

Improved Heat Exchanger Lifecycle Prognostic Methods for Enhanced Light Water Reactor Sustainability

Zachary Welz¹, Alan Nam¹, Michael Sharp¹, J. Wesley Hines¹, and Belle R. Upadhyaya¹

¹Department of Nuclear Engineering, University of Tennessee, Knoxville, TN 37996, USA

zwelz@utk.edu
anam@utk.edu
msharp6@utk.edu
jhines2@utk.edu
bupadhyaya@utk.edu

ABSTRACT

As the licenses of many nuclear power plants in the US and abroad are being extended, accurate knowledge of system and component condition is becoming increasingly important. The US Department of Energy (DOE) has funded a project with the primary goal of developing lifecycle prognostic methods to generate accurate and continuous Remaining Useful Life (RUL) estimates as components transition through unique stages of the component lifecycle. Specific emphasis has been placed on creating and transitioning between three distinct stages of operational availability. These stages correspond to Beginning Of Life (BOL) where little or no operational information is available, early onset operations at various expected and observed stress levels where there is the onset of detectable degradation, and degradation towards the eventual End Of Life (EOL). This paper provides an application overview of a developed lifecycle prognostic approach and applies it to a heat exchanger fouling test bed under accelerated degradation conditions resulting in an increased understanding of system degradation. Bayesian and Bootstrap Aggregation methods are applied to show improvements in RUL predictions over traditional methods that do not utilize these methods, thereby improving the lifecycle prognostic model for the component. The analyses of results from applying these lifecycle prognostic algorithms to a heat exchanger fouling experiment are detailed.

1. INTRODUCTION

The field of prognostics for system reliability focuses on the determination of both component and system health and RUL to provide safety, reliability, and financial benefits. Interest in this field is growing as more commercial reactor

licenses seek to extend operations past original 40 year lifetimes, causing increased concern for the reliability and safety of the system components. A primary goal of prognostic models is to lessen plant down time and the resulting loss of revenue, which can lead to a reduction in the number of total nuclear power plant shutdowns (Meyer, Bond, & Ramuhalli, 2012). To this end, the development of online prognostic models estimating RUL of components can lead to more efficient maintenance scheduling, and when used for on-line monitoring, can reduce sudden loss of operations from unexpected component failure (Coble, Ramuhalli, Bond, Hines, & Upadhyaya, 2012).

Current research focuses on the development of prognostic methods and models for estimating RUL throughout the lifetime of a component. By utilizing several different methods and combining them in calculated ways, a lifecycle prognostic model for a component or system can be developed that handles the predictive maintenance from data acquisition to RUL prediction. To validate the lifecycle prognostics methods developed during this project, an accelerated degradation test bed was constructed that simulates fouling in a shell and tube style heat exchanger. A Nuclear Power Plant (NPP) contains many heat exchangers, which are crucial to the overall performance of the plant, making accurate monitoring and modeling of the RUL for these heat exchangers vitally important. Possibly the most important heat exchanger for maintenance purposes is the main turbine condenser. Failure to remove waste heat in the system by the condenser can significantly reduce plant capability to maintain vacuum, which has occurred during hot summer months at several NPPs such as Watts Bar, resulting in a derating from loss of efficiency (Buecker, 2009). Between 2008 and 2010, the North American Electric Reliability Corporation (NERC) stated that condenser associated performance issues were responsible for the removal of over 9.1 million megawatt hours from the energy grid (Fayard, 2011). This loss of energy is partly

Zachary Welz et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

from loss of efficiency in the condenser, which can reach 20% loss from typical seawater fouling (Ibrahim & Attia, 2015). In an effort to reduce the effects of this efficiency loss for NPPs, the algorithms detailed in this paper are implemented on the data collected from the small-scale heat exchanger fouling experiment onsite at the University of Tennessee. This paper presents the development of data-driven models capable of degradation detection, evaluation of system health indicators, and finally lifecycle prognostic RUL prediction.

The structure of this paper starts with a brief discussion of the background for heat exchanger fouling research and the steps necessary to develop a lifecycle prognostic model for a heat exchanger system with a short explanation of each step. Next is a description of the heat exchanger setup and operating procedure used to generate the accelerated degradation data for lifecycle prognostic model generation. This will be followed by a detailed report of the steps taken to develop the lifecycle model including signal/feature selection, auto-associative kernel regression model development, prognostic parameter generation, general path model generation, Bayesian updating application, and bootstrap aggregating (bagging) implementation. The paper ends by presenting the results of a developed lifecycle prognostic model and concluding remarks.

2. BACKGROUND

There has been much research into physics driven heat exchanger modeling such as plate heat exchangers with mathematical first principles methods (Gut & Pinto, 2003), as well as physics of failure models for heat exchanger degradation due to milk fouling (Georgiadis & Macchietto, 2000). Unlike a physical heat exchanger test bed, first principles models provide the ability to quickly generate large sample data sets with multiple failure modes. Ardsomang, Hines, and Upadhyaya (2013) utilized physics models for heat transfer and effectiveness to estimate the RUL of a developed heat exchanger first principles model. Physics based methods for detecting fouling in heat exchangers, such as Kalman filtering utilizing first principles models, are also currently used. These physics based models have been historically proven to be effective at predicting moderate fouling (Jonsson, Lalot, Palsson, & Desmet, 2007). Heat exchanger first principles modeling is presently used mainly for on-line monitoring, diagnostics and fault detection (Upadhyaya & Hines, 2004).

While there are several advantages to first principles models, a physical test bed allows for validation of the degradation models with real world signals collected from the heat exchanger. First principles models must be designed to include a robust set of different conditions and failure mechanisms, whereas with real world experimentation different natural failure mechanisms,

operations, and natural data fluctuations (process noise) are inherent to the experimental setup. An advantage of test beds is that unexpected developments in testing need not be known unlike when designing first principles models. For example, if a simulation of an induction motor system is developed to model the conditions of onset bearing failure, there may actually be several different failure modes, such as insulation degradation, shaft warping or bearing faults, which the model will not implement. Using a physical test bed reduces the concern to preemptively model every possible form of degradation and their cross-signal cumulative effects. However, physical systems do create the additional burden of obtaining high fidelity “unfaulted” baseline data with which to build comparison models. Unlike many first principle models, empirically driven models can be developed almost exclusively on historic unfaulted data. Once a robust baseline reference model has been established using unfaulted data, incoming real-time data can be passed through to these models and monitored for deviations from expected normality.

One type of empirical modeling technique to measure the deviations discussed is the Auto-Associative Kernel Regression (AAKR) (Wand & Jones, 1995). AAKR models are typically built using vector selection techniques on unfaulted data to construct a memory matrix representative of the expected operating space of the system. The AAKR model used in this study is an error correction model constructed with fault-free data and built from methods developed by Garvey and Hines (2006). AAKR was chosen because it is non-parametric, making it flexible and adaptive through additions to the memory matrix. This is important when the operating space is not known *a priori*. When faulted data is input to the model, the output is a corrected version of the faulted input data. When the corrected data is compared to the actual data, the difference between them is termed residuals. As a component degrades, the residuals are expected to increase until failure. Figure 1 shows the basic arrangement of the AAKR based prognostic system. Operational data is input and residuals are calculated. These residuals can then be combined into a prognostic parameter, which is related to the health of the system. A prognostic model is developed to predict the degradation process and the system RUL. The remaining three steps, residual calculation, prognostic parameter generation and prognostic modeling, are discussed in subsequent sections.

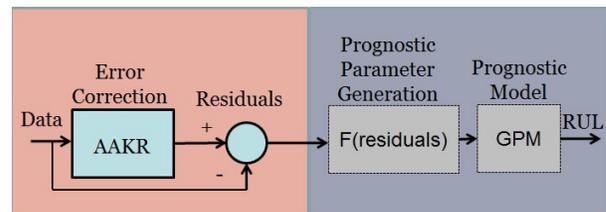


Figure 1. Basic arrangement of an AAKR based prognostic system.

Prognostic models can be classified into three types based on the category of data used in the model (Hines, Garvey, Preston, & Usynin, 2007). The first of these, Type I, or simple time-to-failure distribution models, are used to estimate the failure times of a system, generally before operation begins or if there is no information available from the query system other than failure times. Type I prognostics is very common for preventative maintenance, and often takes the form of a Weibull distribution of the failure times. Type I models can be prone to high uncertainty for systems with multiple operating conditions, or other operational variables that broaden the range of the system's failure times. In Type II prognostic models, stressor information such as the flow rates for heat exchangers can be used to improve the estimates starting at the early stages of operation when expected or continuing stress levels are known. This can be thought of like different Type I models for each operating condition. We expect to see improvements in RUL predictions when transitioning from Type I to Type II models as a result of an increase in understanding of system degradation. When quantifiable measured or inferred degradation is detected in the system, Bayesian techniques can be used to further transition to a Type III model, for more accurate RUL estimates. These Type III models use directly measured signals to track the system/component degradation and are expected to offer the highest accuracy RUL predictions. Use of the general path model for extrapolating an empirical model to failure is the most common of the Type III models (Coble, Ramuhalli, Bond, Hines, & Upadhyaya, 2012), and was the used in the development of this work's lifecycle prognostic model.

The General Path Model (GPM) was first proposed by Lu and Meeker (1993), and was initially used for prognostics by Upadhyaya, Naghedolfeizi, and Raychaudhuri (1994). The GPM is commonly used to extrapolate some measure of system health, called the prognostic parameter, built from degradation data by means of a regression fit. For prognostics, past degradation cycles can be analyzed, and an appropriate functional fit type (linear, quadratic, etc.) can be determined and then applied to an unfailed case with detectable levels of degradation. The regression model is then extrapolated to some failure threshold and the time to failure (TTF) is calculated. Using the GPM attempts to capture the overall shape of the degradation. In this experiment, the heat exchanger undergoes constant degradation. While there is some self-healing, and deviations from the path, the overall the increase in degradation is clear. In most cases, the paths have linear or near linear shapes that can be accurately modeled using quadratic fits within the GPM. This method of utilizing GPM, along with Bayesian inference, was applied to the heat exchanger test bed.

Bayesian methods for including prior information are based on Bayes' theorem and can be used for regression problems. It has been shown by Coble and Hines (2011) that Bayesian inference for application in prognostics problems can be successfully used to update GPM regression weights based on prior information. By appending weighted inputs to the data matrices, GPM regression can be purposefully biased towards historical paths or failure times. For data with high variability in the BOL, this method may drastically improve RUL predictions during this period. This method of Bayesian updating was chosen for use on the heat exchanger experiment data, which suffers from rapidly inconsistent faulted data during the BOL, and is discussed in section 4.4.

Bootstrap aggregating, also known as bagging, is a variance reduction technique that combines outputs from several predictive models and merges them into a new prediction. The most general form of this merging technique is referred to as ensemble learning (Bryll, Gutierrez-Osuna, & Quek, 2003), with bagging and boosting being the most widely used specific methods. The purpose of ensemble learning is to reduce the output variance through improved classification accuracy of the different machine learning models. Bagging is a method of ensemble learning that produces a number (n) of training sets by randomly sampling an original training data set n times. By significantly perturbing the n training sets (through random sampling) with respect to the original set, the resulting predictors can be used to improve accuracy (Breiman, 1996). Another form of ensemble learning related to bagging is subgroup bagging, or subagging. This separate method is denoted for techniques used when conducting empirical research involving bagging as in the case of decision trees. This subagging method has been shown to significantly reduce the variance of empirical data sets in a more computationally efficient way (Büchlmann & Yu, 2002). The general subagging scheme used for this research is shown in Figure 2.

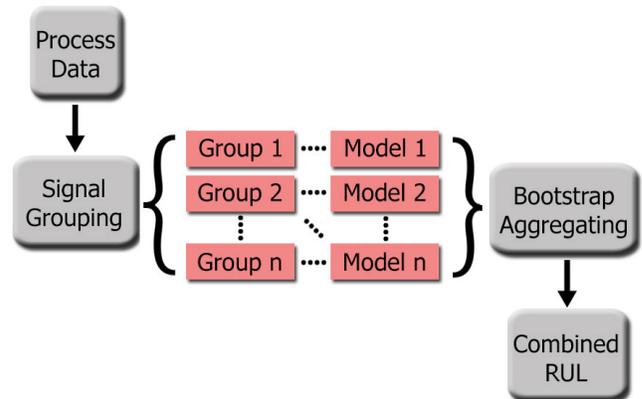


Figure 2. Bootstrap aggregation scheme using multiple subgroups (Baraldi, Gola, Zio, Roverso, & Hoffman, 2011)(Baraldi, Roozbeh, & Zio, 2010)

Figure 2 shows the path taken to implement the subbagging routine. It starts with the process data, which is then grouped into signal sets. Each of the n subgroups is a different combination of signals or process data features. The different subgroups are then used to build models (training group), and the model results are combined to reduce the variance (Büchmann, 2003). Specifics of the subbagging method used for this research are discussed in Section 4.6.

3. EXPERIMENTAL SETUP AND DATA ACQUISITION

The heat exchanger fouling test bed experiment was designed to increase the rate of fouling degradation of a shell and tube heat exchanger by accelerated process-side fouling. This structure contains 8 sensors to monitor temperature, flow, and pressure within the 64 tube cross-flow heat exchanger, shown in Figure 3 and summarized in Table A1.

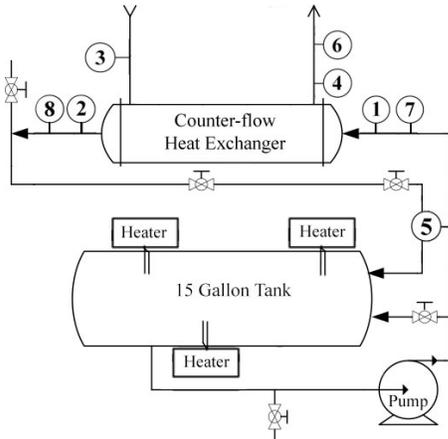


Figure 3. Schematic of heat exchanger physical setup

As seen in Figure 3, there are thermocouples at each of the four entrances and exits of the heat exchanger used to measure the incoming and outgoing temperature of the hot and cold legs: sensors 1, 2, 3, and 4, respectively. Pressure transducers are at both ends of the heat exchanger hot leg to measure the pressure variation (sensors 7 and 8). There are two turbine style flow meters to measure flow velocity of the hot and cold legs (sensors 5 and 6, respectively). A LabVIEW data acquisition (DAQ) system is used to sample and record the signals at 0.1 Hz. Three 250 watt heaters are used to heat the reservoir water for the hot leg supply, and a 0.5 horsepower (HP) pump is used to facilitate flow. The heat exchanger used for this test bed is the Basco 64 shell and tube. Each hot leg tube is 0.25 inches in diameter and 24 inches in length. A full list of system components is given in Table 1.

	Component	Brand	Location
Sensors	Thermocouple	Omega	Hot Leg Inlet Hot Leg Outlet Cold Leg Inlet Cold Leg Outlet
	Turbine Flow Meter	Blancett	Hot Leg Inlet Cold Leg Outlet
	Pressure Transducer	Dwyer	Hot Leg Inlet Hot Leg Outlet
	Data Acquisition System	Texas Instruments	N/A
Components	Heat Exchanger	Basco	N/A
	250 Watt Heater	Tempco	Two on top of tank One on bottom of tank
	15 Gallon Reservoir Tank	McMaster-Carr	Hot Leg - Below Heat Exchanger
	0.5 HP Pump	Berkeley	Below Tank

Table 1. Major systems components, brand, and location

Shell and tube heat exchanger degradation occurs most commonly as continuous fouling within the tubes, which results in a reduction in heat transfer to the point where it no longer meets specifications (Upadhyaya, 2004). For the scope of this experiment, this reduction in heat transfer is due to particulate fouling inside the process side tubes. To accelerate fouling of the test bed experiment, kaolin (china clay) is added to the hot leg water. At startup, a mixture of water and 105 grams of clay is added to the system, with additions of 75 grams of clay in solution every 48 hours during the cycle. This regular addition of clay helps to maintain a consistent clay density in solution within the system. Without these regular additions, the clay has a tendency to fall out of solution and settle in the reservoir tank. In commercial reactors, and more specifically condensers, fouling occurs gradually over time, therefore the consistent clay concentration within the heat exchanger tubes is critical to produce realistic degradation data that simulates real behavior. A typical cycle is 14 days of continuous operation at 1 gallon-per-minute in the hot and cold legs (excluding down time during clay addition).

Operational data has been collected for eight cycles run at one gallon-per-minute. For the purposes of this paper, the average flow rate can be considered a stress related variable as it is directly related to the fouling rate. The flow rate is important for the stressor-based prognostic algorithms, and in future research will be varied during a data collection cycle; for the extent of this paper, each cycle is held at near constant flow rate.

4. MODEL DEVELOPMENT

To study the difference in RUL predictions based on signal selection, multiple competing models are created which rely

on different signal sets. It is expected that there will be significant differences in each model's prediction accuracy.

4.1. Signal and Feature Sets

From the data, certain features such as log mean temperature difference (LMTD), heat rate, and delta temperatures are calculated. The two features used in the prognostic models are heat rate and overall heat transfer coefficient given by Eq. (1) and (2b) respectively.

$$\dot{Q}_{h/c} = \dot{m}C_p(T_1 - T_2) \quad (1)$$

$$\text{LMTD} = \frac{(T_{h1} - T_{c2}) - (T_{h2} - T_{c1})}{\log\left(\frac{T_{h1} - T_{c2}}{T_{h2} - T_{c1}}\right)} \quad (2a)$$

$$U_{h/c} = \frac{\dot{Q}_{h/c}}{\text{LMTD} * A} \quad (2b)$$

where A is the surface area of heat transfer.

These signals and features define the state of the system and are selected for inclusion into the AAKR models. When cleaning the training data for the AAKR model, it is important that the data is fault-free and the test cases operate in the same conditions. To reduce system noise, especially for the mass flow rates, a median filter was applied to remove outliers exceeding three standard deviations. This procedure removed many of the anomalously large spikes that had been seen in the mass flow rate signals, which would be unexpected given that the signal was in a near steady state condition.

One way to improve performance is to ensure that AAKR models are created with groups of related variables. In order to assist in creating related groupings of signals, the linear relationships between the signals and features were analyzed via correlation coefficients. As a rule of thumb, absolute coefficient values of greater than 0.7 correspond to strong correlations between signals, and coefficients of 0.25 and below are considered to show no significant linear correlation. Figure 4 shows a plot of the correlation coefficients for the raw data and calculated feature indices, where variable indices 1-8 are signals and 9-15 are features. The variable indices are summarized in Table A1 in the Appendix.

Figure 4 shows that there is a strong correlation between signal indices 1 to 4 (measured temperatures). There is also a strong correlation between signals 1 and 2 and features 13 to 15 (LMDT and heat transfer signals). There are moderate correlations between signals 1 to 6 (5 and 6 are flow rates) and 13 to 15.

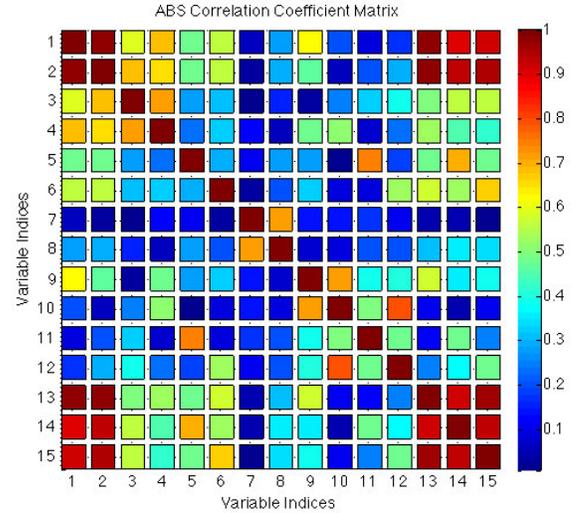


Figure 4. Correlation coefficients of signals and features

Four sets of related variables were chosen based on correlation coefficients and understanding of the system processes. Other signal sets were tested during initial modeling attempts, but did not return desirable residual values and trends, and therefore were not considered for final lifecycle prognostic methods. The selected signals and features were chosen either for being moderately-to-highly correlated to one another or for the strong trend observed in them, such as increasing trends of the hot leg temperatures and decreasing trends of the heat transfer coefficients. The indices chosen for each signal set are given in Table 2.

Signal Set	Signal/Feature Indices Used
1a	2, 3, 11, 12, 14, 15
2a	1, 2, 3, 4, 11, 12, 14, 15
3a	1, 2, 3, 4
4a	1, 2, 3, 4, 14, 15

Table 2. Signal sets used for modeling (with the label 'a' to distinguish these from future signal sets)

The flow rate was designed to remain relatively constant, and both flow signals are used in the development of the heat transfer coefficients (variable indices 14 and 15), which are included in the signal sets. Therefore, flow rates signals (variable indices 5 and 6) were not selected for explicit use in the signal sets. Similarly, the LMTD was not included in the signal sets because it was believed that the heat transfer coefficients would more accurately contain the thermal transfer information. In signal sets 1a, 2a, and 4a, the heat transfer coefficients, heat rates, and temperature signals are used. Since the overall heat transfer coefficients (indices 14-15) are calculated from first principles models that are

dependent on temperature signals (Schmidt, Henderson, & Wolgemuth, 1993), including them in an empirical AAKR model has the effect of increasing both the model's and prognostic parameter's weightings toward the temperature signals. This may improve modeling attempts when the temperature signals have strong increasing trends, and is expected to be more effective than other methods of artificially increasing the weightings, as it collapses signals to known, important dimensionalities.

4.2. Auto-Associative Kernel Regression

After feature selection is completed, the unfaulted heat exchanger data is divided into three data sets termed training, testing, and validation. Training data is used to train the model through historic memory vector selection and should consist of unfaulted data that covers the range of operating values. Testing data is used for bandwidth optimization, and validation data is used to quantify the model performance. AAKR models for the heat exchanger were developed and evaluated with the PEM toolbox (Hines & Garvey, 2006). Kernel regression requires a parametric kernel function, in this case a Gaussian function, defined by a bandwidth that specifies the region of localized weighting for an input vector to the memory matrix output. An optimal bandwidth can be selected by altering it to minimize the error between known unfaulted observations and the model output. This method of determining the bandwidth increases the accuracy of the kernel regression model (Wand, 1995). Example training residuals from an AAKR model of signal set 2a are shown in Figure 5.

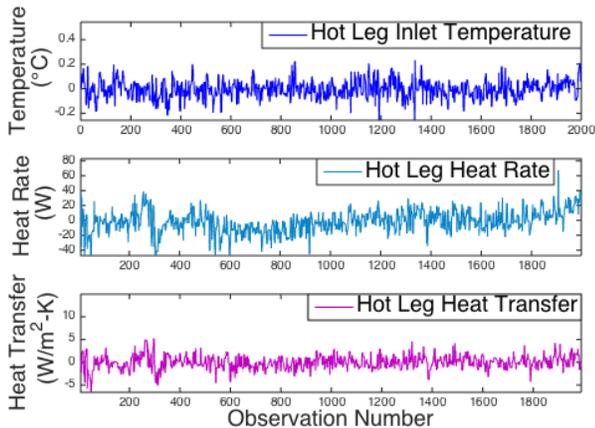


Figure 5. Example training residuals for signal set 2a.

Only one residual for each signal type is shown to improve visibility. For this experiment, the training residuals of the temperature signals are desired to be less than 1°C since the temperature signals change less than 10°C over the faulted range. The training residuals of the heat rate should optimally be less than 50 W , and the heat transfer coefficient residuals should be less than $10\text{ W/m}^2\text{K}$. These levels were chosen based off knowledge of signal and feature operating

ranges over normal cycles as well as the expected signal-to-noise ratios for the heat exchanger system. The expected signal deviation due to normal process and signal noise is determined by visual inspection. After the model is built, faulted data is passed through and residuals for each faulted cycle are calculated. An example of faulted residuals for the temperature sensors in signal set 2a is plotted in Figure 6.

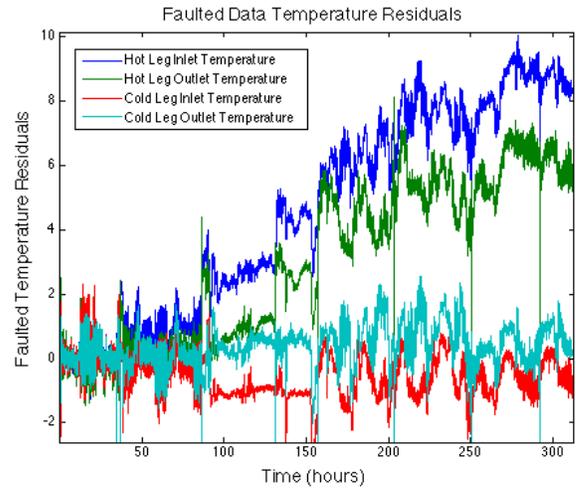


Figure 6. Faulted residuals of temperature signals (indices 1-4) using the signal set 2a model

From the failure residuals shown, strong increasing trends can be seen for the hot leg temperature signals. Dominantly monotonic trends are important when combining residuals to make a prognostic parameter. When combining the residuals, the objective is for the resulting health indicator to increase or decrease over the lifecycle to help indicate the degree of system or component degradation. If the observed trends of the residuals show a strong increasing/decreasing trend then the resulting prognostic parameter will also have a strong trend and be more useful for RUL predictions.

4.3. Prognostic Parameter Generation

A prognostic parameter is a metric that signifies the amount of deviation from normal behavior of the system, and is ideally linked to the overall health of the system. In this project, it is calculated as a linear combination of the residuals from the AAKR model. While the genetic algorithm was used to find a linear combination of weights for the residuals was investigated, it was found to be too computationally expensive. Instead, an ordinary least squares (OLS) regression is applied that mimics the optimization and is less computationally intense for smaller data sets. The model residuals of multiple runs to failure are collected into a single matrix by concatenating each test case. This creates an $\mathbf{n} \times \mathbf{s}$ matrix, \mathbf{X} , where \mathbf{n} is total data points in all test cases, and \mathbf{s} is the number of signal in the model. This \mathbf{X} matrix is regressed against an $\mathbf{n} \times \mathbf{1}$ vector “ \mathbf{y} ” where each y_i corresponds to the percent of the total

unit life at that observation. This means that the residuals of each test case are fitted to a linear curve from 0 to 1. The linear weights are then

$$\hat{\beta} = (X^T X)^{-1} X^T y \quad (3)$$

where $\hat{\beta}$ is an $s \times 1$ vector.

4.4. General Path Model and Bayesian Updating

When using the GPM approach, a parametric function is fit to the degradation parameter, and extrapolated until it crosses a predefined failure threshold. Typically, the failure threshold is based on historical failures but need not directly indicate a point of catastrophic failure. The failure threshold can be set as any point where a system no longer conforms to the necessary specifications and demands placed upon it.

Because of the limited number of test cases, the GPM and resulting predictions are created by the use of a Leave One Out Cross Validation (LOOCV) technique. Hence, when creating a model to calculate the RUL of a specific case, the information from that case is removed and information from every other case is used to build the model. This avoids invalidating a model by keeping training and validation data separate, yet general enough to compare over all cases. With more data, an alternative approach could be to simply divide the cases in half and build one model.

Individual failure thresholds for the LOOCV method were generated with respect to each of their corresponding data subsets. The values were chosen as a reflection of an unacceptable amount of degradation, limited by the least degraded cycle for any given model. Any data collected after this point was considered past the point of failure and was removed from the data analysis. A histogram plot of failure times for the lifecycle prognostic models is shown in Figure 7.

As stated above, each time a prediction is made, the run being predicted is the one left out of the model development. Therefore, in order to make a prediction for 'n' number of runs, the LOOCV method must be iterated 'n' times.

Once the failure threshold from the LOOCV subset is selected, the RUL estimations for the target case can be calculated at each time step by extrapolating the current path to that threshold. To do this, a suitable parametric fit must be chosen. With the OLS method, the fit can be of any linearly separable form such as linear, quadratic, exponential, etc.

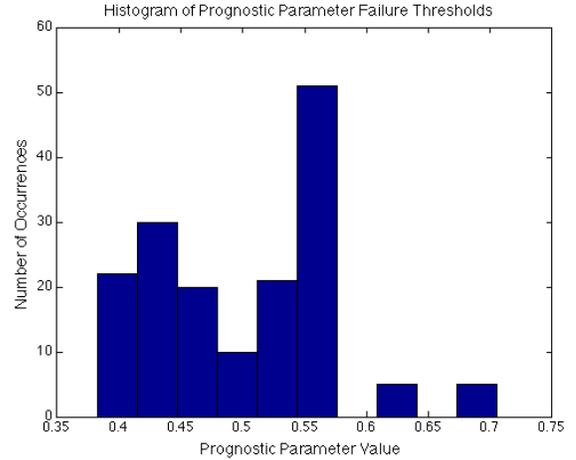


Figure 7. Histogram of failure thresholds

The OLS method is used for regression of the parametric fittings because the OLS regression on a joint Gaussian distribution of parameters gives the maximum likelihood estimate. This method assumes that the error is normally distributed around zero. The OLS solution can be found using the pseudo-inverse given in equation 3. By adjusting the functions in the columns of the input matrix \mathbf{X} , parametric fits can be applied to the test paths. It is assumed that for a certain failure mode the degradation paths will follow similar fits. Therefore, once a suitable fit is chosen for the failed data, it is assumed the censored faulted data will follow the same fit.

Bayesian priors can also be incorporated into the OLS model (Gelman, Carlin, Stern, & Rubin, 2004) to reduce the uncertainty and increase the stability of RUL estimates. Bayesian statistics combines prior distributions with sampling data to create a posterior distribution. When few data points are available, without incorporating any form of Bayesian prior estimations, the model can easily be affected by noise and give widely varying predictions of Time To Failure (TTF). Coble and Hines (2011) use Bayesian methods to incorporate prior knowledge of regression parameters in the GPM. This approach requires historical run-to-failure data in order to evaluate the prior distributions of regression parameters. An alternative approach instead uses RUL estimates from Type I prognostic models as prior information (Nam, Sharp, Hines, & Upadhyaya, 2013). In this approach, the Type I RUL distribution is treated as an additional data point in the OLS regression. The measured data is augmented with the distribution according to Eq. (4):

$$Y = \begin{bmatrix} y \\ thresh \end{bmatrix}, X = \begin{bmatrix} X \\ MTTF \end{bmatrix}, \Sigma = \begin{bmatrix} \Sigma_y & 0 \\ 0 & \Sigma_{RUL} \end{bmatrix} \quad (4)$$

where y is the observed prognostic parameter, $thresh$ is the failure threshold, x is the timestamps (or appropriate transformation thereof), $MTTF$ is mean failure time from the

Type I distribution (or appropriate transformation thereof), Σ_y is the noise or uncertainty associated with the observed prognostic parameter, and Σ_{RUL} is the uncertainty in the Type I RUL estimate. The OLS regression is then solved according to Eq. (5)-(7):

$$\hat{\beta} = (X^T \Sigma^{-1} X)^{-1} X^T \Sigma^{-1} y \quad (5)$$

$$V \sigma^2 = (X^T \Sigma^{-1} X)^{-1} \quad (6)$$

$$\sigma^2 = \frac{1}{n-k} (y - X\hat{\beta})^T \Sigma^{-1} (y - X\hat{\beta}) \quad (7)$$

where k is the degree of the parametric function used in the GPM.

The weight of the prior information in the OLS regression depends on two main factors: the uncertainty of the prior relative to the uncertainty of the measurements, and the number of observations collected (Ghosh, Delempady, & Samanta, 2006). If the uncertainty of the prior is small compared to the uncertainty of the data, the prior β_0 will be weighed more heavily. However, no matter the difference in uncertainties, with enough observations, the data should eventually swamp out the prior in calculating the posterior.

4.5 Bayes Method Implementation

For each of the four AAKR models, two prognostic modeling methods are used:

GPM Method 1: No Bayesian updating

GPM Method 2: Type I Bayes priors

To compare the two methods, plots of the predicted TTF versus the actual TTF are examined. In each plot, the multiple blue lines correspond to the determined TTF of each cycle over time. Figure 8 shows a plot of the TTF comparison when no Bayesian updating is used.

Without Bayesian updating, TTF prediction times have large spikes, and prediction accuracy is reduced. While some peaks are due to the noise and artifacts in the heat exchanger data acquisition system, the somewhat larger and broader peaks at regular intervals are most likely the result of the regular additions of clay into the hot fluid. The extra clay would change the thermodynamics as well as mass flows of the otherwise closed system. In an attempt to improve TTF estimation, past cycle failure times are incorporated as prior information (Type I) as shown in Figure 9. Using Eq. (8), the variance of the regression model can be quantified. This variance decreases due to the inclusion of the prior. Using Monte Carlo sampling from the regression coefficient distribution, probability distributions of the RUL can be made. These distributions will have a smaller confidence interval and standard deviation, due to

the decreased regression variances. The result is a decrease in prediction uncertainty when Bayesian priors are included.

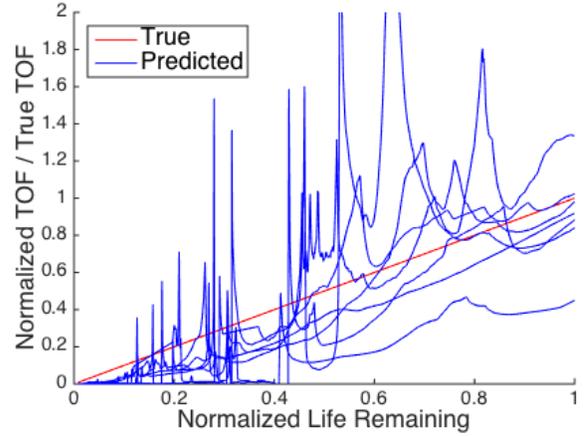


Figure 8. Plot of the GPM method 1 TTF predictions across cycles without Bayesian updating

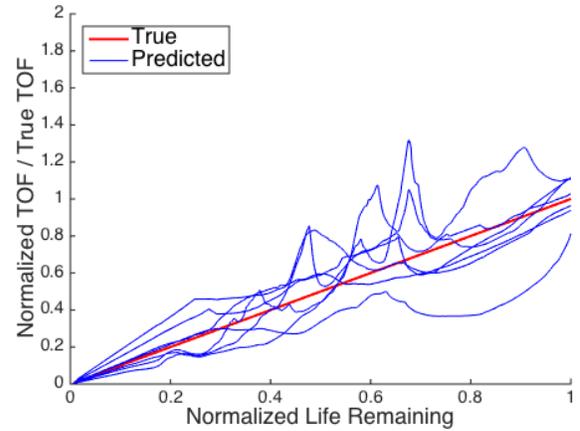


Figure 9. Plot of the GPM method 2 TTF predictions across cycles with Type I Bayesian updating

The predictions using Type I prior information show visual improvement over those with no Bayesian updating. Improvements in model uncertainty will be discussed in following sections.

4.6 Bootstrap Aggregating (Bagging) Implementation

As discussed previously, bagging is a method of ensemble learning that utilized bootstrapped training sets to develop perturbed models, which can then be combined to facilitate variance reduction. The general form of bagging involves randomly sampled training sets to form the predictors. For the purpose of this lifecycle model, subgroups of signal sets based on understanding of the heat exchanger system are chosen rather than randomly sampled. These subgroups are similar to those in Table 2, but were improved upon to illustrate the effectiveness of bagging while retaining the

system information from the original 4 subgroups. The five signal set subgroups for the heat exchanger bagging models are given in Table 3.

Signal Set Subgroup	Signal/Feature Indices Used
1b	1, 2, 3, 4, 13
2b	1, 2, 3, 4, 11, 12
3b	1, 2, 3, 4, 14, 15
4b	1, 2, 3, 4, 11, 12, 14, 15
5b	2, 3, 11, 12, 14, 15

Table 3. Subgroup signal set combinations for bagging (with the label 'b' to distinguish these from the subgroups listed in Table 2)

The signal groups contain all four temperature signals, the log mean temperature difference, and the overall heat transfer coefficients. After the subgroups are created, they are used to train kernel regression models as shown earlier in Figure 1 using the LOOCV. Once the residuals from the AAKR model are calculated, they are combined into a prognostic parameter using OLS. This means that each sub-data set will have a unique prognostic parameter result. Finally, the prognostic parameters are passed to the GPM and used to create RUL estimations. After the RULs are calculated for each observation in the validation data set, they are aggregated for each observation across the signal subsets. The result is a single RUL prediction for each observation.

After this initial bagging method was tested, alternative bagging methods were also investigated on the heat exchanger data set. One such method involved directly aggregating the residuals from the individual AAKR models to create a singular prognostic parameter. This was then passed to the standard GPM algorithm to determine the RUL. This method resulted in similar results to those for the RUL bagging method but was not included due to ineffective variance reduction. Additionally, subbagging was conducted both with and without Bayesian updating. It was found that there was no benefit to conducting bagging without the Bayesian updating discussed earlier. Therefore, the RUL prediction results will only be given for Bayesian updating with subbagging and Bayesian updating without subbagging.

It is expected that the RUL bagging will significantly reduce the variance in the Middle Of Life (MOL) for the heat exchanger. RUL bagging is expected to reduce the variance in the BOL, but incorporating the Bayes Type I priors into the model during testing reduced the subbagging effect on the

BOL RUL estimates. We do not expect the subbagging technique to effect the EOL predictions because in this period of life there is significantly less variance in the RUL predictions. To provide an easy to interpret output, each cycle will be divided into BOL, MOL and EOL to compare the RUL error with and without bagging. Each period of life represents a third of the cycle length. The error reported for the case without bagging is the average total error of the best performing individual model as a conservative comparison.

5. RESULTS AND DISCUSSION

While it was previously discussed that both linear and quadratic fits were used for the GPMs, initial modeling attempts revealed that using a quadratic parametric fit is more accurate than using a linear parametric fit; therefore results will be confined to quadratic fit models. The different GPM methods and signal sets (models) are compared using several performance metrics.

The first model comparison metric used is the absolute error mean (AEM), which returns the average absolute difference between the predicted RUL and the true RUL in real time units, shown in Figure 10, where GPM 1 and GPM 2 correspond to the models without and with Bayesian updating respectively. Signal sets 1a and 3a have the lowest AEM, and GPM method 2 further improves the predictions. Signal set 1a with GPM method 2 results in the most accurate RUL predictions for this data set.

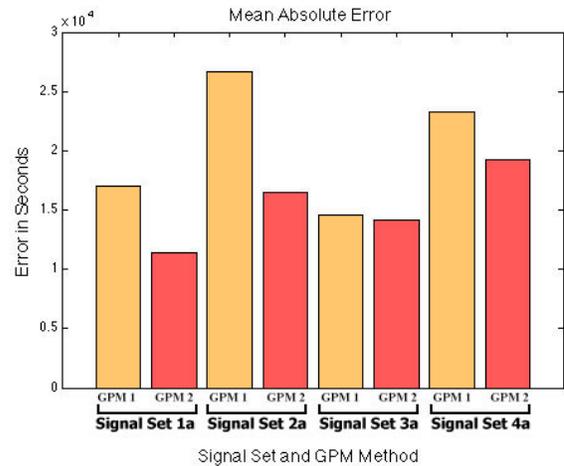


Figure 10. Absolute error mean for four signal set (1a-4a) models and two GPM methods using LOOCV

The second metric used to evaluate the prognostic models is the absolute error standard deviation (AES), which is a measure of the variation in error through time of each model and GPM method, shown in Figure 11. Again, the model built with signal set 1a and GPM method 2 shows the best performance, with highest precision in estimating the RUL.

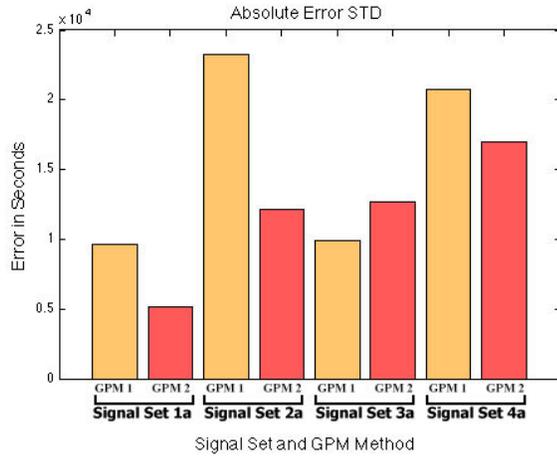


Figure 11. Absolute error standard deviation for four signal set (1a-4a) models and two GPM methods using LOOCV

To quantitatively compare the different GPM methods, the AEM, AES, spread, and coverage metrics are used (Sharp, 2013)(Saxena, Celaya, Saha, Saha, & Goebel, 2010). A plot showing the results of these metrics for each GPM method for signal set 1a is shown in Figure 12 and the un-normalized metric scores are provided in Table 4.

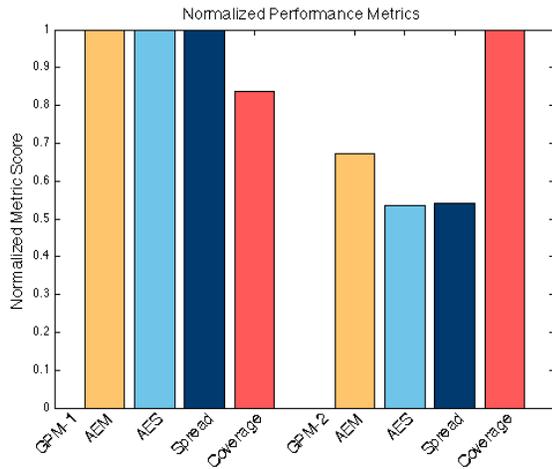


Figure 12. Plot of normalized performance metrics for two GPM methods and signal set 1 using LOOCV

The figure above compares the AEM, AES, spread, and coverage of the two models. The spread is a measure of the prediction uncertainty, weighted at each percent of life. The weight is based on a Gaussian kernel centered on the EOL with a bandwidth of half the total life, giving importance to EOL predictions. The associated confidence intervals, in this case 95%, are taken from distributions of prediction errors. The overall metric is the difference between the upper and lower error intervals, and shows how widely the accuracy can vary. The coverage is the fraction of binned true RUL that fall within the confidence interval. A coverage of 100% shows the model is reasonably accurate,

dependent on the size of the confidence interval, over the entire life. By incorporating Bayesian updating, the spread of the model, which is a measure of variability, has decreased. This is a positive result that indicates that the uncertainty in our model predictions has been reduced. Together, these metrics indicate that the Bayesian updating method (GPM Method 2) is more accurate than vanilla GPM for predicting RUL for this data set.

GPM-1	AEM	1.70E4
	AES	9.62E3
	Spread	131
	Coverage	83.0
GPM-2	AEM	1.14E4
	AES	5.14E3
	Spread	70.8
	Coverage	99.0

Table 4. Performance Metrics Scores

The comparison of model uncertainty between GPM without and with incorporating Bayes priors (GPM-1 and GPM-2 respectively) is also of concern. The following table shows the standard deviation (STD) of the quadratic regression parameters a, b, and c for each period of cycle life (BOL, MOL, and EOL).

Period	Model	a-STD	b-STD	c-STD
BOL	GPM-1	7.30E-10	2.94E-06	2.48E-03
	GPM-2	2.04E-11	7.62E-07	3.21E-10
MOL	GPM-1	8.36E-11	1.25E-06	2.59E-10
	GPM-2	4.36E-11	7.68E-07	1.04E-10
EOL	GPM-1	1.93E-11	5.21E-07	5.98E-11
	GPM-2	1.88E-11	5.14E-07	5.74E-11

Table 5. Quantified model uncertainties

This data shows that the uncertainty in the BOL has the highest reduction between the periods of life, which is what was expected based on the concept of incorporating the Bayes priors. Another useful thing to note is that the use of the prior information does not increase uncertainty in the MOL and EOL periods. The results show that it is appropriate to apply Bayes priors for similar applications across the operating life.

To determine the extent to which subbagging improves the RUL predictions, the results in Figure 13 are given as average total percent error for the three periods of life: beginning, middle, and end of life.

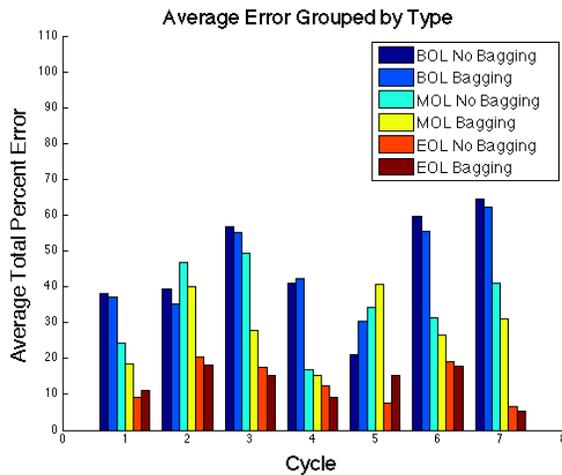


Figure 13. Total percent error for each method and time of life grouped by cycle using LOOCV

From the figure it can be seen that the average total error is reduced for 5 of the 7 cycles in the BOL, 6 of the 7 in the MOL, and 5 of the 7 in the EOL. While the reductions and increases for the BOL and EOL are relatively small, there is an error reduction of up to ~22% in the MOL when subbagging is conducted compared to the best individual model without bagging. The one cycle with increased error also showed a uniquely low signal variance across perturbed subsets before subbagging. This resulted in an increase in the total error, which is believed to have resulted from model error when conducting the AAKR model and GPM.

6. CONCLUSION AND FUTURE WORK

In this paper, a method of lifecycle prognostic modeling is presented and applied on data collected from a heat exchanger test bed setup at the University of Tennessee. This method relies upon predicting the progression of system residuals and how they relate to the overall system health to create estimations of the system's Remaining Useful Life (RUL). Due to restrictions on available data, a Leave One Out Cross Validation (LOOCV) method was used to evaluate and validate the effectiveness of the modeling techniques. Several methods to reduce the variance and improve the accuracy of the resulting RUL estimations were explored and analyzed within this work. The results of this analysis show that across all test cases the Bayesian transition using Type I priors outperformed the GPM with no Bayesian updating, and resulted in up to a 99% reduction in regression parameter standard deviation. From the results of incorporating subgroup bootstrap aggregating (subbagging), it is concluded that subbagging has

a positive impact on error reduction for Middle Of Life (MOL) RUL predictions, and should be used in tandem with the Bayesian techniques discussed, which drastically reduce the beginning of life average total error.

The authors have identified areas of potential future work to improve the prognostics method presented here in several ways. First, it is recommended that investigations into optimally reducing the noise of the prognostics parameter through improved filtering be pursued for prognostics parameter optimization. Second, with the addition of more sample cases, a more robust prognostics parameter with a well-defined degradation threshold could be created, both increasing the prognosability and decreasing the End Of Life (EOL) RUL and Time-To-Failure (TTF) prediction errors. Third, a crucial future implementation is the application of a fault detection method to remove beginning of life test data before a fault is detectable from any GPM style extrapolation model. Trying to trend and extrapolate a progressive degradation parameter prior to any detectible degradation will cause poor predictions and undue errors in the overall model. This was left out of this particular investigation due to the fact that all test cases are seeded with clay in the system from the initial data collection time.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the DOE Nuclear Engineering University Program (NEUP) under prime contract number DE-AC07-05ID14517 and Battelle Energy Alliance LLC Standard Contract No. 00118294, for funding.

The authors would also like to acknowledge the Lloyds Register Foundation and the International Joint Research Center for the Safety of Nuclear Energy for partial funding of this research. Lloyd's Register Foundation helps to protect life and property by supporting engineering-related education, public engagement, and the application of research.

REFERENCES

- Ardson, T., Hines, J.W., & Upadhyaya, B.R. (2013). Heat exchanger fouling and estimation of remaining useful life, Annual Conference of Prognostics and Health Management Society, September 23-27, Minneapolis, MN.
- Baraldi, P., Gola, G., Zio, E., Roverso, D., & Hoffman, M. (2011). A Randomized Model Ensemble Approach for Reconstructing Signals from Faulty Sensors, Expert Systems with Applications, vol. 38, Iss. 8, pp. 9211-9224.
- Baraldi, P., Roostbe, R.F., & Zio, E. (2010). Classifier-ensemble incremental-learning procedure for nuclear transient identification at different operational conditions, Reliability Engineering and System Safety, Elsevier, 2011, 96 (4), pp. 480-488.

- Breiman, L. (1996). Bagging Predictors, *Machine Learning*, vol. 24, pp. 123-140.
- Bryll, R., Gutierrez-Osuna, R., & Quek, F. (2003). Attribute bagging: Improving accuracy of classifier ensembles by using random feature subsets, *Pattern Recognition*, vol. 36, Elsevier Ltd., pp. 1291-1302.
- Büchlmann, P. (2003). Bagging, subbagging and bragging for improving some prediction algorithms, Seminar für Statistik, ETH Zürich, Switzerland.
- Büchlmann, P., & Yu, B. (2002). Analyzing Bagging, Published within *The Annals of Statistics*, vol. 30, no. 4, pp. 927-961, Institute of Mathematical Statistics, JSTOR.
- Buecker, B. (2009). Save big bucks with proper condenser performance monitoring, *Energy Tech Magazine*, April Issue.
- Coble, J.B., & Hines, J.W. (2011). Applying the General Path Model to Estimation of Remaining Useful Life, *International Journal of Prognostics and Health Management* vol. 2, pp. 71-82.
- Coble, J.B., Ramuhalli, P., Bond, L.J., Hines, J.W., & Upadhyaya, B.R. (2012). Prognostics and health management in nuclear power plants: A review of technologies and applications, Pacific Northwest National Laboratory Report, PNNL-21515.
- Fayard, E.C. (2011). Improving condenser reliability and availability through effective offline cleaning and nondestructive testing, *Proceedings of Electric Power Research Condenser Technology Conference*, August 3-4, Chicago, IL.
- Garvey, D., & Hines, J.W. (2006). Development and Application of Fault Detectability Performance Metrics for Instrument Calibration Verification and Anomaly Detection, *Journal of Pattern Recognition Research (JPRR)*, vol. 1, pp. 2-15.
- Gelman, A., Carlin J., Stern, H., & Rubin, D. (2004). *Bayesian Data Analysis*. vol. 2, Chapman & Hall/CRC, Boca Raton, USA.
- Ghosh, J.K., Delampady, M., & Samanta, T. (2006), *An Introduction to Bayesian Analysis: Theory and Methods* 1st ed. Springer Science & Business Media, New York, USA, pp. 29-59.
- Georgiadis, M. C., & Macchietto, S. (2000). Dynamic Modelling and Simulation of Plate Heat Exchangers Under Milk Fouling, *Chemical Engineering Science*, vol. 55, no. 9, pp. 1605-1619.
- Gut, J., & Pinto, J. (2003). Modeling of Plate Heat Exchangers with Generalized Configurations, *International Journal of Heat and Mass Transfer*, vol. 46, Iss. 14, pp. 2571-2585.
- Hines, J.W., & Garvey, D. (2006). Traditional and robust vector selection methods for use with similarity based models, 5th International Topical Meeting on Nuclear Plant Instrumentation, Control and Human Machine Interface Technology (NPIC&HMIT), November 12-16, Albuquerque, NM.
- Hines, J.W., Garvey, D., Preston, J., & Usynin, A. (2007). Empirical methods for process and equipment prognostics, Tutorial presented at the IEEE Reliability and Maintainability Symposium (RAMS).
- Hines, J.W., & Garvey, D. (2006). Process and equipment monitoring toolbox tutorial, Nuclear Engineering Department, University of Tennessee.
- Ibrahim, S., & Attia, S. (2015). The Influence of Condenser Cooling Seawater Fouling on the Thermal Performance of a Nuclear Power Plant, *Annals of Nuclear Energy*, vol. 76, pp. 421-430.
- Jonsonn, G.R., Lalot, S., Palsson, O.P., & Desmet, B. (2007). Use of Extended Kalman Filtering in Detecting Fouling in Heat Exchangers, In the *International Journal of Heat and Mass Transfer*, vol. 50, iss. 13-14, pp. 2643-2655.
- Lu, C.J., & Meeker, W.Q. (1993). Using degradation measures to estimate a time-to-failure distribution, *Technometrics*, vol. 35, no. 2, pp. 161-174.
- Meyer, R.M., Bond, L.J., & Ramuhalli, P. (2012). Online Condition Monitoring to Enable Extended Operation of Nuclear Power Plants, *International Journal of Nuclear Safety and Simulation*, vol. 3, no. 1, pp. 31-50.
- Nam, A., Sharp, M., Hines, J.W., & Upadhyaya, B. (2013). Lifecycle Prognostic Algorithm Development and Application to Test Beds, *Chemical Engineering Transactions*, vol. 33, pp. 901-906.
- Saxena, A., Celaya, J., Saha, B., Saha, S., & Goebel, K. (2010). Metrics for Offline Evaluation of Prognostics Performance, *International Journal of Prognostics and Health Management*, vol. 1, pp. 4.
- Sharp, M. (2013). Simple metrics for evaluating and conveying prognostic model performance to users with varied backgrounds, *Annual Conference of the Prognostics and Health Management Society*. October 14-17, New Orleans, LA.
- Schmidt, F.W., Henderson, R.E., & Wolgemuth, C.H. (1993). *Introduction to Thermal Sciences: Thermodynamics*, Fluid Dynamics, Heat Transfer. Canada: John Wiley & Sons, Inc.
- Upadhyaya, B.R., Naghedolfeizi, M., & Raychaudhuri, B. (1994). Residual Life Estimation of Plant Components, *P/PM Technology*, vol. 7, no. 3, pp. 22-29.
- Upadhyaya, B.R., Hines, J.W. (2004). On-line monitoring and diagnostics of the integrity of nuclear plant steam generators and heat exchangers, Final Report: vol. 1, Experimental and Hybrid Modeling Approach for Monitoring Heat Exchanger System Performance, prepared for the DOE-NEER Program by the University of Tennessee, Knoxville, Report No. DE-FG07-01ID14114/UTNE-07.
- Wand, W.P., & Jones, M.C. (1995). Kernel smoothing, London: Chapman & Hall.

BIOGRAPHIES

Zachary A. Welz is currently a graduate research assistant at the University of Tennessee in Knoxville, TN USA pursuing his Ph.D. in nuclear engineering. His research is focused on prognostics applications, such as algorithm development. He is a member of the American Nuclear Society and holds the student executive committee position for the ANS Human Factors Instrumentation and Controls Division.

Alan Y. Nam was born and raised in Los Angeles, Ca. With a B.S. in Chemical Engineering from the University of California - San Diego and an M.S. in Nuclear Engineering from UT, he is currently pursuing his Ph.D. in the same department.

Dr. Michael E. Sharp is an Assistant Research Professor at the University of Tennessee Nuclear Engineering Department. He received his PhD in nuclear engineering from the University of Tennessee in 2012. Research interests include empirical modeling, system signal processing, optimization methods, and automated pattern matching techniques. Current research includes optimization of generic prognostic extrapolation techniques and information merging.

Dr. J. Wesley Hines is Professor and Head of Nuclear Engineering at the University of Tennessee (UT). He received the B.S. degree in Electrical Engineering from Ohio University and was a nuclear qualified submarine officer in the Navy. He has received an MBA, an M.S. in Nuclear Engineering, and a Ph.D. in Nuclear Engineering from The Ohio State University. Dr. Hines teaches and conducts research in process diagnostics, prognostics and condition-based maintenance. His most recent research projects explore the use of advanced technologies for sensor calibration monitoring, plant fault detection, and plant performance optimization for nuclear, fossil fuel, and chemical process facilities.

Dr. Belle R. Upadhyaya is a Professor of Nuclear Engineering at the University of Tennessee, Knoxville. He received the Ph.D. degree in Engineering Sciences from the University of California at San Diego. He is a registered professional engineer in the State of Tennessee. Dr. Upadhyaya's teaching and research areas include nuclear plant dynamics, instrumentation & controls, monitoring & diagnostics, modern digital signal processing, small modular reactors, reliability and maintainability engineering, and next generation reactors.

APPENDIX

Signal Index	Signal/Feature
1	Hot Leg Inlet Temperature
2	Hot Leg Outlet Temperature
3	Cold Leg Inlet Temperature
4	Cold Leg Outlet Temperature
5	Hot Leg Flow Rate
6	Cold Leg Flow Rate
7	Hot Leg Inlet Pressure
8	Hot Leg Outlet Pressure
9	Delta Hot Leg Temperature
10	Delta Cold Leg Temperature
11	Hot Leg Heat Rate
12	Cold Leg Heat Rate
13	Log Mean Temperature Difference
14	Hot Leg Overall Heat Transfer Coefficient
15	Cold Leg Overall Heat Transfer Coefficient

Table A1. Measured signals and calculated features and their indices