}}+p_{n-2}\frac{d^{n-2}r}{dt^{n-2}}+...+p_0 r \tag{7}$$

Its discrete form can be expressed as,

$$y(k-n)+q_{n-1}y(k-n+1)+...q_0 y(k)=p_{n-1}r(k-n+1)+p_{n-2}r(k-n+2)+...+p_0 r(k) \tag{8}$$

It can be further expressed as,

$$y(k)=p_{n-1}r(k-n+1)/q_0+...+p_0 r(k)/q_0-y(k-n)/q_0-...-q_1 y(k-1)/q_0 \tag{9}$$

That is, the output of the system at time *k* is determined by the current input and the historical input and output. Research on the mapping method can be carried out to get an equivalent model, which can be used to prognosticate the solution of the differential equations in a certain initial conditions, namely,

$$y(k)=f_{neuro}\left(y(k-n),...,y(k-1),r(k-n),...,r(k)\right) \tag{10}$$

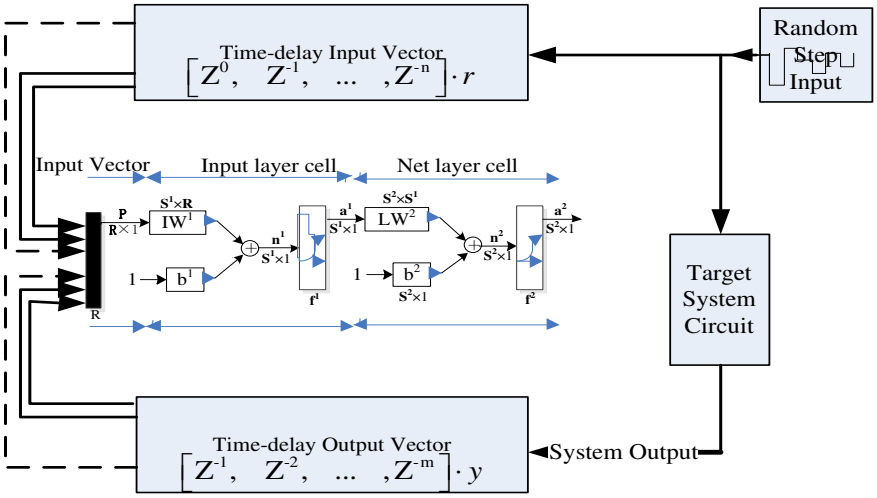


Figure 3 Response approximation modeling method

Because the step signal contains a set of full spectrum signals, the component can be fully motivated. A set of random step signals is used to motivate the circuit simulation model (the target component) to get sufficient amount of the system response data. The output is rich in health information of the circuits. Then, the input vector, its *n*-order time-delay, and the *m*-order output time-delay vectors of the circuit are used as training sample space. The BP neural network, which is chosen to learn the map relationship between the random step signals and their system responses, can be used as the response approximation core algorithm. The response approximation modeling method is shown in Figure 3.

IV. Example

A vehicle longitudinal control system, which is composed by a pitch angle sensor, a controller and an elevator, is shown as Figure 4. In order to simplify the problem, take the pitch angle sensor as an ideal model, the controller as a second-order model in transfer function form. However, a circuit model, which isdescribed by Proteus, is used as the controller. The function of controller is just to produce the PID control signal of elevator based on the difference signal between the command and actual values of pitch angle.

As shown in the Figure 5, the controller contains several elements such as the onboard computer 8051, the crystal, digital/analog, analog/digital converters, the power amplifier circuit, etc.

With orderly health simulation, the health mode and the effect analysis of every element are carried out to form the health behavior table of the controller [11].

Assuming four kind of health mode ($HM_i, i = 0,4$) are considered, in which HM_0 represents the nominal health behavior of the controller. $HM_1 \sim HM_4$ represent the health behavior with some elements performance degradation. As shown in Table 1, two kinds of element health degradation are studied, one corresponds to three parameters of PID of the embedded software in 8051; the other is the variation of the crystal frequency. HM_0 's values are nominal values, while $HM_1 \sim HM_4$'s will change one of the nominal values. With the simulation of the controller circuit by changing these parameters value under different health behavior, a large number of health simulation data of this vehicle control system could be obtained based on the health information of the component. Then the circuit health behavior response approximation model (RAM) can be achieved based on the health simulation database by the method described in the Section 3.

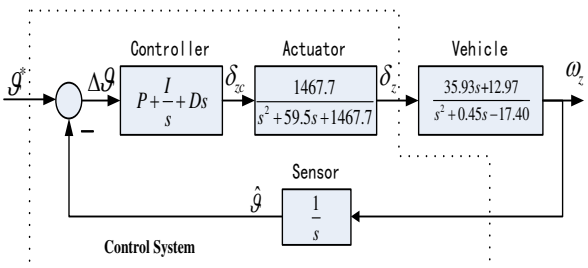


Figure 4 Block diagram of vehicle longitudinal system

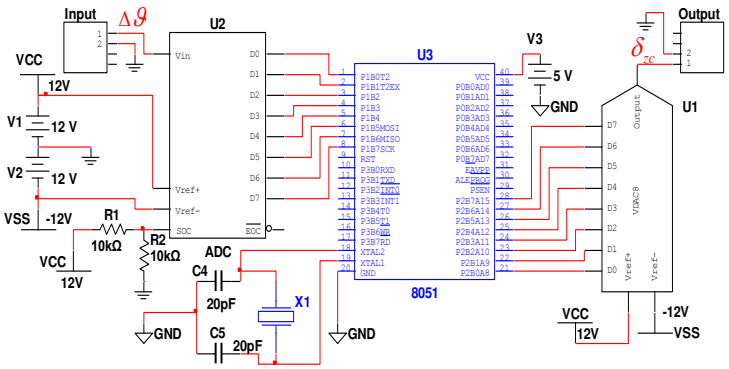


Figure 5 Controller circuit Proteus schematic diagram

Table 1 Controller health behavior table

Case (i)	Health Behavior	HM0	HMi	Corresponding Element
1	Proportion coefficient change	1	0.8	Software module of controller
2	Integral coefficient change	1	0.8	
3	Differentiation coefficient change	0.2	0.5	
4	Crystal frequency drift	12MHz	11.5MHz	Computer Crystal

Limited by the pages, only case 1(HM_3) is shown in this paper. Assuming differentiation coefficient value is changed from the nominal value of 0.2 to degradation value of 0.5 in the controller circuit simulation system, the other parameters are fixed to the nominal value(HM_0). This is implemented by modifying the program of the onboard computer 8051. After the simulation database of HM_0 and HM_3 is obtained, the response approximation models (RAM) have learned the dynamic characteristics of the PID controller with the parameters of [1,1,0.2] and [1,1,0.5]. In order to check the validity of the response approximation model, the transfer function model of PID, which is called target system model (TS model), and RAM are used in the vehicle longitudinal system model respectively. Figure 6 gives the time response comparison between the parameters of [1,1,0.2], [1,1,0.5] and their corresponding TS models. Furthermore, the consistency within a specific frequency band between TS model and its RAM model with parameters [1,1,0.2] is studied, as shown in the Figure 7. The above results shows that the model errors are so small that RAM can be applied into the system-level simulation by replacing the traditional model.

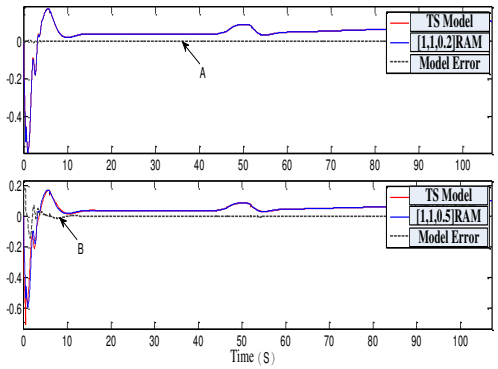


Figure 6 Comparison with RAM model and TS model

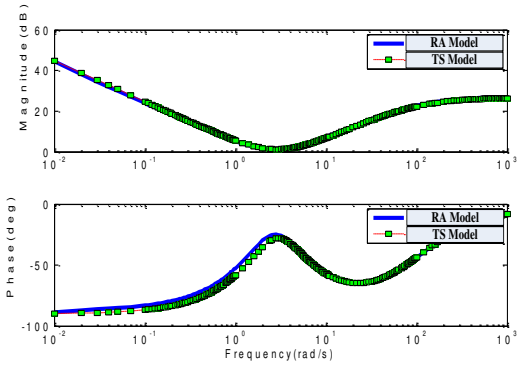


Figure 7 Comparison of the frequency domain characteristics

V. Conclusion

The different health behaviors of a certain system are shown by the different fault dynamic data. This method provides a way to extract health data from a circuit. With the aid of response approximation modeling method, the map relationship between fault data and the system input can be found. All the different health behaviors of components are capable to be consolidated into a system-level simulation in a consistent expression. The different health behaviors of systems with different mechanisms can be simulated on the same platform.

In order to obtain a better approximation model of the system, the time-delay order m and n and the sampling frequency of training data should be adjusted according to the order of the target system. The adequacy of the training information is impacted by the random step input. The width of each random step should be adapted to the adjusting time of the target system. Since the amplitude of the random step input is limited, the input and the output of the response approximation model are bounded. It means that the response approximation model can work effectively within a bounded input range. When the health behaviors are regarded as a set of different dynamic characteristics of a system, the dynamic response data (or health data) of the system becomes the bridge to connect the response approximation model and the system health behaviors. Therefore, with the circuit fault simulating data, this response approximation method is able to find its way to make a contribution to the IVHM system.

This paper discusses the core dynamic learning algorithm of BP neural network. The use of other non-linear learning algorithm may get better performance, but the basic principle is the same.

VI. Acknowledgments

The authors would like to acknowledge the support from the National Natural Science Foundation of China (61174120).

References

[1] E. Baroth, W.T. Powers. IVHM Techniques for Future Space Vehicles. AIAA 37th Joint Propulsion Conference & Exhibit 8-11, July 2001, Salt Lake City, Utah. AIAA-2001-3523.

[2] Belcastro C M . Aviation safety program: integrated vehicle health management technical plan summary. NASA Technology Report , 2006: 1-53.

[3] Fox J J, Glass B J. Impact of integrated vehicle health management (IVHM) technologies on ground operations for reusable launch vehicles (RLVs) and spacecraft. NASA, 2000: 179-186.

[4] Richard B, Mc Sharry M. K-1 integrated vehicle health management system: design for operational efficiency,Proc. of IEEE Areospace Conference, 2000, 4: 323-327.

[5] Zhao Guangyan, Sun Yufeng, Kang Rui, Wu Yue. Fault circuit fault simulation modeling, injection and judgment

method.Microelectronics and Computer, 2007,24(1): 143-146.
[6] Mark Schwabacher, Jeff Samuelsa, Lee Brownstonb. The NASA Integrated Vehicle Health Management technology experiment for X-37. Proceedings of the SPIE AeroSense 2002 Symposium.
[7] Dimitry Gorinevsky, John R. Bain, Gordon Aaseng. Parametric Diagnostics of Flight Control and Propulsion for Rocket Ascent. AIAA Journal of Guidance, Control, and Dynamics, June 2004.
[8] Jean-Nicolas Paquin, Wei Li. A Modern and Open Real-Time Digital Simulator of All-Electric Ships with a Multi-Platform Co-Simulation Approach. IEEE Conference 28-35, 2009.
[9] Tom M.Mitchell, Zeng HuaJun. Machine Learning. China Machine Press, 2003.
[10] Fang ChongZhi. Process Identification. TsingHua University Press, 2006.
[11] Wu Maoxing, Zeng Qinghua. Eelectric servo based health Simulation and Evaluation Method. Aeronautical Computing Technology, 2012,42(3): 38-41.

AUTHOR BIOGRAPHY

Zeng Qinghua, Received an MSc in control engineering and a PhD in aircraft design from the National University of Defense Technology, in 1991 and 2004 respectively. From 1991 worked as teaching and scientific research personnel in the field of aircraft control system design, simulation and fault diagnosis, researched as senior academic visitor at Cranfield University in 2013, presently working as a professor with the College of Aerospace Science and Engineering, NUDT.

Zhao Wei, Received the PhD in Electronic Engineering from the University of Surrey, in 2013. Now, he is a lecturer in the National University and Defense Technology (NUDT). The research interest includes the Flight Dynamics, UAV Flight Control, Robust Control, and Muti-Disciplinary Modeling.

Jia Tao, Master, China Aerodynamics R&D Center, graduated from NUDT, 2013, and studied in Harbin Institute of Technology as an Undergraduate student 2007-2011.

Life Prediction and Classification of Fault-Modes in Solid State Lamps Using Bayesian Probabilistic Models

Pradeep Lall, Junchao Wei, Peter Sakalaukus
Auburn University
Department of Mechanical Engineering and
NSF CAVE3 Electronics Research Center
Auburn, AL 36849
Tele:(334)844-3424
E-mail:lall@auburn.edu

***Abstract* — A new method has been developed for assessment of the onset of degradation in solid state luminaires to classify failure mechanisms by using metrics beyond lumen degradation that are currently used for identification of failure. Luminous Flux output, Correlated Color Temperature Data on Philips LED Lamps has been gathered under 85°C/85%RH till lamp failure. The acquired data has been used in conjunction with Bayesian Probabilistic Models to identify luminaires with onset of degradation much prior to failure through identification of decision boundaries between lamps with accrued damage and lamps beyond the failure threshold in the feature space. In addition luminaires with different failure modes have been classified separately from healthy pristine luminaires. It is expected that, the new test technique will allow the development of failure distributions without testing till L70 life for the manifestation of failure.**

Keywords—Solid State Lighting, Life Prediction, Luminous Flux, Correlated Color Temperature, Bayesian Probabilistic Models.

I. INTRODUCTION

The lighting industry is undergoing a change from the incandescent lamps and compact fluorescent lamps (CFL) to light emitting diodes (LED). Mercury widely used in CFLs has the potential of contaminating large amounts of drinking water to beyond drinkable levels even in trace amounts. Transition to LEDs can impact energy efficiency tremendously because nearly 17% of the annual energy consumption is used for lighting [US EIA 2012]. LEDs are being used in a wide variety of applications including automotive lighting, LED displays, street and home lighting. Traditional methods of failure-detection often used for identification of failure in incandescent lamps may not be applicable to LEDs. Traditional light sources "burn out" at end-of-life. For an incandescent bulb, the lamp life is defined by B50 life or the time by which 50% of the population will fail. However, the LEDs have no filament to "burn". The LEDs continually degrade and the light output decreases eventually below useful levels causing failure. LED failure is characterized by L70 life or 70% degradation of the lumen output [IES LM-80-08; IES TM-21-11]. Currently, it is not possible to qualify SSL lifetime of 10-years and beyond often necessary of high reliability applications, primarily because of lack of accelerated test techniques and comprehensive life prediction models. SSL comprises of several length scales with different failure modes at each level. Interactions between optics, drive electronics, controls and thermal design. Accelerated testing for one sub-system may be too harsh for another sub-system. New methods are needed for predicting SSL reliability for new and unknown failure modes. Presently, there is scarcity of life distributions for LEDs and SSLs which are needed to assess the promised lifetimes. Several cities are experimenting with large scale deployment of luminaries. In order to keep high availability of the system, it is essential that the onset of damage in form of color shift, luminous output degradation, and change in CCT be detected early.

Previous researchers have studied the failure modes of luminaires. Junction temperature of a luminaire plays a substantial role on its lifetime. The degradation rate of the plastic encapsulation material (PEM) on the diode is predominately affected by junction temperature causing the attenuation of the light output [Narendran 2004, Baillot 2010]. Excessive temperatures inside the LED package or the ingress of moisture can produce thermal-mechanical and hydro-mechanical stresses between the various material layers of LED packages causing delamination [Lumileds 2006, Luo 2010]. Elevated temperatures and humidity can produce delamination between the die and silicone encapsulant [Lumileds 2006] and between the encapsulant and packaging lead frame [Luo 2010]. The stresses can also produce a hairline cracks known as lens cracking, which occurs due to thermal expansion at various operating temperatures [Lumileds 2006, Hsu 2008], as well as when a long-term exposure to moisture [Hewlett Packard 1997]. Ahn [2014] has examined the effect of

LED lighting on the heat loads in office buildings and showed that the energy consumption related to cooling of the building reduced in the neighborhood of 11-percent. Chan [2011] examined the accelerated life test of high power white light emitting diodes under exposure to temperature and humidity and found the degradation mechanisms of optical power reduction and degradation of blue wavelengths of emitted light due to bubbling and discoloration of silicone encapsulating material of the package. Choi [2014] studied the correlation between reliability and pH changes of phosphors for white light emitting diodes and showed that the reactivity of phosphor with water depended on the host material and could be identified through pH measurement. Fu [2012] investigated the dynamic color change mechanisms in high power light emitting diodes and found yellowed reflector and damaged silicone caused color shift. Jang [2014] studied the optimum design of radial heat sink for high-power LED lighting applications and found that pin-fin arrays with tallest pin heights in the outermost array exhibited the best performance. Meneghini [2012] studied the effect of DC and pulse wave modulation on the reliability in GaN light emitting diodes and found that in both cases higher stress levels caused higher degradation with PWM exhibiting a higher values of degradation. Yang [2010] investigated the failure and degradation mechanisms of high-power white light emitting diodes under exposure to high temperature and electrical currents and found that degradation rates of luminous flux increased with electrical and thermal stress. High electric stress induced surface and bulk defects in LED chip during short term aging.

In this paper, wet high temperature operating life environmental conditions of 85°C and 85%RH have been used to understand the reliability of solid state luminaires. SSL failure is quantified by the deterioration of luminous flux output and correlated color temperature (CCT) with respect to the time during accelerated testing. The Illuminating Engineering Society test standards LM-80-08 and IES TM-21-11 define the lifetime of an LED for lighting as the degradation to 70-percent of the original luminous flux output at room temperature [IES LM-80-08; IES TM-21-11]. Bayesian generative models have been used for classification of damaged assemblies and Bayesian regression models have been used to model the damage progression in solid state luminaires. The luminous flux, CCT, and the color shift have been used as input variables for identification of the onset of damage and separation of the healthy SSLs from those with significant accrued damage. Discriminant functions have been used to identify the class boundaries and classify SSLs significantly prior to the development of complete failure distributions. The models have been used to estimate the remaining useful life for each sample under test and the model predictions validated versus experimental data. It is expected that, the new test technique will allow early identification of failure distributions.

II. Test Vehicle

One of the original off-the-shelf 60W LED Lamps has been used as the test vehicle (Figure 1). The lamp has a total of 9 LUXEON Royal-Blue LEDs which are divided into three systematic lamp housings with a yellow cerium doped yttrium aluminium garnet phosphor shell.

Figure 1 Ambient LED 60W Lamp

The lamp produces white light through the color mixing of the blue LEDs and the yellow phosphor. The luminaires have been subjected to temperature-humidity at 85°C/85RH in an accelerated test chamber. The luminaires were non-functional during the accelerated test and placed upright inside a lamp holder to prevent movement inside the test chamber. Each of the lamps were extracted on a weekly basis to exam the spectral data for luminous maintenance, chromaticity shift, and correlated color temperature. All of the lamps in the test set were aged in the temperature-humidity condition for a total duration of 2537 hours.

III. Experimental Set-Up

The LED lamp measurement has been accomplished using an integrating sphere. Typically, the lamp measurement system contains parts: (1) Light Emitting Device. (2) Light Gathering System. (3) Light Transmitting System. (4) Light Analyzing System. The light emitting device provides the AC voltage power connection to the LED lamp that is producing continuous measureable light. The Light Gathering System, in this case the integrating sphere redistributes and collects the entire light beam emitting from the LED lamp. The integrating sphere is an optical component that uniformly scatters the light, which has a special coating on its surface of inside sphere. With a small exit ports on the side of sphere, the LED lights can be transmitted through the cosine diffuser, which is a detector, filtering and transferring the distributed light to the cable optical fiber. Then, the light is carried into the Labsphere 'USB4000' Spectrometer. Data on the Lumen Flux and CCT is collected using SpecraSuite Software. The total spectral radiant flux, $\Phi_{test}(\lambda)$, of the LED lamps under test was obtained by comparison of the total spectral radiant flux of the test lamp, $\Phi_{TEST}(\lambda)$, to the total spectral radiant flux of a reference standard, $\Phi_{REF}(\lambda)$. The following equation was used to compute the total spectral radiant flux:

$$\Phi_{TEST}(\lambda) = \Phi_{REF}(\lambda) * \frac{y_{TEST}(\lambda)}{y_{REF}(\lambda)} * \frac{1}{a(\lambda)} \quad (1)$$

Where $y_{TEST}(\lambda)$ is the are the spectrometer readings for the lamp under test, $y_{REF}(\lambda)$ is the spectrometer readings for the reference-lamp, respectively, and $a(\lambda)$ is the self-absorption factor measured using an auxiliary lamp as

described in LM-79. From the measured total spectral radiant flux $\Phi_{TEST}(\lambda)$ [W/nm], the total luminous flux $\Phi_{TEST}(\lambda)$[lm] is obtained by

$$\Phi_{TEST} = K_m \int_{\lambda} \Phi_{TEST}(\lambda) * V(\lambda) d(\lambda)$$
$$K_m = 683 lm/W \quad (2)$$

Where $V(\lambda)$ is the photopic sensitivity as a function of the wavelength. Self-absorption is the effect, in which the response of the sphere system is affected due to the absorption of light by the lamp itself in the sphere. Errors can also occur if the size and shape of the test light source are significantly different from those of the standard light source. The self-absorption factor is given by,

$$a(\lambda) = \frac{y_{aux,TEST}(\lambda)}{y_{aux,REF}(\lambda)} \quad (3)$$

Where $y_{aux,TEST}(\lambda)$ is the spectrometer readings for the auxiliary lamp with the LED lamp in the sphere, and $y_{aux,REF}(\lambda)$ is the spectrometer readings for the auxiliary lamp with the reference standard in the sphere.

IV. Decay Rate And Failure Threshold

The decay rate of the luminous flux and the correlated color temperature in the LED lamps and the LUXEON LEDs inside the LED lamps has been calculated from the exponential model. The degradation model for lumen maintenance and correlated color temperature has been selected because the degradation of light emitting diodes has been modeled using the exponential model in IES TM-21-11 for life prediction of light emitting diodes in operation. Furthermore, the authors have used the exponential model for L70 life prediction in conjunction with Kalman Filter and Extended Kalman Filter [Lall 2013]

$$\Phi = \beta \cdot e^{\alpha \cdot t} \quad (4)$$

Where, β is the pre-decay factor, α is the decay rate, t is the test time, and Φ is either the luminous flux output or the correlated color temperature depending on the decay rate being calculated. The decay rate is a function of temperature and represented by:

$$\alpha = A \cdot e^{-\left(\frac{E_A}{k_B \cdot T}\right)} \quad (5)$$

Where T is the temperature in kelvin, K_B is the Boltzmann constant, and E_A is the activation energy. The Method of Least Square (LS) has been used to compute the decay rate for both CCT and Lumen Maintenance. The data for the LED lamps has been taken from accelerated test data under 85°C/85%RH. The LUXEON LEDs data from [DR05-1-LM80; Philips 2012] is under conditions of 55°C/65%RH at 1A. The two test conditions used for calculation of the activation energy include 55°C and 85°C. Measured values of both the luminous flux output and the correlated color temperature have been normalized with respect to the measured value at time zero. LED lamp data shows the degradation of CCT to 96% from the initial value of 100% after 2500 hours. Similarly, the Lumen Maintenance shows degradation to 68% after 2000 hours of accelerating test.

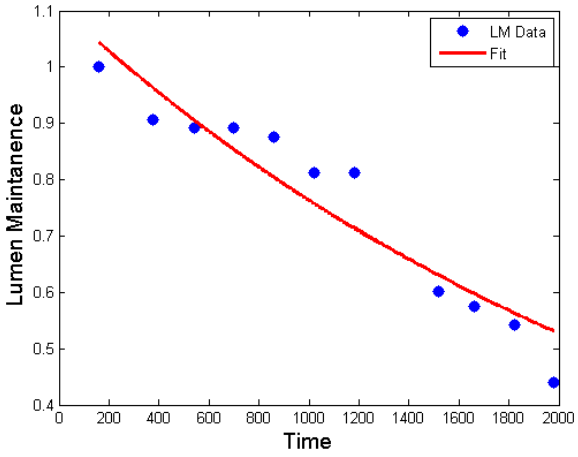


Figure 2: Lumen Maintenance Evolution for the LED Lamp in 85°C/85RH

Typically, L70 (70% Lumen Maintenance) life has been treated as the failure threshold for the luminous flux output of the solid state luminaire. Further, the 7-step MacAdam ellipse states that the target 'Duv' and its tolerance is ±0.006, and the corresponding target CCT and tolerance is 3000±175K for a nominal 3000K lamp [ANSI C78.377-2008 Specification]. One can therefore conclude that variation of CCT of greater than 94.17% of the original CCT values are deemed as unacceptable. The 94.17% value for a 3000K lamp is 2825K for the LED lamp. The lumen decay is more significant than the CCT decay. For the purpose of computing the remaining useful life of the luminaire and LEDs, the degradation of lumen maintenance was used. Figure 2 and Figure 3 show the decay rates for luminous flux output and correlated color temperature.

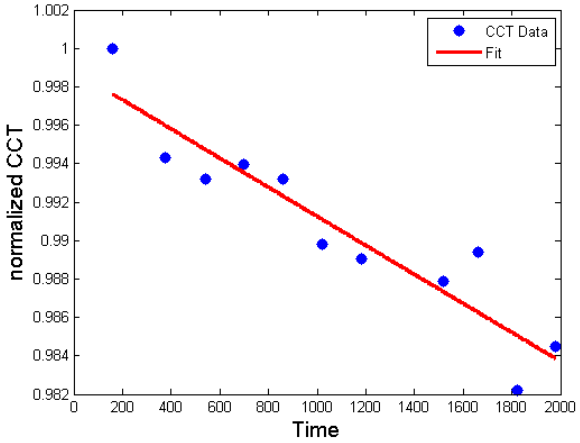


Figure 3: Correlated Color Temperature Evolution for the LED Lamp in 85°C/85RH

V. Decay Rate Distribution

Decay rate has been computed based on testing of LED Lamps. The testing length of time is 2536.85 hours. The mean value of L70 hours is 1673.3 hours. Figure 4 and Figure 5 show the lumen degradation and Correlated Color Temperature (CCT) versus time. We can see that there is distinct lumen degradation throughout the testing history. The CCT has dropped 2.49% compared to the pristine value at the end of testing for the L-prize lamp. It could have combined two types of degradation pattern in the

lumen. One is decelerating decay, and the other is accelerating decay. We will evaluate this behavior using one cubic polynomial model, which could better present this degradation curve than parabolic model.

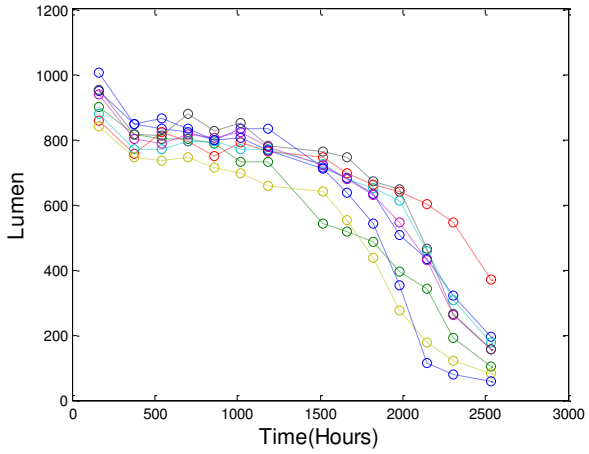


Figure 4: Lamp Lumen Depreciation in History

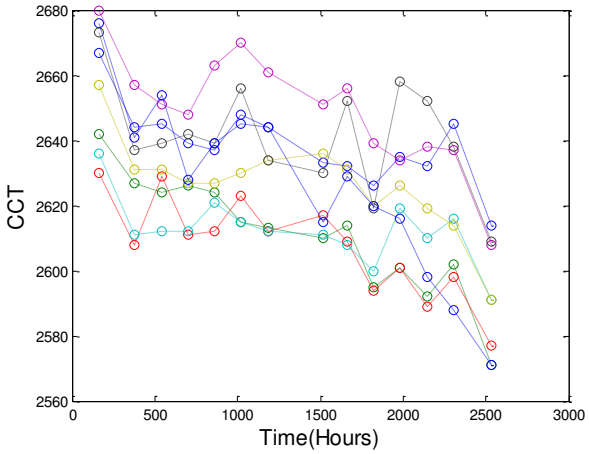


Figure 5: Lamp lamp CCT depreciation in History

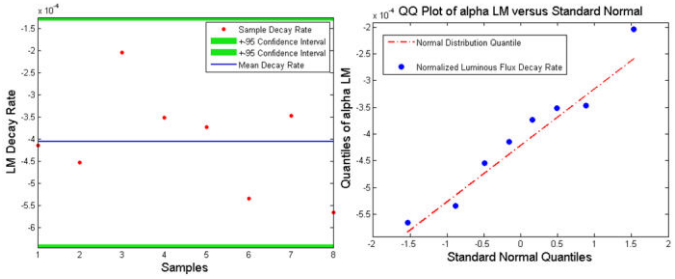


Figure 6: Fitted LM Decay Rate Distribution and QQ plot.

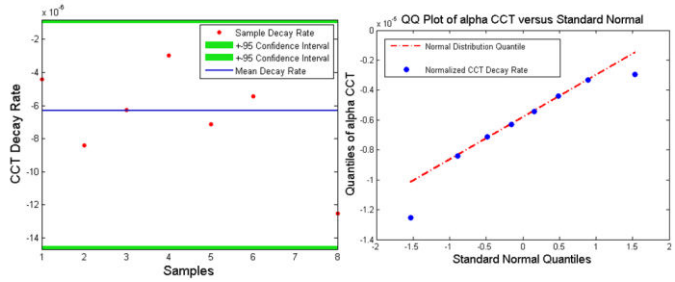


Figure 7: Fitted CCT Decay Rate Distribution and QQ plot

Similarly, the LM and CCT decay rate distribution for the L-prize is also normally distributed. Therefore, both the mean value and variance can be represented as the main decay character. The mean decay rate of normalized luminous flux output for the LUXEON LEDs has been calculated to be -5.14e-06 per hours, and mean decay rate of the normalized CCT for the LUXEON LEDs has been calculated to be -1.28e-06 per hours. Normality of the luminous flux output and the correlated color temperature distributions has been checked using the QQ-plot (Figure 6 and Figure 7). The red line shows the quantile of normal distribution, the blue dots show the decay rate of normalized luminous flux output and correlated color temperature data. Analysis results indicate that the data is normally distributed with only two or three outliers.

The decay rates for the normalized luminous flux output and the correlated color temperature of the lamp has been similarly computed to be -4.05e-04 per hours, and -6.30e-06 per hour respectively. The Arrhenius model has been used to calculate the activation energy for the normalized luminous flux output and the normalized correlated color temperature. The computed activation energy has been used to evaluate the effect of temperature on the luminous flux output and the correlated color temperature.

$$E_A = -\left(\frac{\ln \frac{\alpha_1}{\alpha_2}}{\left(\frac{1}{T_1} - \frac{1}{T_2} \right) \cdot K_B} \right) \quad (6)$$

The activation energy for the normalized lumen degradation is 1.47 eV and the activation energy for the normalized correlated color temperature is 0.53 eV. The failure threshold for the normalized luminous flux and the normalized correlated color temperature was identified by computing the 95% confidence bounds. Data that fell below the failure threshold at any time during the life test was deemed as a failure. Remaining useful life predictions were done for samples that did not fall below the failure threshold. The failure criterion is the curve of the maximum normalized decay rate, which will envelop all the degradation lines in the tested sample-set. For the LUXEON LEDs, the maximum decay rate ($\alpha_{max,LM}$) for the normalized luminous flux output is -6.77e-06 hour^{-1} and maximum decay rate ($\alpha_{max,CCT}$) for the normalized CCT is -2.8e-06 hour^{-1}. The 95% Confidence Interval for the maximum decay rate has been used to compute the highest possible decay rate, i.e. the lower boundary, for formulating the failure criterion. The maximum degradation rate has been used because it will encompass the worst possible parts in the population and prevent the labelling of parts within the ±1.96σ bounds as failures. If the mean lumen degradation and correlated color temperature had been used – then parts lower than the mean would have been labeled as a false-positives

$$\alpha_{LM}^{failure} = \alpha_{max,LM} - 1.96\frac{\sigma}{\sqrt{N}} \quad (7)$$

$$= -7.04 \times 10^{-6} \text{ hour}^{-1}$$

$$\alpha_{CCT}^{failure} = \alpha_{max,CCT} - 1.96\frac{\sigma}{\sqrt{N}} \quad (8)$$
$$= -3.01\,x10^{-6}\ hour^{-1}$$

Where $\alpha_{LM}^{failure}$ is the failure threshold of the decay rate for the normalized luminous flux output, and $\alpha_{CCT}^{failure}$ is the failure threshold of the decay rate for the normalized correlated color temperature.

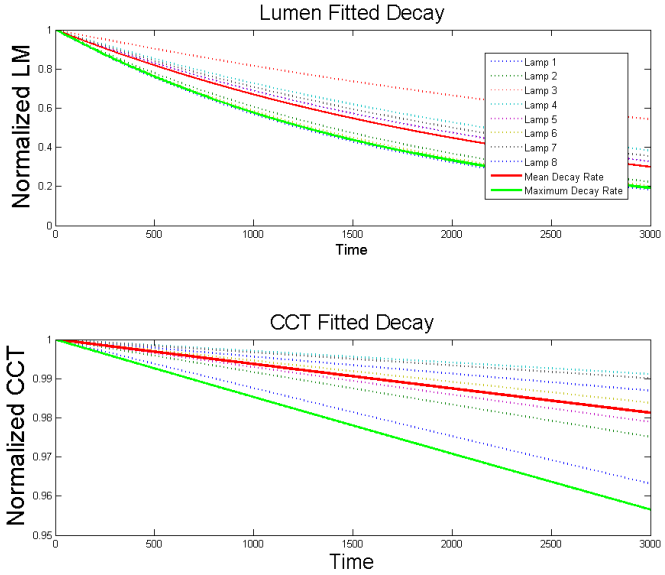


Figure 8: Characterized Decay Rate Curve for the LED Lamp

The decay rate values from the LUXEON LEDs have been used to compute the LED related decay rate for the lamp. The failure threshold decay rate has been calculated using an Arrhenius model:

$$L\alpha_{LM}^{failure} = \alpha_{LM}^{failure} \cdot e^{-\frac{E_A}{K_B}\left(\frac{1}{T_1}-\frac{1}{T_2}\right)} \quad (9)$$
$$= -5.55\ x\ 10^{-4}\ hour^{-1}$$

$$L\alpha_{CCT}^{failure} = \alpha_{CCT}^{failure} \cdot e^{-\frac{E_{act}}{k_B}\left(\frac{1}{T_1}-\frac{1}{T_2}\right)} \quad (10)$$
$$= -1.48\,x10^{-5}\ hours^{-1}$$

The time dependent decay curve for the LED Lamp is shown in Figure 8a where the bold red line is the mean decay rate. The lower-bound is depicted as a green line in a Figure 8a. The CCT decays almost linearly. However, the Lumen Maintenance exhibits much more exponential pattern. The green dash line in Figure 10b shows the failure criterion for the LED Lamp. Figure 8b shows that all of CCT degradation lines are higher than the failure threshold. Only one of the tested samples lies on the failure threshold for Lumen Maintenance.

VI. Bayesian Probabilistic Model

In this paper, Bayesian Probabilistic Generative Models [Bishop 2006] have been used to classify and separate damaged solid state luminaire assemblies from healthy assemblies. The goal of classification is to analyze input vector, x consisting of CCT, Color-Shift and Luminous Flux Output and to assign it to one of the classes, C_k. There are two possible classes including damaged or healthy. The classes are taken to be disjoint, so that each input is assigned to only one class. The input space is divided decision regions whose boundaries are called the decision boundaries. The target variable has been represented as a binary variable such that t=1 represents class C_1 and t=0 represents class C_2. The value of 't' is the probability that the class is C_1 with the values of probability taking only extreme values of 0 and 1. The conditional probability distribution, $p(C_k \mid X)$, has been modeled in the inference stage and then the distribution has subsequently been used to make optimal decisions of classification. A generative approach has been adopted for computing the conditional probability distribution $p(C_k \mid X)$. In this procedure, the class conditional probabilities, $p(X \mid C_k)$ have been modeled as well as the class priors, $p(C_k)$, and then used to compute posterior probabilities through Bayes Theorem. For the purpose of the analysis, it was assumed that the class conditional probability density function is Gaussian, represented by:

$$P(X \mid C_k) = \frac{1}{(2\pi)^{d/2}} \cdot \frac{1}{|\Sigma_k|} \cdot \quad (11)$$
$$\exp\left(-\frac{1}{2}(X-\mu_k)^T \Sigma_k^{-1}(X-\mu_k)\right)$$

Where the input vector, x, is a d-component column vector, μ is the d-component mean vector, Σ is the d-by-d covariance matrix. The class-prior $p(C_k)$ utilizes the weighted group form, for classifying two groups, where the probabilities of each group are given by:

$$p(C_1) = \frac{N_1}{N_1+N_2}; p(C_2) = \frac{N_2}{N_1+N_2} \quad (12)$$

$$p(C_1)/p(C_2) = N_1/N_2 \quad (13)$$

Where, N_1 is the number of samples in the first group, and N is the total numbers of samples. The conditional probability distribution for classifying the 'k' group has been normalized, based on the weighted value between its posterior and the sum of posteriors from all the groups. Minimum error rate classification has been achieved through the use of discriminant functions, $p(C_k \mid X)$:

$$g_i(x) = p(C_k \mid X) = \frac{p(X \mid C_k)p(C_k)}{\sum p(X \mid C_j)p(C_j)} \quad (14)$$

Where, $g_i(x)$ is a discriminant function which is used as a classifier. Thus, the discriminant function for the multiple-class classification is defined as:

$$g_i(x) = p(X \mid C_k)p(C_k) \quad (15)$$

Alternatively, the discriminant may be represented in log-form as,

$$g_i(x) = \ln p(X \mid C_k) + \ln p(C_k) \quad (16)$$

In a general multivariate normal case, the covariance matrices are different for each category. The discriminant function can be computed by substituting Equation (11) for the class conditional probability density function into Equation (16) for the long-form of the discriminant function as follows:

$$g_i(x) = -\frac{1}{2}(x-\mu_i)^T \Sigma_i^{-1}(x-\mu_i) - \frac{d}{2}\ln 2\pi - \frac{1}{2}\ln|\Sigma_i| + \ln p($$ (1

Where the input vector, x, is a d-component column vector, μ is the d-component mean vector, Σ is the d-by-d covariance matrix. The resulting discriminant terms are inherently quadratic:

$$g_i(x) = X^T.W_i \cdot X + w_i^T \cdot X + \omega_{i0} \quad (18)$$

Where the quadratic coefficients are solved as:

$$W_i = -\frac{1}{2}(\Sigma_i^{-1}) \quad (19)$$

$$w_i = (\Sigma_i^{-1} \cdot \mu_i) \quad (20)$$

$$w_{i0} = -\frac{1}{2}(\mu_i^T \Sigma_i^{-1} \mu_i) - \frac{1}{2}(\ln|\Sigma_i|) + \ln p(C_k) \quad (21)$$

The discriminant functions have been computed for all the samples and the samples assigned to the class corresponding to the highest discriminant. The decision boundaries have been computed by setting

$$g_1(x) = g_2(x) \quad (22)$$

VII. Feature Space Creation

A two dimensional feature space has been created for classification of the test data. The two dimensions include the normalized luminous flux output and the correlated color temperature. The decay rate failure thresholds for the solid state luminaires which have been computed previously (Equations (9) and (10)) are used to construct the failure boundary for luminous flux output and a second boundary for the correlated color temperature. Lamps could fail because they breach the failure boundary for either the luminous flux output, correlated color temperature or both. The time at which the lamp breaches either boundary is termed as the failure time and represented by T_{CF}. The luminous flux output and the correlated color temperature at the failure boundary for failure time (T_{CF}) has been computed based on the previously calculated maximum decay rate.

$$\Phi = \beta \cdot e^{\alpha \cdot t} \quad (23)$$

$$LM_{max\ failure} = (100) \cdot e^{L\alpha_{LM}^{failure} \cdot T_{CF}} \quad (24)$$

$$CCT_{max\ failure} = (100) \cdot e^{L\alpha_{CCT}^{failure} \cdot T_{CF}} \quad (25)$$

Where the multiplier of '100' in Equations (24) and (25) was used to change the computed normalized values into percentages.

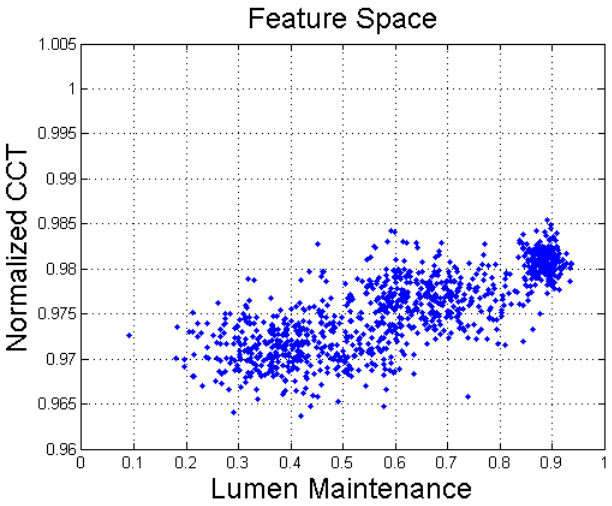


Figure 9: Feature space containing data of lamp's current state, lamp's failure threshold, and pristine healthy distribution of lamps prior to classification.

The computed locations in the feature space allow the location of the failure threshold versus the current state of the lamp in the feature space. The mean and variance of the failure threshold and location at the failure time has been computed for all the devices under test. The classification of the healthy lamps versus the damaged lamps was accomplished using a decision boundary computed based on the discriminant function (Equation (18)). The test lamps have been classified as belonging to the failure threshold distribution or the healthy distribution. The correlation between the luminous flux output and the correlated color temperature has been removed by computing the principal directions of the variance to yield uncorrelated x-axis and y-axis variances. The covariance matrix for the 85C/85%RH dataset is presented by,

$$\Sigma = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix} \quad (26)$$

The covariance matrix has been decorrelated by computing the principal components, thus rendering the correlation matrix in the form,

$$\Sigma = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{pmatrix} \quad (27)$$

Where the subscripts '1' and '2' indicate the principal directions. The distributions of the lamp-state and the lamp's failure threshold have been transformed into the decorrelated principal component feature space for the purpose of classification. The data groups plotted in Figure 9 include the lamp's current state, lamp's failure threshold, and the pristine healthy distribution of lamps prior to classification.

VIII. Bayesian Regression Model

The response variables of luminous flux output, and CCT are the target variables (t) for the Bayesian regression models. Input parameters (w) include weights for the input parameters of time. The posterior probability has been computed based on the conditional probability:

$$p(t \mid w) = \frac{p(w \mid t)p(t)}{P(w)} \quad (28)$$

Where, $p(t \mid w)$ is the normalized conditional posterior of the target variables, and $p(t)$ is the prior distribution of the target variables. The Bayesian conjugate prior Gaussian probability is represented as follows:

$$P(W) = N(W \mid M_0, S_0) = N(W \mid 0, \alpha^{-1}I) \quad (29)$$

Where, α is the precision parameter of the weight distribution. The real-valued input variable column vector is:

$$X = [x_1 \quad x_2 \quad \ldots \quad x_n]^T \quad (30)$$

The real-valued predict target column vector is represented as,

$$T = [t_1 \quad t_2 \quad \ldots \quad t_n]^T \quad (31)$$

Candidate basis functions used may include polynomial functions (ϕ) with weights (w). The basis function 1×M matrix can be shown as the following:

$$\Phi(x) = [1 \quad x \quad x^2 \quad \ldots \quad x^M]^T \quad (32)$$

The weights M×1 matrix is represented as:

$$W = \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \ldots \\ w_M \end{bmatrix} \quad (33)$$

The future degradation of the luminaire can be calculated from the estimation matrix as follows:

$$t_i = W^T \cdot \Phi(x_i) \quad (34)$$

The likelihood function will be represented with a Gaussian probability distribution as follows,

$$P(t \mid x, W, \beta) = N(t \mid W^T \cdot \Phi(x_i), \beta^{-1}I) \quad (35)$$

Where W is the weight vector and β is the precision of the target variable distribution, t. The n- set of observations $t_1, \ldots, t_N$, have been combined into a matrix T of size N × K such that the nth row is given by t_{Tn} . Similarly, we can combine the input vectors $x_1, \ldots, x_N$ into a matrix X. The log-likelihood of the data-set is given by:

$$\ln P(T \mid X, W, \beta) = \sum_{n=1}^{N} \ln N(t \mid W^T \cdot \Phi(x_i), \beta^{-1}I) \quad (36)$$

The likelihood represented by Equation (36) that the target, t corresponds to the input variable sets being considered is maximized with respect to β. The target parameter's variance is represented by:

$$\frac{1}{\beta_{ML}} = \frac{1}{N}\sum_{i=1}^{N}(y_i - t_i)^2 = \frac{1}{N} \cdot (Y - T) \cdot (Y - T)^T \quad (37)$$

The variance computed from equation (37) corresponds to the maximum value of the likelihood function. We can substitute the β_{ML} into Equation (36) for $P(T \mid X, W, \beta)$, which gives:

$$P(T \mid X, W, \beta) = \prod_{i=1}^{n} N(t \mid W^T \cdot \Phi(x_i), \beta_{ML}^{-1}) \quad (38)$$

The weight vector will be updated using the Bayesian posterior conditional probability represented as follows:

$$P(W \mid X, T, \alpha, \beta) \propto P(T \mid X, W, \beta) \cdot P(W \mid \alpha) \quad (39)$$

$$= N(W \mid M_N, S_N)$$

$$M_N = S_N(S_0^{-1}M_0 + \beta\Phi^T T)$$

$$S_N^{-1} = S_0^{-1} + \beta\Phi^T\Phi$$

Where, M_N is the mean and S_N is the covariance of the Bayesian posterior conditional probability

$$M_N = S_N(S_0^{-1}M_0 + \beta\Phi^T T) \quad (40)$$

$$S_N^{-1} = S_0^{-1} + \beta\Phi^T\Phi$$

The prediction of the target vector at the next time step is represented as:

$$P(t \mid T, \alpha, \beta) = \int P(t \mid W, \beta) \cdot P(W \mid T, \alpha, \beta) dW \quad (41)$$

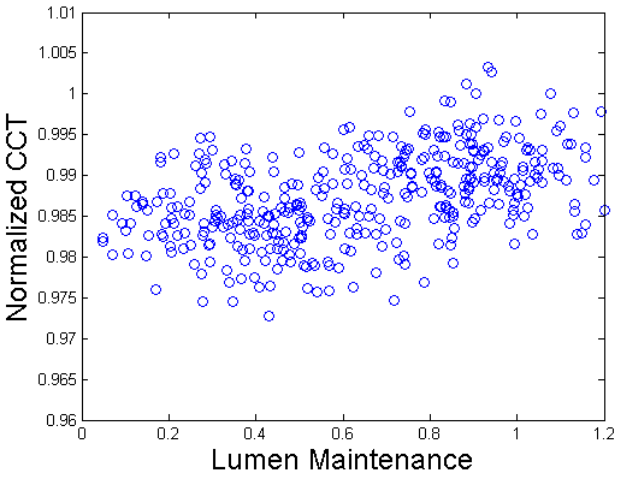
The condition distribution $P(t \mid T, \alpha, \beta)$ has been calculated out as the distribution and probability with its mean and variance depending on the variable 'x'; Therefore, we can finally predict each output 't' including luminous flux and correlated color temperature from each time series input 'x', such as:

$$P(t \mid x, T, \alpha, \beta) = \quad (42)$$

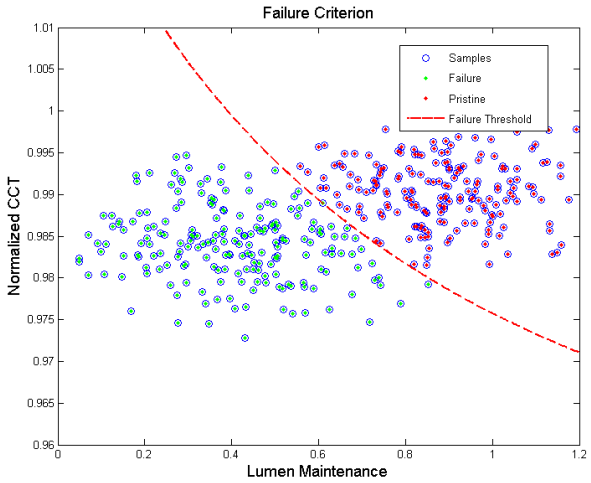
$$N(t \mid M_N^T\Phi(x), \beta^{-1} + \Phi^T(x)S_N\Phi(x))$$

IX. Failure Analysis Results

Once the Bayesian classifier has finished the training process, the data mapped onto the feature space is classified. The discriminant function has been used to classify the samples in the feature space and formulate a decision boundary between the lamps with accrued damage and pristine samples. The lamps migrate in the feature space from the top right to the bottom left with the increase in the amount of accrued damage. In Figure 10, the red data points are the healthy samples and the green data points are the samples with accrued damage. The red dash line shows the failure threshold between the healthy lamps and lamps with accrued damage. The decision boundary has been calculated such that the discriminant for the two classes has an equal value along the boundary.



(a)



(b)

Figure 10: (a) End of Life distribution and pristine for LED Lamp (b) Failure Criterion for Distribution Classification. LEGEND: The red dash line shows the failure threshold between the healthy lamps and lamps with accrued damage; red data points are the healthy samples and the green data points are the samples with accrued damage.

The values of the coefficients of the polynomial that describes the decision boundary have been calculated using the following equations,

$$F(x,y) = I^T W_2 I + W_1^T I + W_0 \quad (43)$$

$$I = \begin{bmatrix} x \\ y \end{bmatrix} \quad (44)$$

$$W_2 = -\frac{1}{2}\Sigma_i^{-1} \quad (45)$$

$$W_1 = \Sigma_i^{-1}\mu_i$$

$$W_0 = -\frac{1}{2}\mu_i^T\Sigma_i^{-1}\mu_i - \frac{1}{2}\log|\Sigma_i| + \log(p(C_i))$$

The matrix expression of $F(x,y)$ has been expanded as the quadratic area function:

$$F(x,y) = W_2(1,1)x^2 + 2W_2(1,2)xy + W_2(2,2)y^2 \quad (46)$$
$$+ W_1(1,1)x + W_1(2,1)y + W_0$$

The classification decision boundary for the failure threshold can be calculated from previous analysis as equating the PDFs for the classes on either side of the decision boundary:

$$G(x,y) = F_1(x,y) - F_2(x,y) = 0 \quad (47)$$
$$= 3.72x^2 - 157.74xy + 131.11x +$$
$$1823.78y^2 - 3984.71y + 2170.75$$

From the classification, the calculated coefficients of the polynomial are:

$$W0 = -2.7019 \cdot 10^4 \quad (48)$$

$$W1 = \begin{bmatrix} 0.0305 \\ 5.4802 \end{bmatrix} \cdot 10^4 \quad (49)$$

$$W2 = \begin{bmatrix} -0.0016 & -0.0149 \\ -0.0149 & -2.7783 \end{bmatrix} \cdot 10^4 \quad (50)$$

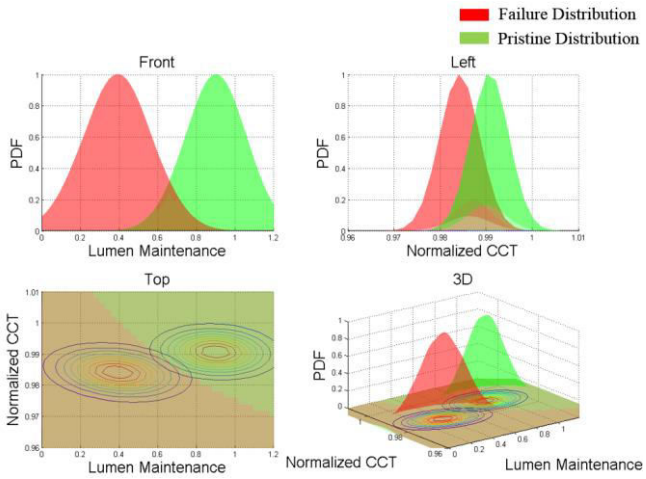


Figure 11: Classification of the lamps with accrued damage and pristine lamp PDFs. LEGEND: red PDF is corresponding to the lamps with accrued damage while the green PDF are the healthy lamps

Overall, the Bayesian unsupervised classifier is powerful classification tool. Even though two groups have been classified, the technique presented is applicable to multiple groups. The distributions corresponding to the healthy group with significant accrued damage has been plotted. The red PDF is corresponding to the lamps with accrued damage while the green PDF are the healthy lamps. The overlapping area displays the transition failure area between the healthy and the lamps with accrued damage. Typically, we want this overlapping PDF region to be as small as possible. The decision boundary has been updated as more data becomes available for the different classes. Figure 12 shows the three groups classification, which we assign the initial parametric distributions for the (a) failure threshold (b) the pristine LED Lamp group and (c) damaged LED Lamp group. The testing data has been grouped, and Bayesian Classifier calculates the mean and variance numerically.

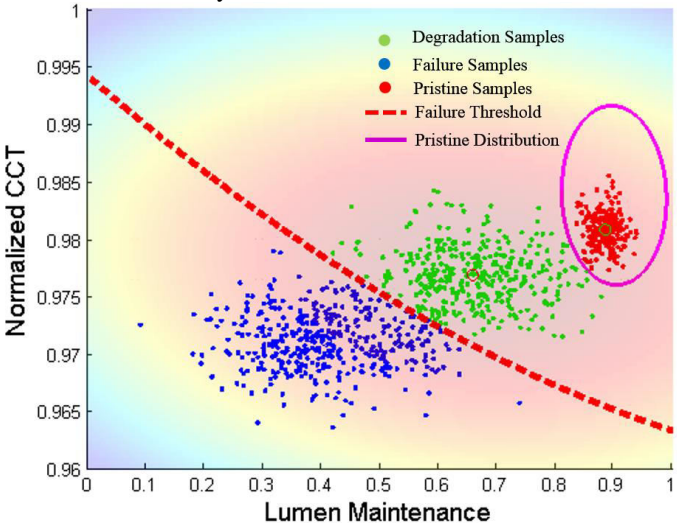


Figure 12: L70 Time Decision Boundary. LEGEND: RED dots – pristine LEDs; GREEN dots – LEDs with

significant accrued damage; BLUE dots – failed LEDs in which the lumen flux degradation has dropped below 70-percent of the initial luminous flux output in nearly all the tested samples.

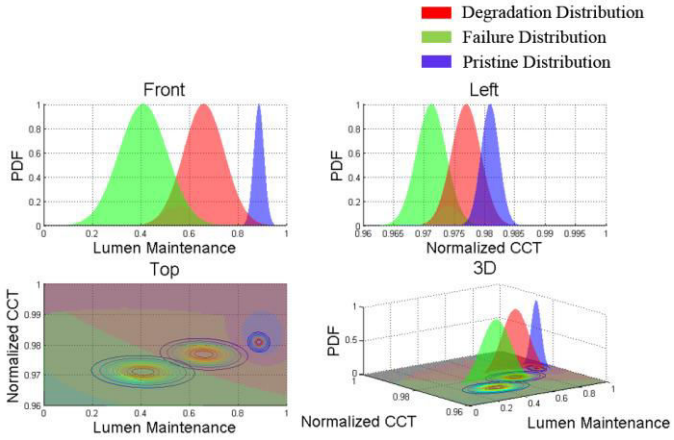


Figure 13: Critical Failure, LED Lamp Failure and Pristine Lamp PDF Distribution. LEGEND: BLUE– pristine LEDs; RED– LEDs with significant accrued damage; GREEN– failed LEDs in which the lumen flux degradation has dropped below 70-percent of the initial luminous flux output in nearly all the tested samples.

Figure 12 shows the migration of decision boundary. The decision boundary between pristine lamp group and failure lamp group is shown with a solid magenta ellipse, and the decision boundary between the failure threshold and the lamps with accrued damage is shown with a dashed red line. The Figure 13 demonstrates the three PDFs for the pristine LED Lamp, damaged lamp group as well as lamp group beyond the failure threshold. The decision boundary between the damaged lamp group and the group beyond the failure threshold has been termed as the critical failure boundary, which should not be breached to avoid failure.

X. Remaining Useful Life

Bayesian regression method has been used to determine the Remaining Useful Life (RUL) for every test lamp. Lumen Maintenance (LM) degradation has been used as the main indicator of system decay, by fitting the Lumen Maintenance degradation curve (Figure 14). Figure 15 shows the training of the Bayesian regression model through the maximum likelihood function and prediction of the posterior distributions. The process discussed previously in the Bayesian regression section, has been used for the future state prediction of the lamp's luminous flux and the remaining useful life. Figure 16 shows the Bayesian linear regression for the third order polynomial model with four weights. The green dots are the measured data points, and the red dots show the predicting decay curve. The testing length is up to 2537 hours. The Remaining Useful Life (RUL) has been calculated by predicting the future luminous flux output state till the L70 threshold.

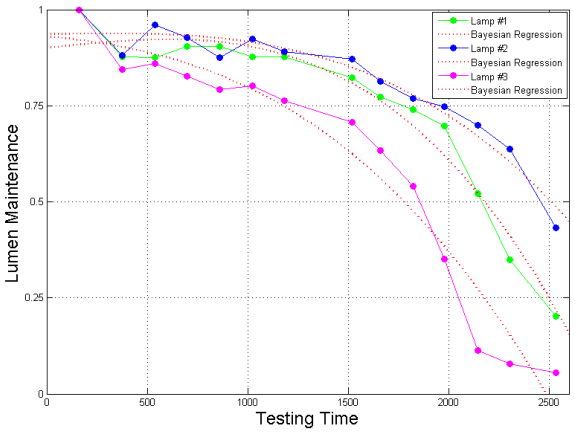


Figure 14: Lumen Maintenance Regression for LED Lamp

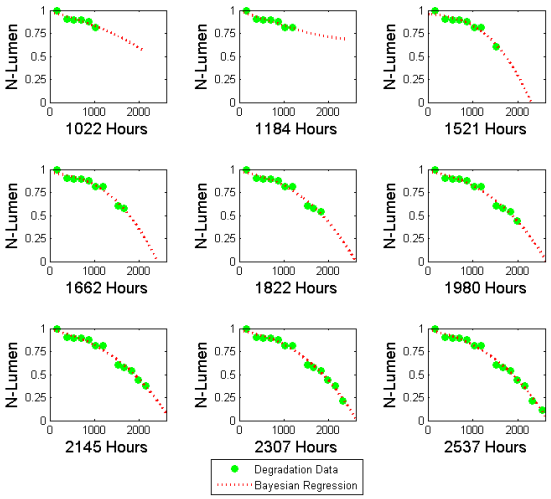


Figure 15: Bayesian Regression Learning Process. LEGEND: GREEN dots are the actual measured data; RED lines are the model prediction of the lumen flux output using Bayesian Regression.

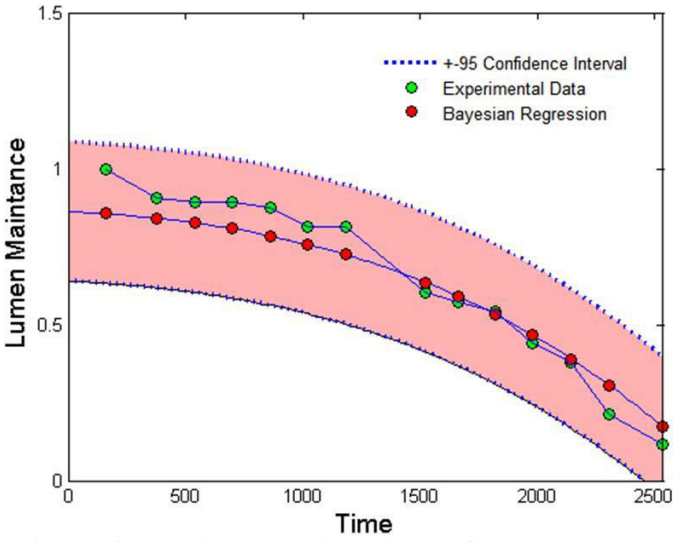


Figure 16: Bayesian Regression with Confidence Interval

The predicted RUL is known as the predicted End of Life (EoL) minus the sampling time, represented by:

$$\mathbf{T}_{\text{predict}} = \mathbf{P}_{\text{EoL}} - \mathbf{T}_{\text{sample}} \qquad (51)$$

The real RUL is known as the actual EoL minus the sampling time. So the algebra equation presents as following:

$$T_{actual} = A_{EoL} - T_{sample} \quad (52)$$

A two parameter weibull distribution has been used to model the lamp failures. The probability density function for the two parameter Weibull distribution has the following form:

$$f(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} e^{-\left(\frac{t}{\eta}\right)^{\beta}} \quad (53)$$

Where β is the shape parameter, η is the characteristic life. The estimated shape parameter and the characteristic life are: $\beta = 7.1$ and $\eta = 1790$ hours. Since the $\beta > 1.0$, it indicates that the failures are wear out failures. The Weibull cumulative distribution, the population fraction failing by time t is given as following CDF:

$$F(t) = 1 - e^{-(t/\eta)^{\beta}} = 1 - e^{-(t/1790.1)^{7.1}} \quad (54)$$

The reliability function is thereby given by the 1-F(t). The CDF indicates that once time reached 1700.2 hours the LED lamp reliability dropped to 50%. The characteristic life (B63.2 life) is 1790.1 hours, which says 63.2% LED lamps have failed at this time in the accelerated test condition of 85°C/85%RH.

$$R(t) = 1 - F(t) = e^{-(t/\eta)^{\beta}} = e^{-(t/1790.1)^{7.1}} \quad (55)$$

Figure 17 shows the representative samples of the failed lamps with and without the lens. Note that encapsulant of several of the LEDs in the failed lamps shows distinct discoloration. It is hypothesized that the discoloration of the encapsulant was a major contributor to the degradation in the luminous flux output and the color shift during the 85C/85%RH accelerated test.

Figure 17 Representative Samples of the Failed Lamps with and without the Lens.

XI. Summary and Conclusions

The 60W LED lamps have been studied under the accelerated test conditions of 85°C/85%RH for both luminous flux output and the correlated color temperature. A Bayesian framework for early classification of the failed lamps in the luminous flux and correlated color temperature feature space has been formulated and demonstrated on the test-population of the lamps. Failures have been identified because of problems of luminous flux degradation or color shift or both. In addition the Bayesian regression model has been developed to predict the luminous flux degradation till the L70 threshold widely used as definition of failure for the solid state luminaires. The proposed methodology allows the early identification of the onset of failure much prior to development of complete failure distributions and can be used for assessing the damage state of SSLs in fairly large deployments. The α-λ plots have been used to evaluate the robustness of the proposed methodology. Results show that the predicted degradation for the lamps tracks the true degradation observed during 85°C/85%RH during accelerated life test fairly closely within the ±20% confidence bounds. Failure modes of the test population of the lamps have been studied to understand the failure mechanisms in 85°C/85%RH accelerated test. Results indicate that the dominant failure mechanism is the discoloration of the LED encapsulant inside the lamps which is the likely cause for the luminous flux degradation and the color shift.

Acknowledgments

The work presented here in this paper has been supported by a research grant from the Department of Energy under Award Number DE-EE0005124.

References

[1]. Baillot, R., Deshayes, Y., Bechou, L., Buffeteau, T., Pianet, L., Armand, C., Voillot, F., Sorieul, S., Ousten, Y. "Effects of Silicone Coating Degradation on GaN MQW LEDs Performances Using Physical and Chemical Analysis." *Microelectronics Reliability* 50(2010): 1568-1573.

[2]. Bishop, C. M., Pattern Recognition and Machine Learning, Springer Science-Business Media, LLC, 2006.

[3]. Byung-Lip Ahn, Cheol-Yong Jang, Seung-Bok Leigh, Seunghwan Yoo, Hakgeun Jeong, Effect of LED Lighting on Cooling and Heating Load in Buildings, Applied Energy, Volume 113, Pages 1484–1489,

January 2014.
[4]. Chan, S.I., W.S. Hong, K.T. Kim, Y.G. Yoon, J.H. Han, J.S. Jang, Accelerated life test of high power white light emitting diodes based on package failure mechanisms, Microelectronics Reliability 51 (2011) 1806–1809.
[5]. Choi, M., Ki Hyun Kim, Changhun Yun, Dai Hyoung Koo, Sang Bin Song, Jae Pil Kim, Direct correlation between reliability and pH changes of phosphors for white light-emitting diodes, Microelectronics Reliability 54 (2014) 2849–2852.
[6]. Cree. "Cree Xlamp XR Family LED Reliability, CLD-AP06 Rev. 7." Cree Inc. 2009.
[7]. Duda, R.O., Peter E. Hart, David G. Stork. Pattern Classification, Second Edition, Wiley & Sons, ISBN: 978-0-471-05669-0, 2001
[8]. Fu, Han-Kuei, Chin-Wei Lin, Tzung-Te Chen, Chiu-Ling Chen, Pei-Ting Chou, Chien-Jen Sun, Investigation of dynamic color deviation mechanisms of high power light-emitting diode, Microelectronics Reliability 52 (2012) 866–871.
[9]. Hewlett Packard. "Reliability of Precision Optical Performance AllnGaP LED Lamps in Traffic Signals and Variable Message Signs." *Application Brief I-004* (1997)
[10]. Hsu, Y., Lin, Y., Chen, M., Tsai, C., Kuang, J., Huang, S., Hu, H., Su, Y. and Cheng, W. "Failure Mechanisms Associated with Lens Shape of High-Power LED Modules in Aging Test." *IEEE Trans. on Electron Devices* 55(2008): 689-694.
[11]. IES (Illuminating Engineering Society), IES LM-80-08. "Approved Method: Measuring Lumen Maintenance of LED Light Sources." 2008
[12]. IES (Illuminating Engineering Society), IES TM-21-11. "Projecting Long Term Lumen Maintenance of LED Light Sources" 2008
[13]. Jang, Daeseok, Se-Jin Yook, Kwan-Soo Lee, Optimum design of a radial heat sink with a fin-height profile for high-power LED lighting applications, Applied Energy, Volume 116, 1, Pages 260–268, March 2014
[14]. Lall, P., Sakalaukus, P., Davis, L., Prognostics of Damage Accrual in SSL Luminaires and Drivers Subjected to HTSL Accelerated Aging, Proceedings of the ASME 2013 International Technical Conference & Exposition on Packaging and Integration of Electronic and Photonic Microsystems, InterPACK2013-73250, pp. 1-8, Burlingame, CA, 2013.
[15]. Lall, P., Wei, J., & Davis, L., Prediction of L70 lumen maintenance and chromaticity for LEDs using extended Kalman filter models, SPIE Optical Engineering+ Applications (pp. 88350M-88350M-19), San Diego, CA, 2013.
[16]. Lall, P., Wei, J., Davis, L., L70 Life Prediction for Solid State Lighting Using Kalman Filter and Extended Kalman Filter Based Models, Electronic Components and Technology Conference, ECTC, 63rd, pp.1452-1465, 2013.
[17]. Lall, P., Wei, J., Davis, L., Solid State Lighting Life Prediction Using Extended Kalman Filter, Proceedings of the ASME 2013 International Technical Conference & Exposition on Packaging and Integration of Electronic and Photonic Microsystems, InterPACK2013-73288, pp. 1-11, Burlingame, CA, 2013.
[18]. Lall, P., Zhang, H., Davis, L., Assessment of Lumen Degradation and Remaining Life of LEDs using Particle Filter, Proceedings of the ASME 2013 International Technical Conference & Exposition on Packaging and Integration of Electronic and Photonic Microsystems, InterPACK2013-73305, pp. 1-13, Burlingame, CA, 2013.
[19]. Lumileds. "Luxeon Reliability." *Philips Lumileds Reliability Datasheet RD25* (2006).
[20]. Luo, X., Wu, B. and Liu, S. "Effects of Moist Environments on LED Module Reliability." *IEEE Trans. on Device and Materials Reliability* 10(2010): 182-186.
[21]. Meneghini, M., M. Dal Lago, L. Rodighiero, N. Trivellin, E. Zanoni, G. Meneghesso, Reliability issues in GaN-based light-emitting diodes: Effect of dc and PWM stress, Microelectronics Reliability 52 (2012) 1621–1626.
[22]. Meneghini, M., Trevisanello, L., Meneghesso, G. and Zanoni, E. "A Review on the Reliability of GaN-Based LEDs." *IEEE Trans. on Device and Materials Reliability* 8(2008): 323-331.
[23]. Narendran, N., Gu, Y., Freyssinier, J.P., Yu, H. and Deng, L. "Solid-State Lighting: Failure Analysis of White LEDs." *Journal of Crystal Growth* 268(2004): 449-456.
[24]. Nichia. "Specifications for Nichia Chip Type White LED Model: NCSW119T-H3. Nichia STS-DA1-0990A." Nichia Corp. 2009
[25]. Sawyer, S., Rumyantsev, S., & Shur, M. (2008). Degradation of AlGaN-based ultraviolet light emitting diodes. Solid-State Electronics, Elsevier, Pages 968-972.
[26]. Saxena, A., Celaya, J., Saha, B., Goebel, K., Evaluating algorithm performance metrics tailored for prognostics. IEEE Aerospace conference, 2009[a].
[27]. Saxena, A., Celaya, J., Saha, B., Saha, S., Goebel, K., On Applying the Prognostics Performance Metrics, Annual Conference of the PHM Society, vol 1, San Diego, CA, 2009[b].
[28]. Saxena, A., J. Celaya, B. Saha, S. Saha, and K. Goebel, Evaluating Algorithm Performance Metrics Tailored for Prognostics, IEEE Aerospace Conference, Big Sky, MT, pp. 1-11, March 2008[a].
[29]. Saxena, A., J. Celaya, E. Balaban, K. Goebel, B. Saha, S Saha, and M. Schwabacher, Metrics for Evaluating Performance of Prognostic Techniques, Intl. Conf. on Prognostics and Health Management, Denver, Colorado, pp. 1-17, October 2008[b].
[30]. US Energy Information Administration, Retrieved from U.S Energy Information Administration, Page accessed June 27, 2014, http://www.eia.gov/tools/faqs/faq.cfm?id=99&t=3, 2012
[31]. Wang, F.-K. (2012). Lifetime predictions of LED-based light bars by accelerated degradation test.

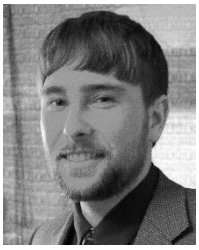
Microelectronics Reliability, Elsevier, Pages 1332-1336.

[32]. Yang, Shih-Chun, Pang Lin, Chien-Ping Wang, Sheng Bang Huang, Chiu-Ling Chen, Pei-Fang Chiang, An-Tse Lee, Mu-Tao Chu, Failure and degradation mechanisms of high-power white light emitting diodes, Microelectronics Reliability 50 (2010) 959–964.

AUTHOR BIOGRAPHY

Pradeep Lall (M'93—SM'08—F'12) is the Thomas Walter Professor with the Department of Mechanical Engineering, and the Director of the NSF Center for Advanced Vehicle and Extreme Environment Electronics at Auburn University. He received the B.E. degree in mechanical engineering from the Delhi College of Engineering, Delhi, India, in 1988, the M.S. and Ph.D. degrees in mechanical engineering from the University of Maryland, College Park, in 1989 and 1993, respectively, and the M.B.A. degree from Kellogg School of Management, Northwestern University, Evanston, IL, in 2002. He was previously with Motorola's Wireless Technology Center. He has published extensively in the area of electronic packaging with emphasis on modeling and predictive techniques. He is author and co-author of 2-books, 14 book chapters, and over 400 journal and conference papers in the field of electronic packaging with emphasis on design, modeling, and predictive techniques. Dr. Lall is Fellow of the IEEE, a Fellow of the ASME, and a Fellow of the Alabama Academy of Science. Dr. Lall is the recipient of the IEEE Exceptional Technical Achievement Award, ASME's Applied Mechanics Award, SMTA's Member of Technical Distinction Award, Auburn University's Creative Research and Scholarship Award, SEC Faculty Achievement Award, Samuel Ginn College of Engineering Senior Faculty Research Award, Three-Motorola Outstanding Innovation Awards, Five-Motorola Engineering Awards, and Twenty Best-Paper Awards at national and international conferences. He holds three U.S. Patents. Dr. Lall has served in several distinguished roles at national and international level including serving as member of National Academies Committee on Electronic Vehicle Controls, Member of the IEEE Reliability Society AdCom, IEEE Reliability Society Representative on the IEEE-USA Government Relations Council for R&D Policy, Chair of Congress Steering Committee for the ASME Congress, Member of the technical committee of the European Simulation Conference EuroSIME, and Associate Editor for the IEEE Transactions on Components and Packaging Technologies. He is a Six-Sigma Black-Belt in Statistics. He is the founding faculty advisor for the SMTA student chapter at Auburn University, and a member of the editorial advisory board for SMTA Journal.

Junchao Wei received the B.E., degree from Hunan University of Technology, Zhuzhou, China, in 2011. Currently, he is Ph.D. student in Mechanical Engineering at Auburn University, and works as a research assistant for the CAVE3 (The NSF Center for Advanced Vehicle and Extreme Environment Electronics). His research areas include LED Reliability models, X-ray CT scan detection and three dimensional reconstruction; deformation and strain measurement; electronics reliability and prognostics; algorithm development; computational models for reliability of thermal, shock, drop, and vibration.



Peter J. Sakalaukus, Jr. (M'12) received the B.S. degree in mechanical engineering from Mississippi State University, Starkville, MS, in 2006 and the M.S. degree in mechanical engineering from the University of South Alabama, Mobile, AL, in 2011. He is currently pursuing the Ph.D. degree in mechanical engineering at Auburn University, Auburn, AL.

Data Driven Approach for Drill Bit Monitoring

[1] Nishchal K. Verma, [1] Rahul K. Sevakula, [1] Sonal Dixit and [2] Al Salour
[1] Department of Electrical Engineering,
Indian Institute of Technology Kanpur, India
[2] The Boeing Company, St. Louis, MO, USA
E-Mail: {nishchal, srahulk, dsonal}@iitk.ac.in and al.salour@boeing.com

Abstract — **Drill Bit Monitoring has become an important part of Automation in manufacturing industries, especially when product quality and efficiency are highly demanded. Conventional drill monitoring methods use imaging and process parameters like thrust, cutting force, torque, feed rate etc. for detecting presence of wear and tear in drill bits. This paper presents a study where various Pre-processing, Feature Extraction, Dimensionality Reduction and Classification strategies are used with a target to find the most accurate and time efficient set of strategies for vibration based drill bit monitoring. The entire fault recognition process has been made simple by developing and implementing all the selected strategies onto a smartphone application (App). The App is able to quickly perform the recognition process of recordings present on Cloud and internal storage. A case study has been performed on a Computerized Numerical Control (CNC) machine having drill bits of 9mm diameter in 4 states namely Healthy and three faulty states. The proposed model is able to correctly predict the drill bit state by analyzing a one second vibration recording with accuracy of 95.5%.**

Keywords — drill bit monitoring; condition monitoring; vibration analysis; tool wear monitoring; spectral analysis; SVM

I. Introduction

Machine Tool automation has great impact in reducing human efforts and also in improving precision of the work done. Tool wear monitoring forms an integral part of Machine Tool automation which also makes production line systems more reliable. Literature Survey [1] states that Drilling accounts for up to 50% of all machining operations in the United States of America; thus making Drilling process one of the most common machining processes in industries [2]. Due to its utter importance in machining operation, tool wear monitoring of cutting tools have been quite popular among researchers [3-6].

Literature study gives multiple investigations for drill bit wear detection based on acoustic, vibration, thrust, torque, power and current measurements. Patra [7] and Everson et al. [8] presented Tool wear monitoring systems using acoustic emission (AE) signature analysis. Patra's scheme used time and wavelet domain based features with back propagation algorithm based artificial neural network (ANN) for tool wear prediction. Everson's experiments demonstrated a relation of AE signals with hole size and lip height variation. However due to sensor mounting and noise attenuation issues, AE based methods are generally deemed to be slightly inaccurate [6]. Ertunc et al. [9] worked with cutting force signals namely thrust and torque using dynamometer to detect the drill wear. Among the four tested methods, HMM based approach was found to be most feasible and reliable solution. Cuppini et al. [10] derived a mathematical relation between wear and cutting power, and used this measure for detection of wear. Patra et al. [11] developed current signal based back propagation network to predict the wear on high speed steel (HSS) drill bit. During drilling, thrust and torque force causes vibration on the surface. Considering this, Wardany et al. [12] presented a threshold controlled based approach with vibration signatures to detect drill wear. Isaan [13] further presented a vibration based approach using harmonic wavelets, entropy and spectrum features with ANN for drill wear detection.

This paper presents a study of drill bit monitoring with accelerometer recordings by describing how the authors validated various pre-processing, feature extraction, dimensionality reduction and classification strategies before deciding the final strategy. After identifying the best strategy, a Smartphone application has been developed, which is able to recognize the presence of faults in the drill bit within 10 seconds. The contribution of this paper is that as compared to previous strategies, a much simpler and faster method has been presented and developed for convenience, with recognition accuracy of up to 95.5%.

The paper further proceeds as follows. Section II throws some light regarding drill bit monitoring. Section III gives details of the variations tested in preprocessing, feature extraction, dimensionality reduction and classification steps. Section IV presents results and a thorough analysis of the effectiveness of features, Principal Component Analysis (PCA) and classifiers. Section V talks about the model that was finally used in making real time drill bit monitoring system for Android based smartphone and finally Section VI concludes our presented work.

II. Drill Bit Monitoring

A typical drilling operation requires drilling machine, work piece, fixture and cutting tool. Fig. 1 shows typical geometry of a drill bit. It has a cone like structure consisting of chisel edge, cutting lips, web, flute, heels, body and shank. As soon as drill bit comes in contact with the work piece, it starts penetrating via chisel edge. The cutting edges then chips off the material during penetration in first stage, also known as penetration stage. At second level, as drill bit moves into the material, flute edges do similar work as cutting edges, but in a more refined manner which makes the drilled holes look clearer. This stage is known as steady stage.

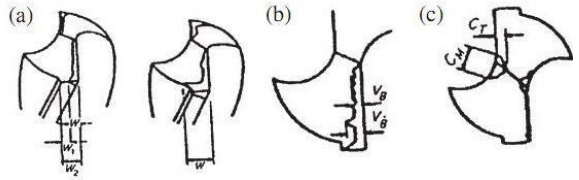
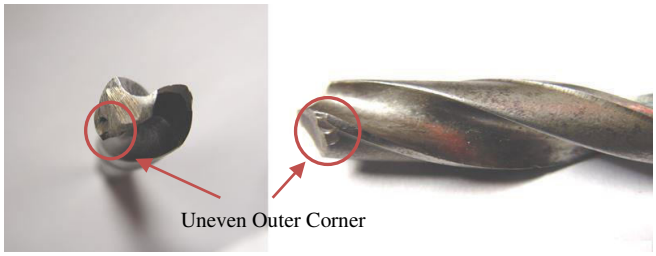
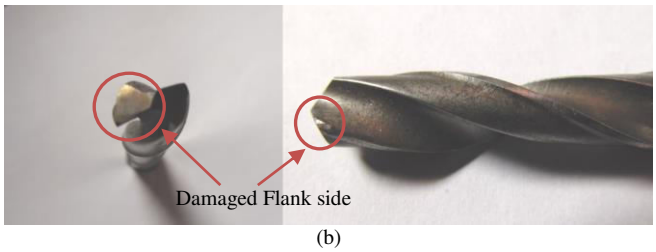


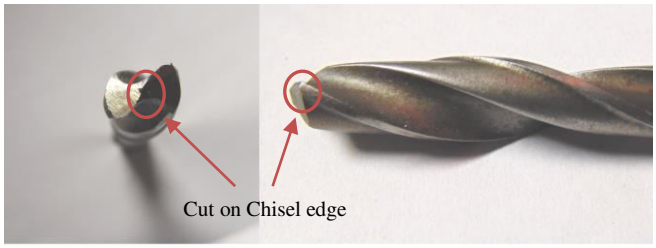
Fig.1 Different types of wear (a) Outer corner wear, (b) Flank wear, (c) Chisel wear. [14]



(a)



(b)



(c)

Fig. 2. Actual images of a Drill bit having following faults: (a) Outer Corner wear, (b) Flank wear, (c) Chisel wear [30]

As drilling continues, deformations due to physical and mechanical reactions begin to occur in the drill bit. In the case study presented here, four states of drill bit namely healthy state and three faulty states have been studied and analyzed. The three faults that have been studied here are as follows:

- Chisel wear: Drilling process starts as soon as chisel point penetrate into work piece. Due to high shear and stress, the temperature at chisel point gets raised and results into a blunted chisel edge as shown in Fig. 1(a).
- Flank Wear: Due to friction between the work piece and flank of drill bit, abrasive wear occurs. As shown in Fig. 1(b), v_b and v_b' flank surfaces get abrasive. This abrasion increases as cutting speed increases.
- Outer corner wear: Due to high impact forces and friction between drill bit and hole's inner head wall, the outer corner of the drill erodes or gets chipped off. This defect causes wear on one or both outer corners of the drill point as shown by points C_m and C_t in Fig 2(c).

III. Fault Recognition Model

For detection of faults, a supervised learning approach has been modeled and implemented. Supervised learning approach consists of two phases namely training phase and testing phase. Training phase includes development of robust model/s using pre-recorded training samples and in testing phase, the developed model is used for finding out the class that test samples belong to. In Training phase, firstly necessary training data are acquired to form a set of recordings to learn from. These recordings are pre-processed to improve its quality and later features i.e. signal characteristics are extracted from the same. Having a reduced set of important features can improve the generalization capability of classifiers. For this reason, dimensionality reduction techniques are used to learn/recognize the reduced set of features. Finally, a classifier is made to learn relationship between samples' reduced features and their respective class label. In Test phase, whenever a test recording arrives, the recording is first pre-processed, then features are extracted from it, then dimensionality is reduced based on earlier learning, and finally using the learnt classifier model, class/category of the test sample is recognized. Details for each step have been given below.

A. Data Acquisition

To capture the vibration signature of machine, uniaxial accelerometer, PCB 63001 was used. The sensor was mounted on the work piece and using NI data acquisition (DAQ) [15] system, analog vibration sensor signals were converted into sampled signals at a sampling rate of 32,768 Hz. NI DAQ system consists of NI 9234 Signal Processing unit in conjunction with NI 9172 chassis connected to a Desktop PC.

B. Preprocessing

Various kinds of noise due to environmental conditions and sudden jerks may also become part of the recording. To reduce effects of such noise, recordings are preprocessed. In pre-processing, 4 steps were incorporated in a consecutive fashion, namely frequency filtering, clipping, smoothing and 0-1 scaling. Frequency filtering step has a low pass 20 order Butterworth filter with cut-off frequency of 12 kHz. Clipping step is used for making systems robust to sudden jerks. For clipping, the entire recording is divided into multiple overlapping windows, for e.g. in this experiment, each recording was divided into 8 overlapping windows having 8192 samples each. The window having least standard deviation was selected for further process. Such window with low standard deviation is expected to give the stable part of recording that is free from sudden jerks. The third step of smoothing is meant to smoothen the recording and

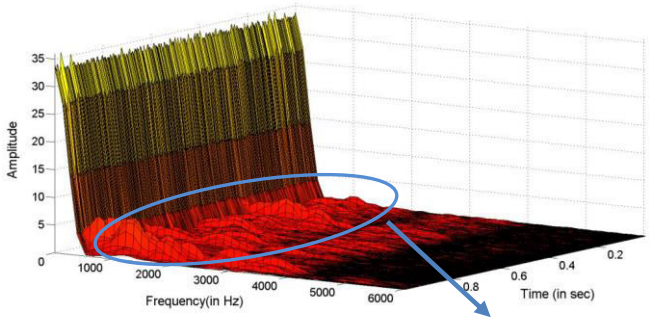
reduce effects of outliers in the recording. A moving average filter has been used for the same. The fourth step is scaling where smoothed recording is scaled from 0 to 1 using max-min normalization.

C. Feature Extraction

One of the most important tasks in analyzing signals is identification of some important signal characteristics which are capable of representing the signal in entirety, at least up to great extent. Such signal characteristics are termed as features here. Here features have been extracted using Statistical parameters in time domain, Fast Fourier Transform (FFT) coefficients, Discrete Cosine Transform (DCT) coefficients, Morlet Wavelet Transform coefficients and Wavelet Packet Transform node energies.

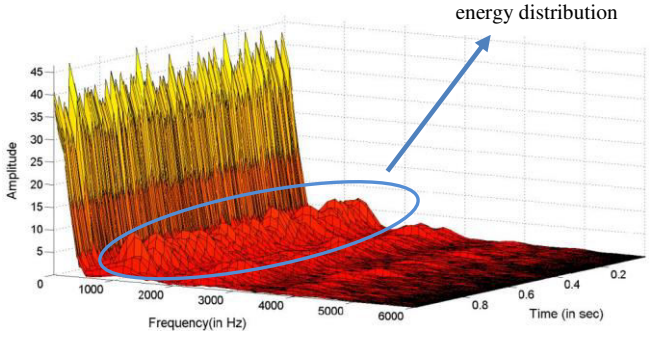
1) Time Domain : In general, signal is stored in the form of sampled signal values, showing how amplitude of the signal changes w.r.t time. Popular statistical parameters were used to extract 14 features in this domain and hence give a fair idea of the distribution and peakedness of data in signal. These 14 statistical parameters [16] inlcude non-parametric features namely absolute mean, root mean square, shape factor, crest factor and parametric features namely median, mean of peaks above upper quartile, variance, skewness, kurtosis, upper quartile, interquartile, negentropy found by approximation [17], Hjorth's complexity and Hjorth's mobility [18].

2) FFT and DCT : Spectral analysis has always been popular in vibration signal analysis [19]. For non-stationary signals, FFT gives an average picture of the signal energy distribution across frequency components. FFT gives both, phase and amplitude information of the signal. Before extracting features, spectogram plots of pre-processed signal samples from all four states were found using Short Time Fourier Transform (STFT). These plots have been shown in Fig. 3. Two inferences could be made from the figures. First

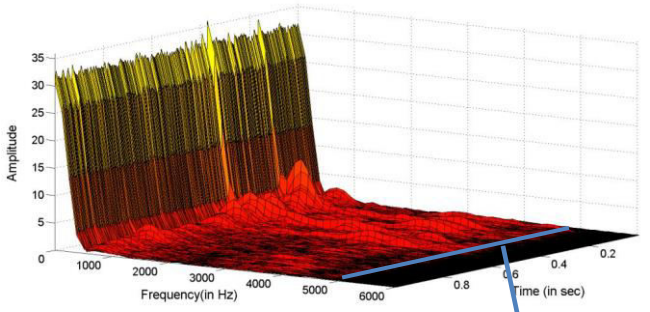


(a)

Change in spectral energy distribution

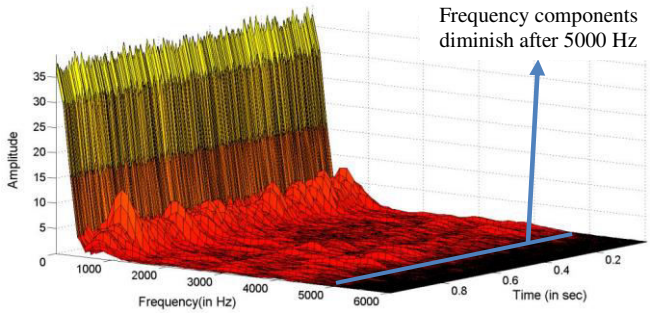


(b)



(c)

Frequency components diminish after 5000 Hz



(d)

Fig. 3 Spectrogram Plots of Accelerometer Recordings while drilling is in Steady state and Drill bit is in following states - (a) Flank wear fault, (b) Chisel wear fault, (c) Outer Corner wear fault and (d) Healthy

is that there are clear differences in the four plots; thus showing that spectral analysis is sufficient to discriminate recordings of the four states. Second inference is that there are no significant frequency components after 5kHz. For this reason, only spectrum of 0-5kHz was considered while extracting features. After performing FFT, the obtained spectrum from 0-5Khz was divided into 14 equal segments, called bins. The ratio of individual bin energy to the total energy of all 14 bins gave 14 features in frequency domain

DCT transform represents signal into sum of small cosine functions with different frequencies. More than representing the signal in spectral domain, it is primarily used for compactly representing the signal with few DCT coefficients. As most of the energy is concentrated in earlier coeffecients itself, starting from the first DCT coefffecient, ceffecients upto which 99.5% energy is present, are considered while extracting features. The considered DCT coeffecients are divided into 14 equal bins. Similar to FFT features, ratio of individual bin energies to total energy were calculated to give 14 DCT features.

3) Morlet Wavelet Transform (MWT) : Every Wavelet Transform has a basis function termed as mother wavelet. When signal is convolved with the translated and scaled versions of mother wavelet, a variety of signal characteristics i.e. time-frequency information become prominent, observable and measurable. Due to its simplicity and similarity with periodic impulses, Morlet wavelet is well known for use in vibration signal analysis. Drill wear causes impulse variations in vibration recordings and are important to be captured for recognizing the drill bit's current state. MWT coeffecients have therefore been considered here for giving useful features.

$$y_{(a,b)} = e^{\frac{-b^2(t-b)^2}{a^2}} \cos\left(\frac{\pi(t-b)}{a}\right) \quad (1)$$

Morlet wavelet is mathematically defined in (1). Two scales of Morlet wavelet have been considered here, one with a = 8 and other with a = 16. These values were decided by experimenting with various possible values and finding which pair amongst them gave best results. Signal is then convolved with both wavelets to give two sets of MWT coeffecients. From each set, 7 features have been extracted by finding the coeffecients' statistical parameters namely standard deviation, variance, skewness, kurtosis, sum of peaks, variance and zero crossing rate. In total, thus 14 features are got.

4) Wavelet Packet Transform (WPT) : WPT is considered an important tool while studying non stationary signals [21]. It is fast to compute and works on a similar principle as in Discrete Wavelet Transform (DWT). Whereas decomposition in DWT generally occurs only with approaximation coeffecients, in WPT they occur on both sides i.e. with approximation coeffecients as well as with detailed coeffecients. The transform can be considered as a series of low pass and high filters at many scales to give approximation and detailed coefficients at each level. When decomposition occurs upto l level, it can be seen as a balanced binary tree of l levels and $2^{l+1}-1$ nodes, with each node containing either approximation coeffecients or detailed coeffecients. As level increseses, time resoltion gets poorer while frequency resolution gets better. For extracting features, WPT coeffecients are found upto 3rd level. Not considering the topmost node, a total of 14 nodes with coeffecients are found. Calculating the node energies of all 14 nodes gives 14 features.

D. Dimensionality Reduction

As known from the Curse of Dimensionality issue, having more features with less training samples is very bad from generalization point of view. The situation worsens when many of the features are redundant in information. To avoid this situation, two ways were used to reduce the number of features. Firstly domain/transform wise features were individually used to see if features from a single transform are sufficient for the recognition process. Apart from that, Principal Component Analysis, a well-known dimensionality reduction technique was tested separately. PCA works on the principle of transforming correlated feature space to an orthogonal uncorrelated space and then choose those basis functions which have maximum variance. More details regarding PCA can be found in [22].

E. Classification

After getting the reduced set of features, for learning relationship between features and the respective class labels, a classifier is built over the training samples. Three well known classifiers namely Support Vector Machine (SVM), Bayes Classifier and Artificial Neural Network have been tested here.

1) Support Vector Machine: SVM [23] is a very popular binary classifier which performs extremely well with less training data. In a linearly separable case, SVM finds the hyper-plane which separates the two classes with maximal margin. Maximal margin ensures lower VC dimension; hence higher generalization i.e. better performance on test samples. On similar concept, solutions for non linearly separable cases and non separable cases are also found. In this paper, *C*-SVM with RBF kernel is used. The parameters namely C and γ were found with crossvalidation based performance checks while varying C and γ through various possible values, also known as grid search method. The pair of (C , γ) giving best crossvalidation performance value was finally used for building the SVM classifier.

2) Bayes Classifiers: Bayes classifier [24] is a probablistic classifier based on Bayes theorem which assumes all the features to be linearly independent. Using prior information of individual classes and likelihood of samples, the classifier for a test sample computes posteriori probability of each class using (2). The class having maximum posteriori probability then becomes the assigned class of the sample. When priori probabilities and likelihood are not available, the same are learned from training samples. Priori probability is based on deductive reasoning and not on past beahavior, whereas posteriori probability is given by accounting relevant evidence and background.

$$P\left(c_i|\mathbf{x}_w,\mathbf{X}\right) = \frac{p\left(\mathbf{x}_w|c_i,\mathbf{X}\right)P\left(c_i|\mathbf{X}\right)}{\sum_{c'=1}^{L} p\left(\mathbf{x}_w|c',\mathbf{X}\right)P\left(c'|\mathbf{X}\right)} \quad (2)$$

In (2), c_i refers to the class under consideration for finding posteriori probability and L refers to the number of classes. $\mathbf{X}$ refers to training data used for getting prior information and $\mathbf{x}_w$ refers to the test sample under consideration. Bayes classifier acts as an important benchmark in classification and works best for large training data.

3) Artificial Neural Network: Artificial Neural Network (ANN) is an extremely powerful tool for learning very complex feature-feature and feature/s-class relationships. ANN started by mimicing biological neural network and is generally based on adaptive neurodynamics structure. It is typically defined by three types of parameters : a) different layers of neurons having interconnection pattern, b) learning techniques involved for the updation of interconnection weights and c) conversion of neuron's weighted input to its output by an activation function. Number of neurons in input layer and output layer is equal to number of input features and number of output classes respectively. The network may contain single or multiple hidden layers. As single hidden layer suffices for most purposes, only single hidden layer network has been considered here. There is no proven best criteria to determine the number of neurons in hidden layer. It is either found by empirical formulations like crossvalidation checks or by experience based thumbrules, for eg. [25] recommends number of hidden neurons should fall in between the input and output layer size and [26] suggests that the number should not be more than the size of input layer, etc. In the case study presented here, number of hidden neurons have been equated to the mean of input layer number and output layer number of neurons. As mentioned earlier, ANNs are extremely

powerful and flexible; due to which they are also known as arbitrary function approximator i.e. they can fit any function. This flexibility though makes it prone to overfitting and may largely provide suboptimal results. To avoid such situations, it is encouraged to have more training data with good variety.

IV. Case Study, Results and Discussions

The entire experimentation was performed with 3-Axis CNC EMCO Concept Mill 105. HSS twist drill bit of diameter 9 mm was used for drilling holes in the work piece made of Mild steel. The target of the case study was to make a state recognition system that is able to accurately recognize 4 drill bit states namely Healthy state, Flank wear state, Chisel wear state and Outer corner wear state. For extensive experimentation, given a drill bit state, for each pair of varying feed rates and cutting speed combinations, a single vibration recording of 8 seconds was taken [30]. Feed rate was varied as 4 mm/min, 8 mm/min and 12 mm/min, and Cutting speed was varied as 160rpm, 170rpm, 180rpm, 190rpm and 200rpm; giving a total of 15 combination pairs. To increase training samples, the 8 second recording was then divided into 8 one second recordings. Thus for each state, 15x8 = 120 recordings were taken i.e. a total of 480 recordings for 4 states. Additionally, for analyzing fault recognition capability in steady stage and penetration stage of drilling, separate sets of 480 recordings were taken for both stages.

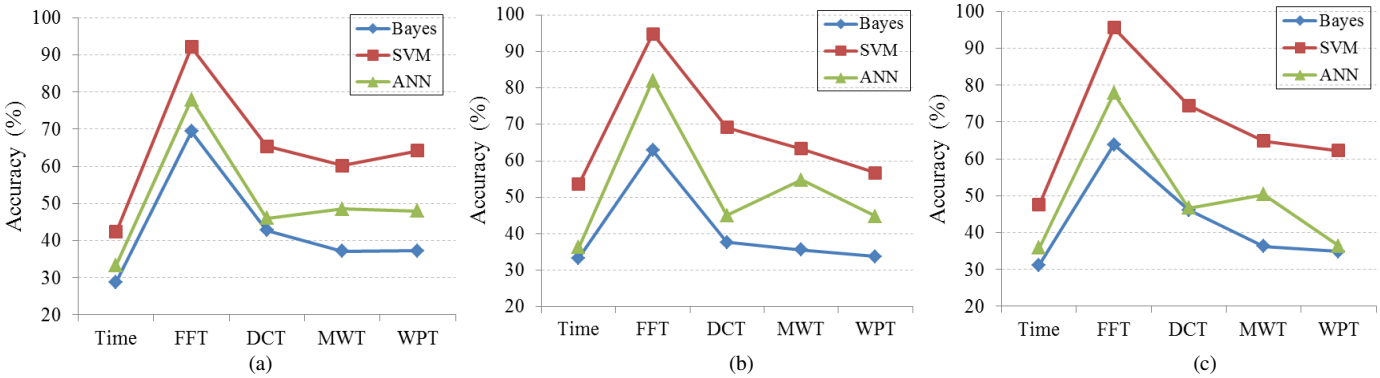


Fig. 4. Performance plot with Domain wise Features for (a) Steady State, (b) Penetration State and (c) Both Steady state and Penetration state combined

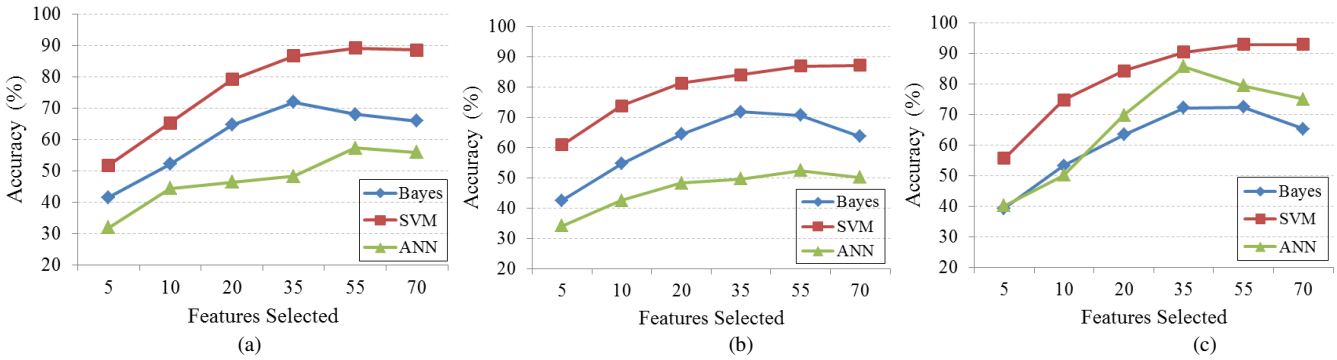


Fig. 5. Performance plots with PCA for (a) Steady State, (b) Penetration State and (c) Both Steady state and Penetration state

Overall 6 experiments were performed whose results are shown by 6 plots given in Figures 4 and 5. The three classifiers namely Bayes, SVM and ANN have been tested separately and their results are shown by three separate performance plots in all figures. In all experiments, average of 4 fold cross validation classification accuracies was used as the performance measure and forms the Y axis of result plots. Contrary to Bayes classifier and SVM, ANN produces different results every time it is run, as the minima it reaches during optimization depends on the random starting point. For this reason, while generating results with ANN, the classifier was tested thrice and average of the three outcomes was noted to make plots. Also as mentioned earlier, 8 recordings of 1 second were taken for each pair of feed rate and cutting speed in each drill bit state. While forming the 4 folds, care was taken such that in each fold, 6 of every set of 8 recordings became part of training set and remaining 2 became part of testing set. This was done to ensure that sufficient and similar variety in training and test datasets is always present.

In first set of three experiments, domain wise features were selected for dimensionality reduction i.e. 14 features from a single domain/transform were used at a time for fault recognition. Their results for Steady stage and Penetration stage recordings are shown in Fig 4(a) and 4(b). A third case was also taken where both Steady stage and Penetration stage data were mixed and used for training and testing. The purpose of this third case was to check if the two stages had any strong dissimilarities which if true could give bad performance with algorithms. The results of third case are shown in Fig. 4(c). In the plots, the five domains/transforms' features namely Time, FFT, DCT, MWT and WPT have been shown separately in X axis. A separate set of three experiments were performed while using PCA for dimensionality reduction.. Similar to Fig. 4, Fig 5(a) shows results for Steady stage recordings, Fig. 5(b) shows results for Penetration stage recordings and Fig. 5(c) shows the results when both Steady stage and Penetration stage recordings were mixed together. In Figure 5, X axes represent the number of features that PCA was made to

reduce. Multiple inferences can be made by analyzing the results. First important inference would be that in all cases, SVM performed better than the other two classifiers. Also among the five domains/transforms considered here, FFT features have significantly outperformed others with all three classifiers. This shows that spectral analysis is most powerful for drill bit monitoring with vibration data. A significant jump can be seen in Bayes classifier after applying PCA. This reconfirms the statement that Bayes classifier works well with linearly independent and uncorrelated features. Another important observation would be that when Steady stage and Penetration stage data were mixed for training and testing, the performance still seemed to be fine; in fact for some cases it was better than individual stages' recordings. This means there are no significant dissimilarities to disrupt data. The reason for improvement in results in some cases would be because when both stages' recordings are combined, training data would be larger and would have slightly more variety, which is good from generalization perspective. Also there is no clear inference as to which stage recordings perform better, as none of them seem to outperform other in most cases. Best performance amongst all cases was found to be 95.52% when FFT features were used along with SVM while both stages' recordings were mixed. From that perspective, domain wise dimensionality reduction fared better than PCA. Thus for final implementation in smartphone, the flow planned to be used was Pre-processing stage followed by Feature extraction stage for extracting FFT features and SVM classifier modeled with both stages' recordings. Summary of model development and operation phase is shown in Fig. 6(a).

V. Smartphone App Development

For allowing easy and regular check on drill bit's health, a smartphone application named "*Drill Bit Monitor*" has been developed on Android platform. By using either wired or wireless accelerometers, vibration data could be collected on a PC, which in turn would allow the data to be accessed on Cloud. For development of the Smartphone App, Eclipse IDE (Indigo version 3.7.2) with Android Development Tool (ADT) 22.0.3 plugin was used. This plugin provides support for development and debugging of App via simulation on Android Virtual Device Emulator. For implementation, external libraries namely *apache commons math* and *LIBSVM* libraries were also used. For deploying the application, user should install the installable named *DrillBitMonitor.apk* onto smartphone and store the *Resource Files* folder in SDcard. This folder contains database of all models/make of Drill Bits and the respective states which the App would support for fault recognition. The Resource File has information about drill model no., type and number of states for which it can be tested.

The App consists of five activity pages namely Drill Bit Input page , Confirmation page, Data Input page, Processing page and Results page. Snapshots of four activity pages running on a smartphone have been shown in Fig. 7. Flow chart presenting summary of the App's workflow is shown in Fig. 6(b). The App first takes input from user and collects information about drill bit and its model from *Resource Files* folder in SDcard. Second page shows stored information about the Drill Bit make and also asks user for confirmation that whether he/she wishes to go ahead for testing. After confirming the inputs, user proceeds to third

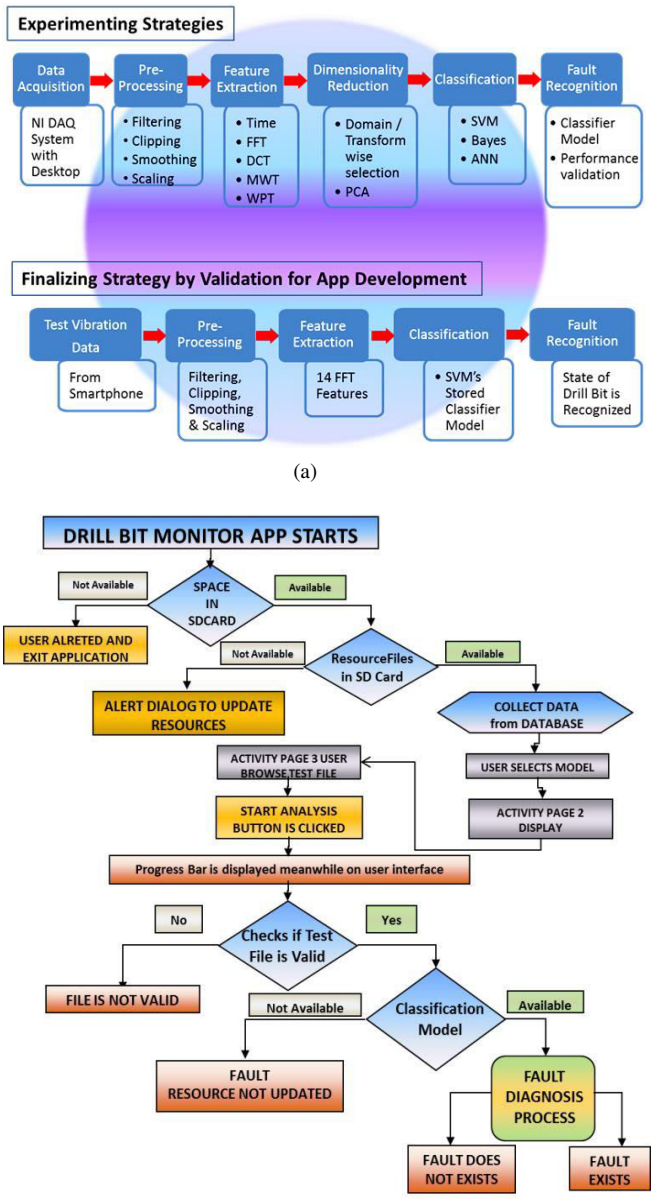
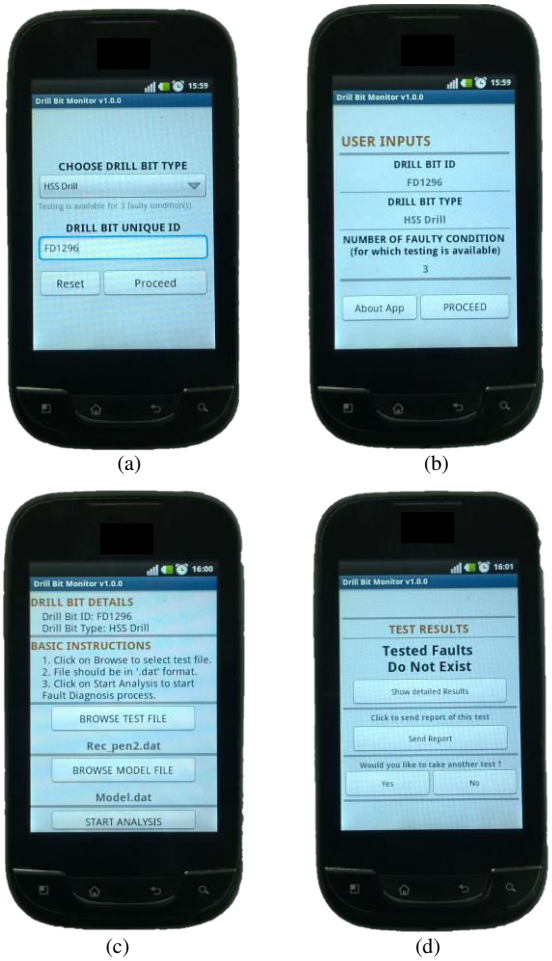


Fig. 6. (a) Summary Flow of Work, (b) Underlying Flow Chart of Smartphone App

page where he/she is asked to browse the input test file that he/she wishes to test. While browsing, the user can select files from either Cloud or internal phone memory or Phone's SDcard. After selecting the file, user selects the fault he/she wants to check for and then clicks the "*Start Analysis*" button. The user proceeds to a Processing page where a progress bar appears on the interface. While progress bar runs in front end, all processes needed during testing phase are run in back end i.e. pre-processing of recording, extraction of FFT features and classification using SVM with needed SVM model/s stored in *Resource Files* database folder. The entire processing in a smartphone with 1.5 Ghz dual core Krait processor, on average took little less than 10 seconds to compute. The user then proceeds to the Results page where based on classification result, determined drill bit state is presented.



(a) (b) (c) (d)

Fig. 7 Android App for Drill Bit Monitoring. (a) Drill Bit Input page, (b) Confirmation page, (c) Data Input page, (d) Results page

VI. Conclusions

An efficient strategy for drill bit monitoring has been presented in this paper. The strategy for test data had following steps: pre-processing with four steps namely filtering, clipping, smoothing and scaling, followed by feature extraction of FFT features and classification with SVM classifier using RBF kernel. This strategy was decided by experimenting with features from five domains/transforms, two dimensionality reduction methods and three classification methods. The final strategy was implemented as a Smartphone Application. The App allows users to recognize the state of drill bit recordings present on Cloud in less than ten seconds and with accuracy up to 95.52%. Future work would be to arrange setup that allows smartphones to directly collect vibration data from sensor/s present on machine/s.

Acknowledgment

Authors express sincere gratitude to The Boeing Company, USA, for sponsoring this research work at IIT Kanpur.

References

[1] Furness R.J., Tsao T.C, Rankin J. S., Muth M. J. and Manes K. W., “Torque Control for a Form Tool Drilling Operation”, *IEEE Trans. Control Syst. Tech.*, vol. 7, pp. 22–30, 1999.

[2] Brinksmeier E., "Prediction of Tool Fracture in Drilling", *CIRP Annals-Manufacturing Technology*, vol. 39, no. 1, pp. 97-100, 1990.

[3] Byrne G., Dornfeld D., Inasaki I., Ketteler G., König W., and Teti R., "Tool condition monitoring (TCM)—The Status of Research and Industrial Application", *CIRP Annals-Manufacturing Technology*, vol. 44, no. 2, pp. 541-567, 1995.

[4] Dimla S. and Dimla E., "Sensor Signals for Tool-Wear Monitoring in Metal Cutting Operations—a Review of Methods", *International Journal of Machine Tools and Manufacture*, vol. 40, no. 8. pp. 1073-1098, 2000.

[5] Liang, Steven Y., Hecker R. L. and Landers R. G., "Machining Process Monitoring and Control: The State–of–the–Art", International Mechanical Engineering Congress and Exposition, American Society of Mechanical Engineers, 2002.

[6] Jantunen E., "A Summary of Methods Applied to Tool Condition Monitoring in Drilling", *International Journal of Machine Tools and Manufacture*, vol. 42, no. 9, pp. 997-1010, 2002.

[7] Patra K., "Acoustic Emission based Tool Condition Monitoring System in Drilling", Proceedings of the World Congress on Engineering, 2011, vol. 3, pp. 6-8.

[8] Everson C. E., and Cheraghi S.H., "The Application of Acoustic Emission for Precision Drilling Process Monitoring", *International Journal of Machine Tools and Manufacture*, vol. 39, no. 3, pp. 371-387, 1999.

[9] Ertunc, Metin H. and Oysu C., "Drill Wear Monitoring using Cutting Force Signals", *Mechatronics*, vol. 14, no. 5, pp. 533-548, 2004.

[10] Cuppini D., D'errico G. and Rutelli G., “Tool Wear Monitoring based on Cutting Power Measurement Wear”, *International Journal on the Science and Technology of Friction, Lubrication and Wear*, vol. 139, pp. 303–311, 1990

[11] Patra K., Pal S.K. and Bhattacharyya K., "Artificial neural network based prediction of drill flank wear from motor current signals", *Applied Soft Computing*, vol. 7, no. 3, pp. 929-935, 2002.

[12] Wardany E., Gao T.L.D. and Elbestawi M.A. "Tool condition monitoring in drilling using vibration signature analysis", *International Journal of Machine Tools and Manufacture*, vol. 36, no. 6, pp. 687-711, 1996

[13] Abu-Mahfouz and Issam. "Drilling Wear Detection and Classification using Vibration Signals and Artificial Neural Network", *International Journal of Machine Tools and Manufacture*, vol. 43, no. 7, pp. 707-720, 2003.

[14] Kanai M. and Kanda Y., "Statistical characteristics of drill wear and drill life for the standardized performance tests" *Annals of CIRP*, vol. 27, no. 1, pp. 61-66, 1978.

[15] Verma N.K., Jagnaatham K., Bhairat A, Shukla T. and Subramanian T., “Statistical Approach for finding Sensitive Positions for Condition Based Monitoring of Reciprocating Air Compressors”, IEEE Control & System Graduate Research Colloquium, 2011, pp.10-14.

[16] Sheskin D. J., "Handbook of Parametric and Nonparametric Statistical Procedures", CRC Press, 2003.
[17] Jones M.C. and Sibson R., "What is Projection Pursuit ?", *Journal of the Royal Statistical Society, Ser. A*, vol. 150, no. 1, pp. 1-37, 1987.
[18] Vidaurre C., Krämer N., Blankertz, B. and Schlögl A., "Time Domain Parameters as a Feature for EEG-based Brain Computer Interfaces", *Neural Networks*, vol. 22, no. 9, pp. 1313-1319, 2009.
[19] S. L. Marple, Digital Spectral Analysis with Applications, 1987 : Prentice-Hall
[20] Qingbo H., Kong F. and Yan R., "Subspace-based Gearbox Condition Monitoring by Kernel Principal Component Analysis", *Mechanical Systems and Signal Processing* , vol. 21, no. 4, pp. 1755-1772, 2007.
[21] Xiaoli Li., "On-line Detection of the Breakage of Small Diameter Drills using Current Signature Wavelet Transform", *International Journal of Machine Tools and Manufacture*, vol. 39, no. 1, pp. 157-164, 1999.
[22] Jolliffe I.T., Principal Component Analysis. New York: Springer-Verlag, 1986.
[23] Vapnik V. and Naumovich V., "Statistical learning theory" New York: Wiley, vol. 2 ,1998.
[24] Meyer E. and Tuthill T., "Bayesian Classification of Ultrasound Signals using Wavelet Coefficients", IEEE National Aerospace and Electronics Conference, 1995, pp. 240–243.
[25] Masri, S. F., Smyth A. W., Chassiakos A. G., Caughey T. K. and Hunter N. F., "Application of Neural Networks for Detection of Changes in Nonlinear Systems", *Journal of Engineering Mechanics*, vol. 126, no. 7, pp. 666-676, 2000.
[26] Goumas S.K., Zervakis M.E. and Stavrakakis G.S., "Classification of Washing Machines Vibration Signals using Discrete Wavelet Analysis for Feature Extraction", *IEEE Transactions on Instrumentation and Measurement,* vol. 51, no. 3, pp. 497-508, 2002.
[27] Google Android Developers, Android Developer Guide, http://developer.android.com/guide/topics/fundamental.html
[28] Commons Math Developers, Apache Commons Math, Release 2.2 Available: http://commons.apache.org/math
[29] Chang C.C. and Lin C.J., "LIBSVM : a library for support vector machines", *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 27-27, May 2011.
[30] Kumar A., Ramkumar J., Verma N. K. and Dixit S., "Detection and Classification for Faults in Drilling Process using Vibration Analysis", IEEE Prognostics and Health Management, June 2014, pp. 1-6.

AUTHORS BIOGRAPHY

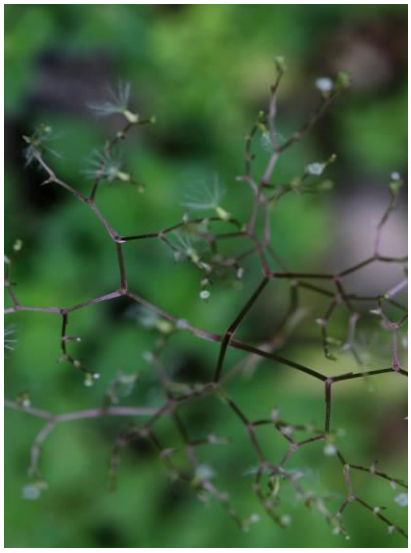
Nishchal K Verma (M'05 & SM'13) received his Ph.D. from IIT Delhi, India, in 2007 in Electrical Engineering. He is currently an Associate Professor with the Department of Electrical Engineering, IIT Kanpur, India. He is Editor of IETE Technical Review, an Associate Editor of Transactions of the Institute of Measurement and Control, UK and Editorial board member for several reputed journals and conferences. His research interests include Big Data, Internet of Things, Intelligent Data Mining Algorithms, Diagnosis and Prognosis of Rotating Machines, Computer Vision and Computational Intelligence.

Rahul Kumar Sevakula (M'12) was born on October 31, 1987. He received B.Tech. degree from National Institute of Technology Warangal in 2009. He is currently a doctoral student at Electrical Engineering Dept. of IIT Kanpur and also the Student Head of IEEE U.P. Section Chapter, Computational Intelligence Society. His research interests include Machine Learning Algorithms, Fuzzy Systems, Intelligent Informatics and Health monitoring.

Sonal Dixit (M'13) was born in India on July 31, 1989. She received her B.E. in Computer Science from RGPV University of in 2009 and Masters in Computer Science from Banasthali University in 2011. She is currently a doctoral student at IIT Kanpur and Secretary of IEEE U.P. Section Chapter, Computational Intelligence Society. Her research interests fall mainly in the field of Machine Learning, Intelligent Informatics, Natural Language Processing and Condition based Monitoring of Rotary Machines.

Al Salour is a Technical Fellow at The Boeing Company's research & technology organization. Al is leading the manufacturing technologies for equipment and process health monitoring as well as automated identification. In these areas he has developed methods and procedures to collect performance data from machine tools and robotic systems and report them in real time to the operation users. He has installed multiple systems to initiate automated manufacturing transactions and alert notifications for quick actions and has focused on eliminating non-value added tasks. Dr. Salour is the Chairman of the Intelligent Maintenance Systems (IMS) industrial advisory board supported by four universities and 60 member companies.

Fault Diagnostic Opportunities for Solenoid Operated Valves using Physics-of-Failure Analysis



N. Jordan Jameson* (jjameson@calce.umd.edu)
Michael H. Azarian (mazarian@calce.umd.edu)
Michael Pecht (pecht@calce.umd.edu)

Center for Advanced Life Cycle Engineering
University of Maryland
College Park, MD

Abstract—**Solenoid operated valves are vital components in many process control systems. They are components that are often critical to safety. Solenoid valve degradation is difficult to detect in situ, leading to failures, which are often sudden and unexpected. This paper reviews some of the common causes of solenoid valve degradation, presents strategies that leverage these mechanisms to detect and diagnose faults before they lead to failure, and discusses research opportunities aimed at improving solenoid valve diagnostics and prognostics.**

Keywords-solenoid valve; diagnostics; electrical coil; fault detection

I. INTRODUCTION

The Transocean Deepwater Horizon disaster in 2010 was a major incident resulting in 11 lives lost and an estimated 4.9 million barrels of oil discharged into the Gulf of Mexico. Tests performed by Transocean Ltd. and Cameron International after the incident revealed that the coil of a solenoid valve failed to energize, suggesting an electrical coil fault. The investigation team found no evidence to suggest that this fault was a result of the incident. Rather, they concluded that the electrical fault(s) likely existed prior to the accident [1]. Had the solenoid valve been working properly, it could have yielded at least a partial closure of the blind shear rams, resulting in a far less serious incident.

Solenoid operated valves (SOVs) are utilized to shut off, discharge, dose, allocate, or combine fluids. This action is accomplished by passing an electric current through a coiled wire, thereby producing a magnetic field, which magnetizes the plunger resulting in a position change. The position of the plunger controls the flow of the process fluid(s).

SOVs are integral components of many systems. Their popularity is primarily due to their simple and rugged construction, and their inexpensive cost. Within the automotive industry, solenoid valves are used to achieve intelligent control in electro-pneumatic braking systems of motor vehicles [2], control in diesel fuel injection systems [3], [4], and the control of automobile transmissions [5]. In the process and nuclear industry, solenoid valves are used for process fulid control and in critical safety instrumented functions (SIF). Approximately 2–4% of all solenoid valves in a typical chemical plant are part of a SIF [6]. Moreover, safety valves are generally the most important components in the safety loop [7]. Thus, their reliability and availability are critical.

Due to the pervasive use of SOVs in a variety of industries, interest continues to grow in estimating their health and remaining useful life (RUL). Understanding the underlying physics of their failure mechanisms can yield insight into the measurement techniques that may produce useful results for health estimation. This is referred to as the physics of failure (PoF) approach to diagnostics and prognostics. The first step in this process is to identify and analyze the hardware of the system. This yields an understanding of how the components connect and their functional relationships, which can be used in identifying loading conditions applied to system components. The loading conditions are a direct result of the life-cycle demands of the system. However, in a system where there is heavy interaction between the components, as seen in the SOV, life-cycle loads and demands can produce stresses that interact among the components. These stresses may be classified as mechanical, electrical, chemical, thermal, or environmental radiation. The presence of any particular load in the life-cycle depends on the specific application of the SOV. The next step is to perform a failure modes, mechanisms, and effects analysis (FMMEA) on the system. This, combined with a criticality analysis, is useful for identifying and prioritizing the failure mechanisms of the system. With an understanding of the failure mechanisms, a system can be designed to monitor key parameters in order to perform system diagnostics and prognostics.

The purpose of this paper is to identify the critical components and failure mechanisms of the solenoid valve system, and then explore existing and potential methods of performing health diagnostics and prognostics.

II. HARDWARE ANALYSIS

Solenoid operated valves are used in many different operating environments and thus can have a variety of designs. The fundamental differences can usually be understood using the following terms: normally open versus

normally closed; direct-acting versus pilot-controlled; and two-way versus three-way versus four-way. Normally open (closed) refers to a valve where the inlet port is open (closed) when the valve is de-energized. A direct-acting valve is one where all flow passes through an orifice that is opened directly by an electromagnet and plunger. Pilot-controlled refers to a solenoid valve that operates by means of a minimum and maximum pressure differential and uses an electromagnet and plunger to open or close a small orifice thus controlling the pressure differential across a piston or diaphragm. A two-way valve, as shown in Figure 2.1, is one where there are two ports and a single orifice that can be opened or closed. Two-way valves are used to control a single working fluid. In a three-way valve, there are three ports and two orifices, similar to what is shown in Figure 2.2. Three-way valves can have several possible functions. They are commonly used to alternately apply pressure to and exhaust pressure from the diaphragm operator of a control valve, single-acting cylinder, or rotary actuator. It can operate with an inlet port, an outlet port, and an exhaust port for operating a single acting cylinder; one inlet port and two outlet ports for selecting or diverting flow; or two inlet ports and one outlet port for mixing fluids. Four-way valves are generally used to operate double-acting cylinders or actuators. They have four or five pipe connections: one pressure, two cylinder, and one or two exhausts. In this paper, a two-way, direct-acting solenoid valve will be analyzed, not because it is necessarily the most common, but because it provides an opportunity to analyze components and loading conditions that are common to the valves previously mentioned.

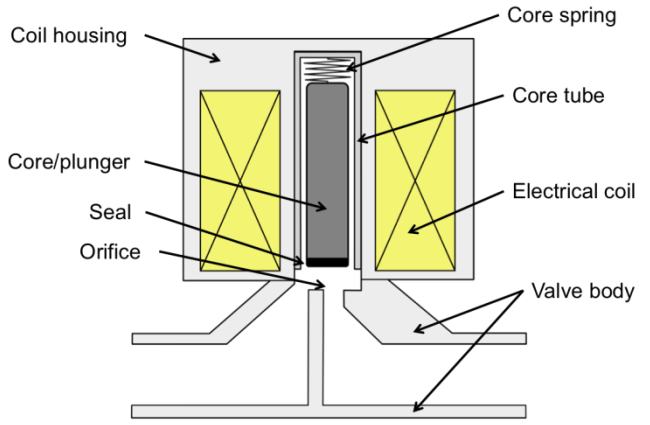


Figure 2.1. Two-way, direct-acting normally open solenoid valve

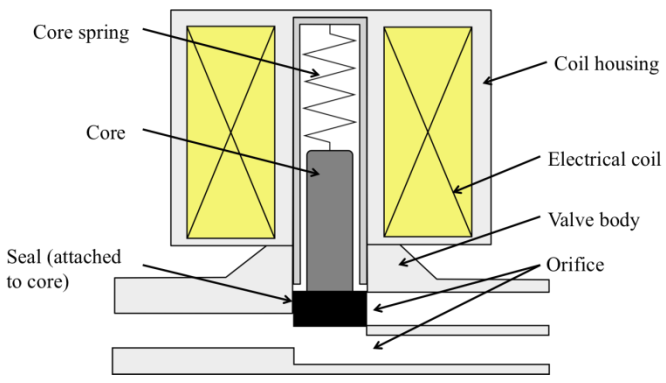


Figure 2.2. Three-way, direct-acting solenoid valve

In a two-way valve design, the spring, core, and core tube are exposed to the process fluid. This type of design is common, though there are available designs where these components are separated from the process fluid by a membrane, used with ultra-pure or extremely aggressive process fluids.

III. Loading Conditions

In order to assess the reliability of a SOV, environmental and operational loads must be understood. For most applications, there are chemical and contamination loads arising from the process fluid and airborne environment, and thermal loads arising from the process fluid and the electrical coil. Further, due to the interaction of the components, friction and impacts will be present during the lifetime of the SOV. Each component in the valve will be subjected to a different combination of loads. A less general load is radiation loading, in the case where a SOV is used in a nuclear facility. In providing an FMMEA for SOVs, motivation is given for further research into diagnostic and prognostic methods for SOVs.

IV. Failure Modes, Mechanisms, and Effects Analysis

Failure modes, mechanisms, and effects analysis (FMMEA) is a systematic methodology of finding root cause failure mechanisms of a given product [8]–[10]. An effect is the observable result of a failure on the product. Failure mode is a way in which a component, system, or subsystem may fail to meet its intended function. Failure mechanism is the mechanical, chemical, thermodynamic or other physical process or combination of processes that result in a failure. FMMEA helps to identify potential failure mechanisms and models for expected failure modes and prioritize them. An important result from FMMEA is an understanding of possible parameters to be monitored for diagnostics and prognostic purposes.

A. Potential Failure Modes

The major components of the solenoid valve are selected from the analysis of hardware given in Section 2 for a two-way, direct-acting solenoid valve. An overview of the potential failure modes, mechanisms, and effects is given in Table 4.1.

1) Valve Body

The valve body is exposed to the process fluid and must therefore be resistant to corrosion and contamination. The materials used to construct the valve body are most commonly brass, bronze, cast iron, or stainless steel. Some alternative materials are used in specialized applications. Polyvinylidene fluoride (PVDF) is suitable for valves in acidic and solvent applications. Polyether ether ketone (PEEK) has desirable mechanical properties but is susceptible to attack by nitric and sulphuric acid.

The failure of the valve body will be evidenced by leakage of the process fluid. This could be caused by loosening of the outside connections or, in extreme cases, plastic deformation of the valve body. In cases where there is a mechanical loading, such as vibration or impacts, applied to the valve body, the failure mechanisms are fatigue

leading to fracture or overstress fracture. However, since contaminants and corrosive media are present in most solenoid valve applications, the expected failure mechanism for the valve body is corrosion fatigue or corrosion fracture, depending upon the ambient environment.

2) Seal

The seal is used to control the flow of the process fluid through the valve. Implied by this function is the requirement to prevent leakage from the input to the output inside the valve, referred to as seal leakage. Since the seal is exposed to the process fluid, there are several different materials used for seal construction. Some examples are: NBR (nitrile butadiene rubber), EPDM (ethylene propylene diene monomer rubber), FPM (fluorocarbon rubber), and FFPM (perfluorinated elastomer).

The seal will experience impact loading from the core and will also experience chemical loading from the process fluid. Further, the process fluid or the electrical coil could cause the temperature of the seal to increase. Some valve designs locate the seal on the tip of the core, exposing the seal to friction. A failure of the seal would result in seal leakage. This could be caused by a combination of mechanisms: corrosion, embrittlement, erosion of the seal material caused by the process fluid, impacts from the core and friction, and fatigue caused by impacts from the core.

3) Core Spring

The function of the core spring is to return the core to its default position when de-energized. In many valve designs, the core spring is exposed to the process fluid and must be resistant to corrosion from the process fluid. Thus, it is generally constructed from paramagnetic stainless steel.

As the spring is subjected to cyclic motions, the stiffness will decrease over time. Further, as the spring is commonly exposed to the process fluid, it could corrode and further fatigue. This loss of stiffness will cause the valve to improperly meter the process fluid, as the orifice will not be properly plugged or fully opened. If allowed to continue in operation, the spring could eventually fracture, resulting in a total loss of function.

4) Core/plunger

The core/plunger is responsible for allowing or preventing the flow of process fluid through the solenoid valve. In common designs, the core is exposed to the process fluid. The core must be a soft ferromagnetic material in order to perform the functions necessary for the valve. The most common material used for this purpose is stainless steel 430F, a low carbon, high chromium stainless steel, which was developed specifically for solenoid plunger applications in corrosive environments.

As the core is often exposed to the process fluid, corrosion frequently acts on the core material. Additionally, the core is in contact with the core tube, which introduces friction, wear, and material loss. This will be evidenced by stick slip behavior or a failure to fully seal the valve when closed. The core is also exposed to the magnetic field created by the electrical coil. Prolonged exposure to this field can result in permanent magnetization of the plunger, resulting in improper behavior of the core, and improper metering of the process fluid.

5) Coil Housing

The coil housing performs three functions for the SOV: it completes the electromagnetic flux path of the solenoid, provides protection from contact with the coil, and protects the coil against environmental conditions. For this reason, it is generally constructed using a soft ferromagnetic stainless steel.

The housing will be directly exposed to the environmental conditions. If the SOV were used in extreme environments, the combination of corrosion and temperature from the process fluid and electrical coil could produce a loss of material resulting in the decrease of magnetic flux. In environments with high hydrogen concentration, hydrogen embrittlement could potentially be a failure mechanism.

6) Core Tube

The core tube functions as a barrier between the core and the electrical coil. It helps to protect the coil from the process fluid and direct the magnetic flux into the core instead of around the core. Most designs call for the core tube to be constructed of aluminum or paramagnetic stainless steel. (A ferromagnetic core tube would provide a shunt path for the magnetic field lines, which would reduce the efficiency of the SOV.)

Aggressive process fluids and friction produced by interaction with the moving core result in wear of the core tube. This produces wear particles that can inhibit the movement of the core.

7) Electrical Coil

The electrical coil is responsible for producing the magnetic field that magnetizes the core and produces the necessary motion of the valve. The wire used is generally referred to as magnet wire and is usually constructed of copper. Within the solenoid valve field, there are three main types of insulation used to coat the wire. Class E insulation is rated for temperatures up to 120°C; class F is rated for temperatures up to 155°C; and class H is rated for temperatures up to 180°C. Electrical coil construction is generally divided into two methods: tape wrapped coils and encapsulated coils. Tape wrapped coils are manufactured by winding wire around a spool or bobbin, and then protecting the winding with insulation tape. Encapsulated coils also have a wire wound around a spool or bobbin, but the wire is then encapsulated or molded over with a suitable resin.

As an electric current is passed through the wire, Joule heating causes an increase in the wire temperature. If the temperature is too great, the dielectric material between the wires could degrade, fail, and two neighboring wires would form an electrical connection, producing a turn-to-turn or layer-to-layer short. These shorts cause the coil resistance to decrease, thus pulling a greater current into the valve. At the location of the short, a hot spot can form, where the local temperature is great enough to cause the wire to burn out, resulting in an open circuit. Corrosion can also play a role in the failure of the electrical coil by causing necking and loss of material in the wire.

B. Prioritization of Potential Failure Mechanisms

In order to prioritize the potential failure mechanisms of the SOV, one must utilize past experience, stress analysis, accelerated tests, and engineering judgment. In 1987, Oak Ridge National Laboratory (ORNL) gathered and analyzed data taken from the Nuclear Plant Reliability Data System

(NPRDS) records of the Institute of Nuclear Power Operations (INPO) for SOVs, covering September 5, 1978-July 11, 1984, and the NRC Licensee Event Reporting (LER) system records for January 26, 1981-July 11, 1984 [11]. The data showed that over 50% of SOV failures resulted from 4 sources: worn or degraded parts, contamination by foreign materials, short circuit in the SOV coil, and open circuits in the SOV coil. The remaining failures were attributed to manufacturing defects, improper installation, incorrect assembly, corroded parts, loose or misaligned parts, or their failure source was unspecified. Overall, the dominant failure source was shorts in the electrical coil, followed by foreign material contamination, and then electrical coil open. Importantly, there was no further breakdown into specified failure sites for the cases of worn, degraded, or broken parts and foreign materials.

Table 4.1. Potential Failure Effects, Modes, and Mechanisms of Solenoid Operated Valves

Failure site	Potential failure effect	Failure mode	Failure mechanism
Valve body	Body leakage	Loosening of connection seals, opening in material	Corrosion, fatigue fracture, overstress fracture
Seal	Improper media flow (e.g. seal leakage), noise	Loosening or deterioration of seal, impacts with core, friction	Polymer embrittlement , erosion, overstress, fatigue
Core spring	Improper media flow	Weakening of spring strength, spring breakage, material defects	Corrosion, fatigue, hydrogen embrittlement
Core/ plunger	Irregular movement, seal leakage	Loss of material, stick slip	Wear, residual magnetism, debris build-up
Coil housing	Disruption of magnetic flux path (reduced magnetic efficiency)	Loss or discontinuity of material in housing from corrosion or overstress	Corrosion, overstress
Core tube	Irregular core movement resulting in seal leakage	Debris build-up, excessive friction	Corrosion, wear
Electrical coil	SOV unable to operate (coil open), leakage resulting from reduced magnetic field strength (coil short)	Fracturing or necking of wire; degraded insulation from temperature, conductor thermal expansion, or electrical transients; material defects	Dielectric breakdown, corrosion, thermal overstress of conductor, fatigue fracture

A study in 2009 by Angadi *et al*. [5], [12] revealed that solenoid valves are susceptible to coupled electrical-thermo-mechanical failure mechanisms. In particular, they emphasized the role of Joule heating in the thermal expansion of the magnet wire, causing the degradation and failure of the insulation between the wires. This mechanism results in a turn-to-turn short, and ultimately in a coil burnout.

With these cases, and the variety of failure mechanisms acting on each component, in mind, it is difficult to draw conclusions on the prioritization of failure mechanisms with respect to failure sites. However, it is clear that electrical faults and valve contamination are prevalent and critical to the health of the SOV.

V. Available Methods of Solenoid Valve Diagnostics

Oak Ridge National Laboratory evaluated several methods of health monitoring for solenoid valves in nuclear facilities [13], [14]. The evaluation methods were chosen based upon what the researchers considered to be the most prominent sources of failure, namely, open-circuited coils, short-circuited coils, worn or degraded mechanical parts, and contamination by foreign materials. They suggested measuring coil temperature via coil resistance or impedance. This is useful for instances where the temperature of the coil is high enough to degrade the insulation. By measuring the resistance, an unsafe operating condition can be detected. This method was deemed to be ready for immediate use. However, this approach only measures the mean coil temperature, which could be attributed to the operating environment or the process fluid. It does not sense hot spots in the coil or the valve body temperature, thus failing to isolate turn-to-turn or layer-to-layer shorts within the electrical coil.

In order to diagnose mechanical binding, failure to shift, or sluggishness as a result of worn or improper parts or the presence of foreign material in the valve, a method was proposed which used coil impedance measurements to indicate the position of the core. As the core becomes magnetized and moves through the magnetic field, its altered position is reflected in increased impedance of the coil. It was shown that the impedance change can be used to determine the position of the core to within a few thousandths of an inch and thus, detect any anomalous movement of the core due to contaminants, deformation, debris build-up, or residual magnetism. This method was deemed as having high promise for in-plant use, though the method was only useful for AC SOVs and required the introduction of a special ramp-voltage power supply. In a subsequent study [15], a methodology was developed for in situ diagnostic testing of DC SOVs by analysis of the characteristics of the transient current waveform accompanying valve actuation. This method was most sensitive to mechanical valve faults such as impeded or incomplete plunger motion, and reduced plunger spring force.

In ORNL's 1990 study [14], to address the electrical failure of a solenoid coil caused by high-voltage turn-off transients in combination with insulation weakened by prolonged operation at high temperatures, it was suggested to measure the characteristics of the electrical transient response generated upon de-energizing a DC SOV. To model the electrical characteristics of the valve, an equivalent circuit was employed. This equivalent circuit is shown in Figure 5.1.

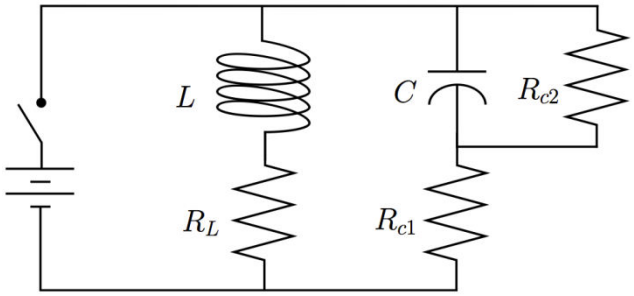


Figure. 5.1. Equivalent circuit model for solenoid operated valve

In this circuit, R_L is the series resistance of the coil, L is the coil inductance, C is the distributed capacitance of the coil, R_{c1} is the series resistance due to the distributed capacitance, and R_{c2} is the parallel resistance due to the distributed capacitance. For their simulations, the resistances R_{c1} and R_{c2} were neglected. They tested a healthy solenoid valve and a solenoid valve with approximately 6 percent of its turns shorted. The test consisted of measuring the voltage response of the valves after stepping down to 0V DC from 30V DC. The equivalent circuit model was able to reproduce the response of the healthy valve, as it behaved like a damped oscillator system. However, the response of the faulty valve was not reproduced using the equivalent circuit model. An expected advantage of this approach was the ability to detect faults in the valve that were undetectable using other methods, since the ECM could take advantage of direct electrical measurements and observe any faults existing only in the coil. As the equivalent circuit could not accurately reproduce the response of the faulty valve, this method was deemed as having low promise in the field.

The state-of-the-art in solenoid valve fault detection is partial stroke testing (PST) [7]. In PST, the plunger is moved a small amount of its total stroke length in order to measure the dynamic response. With the help of a position sensor, a PST can detect sluggishness, leakage, wear of valve seals, stick slip friction, or deposit build-up by comparing the test responses to a reference response [16], [17]. This type of testing can be carried out if three conditions hold: the valve must have a position sensor; the solenoid valve must have a sufficiently long stroke length; and the movement of the valve should not produce a significant disturbance to the process or the safety function of the valve. Otherwise, partial stroke testing can yield little information into the health state of the valve or cannot be utilized.

Visual inspection is occasionally used in solenoid valve diagnostics. This entails inspecting the solenoid valve for damage easily observed in a periodic inspection of the system as a whole, or in some cases, removing the valve for inspection in a workshop. This can be advantageous for some failure modes such as leakage or, using the workshop, contamination of the plunger pathway. However, faults within the electrical coil can be difficult to diagnose without destruction of the valve.

VI. Conclusions

As shown in the Section 5, there are several techniques available for diagnosing SOV faults such as binding, sluggishness, leakage, and wear of seals. ORNL developed a method of measuring the position of the core using coil impedance. They also proposed an in situ method of detecting faults in the core motion by using the current waveform. Partial stroke testing is capable of performing fault detection of the valve. Interestingly, based upon the data from ORNL's 1987 study, two of the top four contributors to SOV failure have been addressed, namely: worn or degraded parts and contamination by foreign materials. Unfortunately, the remaining two sources of SOV failure (coil short and coil open) have not yet been successfully addressed, despite the efforts of ORNL in 1990. Since then there has been a relatively small amount literature studying SOV failures, especially those failures originating from coil faults. There is a significant amount of industry-focused literature addressing the use of PST in increasing safety instrumented function reliability (see e.g. [16]–[23]). Yet, it is surprising that such a significant

problem, as coil faults in SOVs, could remain unaddressed and having no available solutions.

It remains necessary to develop a technique to assess the health of the SOV electrical coil without disrupting the process. The dielectric insulation around the wires can be degraded given the temperature of the wire, the thermal expansion of the conductor, and the presence of environmental contaminants and/or humidity. Moreover, the dielectric is likely subjected to temperature cycles, as the SOV is not generally in consistent use. Thus, a method to measure the health state of the dielectric insulation material could be valuable.

Further work is needed in developing a fundamental physics-based model of the electrical coil. A problem with the equivalent circuit model is that the model parameters of the coil experience minute changes due to local faults in the wire. In the ORNL research, the measured capacitance of the healthy coil was 12.6nF, whereas for the faulty coil it was 2.89nF; the resistance changed from 100.4Ω to 97Ω; and inductance changed from 113.9mH to 106.2mH. Thus, the introduction of these altered parameters does very little to simulate the true response of the SOV to shorts in the coil. Some work has been performed to develop physics models for multiple layer inductors [24]–[26]. There has also been work to develop new ECMs for coils involving fractional derivatives [27], [28]. These models could be adapted for modeling the coils in solenoid valves. Simulations should be performed using these or similar models, in order to determine the best parameters to be measured in order to diagnose localized coil faults.

Research efforts can also be directed to developing sensor techniques for probing the health of the insulation material. Werynski *et al.* used a magnetic field sensor to detect the migration of the bulk capacitance of the windings in an AC machine [29]. Adapting this work for detecting insulation deterioration in solenoid valves would be a valuable asset in the pursuit of improved diagnostic and prognostic methods for SOVs.

Acknowledgements

The authors would like to thank the more than 100 companies and organizations that support research activities at the Center for Advanced Life Cycle Engineering (CALCE) at the University of Maryland annually. Also special thanks go to the members of the Prognostics and Health Management Consortium at CALCE for their support of this work.

References

[1] M. Bly, *Deepwater Horizon Accident Investigation Report*. DIANE Publishing, 2011.

[2] V. Szente and J. Vad, "Computational and Experimental Investigation on Solenoid Valve Dynamics," in *2001 IEEE/ASME International Conference on Advanced Intelligent Mechatronics, 2001. Proceedings*, 2001, vol. 1, pp. 618–623 vol.1.

[3] D. Howe, "Magnetic Actuators," *Sensors and Actuators A: Physical*, vol. 81, no. 1–3, pp. 268–274, Apr. 2000.

[4] H.-H. Tsai and C.-Y. Tseng, "Detecting Solenoid Valve Deterioration in In-Use Electronic Diesel Fuel Injection Control Systems," *Sensors (Basel)*, vol. 10, no. 8, pp. 7157–7169, Jul. 2010.

[5] S. V. Angadi, R. L. Jackson, S.-Y. Choe, G. T. Flowers, J. C. Suhling, Y.-K. Chang, and J.-K. Ham, "Reliability and Life Study of Hydraulic Solenoid Valve. Part 1: A Multi-Physics Finite Element Model," *Engineering Failure Analysis*, vol. 16, no. 3, pp. 874–887, Apr. 2009.

[6] T. Karte and E. Nebel, "Reliability Data and the Use of Control Valves in the Process Industry in Accordance with IEC 61508/61511," *Automatisierungstechnische Praxis*, vol. 47, no. 2, 2005.

[7] J. Yli-Petäys, "The Value of Safety Valves," *Automation*, vol. 3, 2008.

[8] S. Mathew, D. Das, R. Rossenberger, and M. Pecht, "Failure Mechanisms Based Prognostics," in *International Conference on Prognostics and Health Management, 2008. PHM 2008*, 2008, pp. 1–6.

[9] S. Ganesan, V. Eveloy, D. Das, and M. Pecht, "Identification and Utilization of Failure Mechanisms to Enhance FMEA and FMECA," in *Proceedings of the IEEE Workshop on Accelerated Stress Testing & Reliability (ASTR)*, Austin, TX, 2005.

[10] M. G. Pecht, *Prognostics and Health Management of Electronics*, 1 edition. Hoboken, N.J: Wiley-Interscience, 2008.

[11] V. P. Bacanskas, G. C. Roberts, and G. J. Toman, *Aging and Service Wear of Solenoid-operated Valves Used in Safety Systems of Nuclear Power Plants. Volume 1: Operating Experience and Failure Identification*. Oak Ridge National Laboratory, 1987.

[12] S. V. Angadi, R. L. Jackson, S. Choe, G. T. Flowers, J. C. Suhling, Y.-K. Chang, J.-K. Ham, and J. Bae, "Reliability and Life Study of Hydraulic Solenoid Valve. Part 2: Experimental study," *Engineering Failure Analysis*, vol. 16, no. 3, pp. 944–963, Apr. 2009.

[13] R. C. Kryter, *Aging and Service Wear of Solenoid-operated Valves Used in Safety Systems of Nuclear Power Plants. Volume 2: Evaluation of Monitoring Methods*. Oak Ridge National Laboratory, 1992.

[14] R. C. Kryter, "Nonintrusive Methods for Monitoring the Operational Readiness of Solenoid-Operated Valves," *Nuclear Engineering and Design*, vol. 118, no. 3, pp. 409–417, Apr. 1990.

[15] E. D. Blakeman and R. C. Kryter, "Noninvasive Testing of Solenoid-Operated Valves Using Transient Current Signature Analysis," Oak Ridge National Lab., TN (United States), CONF-970591--3, Mar. 1997.

[16] B. Mostia, "Partial Stroke Testing: Simple or Not," *Control*, Nov-2003.

[17] K.-P. Heer, "Increasing Safety: Combining Partial Stroke Testing and Solenoid Valve Testing," *Valve World Magazine*, vol. 69, pp. 69–73, Jun-2009.

[18] L. Stewart, J. Bukowski, and W. Goble, "Improving Reliability & Safety Performance of Solenoid Valves by Stroke Testing," in *9th Global Congress on Process Safety*, San Antonio, TX, 2013.

[19] B. Mostia, "Ins and Outs of Partial Stroke Testing," *Control*, Nov-2001.

[20] T. Karte, "Partial Stroke Testing for Final Elements," presented at the Petroleum and Chemical Industry Conference (PCIC) Europe 2005, Basle, Switzerland, 2005.

[21] T. Karte and K.-B. Schartner, "Partial Stroke Testing of Final Elements to Extend Maintenance Cycles,"

Automatisierungstechnische Praxis, vol. 47, no. 4, 2005.
[22] P. Gruhn, J. Pittman, S. Wiley, and T. LeBlanc, "Quantifying the impact of partial stroke valve testing of safety instrumented systems," *ISA Transactions*, vol. 37, no. 2, pp. 87–94, Apr. 1998.
[23] M. A. Lundteigen and M. Rausand, "The effect of partial stroke testing on the reliability of safety valves," presented at the Risk, Reliabilityand Societal Risk, Volume 3, 2007, pp. 2479–2486.
[24] A. Massarini, M. K. Kazimierczuk, and G. Grandi, "Lumped Parameter Models for Single- and Multiple-Layer Inductors," in *27th Annual IEEE Power Electronics Specialists Conference, 1996. PESC '96 Record*, 1996, vol. 1, pp. 295–301 vol.1.
[25] A. Massarini and M. K. Kazimierczuk, "Self-Capacitance of Inductors," *IEEE Transactions on Power Electronics*, vol. 12, no. 4, pp. 671–676, Jul. 1997.
[26] Q. Yu and T. W. Holmes, "A study on stray capacitance modeling of inductors by using the finite element method," *IEEE Transactions on Electromagnetic Compatibility*, vol. 43, no. 1, pp. 88–93, Feb. 2001.
[27] I. Schäfer and K. Krüger, "Modelling of Coils Using Fractional Derivatives," *Journal of Magnetism and Magnetic Materials*, vol. 307, no. 1, pp. 91–98, Dec. 2006.
[28] I. Schäfer and K. Krüger, "Modelling of Lossy Coils Using Fractional Derivatives," *J. Phys. D: Appl. Phys.*, vol. 41, no. 4, p. 045001, Feb. 2008.
[29] P. Werynski, D. Roger, R. Corton, and J.-F. Brudny, "Proposition of a new method for in-service monitoring of the aging of stator winding insulation in AC motors," *IEEE Transactions on Energy Conversion*, vol. 21, no. 3, pp. 673–681, Sep. 2006.

AUTHOR BIOGRAPHY

N. Jordan Jameson received the B.S. degree in mechanical engineering from Tennessee State University, Nashville, TN. He is currently pursuing a Ph.D. degree in mechanical engineering at the University of Maryland, College Park. His research interests include diagnostic and prognostic methods for health management of mechanical and electrical systems.

Michael H. Azarian received the B.S.E. degree in chemical engineering from Princeton University and the M.E. and Ph.D. degrees in materials science and engineering from Carnegie Mellon University.

He is a Research Scientist with the Center for Advanced Life Cycle Engineering (CALCE), University of Maryland, College Park. Prior to joining CALCE he spent over 13 years in industry. His research focuses on the analysis, detection, prediction, and prevention of failures in electronic and electromechanical products. He is the holder of five U.S. patents.

Dr. Azarian is co-chair of the Miscellaneous Techniques subcommittee of the SAE G-19A standards committee on detection of counterfeit parts. He has previously held leadership roles in various IEEE reliability standards committees and co-chaired iNEMI's Technology Working Group on Sensor Technology Roadmapping. He is on the Editorial Advisory Board of Soldering & Surface Mount Technology.

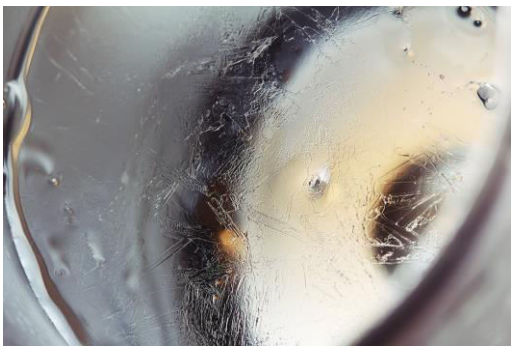
Michael Pecht received the M.S. degree in electrical engineering and the M.S. and Ph.D. degrees in engineering mechanics from the University of Wisconsin, Madison.

He is the Founder of the Center for Advanced Life Cycle Engineering, University of Maryland, College Park, where he is also a George Dieter Chair Professor in mechanical engineering and a Professor in applied mathematics. He has consulted for over 100 major international electronics companies. He has written more than 20 books on electronic-product development, use, and supply chain management and over 400 technical articles.

Dr. Pecht is a Professional Engineer and a fellow of ASME and IMAPS. He is the editor-in-chief of IEEE Access. He was the recipient of the IEEE Reliability Society's Lifetime Achievement Award, the European Micro and Nano-Reliability Award, the 3M Research Award for

electronics packaging, and the IMAPS William D. Ashman Memorial Achievement Award for his contributions in electronics reliability analysis.

Fuel Cells Impedance Estimation Using Regression Analysis



Wlamir Olivares Loesch Vianna,
Ivo Paixão de Medeiros,
Bernardo Santos Aflalo,
Leonardo Ramos Rodrigues,
João Pedro Pinheiro Malère

EMBRAER S.A.
São José dos Campos, Brazil
{wlamir.vianna, ivo.medeiros, bernardo.aflalo, leonardo.ramos, joao.malere}@embraer.com.br

***Abstract* — This paper describes the application of the PHM concept to assess the State of Health (SoH) of a Proton Exchange Membrane Fuel Cell (PEMFC) as part of the IEEE PHM 2014 Data Challenge. A linear regression approach is used as health monitoring algorithm to estimate the impedance of the PEMFC. So that, the linear regression curves were estimated using least-square equations. After that, linear regression is deployed to find future values for impendance for four different frequencies: 50mHz, 789mHz, 5.18mHz and 505Hz.**

Keywords—Regression Analysis, Prognostics, Health Monitoring, Fuel Cells.

I. INTRODUCTION

Prognostics and Health Management (PHM) can be defined as the ability of assessing the health state, predicting impending failures and forecasting the expected Remaining Useful Life (RUL) of a component or system based on a set of measurements collected [1].

In order to accomplish this task, it is necessary to collect a set of data from the component. This dataset is defined on the basis of the type of equipment to be monitored (hydraulic, electronic, mechanic, etc.) and the failure modes that are intended to be covered by the PHM system.

After that, a health monitoring algorithm must be developed for each monitored equipment or failure mode. Each algorithm processes the relevant data and generates a degradation index that indicates how degraded the monitored equipment is.

This paper describes the application of PHM concept to assess the state of health (SoH) of a Proton Exchange Membrane Fuel Cell (PEMFC) as part of IEEE PHM 2014 Data Challenge.

PEMFC are electrochemical systems that convert directly hydrogen energy into electrical energy with high efficiency, and no CO_2 emission.

SoH is addressed by considering frequential domain. So that, EIS (Electrochemical Impedance Spectroscopy) measurements made at separate frequency ranges may be used to diagnose important PEMFC failures, as flooding and dehydration, tackled in [2].

PEMFC relevant physical processes (electron transfer, reactant diffusion, etc.) are controlled by the bulk properties of the different materials and, more importantly, by the interfacial characteristics between materials or phases. At each interface, the material properties change discontinuously and abruptly, and can become the limiting performance factors. EIS is a technique especially suited to characterize interfaces, and it can be used to characterize fuel cell performance non-invasively and in situ [2].

A linear regression approach is used as health monitoring algorithm to estimate the impedance of the PEMFC. This algorithm was developed over data provided by FCLAB Research Federation (FR CNRS 3539, France, http://eng.fclab.fr/) as part of the IEEE PHM 2014 Data Challenge.

Relating to the data, two datasets were provided: 1) fuel cell operated in stationary regime and 2) fuel cell operated under dynamic current. The first one corresponded to 1,155 hours of experimental data of the entire life of the fuel cell and the second one corresponded to 550 hours, without the failure information. Both datasets contained polarization, electrochemical impedance spectroscopy (EIS) and ageing parameters.

The rest of this paper is organized as follows: Section II presents the problem of assessing the State of Health (SoH) of a PEMFC; Section III describes the approach adopted to tackle that problem; followed by results in Section IV and concluding remarks in Section V.

II. PROBLEM DESCRIPTION

The problem addressed in the IEEE PHM 2014 Data Challenge is focused on a PEMPF (Proton Exchange Membrane Fuel Cell) and can be divided into two parts.

- Part 1: Assess the State of Health of a PEMFC
- Part 2: Predict its Remaining Useful Life

In part 1, the dynamic behavior of a fuel cell stack must be assessed based on information regarding its internal physical parameters. The State of Health estimation is addressed by considering frequency domain. The objective is to predict both the real and the imaginary parts of the impedance of a fuel cell stack.

In part 2, the objective is to predict the Remaining Useful Life of a fuel cell. The RUL is defined as the time before a fuel cell stack looses its ability to provide sufficient power. Various power drops are considered in this part of the challenge: 3.5%; 4.0%; 4.5%; 5.0% and 5.5% of the initial power.

In this paper, the authors present the approach used to solve the part 1 of the challenge.

A. Polarization Curves

Polarization curves, which describe the relationship between the cell voltage and the current density, provide information about the static behavior of the fuel cells. These curves have been used in many applications such as fuel cell characterization, ageing studies and diagnosis. In order to perform a polarization test, the fuel cells must be operating in stable conditions (temperature, pressure, humidity, etc.). In this work, before each polarization test, the fuel cells operated under the same operating conditions during 30 minutes.

According to previous research, polarization curves may provide three different internal behaviors of fuel cells, depending on the current density level, as described below [3].

- In low current density levels, polarization curves illustrate charge transfer kinetics.
- In medium current density levels, the form of the polarization curves is influenced by ohmic resistances.
- In high current density levels, polarization curves are influenced by the mass transfer.

Fig. 1 shows the variation of both the form and the slope of a polarization curve during an ageing experiment. The parameterization of polarization curves is helpful in order to provide indications about the evolution of the ageing process.

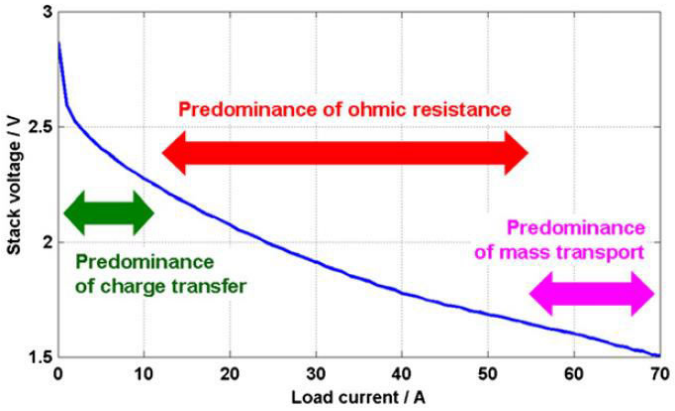


Figure 1. Example of a polarization curve. Source: Ref [4]

B. Electrochemical Impedance Spectrum

The Electrochemical Impedance Spectroscopy (EIS) test, which measures the dielectric properties of a medium as a function of frequency, provides relevant information about the dynamical behavior of the fuel cells during the ageing process [4]. This test has been used in diagnostics and prognostics applications for batteries [5] and fuel cells [3]. EIS allows the characterization of the dynamic processes at different timescales. In this work, fuel cell characterization was obtained by separating the impedance spectrum into two curves that represent the evolution of the real and the imaginary parts of the impedance as functions of the frequency.

Previous research focused on defining some characterization points in the impedance spectrum that could be used for fuel cells diagnosis purposes. The following parameters were chosen as features:

- The polarization resistance value.
- The minimum value of the imaginary part of the impedance, its corresponding real part value and its corresponding frequency.
- The internal resistance and its corresponding frequency.

Fig. 2 shows an example of the identification of the characterization points in the impedance spectrum.

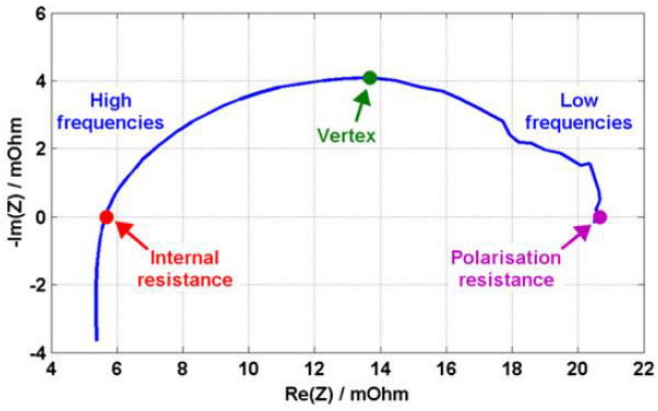


Figure 2. Example of an impedance spectrum curve. Source: Ref [4]

III. PROPOSED SOLUTIONS

Polynomial regression analysis consists of one of the common approaches [6] to model historical data for the purpose of prognostics. Its implementation is simpler compared to other data-driven method requiring less computational resources and fewer setting parameters. In spite of that, the polynomial order must be chosen and often its definition influences on the quality of results. In the present work, a linear regression, which is a 1st order polynomial regression was chosen. That proposal is described on the following topic.

A. Linear Regression

Linear regression has been widely used and compared to other methods. Examples include [7] and [8] which compares linear regression to non-linear methods such as Neural Networks. Although this method does not consider non-linear characteristics of the data, it can provide satisfactory results when the non-linear effects are less relevant compared to the linear ones [9].

The proposed solution consists in finding the coefficients a_{Re}, b_{Re}, a_{Im}, b_{Im} of the following linear equations:

$$Re_Z = a_{Re} t + b_{Re} \quad (1)$$

$$Im_Z = a_{Im} t + b_{Im} \quad (2)$$

Where:

Re_Z is the real part of the impedance in Ohms for a given time and frequency;

Im_Z is the imaginary part of the impedance in Ohms for a given time and frequency;

t is the time.

The challenge consists in finding future values for the impedance for four different frequencies: 50mHz, 789mHz, 5.18Hz and 505Hz. For each of them the coefficients in (1) and (2) are estimated with historical data using the least-squares equation given by (3).

$$B = (X^T X)^{-1} X^T Y \quad (3)$$

Where:

X are the past time historical data

Y are the past time values of impedance

B are the coefficients estimations

Given the results of (3) the future estimations of impedance are given by (4).

$$\hat{Y} = X_f B \quad (4)$$

Where:

$\hat{Y}$ are the future estimations of impedance

X_f are the estimation times

IV. RESULTS

In order to evaluate the method, the score was estimated for times 685h, 823h and 991h using FC1 data. The estimation of both real and imaginary impedance components for frequencies 50mHz, 789mHz, 5.18Hz and 505Hz for the linear regression are shown in Fig. 3 to Fig. 10.

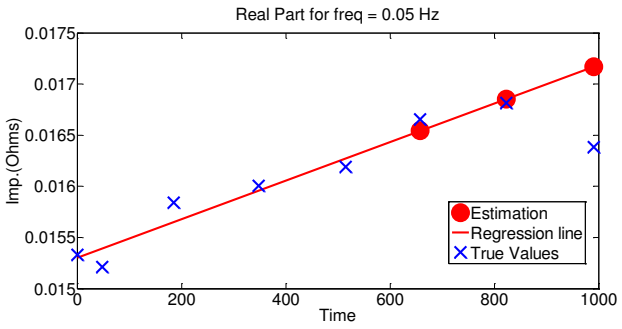


Figure 3. Real part linear regression estimations for FC1 at 50mHz

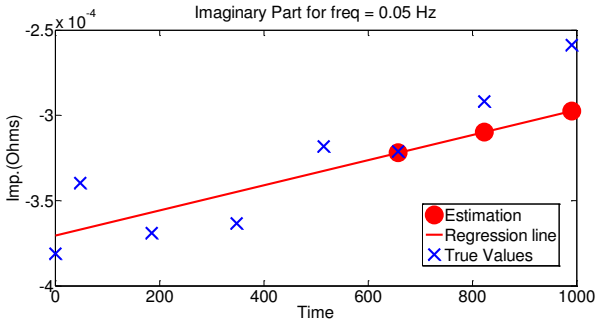


Figure 4. Imaginary part linear regression estimations for FC1 at 50mHz

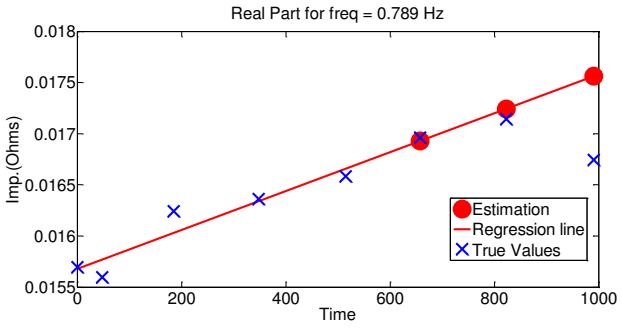


Figure 5. Real part linear regression estimations for FC1 at 789mHz

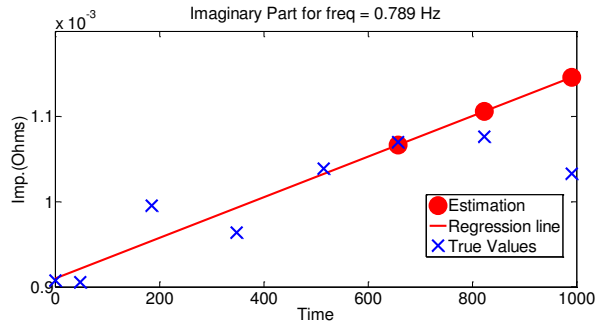


Figure 6. Imaginary part linear regression estimations for FC1 at 789mHz

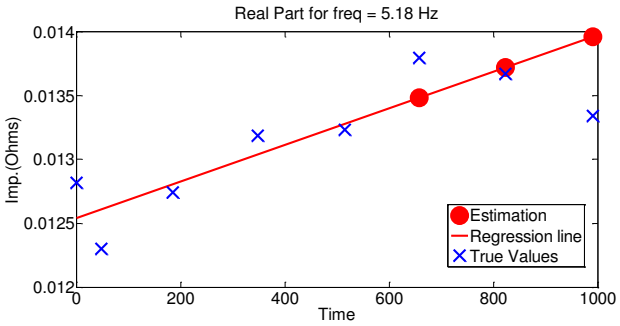


Figure 7. Real part linear regression estimations for FC1 at 5.18Hz

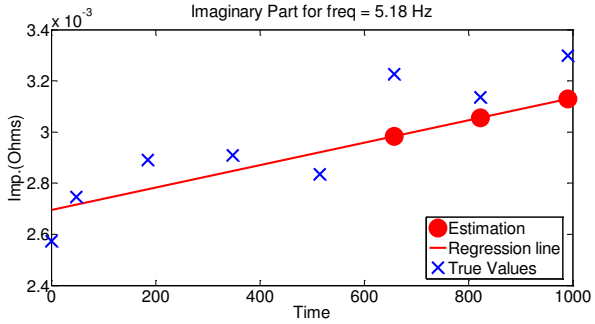


Figure 8. Imaginary part linear regression estimations for FC1 at 5.18Hz

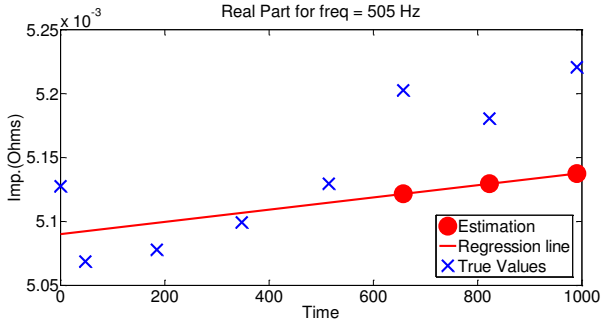


Figure 9. Real part linear regression estimations for FC1 at 505Hz

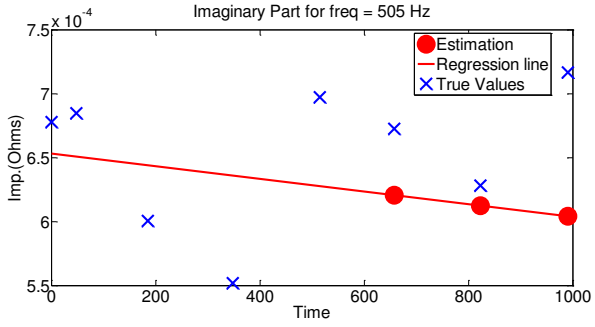


Figure 10. Imaginary part linear regression estimations for FC1 at 505Hz

The points indicated as True Values are the impedances collected for the EIS curves of FC1 data, the Regression curve is the curve estimated from (1) and (2), and the Estimation points are estimations from the linear regression for score estimation. The closest the Estimation points are to the True Values, best is the estimation and consequently the score.

Finally score was estimated for the FC1 data. The linear regression got a score of 4.14.10e-6. The score metrics are detailed on the following web page: http://eng.fclab.fr/ieee-phm-2014-data-challenge/.

The same methods were used to estimate future impedances for the FC2 data. Result is shown in TABLE I.

TABLE I. FC2 LINEAR REGRESSION IMPEDANCE ESTIMATIONS

	50mHz	**789mHz**	**5.18Hz**	**505Hz**
Re/666	0.0178	0.0181	0.0141	0.005
Im/666	-0.0003	0.0013	0.0034	0.0006
Re/830	0.0183	0.0186	0.0145	0.005
Im/830	-0.0002	0.0013	0.0036	0.0005
Re/1016	0.0189	0.0192	0.0149	0.0051
Im/1016	-0.0002	0.0014	0.0038	0.0005

V. CONCLUSIONS

This paper presented a PHM method used to estimate the State of Health of a Proton Exchange Membrane Fuel Cell (PEMFC) as part of the IEEE PHM 2014 Data Challenge.

The linear regression method consisted of linear curves estimated for the real and imaginary impedance components using least-squares equations. These curves were then used to estimate the future fuel cell impedance values.

The impedance for future times was estimated and the score was calculated for the FC1 data.

ACKNOWLEDGEMENT

The authors acknowledge FCLAB Federation (FR CNRS 339, France) for providing datasets exploited in this paper.

REFERENCES

[1] G. Vachtsevanos, F. L. Lewis, M. Roemer, A. Hess and B. Wu, “Intelligent fault diagnosis and prognosis for engineering systems”, 1st ed., New Jersey: Wiley, 2006.

[2] W. Mérida, D.A.Harrington, J.M. Le Canut, and G. McLean, “Characterisation of proton exchange membrane fuel cell (PEMFC) failures via electrochemical impedance spectroscopy”, Journal of Power Sources, pp. 264-274, 2006.

[3] M. Jouin, R. Gouriveau, D. Hissel, M. Péra, and N. Zerhouni, “Prognostics and health management of PEMFC − State of the art and remaining challenges”, International Journal of Hydrogen Energy, vol. 38, pp. 15307–115317, 2013.

[4] R. Onanena, L. Oukhellou, D. Candusso, F. Harel, D. Hissel, P. Aknin “Fuel cells static and dynamic characterizations as tools for the estimation of their ageing time”. International Journal of Hydrogen Energy, 2011, pp.1730-1739.

[5] K. Goebel, B. Saha, A. Saxena, J. R. Celaya, J. P. Christophersen, “Prognostics in battery health management”, IEEE Instrumentation & Measurement Magazine, vol. 11 (4), pp. 33–40, 2008.

[6] K.Goebel, B.Saha, A.Saxena, "A comparison of three data-driven techniques for prognostics", NASA Ames Research Center, 2008.

[7] R. M. Pascoal, W. O. L. Vianna, J. P. P. Gomes, R. K. H. Galvão, "Estimation of APU Failure Parameters Employing Linear Regression and Neural Networks", Proceedings of Annual Conference of the Prognostics and Health Management Society 2013, New Orleans.

[8] A. M. Riad, H. K. Elminir, H. M. Elattar, "Evaluation of a Neural Networks in the Subject of Prognostics as Compared to Linear Regression Model", International Journal of Engineering & Technology IJET-IJENS, 2010, Vol:10 No:06 pp. 50-56.

[9] N. R. Draper, H. Smith. Applied Regression Analysis, 3rd ed., Danvers: Wiley, 1998.

AUTHOR BIOGRAPHY

Wlamir Olivares Loesch Vianna holds a bachelor's degree on Mechanical Engineering (2005) from Universidade de São Paulo (USP), Brazil, and Master Degree on Aeronautical Engineering (2007) from Instituto Tecnológico de Aeronáutica (ITA), Brazil. He is with Empresa Brasileira de Aeronáutica S.A (EMBRAER), São José dos Campos, SP, Brazil, since 2007. He works as a Development Engineer of a R&T group at EMBRAER focused on PHM technology applications in aeronautical systems.

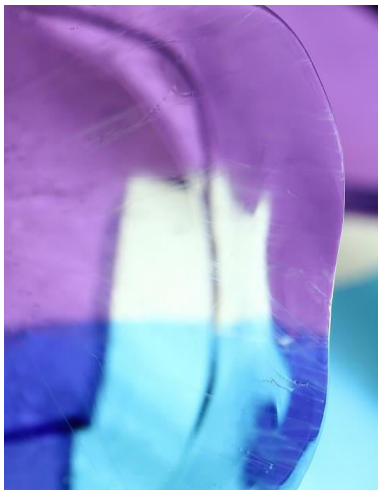
Ivo Paixão de Medeiros holds a bachelor's degree on Computer Engineering (2008) from Universidade Federal do Pará (UFPA), Brazil, and Master Degree on Eletronics and Computer Engineering (2011) from Instituto Tecnológico de Aeronáutica (ITA), Brazil. He is with Empresa Brasileira de Aeronáutica S.A (EMBRAER), São José dos Campos, SP, Brazil, since 2010. He works as a Development Engineer of a R&T group at EMBRAER focused on IVHM (Integrated Vehicle Health Management).

Bernardo Santos Aflalo is a Mechanical Engineer, graduated from Federal University of Minas Gerais, holds a Master Degree in Aeronautical Engineering from ITA (Brazil) and is currently enrolled in a PhD. He works for Embraer S.A. since 2007 and, currently, act as a R&D engineer. Current research areas include machine learning, data mining and analytics applied to enterprise business and processes optimization.

Leonardo Ramos Rodrigues holds a bachelor's degree in Electrical Engineering from Universidade Federal do Espírito Santo (UFES, 2003), Brazil, a Masters's Degree (2008) and a Doctorate's Degree (2013) in Aeronautical Engineering from Instituto Tecnológico de Aeronáutica (ITA, Brazil). He is with EMBRAER S.A. since 2006, working as a Development Engineer focusing on PHM technology applications. His current research interests are the application of health monitoring techniques for electronic components and the usage of PHM information for inventory management optimization.

João Pedro Pinheiro Malère holds a bachelor's degree in Control Engineering from State University of Campinas (Unicamp, 2004), Brazil, and a Master Degree in Aeronautical Engineering from Instituto Tecnológico de Aeronáutica (ITA, 2007), São José dos Campos, São Paulo, Brazil. He is with EMBRAER S.A., São José dos Campos, São Paulo, Brazil, since 2006. He works as a Development Engineer in an R&T group at EMBRAER performing research on IVHM technologies. His current research interest is on integrated health management systems, machine learning, data mining and analytics.

Challenges and Success in the Implementation of a Fleet Wide PHM for Energy Applications



Preston Johnson
National Instruments Corporation
Austin, Texas, 78759, USA
preston.johnson@ni.com

***Abstract*— Energy industry companies, both OEMS and operators, are quickly adopting fleetwide asset monitoring strategies in an effort to improve reliability and increase revenue. To derive efficient business insight and value from monitored equipment, seamless end-to-end solutions are needed - from the sensors to IT infrastructures. And, advanced tools and techniques are required for data transfer, management, and analytics. Further, systems management is needed for the many sensor connected data acquisition systems nodes. This paper details challenges and sample solutions in the design and integration of PHM Systems in the energy industry.**

Keywords- Anomaly Detection; Big Data; Data Management; Data Source Management; Diagnostic Visualization; Feature Extraction; Predictive Analytics; Predictive Maintenance; Sensor Selection

I. INTRODUCTION

The power generation industry is undergoing a transition from traditional power using Nuclear and Coal to more efficient gas turbine combined cycle technologies. While many of the traditional power plants have been in existence for many years, they are aging and require more maintenance [1]. Further, the cost of coal and nuclear fuel has increased significantly in the past several years. While natural gas is cost competitive in comparison to coal and nuclear fuels, the gas turbine combined cycle technology is more complex, and often more costly to repair. As a result of older power plants aging, and newer plants being more complex, a growing need for fleet health monitoring coupled with automated diagnostics and prognostics is needed.

In order to meet these needs, the power generation industry continues to identify and deploy new technologies for fleetwide monitoring. These technologies include information technologies such as data storage and management, expanding sensory data sources, automated sensory data recordings, and improved automation in analysis of sensory data.

II. POWER GENERATION INDUSTRY CHALLENGES

A. Aging and New Power Plants

The United States Power Industry has relied on Nuclear and Coal based power generation for the majority of base load demand for many years, Fig 1. As of March 2011, 51% of all generating capacity is over 30 years old.

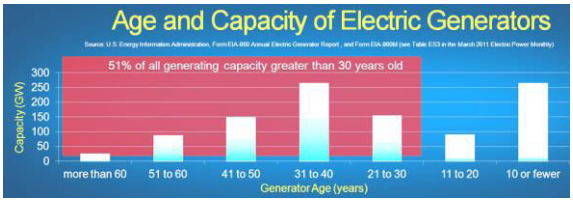


Figure 1. Age and capacity of electric generators

To keep these older generation assets producing power, additional maintenance is required. Adding to the maintenance challenges in power generation, the majority of new assets brought on line in the last 20 years are natural gas based. Combustion turbine and combined cycle power generation plants are more economical to operate, given the lower price in natural gas. However, natural gas plants incorporate newer technology that is more complex and often more costly to repair.

B. Adding Maintenance Intelligence Information Technology

To best understand and plan for maintenance requirements in both old and new power plants, additional sensory data acquisition systems are being deployed to improve visibility of rotating and stationary assets. In addition to adding monitoring sensors, information technology is being expanded to detect degradation and predict future performance and maintenance requirements, Fig 3. Research Organizations, such as the Center for Intelligent Maintenance Systems, and equipment vendors are developing information processing technologies to detect degradation and future health of many asset types, including those used in power generation [2].

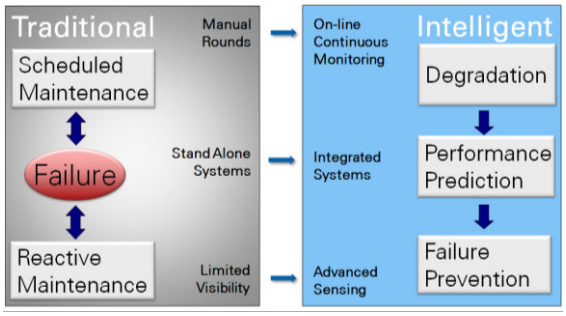


Figure 2. Transformation to intelligent maintenance systems

The transformation from traditional to intelligent maintenance systems requires several key technologies from the information technology domain space. These include embedded signal processing in the sensory data acquisition systems, systems management, data management to store and access the massive amount of new sensory data, and of course improved algorithms to process the new sensory data.

III. Duke Energy Business Drivers

A. Duke Energy, the Largest Regulated Power Producer

Duke Energy (Duke) is the largest power generation holding company in the United States. It operates a mixed portfolio of generation, totaling 58GW with 41GW non-nuclear [3]. The power generation assets are distributed amongst 81 plants, with some additional overseas holdings. When looking at the distribution non-nuclear (fossil) technologies used to produce power, 50% is coal, 40% is combustion turbines and combined cycle. Most of the plants are located in South Eastern United States with some in the Ohio valley.

Duke's traditional equipment condition monitoring has target goals of operating within design limits, identification of and planning resources for maintenance needs, with measurable reduction in outage time and increased availability. Maintenance technicians have historically used manual technologies, with sensory data collection conducted only during physical visits to the machine. Engineering evaluations are historically based on performance and long term trends using a data historian.

In 2010, Duke launched its "SmartGen" program to provide automatic sensory data collection, freeing maintenance personnel to spend time analyzing and reviewing sensory data rather than collecting it. The main objective of the program is to detect operational or equipment problems as soon as possible for the most critical units. This fast identification is expected to yield damage mitigation, shift from forced outages to planned repairs, identification of performance problems, and improved safety all while controlling costs and increasing availability of equipment.

B. Change in Operational Patterns Underscore Reliability

Traditionally, the nuclear and coal fleets provided the baseload of electrical power, running continuously. Duke and other coal fleet operators made large investments to clean emissions. Meanwhile, the smaller coal plants including combine cycle plants and simple combustion turbines were cycled on and off to meet peak demands during the day and summer or winter seasons. This pattern has recently switched, due to the cost of fuel where natural gas prices are low and coal prices are high. Baseload demand is now predominately provided by combined cycle gas turbine and steam turbine operations. Larger coal plants are now used to meet peak demand and smaller coal plants are being decommissioned. The result of this operational change is the combined cycle plants have higher reliability and availability demands. Further, the operating coal plants are experiencing reliability challenges as they operate differently than their design, that is they cycle on and off as compare to continuous operation.

As a result of these increasing reliability demands, the executive team at Duke issued a challenge to leverage new technologies to address increasing reliability demands and workforce optimization. Duke started an internal project in 2010 to explore technology and process changes. It went a step further, to collaborate with the Electrical Power Research Institute (EPRI) to address reliability needs from an industry perspective.

C. Increasing Impact with Industry Collaboration

With EPRI collaboration, technology and business challenge exchanges began to occur across the power generation industry. Exelon, the largest nuclear operator in the United State, meanwhile is experiencing similar challenges [4]. Exelon has limited online vibration monitoring instruments installed, limiting the ability of engineering to identify and diagnose mechanical issues. Plant experts spend too much time manually collecting sensory data, and in many cases equipment is in radiation areas. Exelon's immediate goal is to focus their vibration experts on analysis of equipment as compared to collecting vibration sensory data.

To further collaboration, both EPRI and Duke continue to solicit participation from vendors that provide online monitoring equipment, data analytics, and resource management tools. Fundamentally, EPRI company members are power generation, transmission, and distribution owner operators. Acting as a group, EPRI can help drive standards and invite vendor input on new technologies and best practices.

EPRI for example, hosts both a fleetwide monitoring (FWM) and condition based maintenance (CBM) working group and conference each year. These conferences encourage collaboration with technology and practices exchanges. Both conferences build on the EPRI Equipment Condition Assessment [5] and the Fleet Wide Monitoring for Equipment Condition Assessment [6] reports.

In particular, EPRI and company members are promoting technology innovations in several key areas, Fig 3. These technology innovations range from communication networks, to sensors, to diagnostic / prognostic analytics and data integration / fusion and visualization.

Figure 3. Technology area enhancements

IV. On-line Monitoring Architecture

A. Objectives of the On-line Monitoring System

As stated earlier, the core objectives of an on-line monitoring system are to greatly reduce the time equipment specialists spend collecting the data, and as a result to increase the amount of time specialists spend analyzing sensor data and results from automated analysis, Fig 4. This change from manual sensory data collection is intended to result in improved consistency in diagnostics using automation and standardization. Other improvements include better fusion of technology exam sensory data with process data. The end result is expected to be a more integrated monitoring and diagnostics center with improved visualization, enabling engineering and specialist workforces to perform higher value tasks.

In comparison to manual route based data collection, Fig. 5, on-line monitoring systems overcome several disadvantages. The first disadvantage is sparse data collection schedules. With manual route based exams, specialists visit the machines on schedules perhaps just once per month or once per quarter. These schedules may be interrupted by unplanned higher priority needs of the plant. Duke, for example, makes nearly 60,000 manual exams per month. A second disadvantage is equipment availability for an exam. The equipment may not be in operation during the specialist physical visit. Further, there is a high probability of missing an event, as the symptom of degradation may not adequately show itself during the periodic visit. Fourth, when the technical exam sensor data is collected, it often remains on the specialist's computer, until such time as the specialist determines it is important to report during a face to face meeting. In other words, an individual's limited view of the overall equipment may prevent data from being reported at a face to face meeting. And perhaps most importantly, over 60% of specialist manpower is used to collect sensory data, with limited time left for analyzing and reporting equipment health.

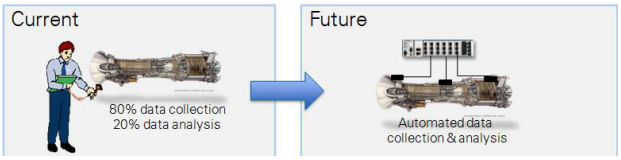


Figure 4. Workforce optimization thru on-line monitoring

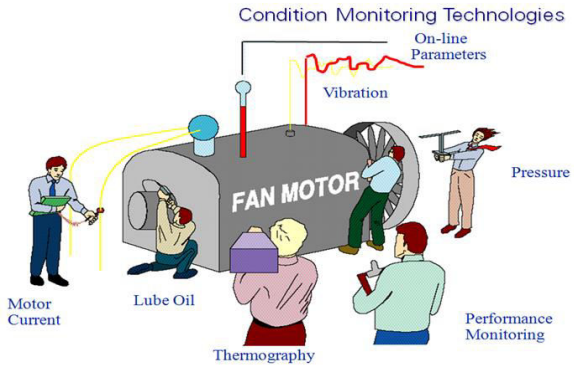


Figure 5. Manual technology exam measurements

B. Architecting the On-line Monitoring System

Over the past several decades, many vendors of data acquisition systems have brought to market and evolved on-line data acquisition systems. Systems have evolved from pure sensory time waveform acquisition and transfer to those with embedded intelligence. Over time, experience has driven embedded intelligence towards data reduction, data quality validation, and data acquisition system health. Further, experience has driven node system management systems that manage configuration and the networks these devices reside on.

To implement an on-line monitoring system at Duke, automatic data collection nodes, capable of measuring sensors from multiple technologies, are added to a computer communications network, Fig. 6. These automated data collection systems are most easily added to the business network, taking advantage of both wired and wireless infrastructures that have appeared in the power generation plants over the last several years. By placing the data acquisition systems on the business computer network, the data acquisition systems avoid interfering with control systems, and face less interference evaluation. A sample data acquisition system, Fig. 7, includes data acquisition hardware, power supplies, fuses, and communication equipment.

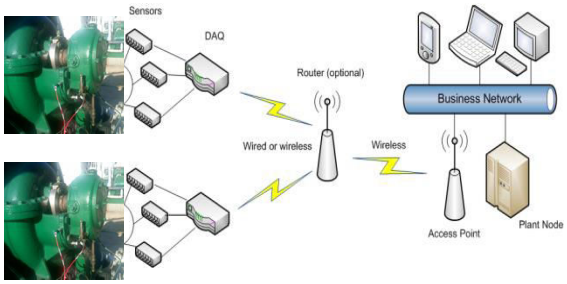


Figure 6. Data acquisition system on the business network

Figure 7. Sample data acquisition hardware cabinet

Data acquisition systems must be able to digitize physical phenomenon from a variety of sensors, both those making dynamic and static measurements. Dynamic measurements are of physical phenomenon that changes rapidly such as vibration, motor currents, and pressure. Static measurements include oil, temperature, flow, and loads. Dynamic measurements may utilize analog to digital sampling rates in the 10's of thousands of measurements per second. These systems are designed to continuously monitor sensors, in order to overcome the problem of missing an equipment degradation indication.

A challenge in continuous monitoring, is managing the large amount of data being acquired. For example, just monitoring two critical feed water pump shafts with two

bearings can produce over one terabyte of data per week with continuous sampling. To overcome this sensory data deluge, the continuous monitoring data acquisition systems must be designed to record and transfer sensory data on either a periodic basis or an event bases, Fig. 8.

As a result of the need to reduce the amount of data recorded and transmitted on the network, these networked systems must be both data acquisition and analysis network nodes (DAAN). The DAAN initializes itself on power up and begins monitoring sensory values as digitized by onboard analog to digital converters. Both dynamic and static measurements are made with their time stamps synchronized. Some sensory values may come from communications to local control systems. As the sensory values arrive in memory, the DAAN analyzes the time stamps and values of sensory measurements to determine a periodic trigger, or an event based trigger. With a trigger identified, sensory time waveform data is validated for quality and recorded to local on-board storage and placed in an out box directory for later transfer onto the network.

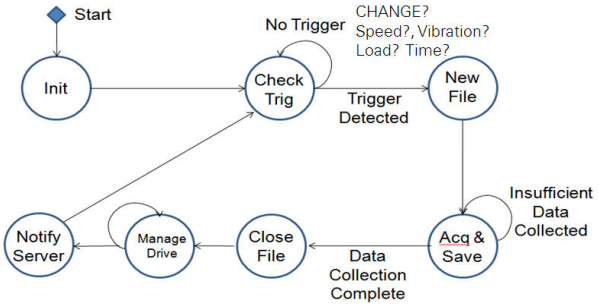


Figure 8. Data acquisition system software diagram

A server class computer, also residing on the business network, is responsible for managing the DAANs and retrieving sensory data recordings from them, Fig. 9. It should be noted that the DAAN, when recording data to its local disk, has provided metadata including equipment hierarchy, sensor calibration information, sensor location, time stamps, and other pertinent information to facilitate data search, off-line analysis, and peer to peer comparisons.

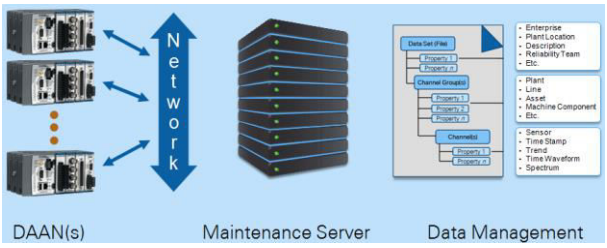


Figure 9. Data acquisition, network, and server

The server has the responsibility of managing all of the DAANs within the network. The server tasks include discovery of monitoring devices, insuring correct configuration of the monitoring devices, managing network and communications security, and monitoring the health of the DAAN as well as the attached sensors. These tasks are performed using both standard and proprietary vendor specific communications protocols to detect, configure, manage, and retrieve data from the DAANs.

The maintenance server has the responsibility of hosting specialist visualization and analysis software. With these tools, vibration analysts in particular can retrieve vibration time waveforms to perform comprehensive visual, graphical, and comparative analytics within a single machine or across machine peers. The maintenance server also has the responsibility of transferring condition indications to the plant historian where DAAN collected condition indicators are later correlated with process and operations data, Fig. 12.

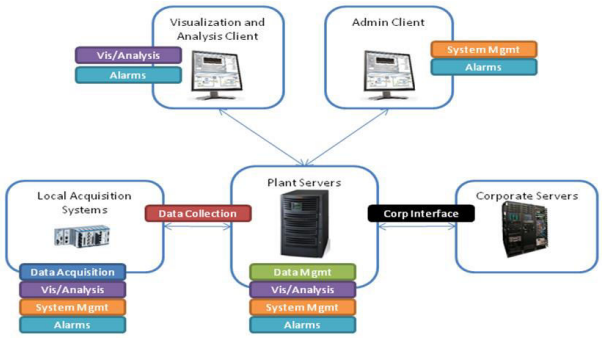


Figure 10. High level architecture of monitoring system

In summary, the networked automatic sensory data collection system performs many tasks. The system resides on a business network, to reduce interference with operations. The DAANs have built in analytics and intelligence to determine when to record sensory data and to determine both sensor and its own operational status and health. The server computer managing the network aggregates sensory data from all DAANs, publishes condition indicators to a plant historian, and provides search, retrieval, and analytics of collected data recordings.

V. DISTRIBUTED ANALYTICS

A. *Overall Analytics Architecture*

Analysis of sensory data from the DAANs, and supporting control systems occurs in multiple locations, and ranges from threshold alarms, to advanced signal processing of time waveforms, to automated diagnostics, and health prediction, Fig. 11. To optimize the overall process of tracking equipment health, reliability, and availability; a distributed analytical architecture provides both advanced calculations capabilities and data fusion opportunities.

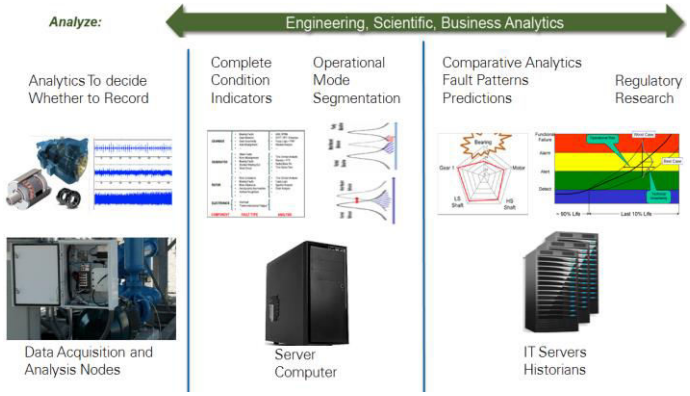


Figure 11. Analytics architecture

The analysis can be broken down into several areas. These include decisions to record and save sensory data recordings, condition indicating calculations, determining the operational state of equipment, peer to peer and fault signature comparisons, prognostics or equipment maintenance projects, and even budget and maintenance schedule optimizations. The location of analytics is often conducted as soon as possible in the sensory information flow, from DAAN to plant historian. If an analysis result can drive a decision to record sensory data, then it should occur in the DAAN. If additional computation power is needed to complete the set of condition indicators for a specific piece of equipment, these calculations can be made on the maintenance server. When comparisons to fault signatures, operational states, and peer equipment are made; these calculations may best be made on either the maintenance server, or after operational data is fused at the plant historian level.

To summarize, analysis of sensory data is made at either the DAAN level, the maintenance server level, or the plant historian and plant wide or corporate level.

B. Calculation of Condition Indicators

To determine when to record data, analysis in the DAAN is conducted in real-time and continuously to determine if a core condition indicator has changed. The analysis in the DAAN includes both vibration analysis and level thresholds. For vibration and similar dynamic sensors, onboard analysis includes true overall, derived overall from a spectrum, crest factor, unbalance harmonics, and mechanical looseness harmonics. Using both dynamic signal analysis and threshold analysis, either dynamic sensory information or static sensory information can trigger a sensory data recording at the DAAN level of the network.

Once the sensory data recording reaches the server, additional condition indicating calculations are performed including more advanced analysis of dynamic sensory time waveforms, Fig 12. These calculations include bearing and gear harmonic analysis, motor current signature analysis, pump vane harmonics, and so on. These calculations can cover a wide range of machinery conditions and indicate a wide range of failure types [7]. The system then supports specific analysis for specific equipment and components, taking advantage of the sensors applied to the specific equipment instance. With the server augmenting the DAAN with additional analysis, a complete set of condition indicators becomes available for each equipment and equipment component.

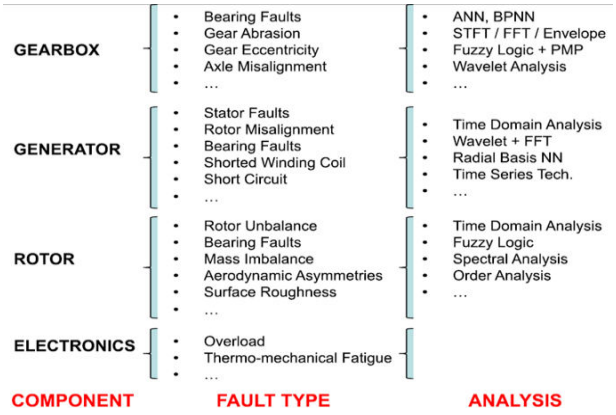


Figure 12. Machinery components, faults, and analysis

C. Finding Patterns Amongst the Condition Indicators

After sorting the condition indicators by equipment type, component type, and operational mode, it is then possible to begin developing patterns from the condition indicators. In many cases, the patterns are first developed for normal or equipment healthy status. Equipment specialists can confirm specific equipment is operating normally with little signs of degradation or reliability. The condition indicators from this normally operating equipment then become an expected pattern of behavior for this equipment under the specific operating conditions.

There are a number of pattern formation tools in the data science and prognostics toolbox, Fig. 12. In practice, the most commonly used technique is the un-supervised tool of anomaly detection. With this tool, a normal set of condition indicators is continuously compared with in-coming sensory data and condition indicators to detect if something unexpected with the equipment is occurring. Anomaly detection can take the form of statistical pattern recognition, quantization errors from expected values, or even logistical regression. The anomaly detection is generally undertaken using condition indicators and process data that have been recorded in the plant historian. If a specific cause is determined, the anomaly then forms a pattern of degradation such as a bearing or gear failure. These patterns are then added to a fault signature database for use in automated diagnostics applications.

D. Automating Diagnostics and Prognostics

With known patterns of specific failure modes, automated diagnostics and predictive maintenance is becoming more possible and probable. Today, there remains a need to validate computer generated recommendations, yet with the advances in data science and predictive analytics, the reliability of automatic diagnostics and prognostics is improving.

For example, given the existence of data driven patterns of bearing, gear, motor, and other rotating component failures; the use of logistic regression (and other techniques) to compare selected condition indicators with fault signature patterns can yield both a specific degradation mode of the equipment, and by trend analysis a likely number of working hours before maintenance is needed.

Visualization tools, including the health radar chart, Fig. 13, provide a visual overview of the current state of health of

specific equipment. Sequential comparison of condition indicators for each fault type, or comparison by other correlation techniques, help distinguish between normal patterns of condition indicators and process parameters and those patterns indicating specific failure modes. There are a number of techniques in use today, both in research and in commercial products. Within the power generation industry, anomaly detection is provided by Instep Software's PRiSM™, General Electric's SmartSignal™, General Physics EtaPro™, and other trend analysis products.

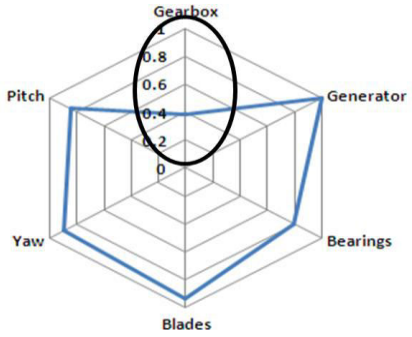


Figure 13. Health radar chart indicating component health

VI. Lessons Learned at Duke Energy

A. Current Implementation Status

Duke Energy has deployed DAANs, condition indicating analytics, as well as anomaly detection and visualization tools within several of their power generation plants in North America, Fig. 14. Each of these plants has deployed 20 or more DAAN nodes per power generation block. Each plant has a server managing the DAANs, calculating condition indicators and reporting these condition indicators to the OSIsoft PI™ Historian. Instep Software's PRiSM™ software is at work building data driven models of normal behavior for anomaly detection.

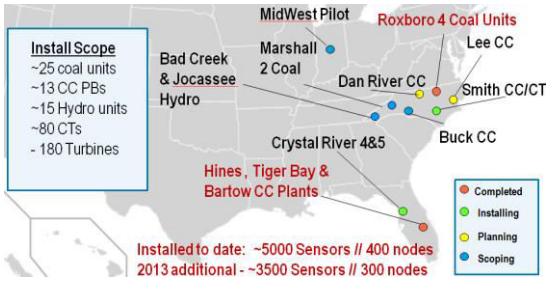


Figure 14. Current fleetwide monitoring at Duke Energy

B. Core Lessons Learned At Duke Energy

Deployment of automated sensory data collection on the fleetwide scale requires significant resources for planning and implementation. Implementation managers are needed at each facility to manage the sequence, personnel resources, and equipment resources that come together to roll out the DAANs, server software, and enterprise connectivity.

Hardware installations can proceed ahead of software installation, especially identification of sensor types and locations and the subsequent installation. Server installations should be timed to coincide with DAAN installation. Once sensors and DAANs are installed, a validation process is needed to validate sensory measurements and calculated condition indicators match traditional manual based activities.

As the anomaly detection occurs on the OSIsoft PI™ Historian, validation on both condition indicators and PI representation of the condition indicators must occur prior to building data driven models. This is similar to any data science or predictive analytics application, data integrity is of high importance.

All sites are excited and bought into the prospect of automated data collection, and assisted and automated diagnostic and predictive techniques. Site persons continue to ask for more automation, more sensory types, and greater equipment coverage. Regular implementation process meetings focused on streamlining the implementation process, and on streamlining feedback are recommended.

The biggest lesson learned is that the system is working as expected. Already visibility of equipment reliability has greatly improved, and plans are now being made to track maintenance savings and availability improvements. With a track record of plant installations, the roadmap for addressing additional plants is established, and can be built upon. Duke Energy is well on its way to complete its fleetwide monitoring and diagnostics center. Duke's efforts promise to result in maintenance savings and availability improvements, while increasing equipment health visibility and optimizing logistics of maintenance.

References

[1] US Energy Information Administration, "Electric generator report", Form EIA 860 and form 860M, Electric Power Monthly, March 2011.

[2] Center for Intelligent Maintenance Systems, "IMS Center Brochure", www.imscenter.net, 2012.

[3] B. Cook, "Duke Energy SmartM&D project", National Instruments Week 2013, August 2013.

[4] M. Yousuf, "Advanced monitoring, technology innovation and work force optimization", National Instruments NIWeek 2013, August 2013.

[5] R. Shankar, "Equipment condition assessment, volume 2, technology evaluation and integration", EPRI publication 1009601, December 2004.

[6] R. Shankar, "Fleetwide monitoring for equipment condition assessment", EPRI publication 1010266, March 2006.

[7] P. Jayaswal, A. Wadhwani, and A. Mulchandani, "Machine fault signature analysis", *International Journal of Rotating Machinery*. Volume 2008, Article ID 583982, 10 pages doi:10.1155/2008/58398

AUTHOR BIOGRAPHY

Preston Johnson is the Principal Sales Engineer for Condition Monitoring Systems at National Instruments (NI) in Austin, Texas. He has worked for National Instruments for over 27 years in roles of Field Sales, Sales Management, Automation Business Development, Sound and Vibration Segment Manager, Platform Manager for Condition Monitoring Systems and Global Program Manager for Asset Monitoring Systems. In his current role as Principal Sales Engineer, Preston works with NI OEM and End User customers to deploy fleetwide asset monitoring systems that lower operation costs, improve machinery reliability, and ultimately increase revenue. His interests lie in embedded signal processing, data acquisition systems and architectures, and prognostics. He earned his BSEE in Electrical Engineering and Computer Science from Vanderbilt University in 1985 and his MBA in Information Systems from the University of Texas in 1987. Preston is experienced in project management and holds a Category III vibration analyst certificate.

Fleet-Wide Prognostic and Health Management Suite: Asset Fault Signature Database

Vivek Agarwal and Nancy J. Lybeck
Department of Human Factors, Controls, and Statistics
Idaho National Laboratory
Idaho Falls, ID, USA
vivek.agarwal@inl.gov, nancy.lybeck@inl.gov

Randall Bickford
Expert Microsystems
Orangevale, CA, USA
rbickford@expmicrosys.com

Richard Rusaw
Nuclear Division
Electric Power Research Institute
Charlotte, NC, USA
rrusaw@epri.com

***Abstract*—Proactive online monitoring in the nuclear industry is being explored using the Electric Power Research Institute's Fleet-Wide Prognostic and Health Management (FW-PHM) Suite software. The FW-PHM Suite is a set of web-based diagnostic and prognostic tools and databases that serves as an integrated health monitoring architecture. The FW-PHM Suite has four main modules: (1) Diagnostic Advisor, (2) Asset Fault Signature (AFS) Database, (3) Remaining Life Advisor, and (4) Remaining Useful Life Database. This paper focuses on the AFS Database of the FW-PHM Suite, which is used to catalog asset fault signatures. A fault signature is a structured representation of the information that an expert would use to first detect and then verify the occurrence of a specific type of fault. The fault signatures developed to assess the health status of generator step-up transformers are described in this paper. The developed fault signatures capture this knowledge and implement it in a standardized approach, thereby streamlining the diagnostic and prognostic process. This will support the automation of proactive online monitoring techniques in nuclear power plants to diagnose incipient faults, perform proactive maintenance, and estimate the remaining useful life of assets.[1]**

***Keywords*—generator step-up transformer; fault signatures; fleet-wide monitoring; diagnostics.**

I. INTRODUCTION

The average age of existing commercial nuclear power plants (NPPs) in the United States is 35 years. As these plants continue to age and their components degrade, it is important to understand their condition and be proactive in maintenance and replacement. The current periodic and condition-based maintenance practices at NPPs result in high maintenance costs and increased likelihood of human error. Additionally, the inability to identify developing faults can lead to either disabling component failure or forced plant outage. Implementing advanced, predictive online monitoring would minimize these limitations and enhance plant safety by enabling plant maintainers to diagnose incipient faults and estimate the remaining useful life (RUL) of their assets, reducing operational costs by optimizing maintenance activities. Predictive online monitoring techniques include advanced diagnostic and prognostic techniques.

The U.S. Department of Energy (DOE) Office of Nuclear Energy funds the Light Water Reactor Sustainability Program to develop the scientific basis to extend the operation of commercial light water reactors beyond the current 60-year licensing period. The Advanced Instrumentation, Information, and Control Systems Pathway under the Light Water Reactor Sustainability Program is collaborating with the Electric Power Research Institute's (EPRI's) Long-Term Operations Program to conduct research and development on technologies that can be used to enhance the long-term reliability, productivity, and safety of aging light water reactors. One of the primary areas of focus for the Light Water Reactor Sustainability and Long-Term Operations Programs is online monitoring of active assets in the nuclear industry.

An important objective of the research is to implement predictive online monitoring for the existing fleet of NPPs. EPRI's Fleet-Wide Prognostic and Health Management (FW-PHM) Suite software was selected for use as a demonstration platform. EPRI and Idaho National Laboratory (INL) are working with nuclear utility partners to develop asset fault signatures in the FW-PHM Suite software for generator step-up transformers (GSUs) and emergency diesel generators (EDGs). The nuclear utility

[1] The research work was supported by the Light Water Reactor Sustainability Program at INL, funded by the U.S. Department of Energy under U.S. Department of Energy Idaho Operations Office Contract DE AC07-05ID14517.

partners include Shearon Harris Nuclear Generating Station (owned by Duke Progress Energy) for GSUs and Braidwood Generating Station (owned by Exelon Nuclear) for EDGs.

This paper presents a detailed description of fault signature development and the attributes of fault signatures. Examples of fault signatures are presented that show different fault signature attributes and their structure in the Asset Fault Signature (AFS) Database. The paper is organized as follows. Section II describes the FW-PHM Suite software architecture. Section III describes different attributes associated with a fault signature and steps involved in developing a fault signature. Section IV presents a representative fault signature for GSUs. Section V presents conclusions and future FW-PHM Suite software research.

II. Fleet-wide Prognostic and Health Management Suite Software

The FW-PHM Suite software, shown in Fig. 1, is a comprehensive asset health management solution that supports Condition-Based Maintenance (CBM) activities common to the power generation industry. The FW-PHM Suite was developed by Expert Microsystems, Inc., for EPRI's Fossil Generation and Nuclear Sectors [1]. The software is designed to support all types of generating and transmission assets, including passive and active assets in nuclear, coal, combined cycle, co-generation, wind, hydro, and switchyard facilities.

The FW-PHM Suite has four primary modules: the AFS Database, the Diagnostic Advisor, the RUL Database, and the Remaining Life Advisor. The AFS Database organizes fault signatures and troubleshooting information for power plant assets. At the most basic level, a fault signature is a structured representation of the information that an expert would use to first detect and then verify the occurrence of a specific type of fault. The Diagnostic Advisor accurately identifies faults and impending failure conditions in plant assets by comparing known asset fault signatures with operating data, sensor values, nondestructive examination (NDE) results, and expert opinions from power plant personnel.

The RUL Database collects and organizes power plant asset life model information, and archives and shares that information to make accurate prognostic projections. The Remaining Life Advisor uses plant information to determine how long an aging asset will provide reliable service. For additional details on the FW-PHM Suite, refer to EPRI's 2012 report and manual [1, 2].

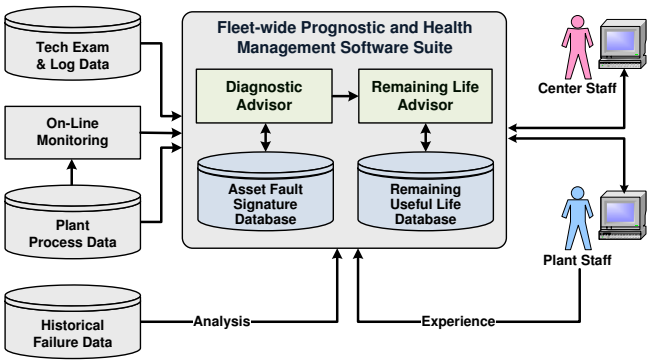


Fig. 1. FW-PHM Suite software architecture [1].

III. Asset Fault Signature Database

Diagnostic fault signature information for high-priority power plant equipment is elicited from industry experts and captured in the AFS Database of the FW-PHM Suite software [2]. A methodical approach is required to develop content for the AFS Database. The following subsections provide a detailed explanation of the terminology and structure of fault signatures, along with a procedure for preparing fault signatures for implementation in the FW-PHM Suite.

A. Asset Fault Signature Database Terminology

The AFS Database contains tables that organize and store reference information for various power plant assets. This information can be represented using an asset subtype and a reference asset-type hierarchy as shown in Fig. 2. Asset subtypes represent a generalized definition of an asset, without specifying contextual information such as the nature of its use in service within a particular kind of plant or the plant application. Asset types represent a more specific definition of an asset subtype, including information related to its use in service within a particular kind of plant or plant application.

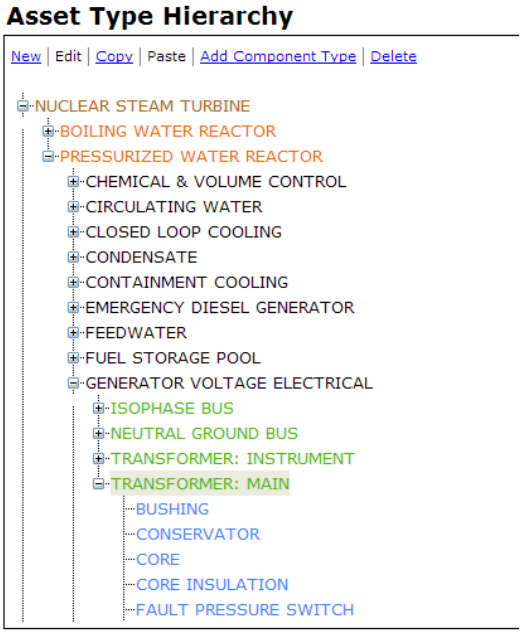


Fig. 2. An example of asset types.

Asset subtypes and asset types are organized into five discrete levels within the AFS Database: plant, unit, system, equipment, and component. Asset types are additionally organized within a set of reference asset hierarchies for various kinds of power generating plants. The collection of reference asset hierarchies is called the reference asset taxonomy.

A fault is a particular mode of degradation that can be detected by analyzing plant information before the asset fails to meet its service requirement. Fault types represent a specific definition of a fault, including contextual information such as its location and the nature of its use within a particular kind of power plant application. Fault information tables are used to organize and store information for the different fault types associated with plant assets. Paper insulation degradation in a transformer winding is an example of a fault type (Fig. 3). Attributes

associated with fault types can be used to tailor how the fault applies. Fault attributes and their values also provide a way to discriminate specific fault causes (e.g., arcing and cellulose degradation are two of the fault attributes associated with paper insulation degradation).

Technology examination tables organize and store information about technology examinations performed for various plant assets. Here, the term technology examination is not limited to its common use, but is broadened to mean any form of examination of plant information, including operator examinations such as inspections, and even maintenance actions that can influence a fault diagnosis, such as on oil analysis that counts the number of particles found in a sampled volume of oil.

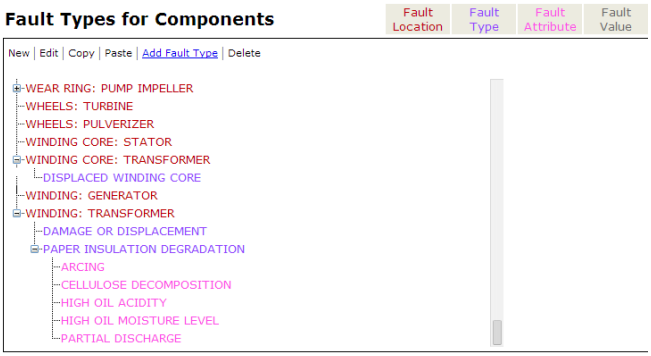


Fig. 3. An example of fault types at the component level for transformer winding.

Technology subtypes represent the various technologies available to observe degradation indications. Examples of technology subtypes include vibration, oil analysis, and temperature. Exam subtypes represent the various exams associated with technology subtypes that can be performed to measure or observe the degradation indicators. An acid number is an exam subtype associated with oil analysis.

Result types define the possible set of outcomes for a technology examination. Each technology examination is associated with a single result type. The term exam result refers to the possible outcome of a technology examination. Exam results can be represented as categorical values or non-categorical values. Non-categorical values can include numeric values, a time series of numeric values, or a vector of numeric values.

B. Asset Fault Signatures

Asset fault signature tables, the backbone of the AFS Database, are used to organize diagnostic fault signature information. At a minimum, a fault signature is comprised of an asset type, a fault type and a set of one or more observable features that may indicate the presence of the associated fault. Optionally, the fault signature can specify one or more fault types that can either cause or be caused by the specified fault. Corrective actions or a list of possible remedies can also be included in a fault signature.

Fault feature tables organize and store information for the fault features associated with various fault signatures. A fault feature definition includes a technology examination type, the location for the technology examination, and the results from the examination that indicate the possible presence of the fault. Additionally, a fault feature describes the effectiveness of the associated technology examination. Finally, a fault feature contains the technology examination result that indicates the presence of the asset fault. For example, a high-temperature measurement is a highly effective fault feature for a bearing damage fault in a pump.

C. Gathering Fault Signature Information

A step-by-step procedure for developing a fault signature includes the following four steps [2]:

- Specify the asset type for which the fault signature(s) is to be developed. For the specified asset type, gather information on its operating range, mechanisms of degradation, and observable features that can be used to detect degrading conditions. The information can be gathered from multiple sources such as EPRI's Preventive Maintenance Basis Database, Fossil Maintenance Application Center and Nuclear Maintenance Application Center reports, and from other sources (e.g., textbooks, equipment guide, and publications).
- Specify the fault type and any narrowing attributes that can be used to make the fault more specific. For example, the most common fault type associated with transformer winding is paper insulation degradation. Fault attributes such as arcing, cellulose decomposition, high oil acidity, high oil moisture level, and partial discharge are commonly used to identify the root cause of paper insulation degradation in a transformer winding.
- For each fault type, specify one or more fault features comprised of information on (1) location where the plant data are collected; (2) technology or technologies used to identify the fault (e.g., oil analysis in transformers); (3) examination (e.g., particle content, gas analysis, or moisture content) and outcome of examination (i.e., the result, whether normal, abnormal, high, low, marginal, or unacceptable); and (4) the effectiveness (e.g., low, medium, high, or very high) of the fault feature in detecting the fault condition. A specific fault feature can be associated with one or more fault signatures.
- Provide a description of the fault condition, possible causes, remedies, and effects on the asset (if left uncorrected).

Several fault signatures have been developed and implemented in the AFS Database as part of a knowledge transfer exercise with utility partners for GSUs [3, 4] and EDGs [3, 5]. Twenty-three fault signatures have been implemented in the AFS Database for GSUs. The following section discusses fault signatures developed for transformer bushings and insulating oil.

D. GSU Fault Signatures

High-voltage transformer bushings are used to connect overhead power lines to transformers. Badly degraded bushings may explode, causing extensive damage to the transformer. One of the most common degradations observed in the majority of bushing types is loss of dielectric strength. There are several causes for loss of dielectric strength that include moisture ingress, loss of oil from the bushing, and oil contamination. Technology examinations such as capacitance tests, power factor tests, tan delta tests, and oil level inspections are widely used to monitor loss of dielectric strength in transformer bushings. Fig. 4 shows the bushing fault signature for low dielectric strength developed and implemented in the AFS Database.

Fig. 4. Transformer bushing low dielectric strength fault signature and associate fault features.

Fig. 5. Insulating oil degradation due to contamination and associated fault features.

Transformer insulating oil functions both as an electrical insulation and as a heat transfer fluid. Transformers are expected to function reliably and efficiently. The quality of the insulating oil plays an important role in performing this function for extended periods. Chemical compounds of the oil disintegrate over time when the transformer is subject to various operating conditions. These chemical changes degrade the quality of the oil, thereby leading to contamination, an increase in oil acidity, a decrease in dielectric strength, and degraded thermal properties. These changes impact the oil's electrical insulating and heat transfer properties. If left to degrade beyond allowed limits, the degradation could cause potentially catastrophic electrical faults inside the transformer. Therefore, it is important to monitor oil quality and diagnose any impending failure in a timely manner. Fig. 5 shows one of four fault signatures (contamination, high acidity, low dielectric strength, and thermal degradation) developed and implemented in the AFS Database for GSUs.

Note that some of the fault features might be common to more than one fault signature. In the case of GSUs, the interfacial tension of the transformer insulating oil is a common fault feature for two fault signatures: contamination and low dielectric strength. In this situation, the effectiveness of the fault feature serves as a differentiating factor and is used by the Diagnostic Advisor to rank the possible diagnoses.

The implemented GSU fault signatures in the AFS Database were verified and validated by EPRI subject matter experts. A demonstration showing the ability of the FW-PHM Suite to diagnose a developing fault is documented in a report [6] and in a video [7]. Emerging faults were simulated by adding drifts to plant data. The demonstration highlighted the software's ability to identify faults based on evolving symptoms, using both online and offline data sources. Additionally, some of these GSU fault signatures were used to demonstrate the prognostic ability

of the FW-PHM Suite to estimate the remaining useful life of transformer winding paper insulation, using two well-established approaches [8].

IV. Conclusions and Future Work

This paper discussed the AFS Database of the FW-PHM Suite that is used to store fault signatures for different asset types. Fault signatures allow a structured representation of the information captured from multiple plant sources. Representative GSU fault signatures implemented in the AFS Database were presented; these fault signatures served as a foundation for demonstrations of the diagnostic and prognostic capabilities of the FW-PHM Suite.

As part of the future work, EPRI and INL will continue to develop and implement fault signatures for different asset types. For example, fault signatures for large electric motors are of great interest in the industry.

V. Acknowledgment

The authors would like to thank Kirk Fitzgerald (INL), Binh T. Pham (INL), and Rahul Palnitkar (Expert Microsystems) for their substantial support. The authors are grateful to Richard Reister (U.S. DOE), Bruce Hallbert (INL), Kenneth Thomas (INL), and Kathryn McCarthy (INL) for championing this research.

VI. Disclaimer

References

[1] Electric Power Research Institute (EPRI). *Fleet-Wide Prognostics and Health Management Application Research*. Report EPRI 1026712. Electric Power Research Institute, Charlotte, NC, 2012.

[2] Electrical Power Research Institute (EPRI). *Asset Fault Signature Requirements. Software Manual*. Electric Power Research Institute, Charlotte, NC, 2012.

[3] V. Agarwal, N. J. Lybeck, B. T. Pham, R. Rusaw, and R. Bickford, "Online monitoring of plant assets in the nuclear industry," Annual Conference of the Prognostics and Health Management Society, October, New Orleans, USA, 2013.

[4] N. J. Lybeck, V. Agarwal, B. T. Pham, H. Medema, and K. Fitzgerald, "Online Monitoring Technical Basis and Analysis Framework for Large Power Transformers," Interim Report INL/EXT-12-27181. Idaho National Laboratory, Idaho Falls, ID, 2012.

[5] B. T. Pham, N. J. Lybeck, and V. Agarwal, "Online Monitoring Technical Basis and Analysis Framework for Emergency Diesel Generators," Interim Report INL/EXT-12-27754. Idaho National Laboratory, Idaho Falls, ID, 2012.

[6] V. Agarwal, N. J. Lybeck, L. C. Matacia, and B. T. Pham, "Demonstration of Online Monitoring for Generator Step-up Transformers and Emergency Diesel Generators," Technical Report INL/EXT-13-30155. Idaho National Laboratory, Idaho Falls, ID, 2013.

[7] http://www.youtube.com/watch?v=8JCchydWlcg&feature=c4-overview&list=UUctcciH1NrAGpwMnKwvnLgQ

[8] V. Agarwal, N. J. Lybeck, and B. T. Pham, "Diagnostic and Prognostic Models for Generator Step-up transformer," INL/EXT-14-33124 Rev. 0, September 2014.

Author Biography

Vivek Agarwal is a research and development scientist at Idaho National Laboratory since 2011. He received his Ph.D. in Nuclear Engineering from Purdue University in 2011, M.S. in Electrical Engineering from The University of Tennessee Knoxville in 2005, and B.E. in Electrical and Electronics Engineering from University of Madras in 2001. His research interest includes signal processing, machine learning, battery modeling, wireless sensor networks, instrumentation and controls, diagnosis and prognosis. He is a Section Editor for the Journal of Pattern Recognition Research.

Nancy J. Lybeck is a data analyst at Idaho National Laboratory. She holds a Ph.D. in Mathematics from Montana State University. She currently serves as the technical lead for INL's Nuclear Data Management and Analysis System. Prior to joining INL, she worked in the prognostics industry.

Randall Bickford is Expert Microsystems' president and chief technology officer. He holds a B.S. in Chemical Engineering from University of California, Davis. He is a recognized worldwide expert in Prognostic and Health Management technology. He is one of the industry's pioneers and holds multiple patents in the areas of pattern recognition, fault detection, diagnostics, and prognostics. Prior to founding Expert Microsystems, he worked in the aerospace industry where he developed advanced diagnostic and digital control technologies for space propulsion systems.

Richard Rusaw is a Senior Project Manager in the nuclear sector at the Electric Power Research Institute. He received his B.S. in Nuclear Engineering from University of Michigan in 1979, and his MBA from University of North Carolina – Charlotte in 1983. He is a registered PE in South Carolina. At EPRI, his responsibilities are

focused on instrumentation and control (I&C) with a high degree of specialization in monitoring technologies and system reliability. Prior to joining EPRI, he spent 25 years at Duke Power as an I&C engineer supporting Oconee, McGuire, and Catawba nuclear power plants.

Submission Instructions

Authors should use the designated IEEE Reliability Digest Manuscript Central Website to summit their papers.

Please refer to the following steps to submit your papers:

1. Login to IEEE Reliability Digest Manuscript Central. If you have no account, sign up for one.
2. Click "Authors: Submit an article or manage submissions".
3. Please click "CLICK HERE" at the bottom of this page, and you will be brought to the five-step submission process.
4. You need to 1) choose the section that you are going to submit your paper to; 2) complete the submission checklist; 3) enter the comments for the editor, which is optional; 4) save and continue.
5. If you have any supplementary files, please upload them in step 4.

Manuscript Types

Manuscripts for regular issues fit within the scope of the magazine, but are not intended for a special issue. Special issue manuscripts cover a specific topic scheduled on our editorial calendar. Please select the appropriate issue (manuscript type) when uploading your manuscript. For more information and to see upcoming special issue topics, see our Editorial Calendar at http://rs.ieee.org/reliability-digest/author-guidelines.html.

Typing Specifications

The manuscript should be written in Times New Roman in a double-column format. The typical length of the submitted manuscript is 4 single-spaced pages. The text portion of the manuscript should be in 10-point font and the title should be in 24-point font, bold.

Manuscript Length

The typical length of the submitted paper is 4 pages, including text, bibliography, and author biographies. Please note that proper citations are required.

Illustrations

The illustrations in the articles must be cited in the text and numbered sequentially. Captions that identify and briefly describe the subject are needed as well. In order to avoid dense and hard-to-read illustrations, graphs should show only the coordinate axes, or at most the major grid lines. Line drawings should be clear. To prevent potential layout problems from happening, related figures described within the same section of text should be grouped together as parts (a), (b), and so on.

References

All manuscript pages, footnotes, equations, and references should be labeled in consecutive numerical order they are mentioned in the text. Figures and tables should be cited in text in numerical order.

Biographical Sketch

A brief biographical sketch should contain the full title of the paper and complete names, affiliations, addresses, and electronic mail addresses of all authors. The corresponding author should be indicated.

Please provide a short biography and a picture for each author of your paper at the end of your paper.

The short biography should contain no more than 150 words

Copyright

IEEE Reliability Society does not own the copyrights of the articles published in Reliability Digest. If you wish to reproduce the copyrighted materials, please contact the copyright owners and seek their permissions. The contents in this website should be copied with use of proper citation.

Special Issue Proposal Submissions

For a special issue in Reliability Digest, experts are welcome to serve as our guest editors. To know more information, please contact Editor-in-Chief on Reliability Digest, Shiuhpyng Shieh: ssp@cs.nctu.edu.tw.