Data-driven Application of PHM to Asset Strategies

Sarah Lukens[1] and Matt Markham[2]

[1,2]*GE Digital, San Ramon, CA, 24013, USA*
sarah.lukens@ge.com
matt.markham@ge.com

ABSTRACT

There are many benefits from implementing a prognostics and health management (PHM) initiative in an industrial facility, such as realizing potentials from reducing unplanned downtime and increased asset efficiency. Many industrial companies would like to take advantage of PHM technologies and algorithms to meet their business objectives, but identifying how to get started can be a daunting challenge. The classical approach is to begin with a Reliability Centered Maintenance (RCM) program supported by failure modes and effects analysis (FMEA) where all possible failure modes, their risks, and mitigating actions are evaluated in the context of asset function. In this framework, application of PHM technologies is viewed as a maintenance strategy effective at mitigating certain failure modes in specific cases that are both feasible and cost-effective. However, there are many challenges and limitations to traditional RCM where data-driven analytics embedded in these work processes can help overcome and/or automate. On the other hand, the use of data-driven approaches introduces new challenges surrounding available data, data quality, and identifying numerical methods that are scalable across large datasets. In this paper, we present a case study applied to historical maintenance data for identifying and prioritizing where to start a PHM initiative, and discuss the work processes and various challenges encountered when embedding data analytics in classical reliability approaches.

1. INTRODUCTION

Prognostics and Health Management (PHM) is a family of industrial work processes which can be viewed under a larger umbrella of Asset Performance Management (APM) work processes. The goals of both PHM and APM initiatives are to satisfy the business objectives of an organization, whether it be increasing profit through reduced spending or increased efficiency, demonstrating safety and compliance, or improving quality of services or goods produced. PHM work

processes are specifically geared towards using asset information and technologies for diagnosis, prognosis, and health management of assets (Rajamani & Bird, 2016).

Reliability-centered maintenance (RCM) is a process for developing an efficient and effective maintenance plan for an asset to ensure that it is able to provide its intended functions in their intended operating contexts by systematically identifying and mitigating risks (Gulati, 2009) (Moubray, 1997) (Nowlan & Heap, 1978). An asset strategy is the collection of all intentional actions taken to mitigate known risks such as preventive maintenance, predictive maintenance, inspections, condition monitoring etc.. An optimal maintenance strategy is one that minimizes the total maintenance expenditure while minimizing risks (Casto, 2010) (Whitt, 2009).

RCM approaches are accompanied by Failure modes and effects analysis (FMEA), which is a standard methodology for identifying and assessing risks and actions that could eliminate or reduce the likelihood of risks, which is used for prioritization towards where to focus on a maintenance strategy. A health monitoring process begins with gathering information on a product's failure mechanisms, modes, environmental conditions, and performance parameters that can be monitored (Kumar, Galar, Parida, Stenström & Berges, 2013). Using this approach, an engineer can identify where to implement a PHM work process to monitor the health and condition of an asset with respect to the asset's intended function.

Systematic processes for identifying and specifying requirements during the design and implementation phase of a PHM initiative have been developed (Goebel, Daigle, Saxena, Sankararaman, Roychoudhury & Celaya, 2017) (Saxena, Celaya, Saha, Saha, & Goebel, 2010). The purpose of these processes are to ensure that a PHM initiative will be beneficial with respect to the goals of an organization. Early key requirements are to identify stakeholders and needs, define the scope, identify how the users will interact with the data and information, and to define data inventory. Once these factors have been identified, methods for prioritization

have been suggested to identify critical components as candidates for PHM strategies (Lee, Wu, Zhao, Ghaffari, Liao, & Siegel, 2014) (Lee, Liao, Lapira, Ni, & Li, 2009).

Risk-based processes for determining appropriate usage of PHM are important. Adding continuous real-time monitoring to all assets is expensive, data management can be overwhelming, and condition monitoring strategies are only significantly effective for monitoring failure modes that can be detected early enough that action can be taken to significantly reduce consequences of failure. Different failure modes have different dynamics, and the time scale and failure patterns of a particular failure mode may dictate which PHM technology to use and what assets to use it on. For instance, for failure modes where degradation may occur over several years on medium or low critical assets, it may be more cost-effective to monitor asset health through manually collected measurements such as visual inspections or route-based spot readings collected with portable instruments. It has been estimated that in practice, PHM strategies are only technically feasible for at most 20% of observed failure modes, and in less than half of those cases it may not be beneficial from a business perspective (Moubray, 1997).

However, when used appropriately, the benefits of a PHM initiative have been proven to be effective. Monitoring the condition of an asset continuously with the goal of early failure detection enables the maintenance engineer sufficient time to plan and make repairs with minimal disruption to the asset, operations, and safety. For example, past studies have shown that an appropriately implemented CBM program have provided an savings average of 10% (7-15%) over a maintenance strategy employing only preventative maintenance tasks (Gulati, 2009).

Due to these benefits, many industrial companies would like to take advantage of PHM technologies and algorithms to meet their business objectives, but are faced with the challenge of how to get started. Theoretically, implementation of RCM to establish optimal maintenance strategies, applying PHM technologies, and leveraging asset data and information in order to satisfy business objectives works perfectly, however, there are many challenges when implemented in practice. Challenges include how to get started and prioritize an initiative, and how to measure what approaches are effective. Fortunately, there are many opportunities to address many of these challenges through embedding data-driven analytics in workflows towards identification prioritization. Automating analytics enhanced workflows can use different data sources and provide information to the user which may quantify certain challenging areas to quantify. However, when tackling data challenges across broad datasets, many new challenges arise around data quality problems, noise, and numerical methods that are robust enough to automate an analysis across a broad dataset.

This paper walks through the steps with a data-driven use case towards prioritizing and getting started with a PHM initiative through an RCM basis for choosing where to apply PHM technology. We focus on a particular example of looking at different age-reliability characteristics to determine candidates for condition and health monitoring based on identifying observed failures with failure patterns physically possible for applying prognostics models. We discuss the steps we took and the challenges we discovered in generalizing a classical approach to a large dataset. This approach can be used to identify how and where to get started in a PHM initiative with respect to business opportunity.

The rest of the article is organized as follows. Section 2 summarizes common industrial data sources as well as reviews traditional RCM approaches and age-reliability analyses and challenges. Section 3 details the numerical approaches proposed for applying an RCM approach to large datasets and numerical challenges. Section 4 presents a case study illustrating the methodology. The paper ends with concluding discussions and suggests future research directions.

2. BACKGROUND

In this section, we review common industrial data sources as well as traditional RCM approaches for determining maintenance strategies and a review of challenges.

2.1. Maintenance data sources and data quality

A major and common source of data generation and storage for assets and maintenance activities across industrial companies are in the Enterprise Asset Management (EAM) or Computerized Maintenance Management Systems (CMMS) (Gulati, 2009). Capabilities of CMMS/EAM systems include work task identification, planning, scheduling, and reporting. Databases from CMMS/EAM systems include records of maintenance activities and costs across asset fleets. A strength in transactional data from the CMMS/EAM is that it contains information about a wide range of assets in the organization, as well as maintenance logs capture any maintenance event performed. The information in these databases can be mined as a starting point for identifying APM and PHM opportunities across an entire company or plant's operation. However, challenges in this approach arise due to the high volume of data and data consistency and quality issues.

Data quality challenges are common across nearly all companies and especially in situations where data is manually entered. Historical records are important for providing valuable insight in past maintenance on existing pieces of equipment and provide valuable asset information and material for reliability analyses. While the value potential from using transactional data is unbounded, the

abundance of missing and inconsistently filled in information limits the possible analysis on data sets. Discussions on different data quality challenges are well reviewed in (Lukens, Naik, Hu, Doan, & Abado, 2017) (Meeker & Hong, 2014) (Hodkiewicz, Kelly, Sikorska, & Gouws, 2006) (Koronios, Lin, & Gao, 2005) (Lin, Gao, Koronios, & Chanana, 2007).

Problems around missing or incomplete failure event information, such as the failure mode or the root cause, are data quality challenges central to conducting reliability analytics (Sikorska, Hammond, & Kelly, 2007). Failure mode data analysis relies on consistent failure mode coding practices such as those standardized for the oil and gas industry by ISO 14224 (ISO 14224, 2004). Failure codes entered in CMMS/EAM systems often take the form of structured fields for which values may be selected from a drop-down menu, but in practice, many of these structured fields may be incorrectly filled in or missing. Even more fundamental, the characterization of a failure event itself often depends on the perception of a craftsman and may be miscoded or missing critical information.

A data quality assessment should precede efforts where data is used to evaluate any quantity – whether it be a simple metric or a sophisticated model. In the case of performance measures, it is not desirable to use a performance measure that is easy to manipulate to make a user 'feel good' (Gulati, 2009) (Kumar et al., 2013) and poor data quality can erroneously alter many common metrics to look good. For instance, not recording failures properly can improve measures of asset reliability. After determining which data is sufficiently-good for analysis and which analyses are possible, you can analyze asset performance as far as it will allow, and start improving processes for the quality of the rest of the data (Naik, 2016) (Naik & Saetia, 2018). Improving data quality, once measured, can be done by changing the process in which the data is created and/or by improving the existing data. Frameworks for assessing and improving data quality for asset performance applications such as evaluating metrics have been developed extensively (Hodkiewicz & Ho, 2016) (Koronios et al., 2005) (He, 2016).

2.2. Assess risks for maintenance tasks through FMEA

Failure modes and effects analysis (FMEA) is a standard methodology for identifying and assessing risks and actions that could mitigate (eliminate or reduce) the likelihood or consequence of risks. An FMEA includes an assessment of risks based on a risk's probability of occurrence and consequence, which is used for prioritization towards where to focus on a maintenance strategy in order to have the biggest impact with respect to the pre-determined business goals.

Consequence is less straightforward to estimate than the likelihood of risk (probability), and focuses on answering the magnitude of how much each failure impacts the operation if or when it occurs. Estimates of consequence may depend on asset function, safety and environmental factors, operations, as well as economic factors such as cost-effectiveness and production losses. Consequence should be evaluated independently from probability. In an FMEA approach, once probability and consequence have been estimated, a risk priority number (RPN) can be calculated. The RPN is the product of the probability and consequence and is used for ordering the priority of risks as candidates for mitigation efforts.

The RCM/FMEA approach for designing maintenance strategies has several limitations in practice. It is a daunting task to exhaustively assess failure modes for every asset in your organization depending on each individual asset's operating context. Assessing an FMEA across all assets in a facility can be impractical when kick-starting an initiative. Criticality Analysis provides a risk-based approach to prioritizing efforts (Whitt, 2009). A critical asset is an asset that has been evaluated and classified as critical due to its potential impact on safety, environment, quality, production, and cost (Gulati, 2009) (SMRP, 2017). The number of assets potentially at risk may outweigh the available resources to manage them and starting with a criticality assessment helps prioritize where to apply the available resources to be cost-effective and efficient.

There are also challenges in using different sources of information for FMEA. Sources of data can include information from the Original Equipment Manufacturer (OEM), out-of-the box FMEA or strategy templates, peer data from other users of similar equipment, work history from the CMMS/EAM, and the knowledge base of the operators and maintainers of the assets themselves (Moubray, 1997). Manufacturer data is challenged by the fact that few manufacturers are involved in the long term usage of the equipment that they manufacture and they often do not obtain failure data after the warranty period ends (Meeker & Hong, 2014).

Templates and peer data may contain levels of root cause information that is inappropriate for a specific case; there can be tremendous variation in usage and operating context within an organization for two assets of identical make and model which will affect optimal maintenance strategies. Operators and maintainers have a deep knowledge of how the assets work, what goes wrong, and how the failure matters to the organization, however, there may be subjectivity and blind spots when used alone.

Historical maintenance records from the CMMS/EAM are an excellent source for data, but should only be used as a supplemental source. Challenges include many of the data

quality issues discussed above, such as incomplete information about a maintenance event. Additionally, when using the free text fields, often it is described what was done to fix a failure, not a description of the failure cause or damage mechanism. And lastly, only failure modes that have been observed can be identified – only a good source of information for failure modes with higher frequency. And lastly, softer to measure consequences to the environment or to the reputation of the firm are more difficult to quantify, so consequences measured from maintenance logs are restricted to economic consequences.

The ideal situation is to use different sources of data that provide independent perspectives on the performance of particular pieces of equipment. Out-of-the box templates are helpful sources of information with respect to providing exhaustive lists of possible failure modes and actions, which can help supplement the perspective of a domain expert. Peer and historical maintenance data may be helpful with providing quantitative starting points for observed frequency and consequence of historical failure patterns. Developing work processes in which analytics can automate and leverage different data sources in a workflow integrated with a human user can overcome many limitations associated with RCM/FMEA approaches.

2.3. Evaluate effectiveness and feasibility of strategies with respect to business goals

Once areas of opportunity have been identified, and failure modes, risks, and mitigating actions have been identified, the remaining step for risk mitigating task identification is to evaluate the effectiveness of the possible risk mitigating actions. Effectiveness is the measure of the ability of an action to mitigate the risk. The effectiveness will depend on the consequence of each risk, as well as the feasibility of each possible task.

The feasibility of a maintenance task is related to the physics of an actual damage mechanism. The classic engineering model for the progression from when a potential failure becomes detectable to when it degrades to a functional failure (when an asset is unable to perform to specifications) is known as the P-F curve, describing the time elapsed between potential failure (P) and functional failure (F) (Gulati, 2009). Condition or health monitoring tasks are only feasible when it is possible to define a potential failure condition for a failure mode, the potential failure conditions can be monitored at intervals that are shorter than the P-F interval, the P-F interval is long enough to take action, and consistent enough in duration to be able to control the monitoring and action steps (Moubray, 1997). For each possible potential failure monitoring condition, we should evaluate which tasks are the most cost-effective for mitigating an anticipated failure mode.

In this paper, we will discuss one such possible analytic workflow which looks at age-related failure patterns in order to assess which failure modes are candidates for PHM, and which failure modes may need other suitable strategies to mitigate.

2.4. Shape assessment of failure patterns

Nowlan and Heap (1978) in their revolutionary 1978 paper "Reliability Centered Maintenance" first described an approach for conducting a shape assessment of age-characteristics of failure patterns in order to determine what types of maintenance strategies would be most appropriate to mitigate different observed dynamic failure patterns. Age-reliability characteristics of an item, component, or asset population under study are defined as the relationship between the operating age and the probability of failure. This approach was applied to failure data for different airplane components, and based on lifetime data analysis and estimation of the survival curve, which is the probability of survival of a component beyond a specific age (Lawless, 2011).

The measure central to characterizing the age-reliability relationship is the hazard rate (or conditional probability of failure). The hazard rate is the instantaneous probability that an item will fail in the interval $(t, t + dt)$ given that it has survived up to time t (Lawless, 2011) (Wang, 2005) (Nowlan & Heap, 1978). The shape of the hazard function describes the age-related failure patterns, including the famous "bathtub curve" in reliability analysis (**Figure 1**).
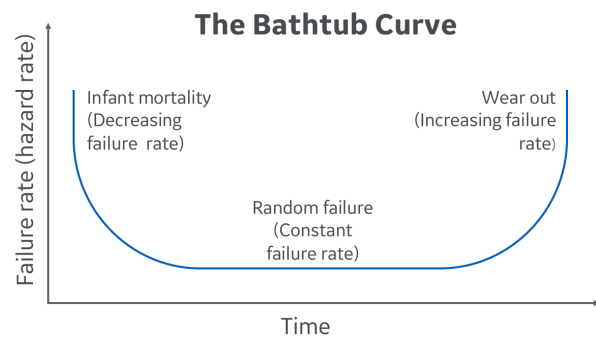

Figure 1. Cartoon depiction of the famous "Bathtub curve", a model of a hazard function over the survival time of a population.

The bathtub curve is a possible model of a hazard function over time for a population of items, and has three distinct regions. The first region is "infant mortality" when there is a relatively high probability of failure after replacement or overhaul that decreases over time. The middle region is a "random failure" region of constant and relatively constant but low probability of failure. The last region is the "wear

out region", when the probability of failure begins to increase rapidly with increasing age.

Various failure modes across varying applications are linked with different age patterns. For an asset population, grouped by similar assets in similar asset populations, the calculated hazard functions show the overall age-based failure patterns for that population. Nowlan and Heap (1978) observed six different possible shapes describing the hazard function, corresponding to different combinations of the three different regions. In **Figure 2**, the six observed hazard function shapes are reproduced along with the frequency of observance across different populations in the Nowlan and Heap (1978) study. The plots are interpreted by the shape of the hazard function (vertical axis) over lifetime (horizontal axis). For example, populations that fall into shape A exhibit a classic bathtub curve failure pattern.
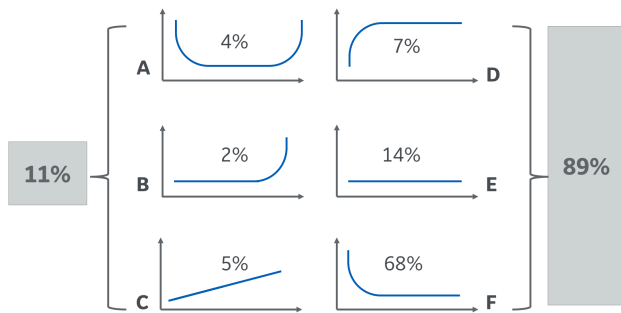

Figure 2. Reproduced results from Nowlan and Heap (1978). In their study, across populations of airplane components, they observed the non-parametric hazard rate took one of 6 shapes, and that only about 11% of components exhibited "wear out" failure patterns.

The intention for evaluating the hazard function curves across equipment populations is to get a view of age-failure patterns to be used to assess the appropriate maintenance strategy in terms of failure mode dynamics. In particular, assets that demonstrate failure modes with wear-out patterns are candidates for strategies with potential performance improvements by imposing age-limits for component replacements or for monitoring for degradation (Nowlan & Heap, 1978). In Figure 2, this corresponds to populations that exhibit shapes A, B, or C, which can be seen by the increase in the hazard function as age increases. Wear out patterns are commonly observed under conditions of wear, fatigue, corrosion, oxidation, and evaporation (Moubray, 1997).

The most cited finding from Nowlan and Heap's (1978) study is that asset populations with a wear-out age-related failure pattern were observed significantly less than asset populations without a wear-out component (11% opposed to 89% of the populations in their study). The significance of this finding was that only a fraction of asset populations were candidates for performance improvements through some

form of maintenance task relying on degradation, such as an effective condition and health monitoring initiative.

Most of the observed failure patterns were due to non-age-related failures, which could be caused by variable stress or complexity in equipment (Moubray, 1997). The greater the complexity in equipment, as technology and electronics grow more sophisticated, the more possible number and varieties of ways that a failure could occur. For these types of observed patterns, it is still necessary to understand the underlying failure patterns in order to devise an effective strategy to mitigate failures. Some possible strategies could be redesign (such as install a screen to prevent jamming or clogging by foreign objects) or evaluating maintenance or operating processes (such as an operator who aggressively uses an asset, or ineffective maintenance practices).

3. METHODS: DATA-DRIVEN WORKFLOW FOR ASSET STRATEGY PRIORITIZATION

The workflow based on Nowlan and Heap (1978) for asset strategy identification and prioritization is shown in **Figure 3**. The input is assumed to already have passed through the data quality and a prioritization workflow such as Criticality Analysis. Analytics to treat data quality challenges and estimate the hazard functions are integrated in the workflow to automate the process across broad datasets and address the limitations discussed in Section 2.2.
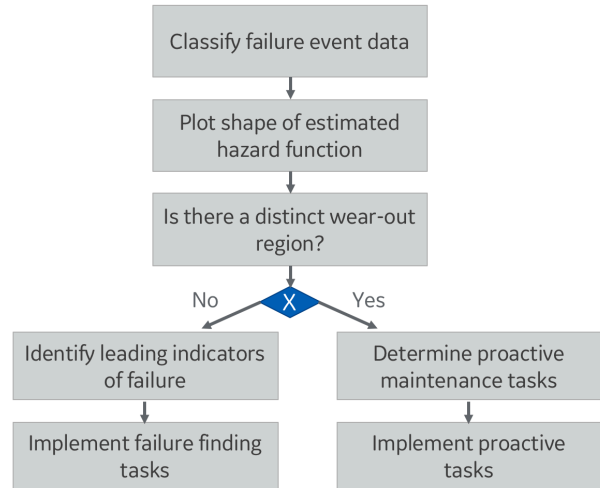

Figure 3. Flow chart for assessing asset strategies based on calculating the hazard function

Running the analysis across general or broad maintenance records presents many challenges around processing the data and ensuring interpretable results. The general steps from (pre-processed) data from the CMMS/EAM to proposing effective strategies based on age-failure patterns are:

1. Identify which historical events are a failure

2. Identify meaningful groupings or populations of assets (by component, type, manufacturer, model, unit type, company, etc.)
3. For each grouping of assets, estimate the hazard function
4. For each hazard function, characterize the shape and drill down to understand and interpret the shape at the failure mode level. Populations with wear-out regions identified are candidates for performance improvements through PHM.
5. Evaluate the effectiveness of a proposed task

For each step, we discuss the challenges, numerical approaches for calculations and the new challenges that arise.

3.1. Identify which historical events are a failure

As mentioned in the data quality background (Section 2.1), one key data quality challenge for conducting reliability analytics across broad datasets are lack of failure event classification, and even more fundamentally, recording whether an event was a failure event or not. Any large-scale data mining effort across historical maintenance records for reliability analytics is fundamentally challenged by missing or inconsistently coded breakdown indicator fields, which records whether an asset failed or not.

In addition to addressing the data entry and generation process, there are analytics opportunities to tackling this problem across historical data that already exists. In many cases, the true nature of the failure cause can be inferred from unstructured fields such as the free text field or service notes. Different approaches for treating this problem, both through analytics to clean existing data or processes to improve data collection have been discussed and explored (Lukens et al., 2017) (Hodkiewicz & Ho, 2016).

The approach we take is to use a machine learning classifier trained on labeled data. Major strengths of deploying classification algorithms are consistency and scalability. Two similar inputs will always have the same classification by a computer model, which is not always the case with human labeling. Further, computer models can run very quickly over large amounts of records (can label thousands of records in minutes). We use commercial software in the GE Digital APM offering, which classifies a repair event as a failure or not a failure based on a machine learning algorithm trained on a large dataset describing repair events across an aggregate of industrial peer companies and reviewed and labeled by a team of subject matter experts (GE Digital APM, 2017). Naïve-Bayes was the algorithm we used for a couple of reasons. A lazy learner was desired because of data-sensitivity issues when training a classifier on aggregated confidential data and applying it elsewhere. Of the lazy learners explored, Naïve-Bayes had the highest accuracy as well as the highest reported satisfaction of the results by the team of subject matter experts.

New sets of challenges arise when applying classification models to broad datasets. While consistency is a strength, it can also be a weakness, as the definition of "functional failure" may differ between identical assets in different operating contexts. Additionally, fundamental challenges in work history descriptions can lead to problems. Models are only as good as the data they are trained on, and edge cases abound in work history descriptions. For example, the information about what was done to repair a failure or what was observed (eg: "replace bearing", "pump is not pumping") may not give the desired level of information, and while a human may figure out intelligently with to do with this information, a computer algorithm will not. It has been in our experience that computer models adequately get most of the labels correct, which significantly reduces the time and effort required for a human, and further, that meaningful patterns, trends, and insights can be obtained even from the data not being 100% labeled and catching every functional failure specifically in every operating condition.

3.2. Identify meaningful sub-groupings or populations of assets

The hazard rate shape assessment is calculated over a population of similar assets operating under similar operating conditions. Population mixtures are mutually exclusive groupings of units within a population, which could result from differences in the manufacturer or the use of the product. Nowlan and Heap (1978) used different aircraft components such as a particular make and model of engine as different subpopulations.

A new challenge that arises from applying analytics to broad datasets is processes and methods for determining optimal ways for stratifying a population. We have observed through experience both first-hand and with applying different statistical methods to CMMS/EAM data that different sites and companies tend to be the most meaningful explanatory variable for differences between subpopulations. For this reason, we group assets of the same type by site in the case study, but remark that there is opportunity for future statistical work here towards automating and improving this workflow.

3.3. Numerical estimation of the hazard rate

In lifetime data analysis, the survival curve or the reliability function is the probability of survival beyond a specific age (Lawless, 2011). If we assume that T is the random variable describing time to failure, the probability of survival past time T is given by the survival function

$$S(t) = \Pr\{T > t\}. \tag{1}$$

The survivor function can be estimated parametrically or non-parametrically. Non-parametric estimates look at the proportion survived over a period of time running over the observed lifetime data. When right-censored data is included in the lifetime dataset (equipment that did not fail over the observation time window), the Kaplan-Meier estimate can be used to adjust the non-parametric survival function.

Parametric models, such as models using the Weibull distribution, are also commonly used to estimate the survivor function. In a probability distribution framework, the probability density function (pdf) $f(t)$ models the distribution times of time to failure and the cumulative distribution function (CDF) $F(t)$ gives the probability of failure at a certain time. The survival function is denoted as the complement of the CDF ($S(t) = 1 - F(t)$).

The measure used for age-reliability is the conditional probability that a component, having survived up to a given time, will fail at the time. This is known as the conditional probability of failure, or the hazard rate. The hazard function is given by

$$h(t) = \lim_{dt \to 0} \frac{\Pr\{t \leq T < t + dt | T > t\}}{dt} . \qquad (2)$$

In probability distribution terms (such as in a Weibull analysis), this limits to

$$h(t) = \frac{f(t)}{S(t)} = \frac{F'(t)}{1 - F(t)}. \qquad (3)$$

The strengths of using a parametric approach are that it is straightforward to implement and often the fitted parameters have meaningful physical interpretation. The parametric models can also be used for other analyses such as to estimate the number of expected component failures over a period of time for planning purposes. Limitations of the parametric approach include factors around the data, data fitting and the model itself. It assumes that the data can be described by a parametric model and that the fit itself is reasonable, which is not always the case. A further limitation to estimating the hazard function in the parametric approach is that the shape of the hazard function is restricted only to pre-determined shapes by the distribution used.

One of the major challenges in numerically conducting a Weibull analysis from historically observed failure data is understanding and cleaning the data up so that a reasonable fit to a Weibull distribution can be made. The Weibull distribution model is a failure-mode specific model, and when mixed failure modes are observed across an asset population, the only way to fit the time to failure data is to have the different similar damage mechanisms or failure

mode for each observed failure event well characterized. If the analysis is conducted on a small data set describing well-known asset failures, this can be effective, however, this method does not scale up well for large datasets spanning many populations of assets across many sites and possibly across several different industrial companies and verticals. Once a Weibull is fit from each failure mode, the hazard function can be estimated. In a Weibull analysis, the 3 possible regimes (infant mortality, random, and wear out) are the only possible shapes the hazard function can take.

The non-parametric approach to hazard rate estimation has added flexibility, is model-free, and data-driven. As a result, more complex shapes can emerge describing a hazard function profile than may be possible to model using a probability distribution and can describe failure dynamics across multiple failure modes. However, limitations of the non-parametric approach include difficulty in using the fits for predictive purposes and for simulation, and numerical challenges in estimating the hazard function. The non-parametric survival curve (such as generated by Kaplan-Meier estimates) may be jagged or noisy, and numerical estimation of the derivative of noisy data is even noisier, while the hazard rate is assumed to be a smooth function (Wang, 2005).

The numerical challenge of smoothing non-parametric hazard estimates was tackled by the biostatistics community in the 1990's. Different possible approaches for smoothing the hazard rate function for continuously observed data, such as using a convolution type estimator with different smoothing methods such as a kernel method, a ratio type, and spline estimators were developed, as well as methods and considerations such as selection of kernel, choice of bandwidth, and left and right boundary effects (Wang, 2005) (Muller, 1994). There are different software packages available, both commercial and open source for estimating smoothed hazard rates using the kernel method such as 'survPresmooth' (López-de-Ullibarri & Jácome Pumar, 2013) and 'muhaz' (Hess & Gentleman, 2010) in R.

3.4. For each hazard function, characterize the shape

Characterizing the shape of the estimated smoothed hazard function is straightforward through visual inspection. Challenges arise when attempting to systematically characterize shapes across hundreds of different sub-populations across broad datasets (hundreds of equipment types across many units or sites), and also for then drilling down into the data for determining a prioritization strategy. To tackle the first problem, the performance measures and criticality assessments performed previously should be used for prioritization of where to start. In order to identify candidates for a PHM initiative, we can identify which populations resulted in age-related failure patterns as candidates to first investigate.

Once a prioritized set of populations are determined, we can start making assessments systematically towards informing the maintenance process. The next step is to drill down into the different failure modes to determine which failure modes are driving the shapes. For the age-related failure patterns for identifying potential PHM opportunities, it would be to assess which failure modes contribute to the observed wear out or degradation patterns.

The data challenges that arise here circle back to the data quality challenges around failure event classification. We take the same approach as for identifying which repair events are failures to characterize failure events by their maintainable item and failure mechanism. For each characterized population, we identify the dominant damage mechanisms and run a Weibull analysis to characterize each failure mode. The outcome is identification of failure modes on prioritized asset groupings that are potential candidates for maintenance tasks aimed specifically at how their failures dynamically have occurred.

3.5. Evaluate the appropriate maintenance tasks and strategies

Once we have established frequencies of failures and failure modes based on their dynamics, the last step is to determine which tasks are appropriate for mitigating risks by looking at both the feasibility and effectiveness. Effectiveness arises from studying the consequence of different failure modes, which is another challenging analytics opportunity and an area of future work.

4. CASE STUDY

In this case study, we walk through an example of how we can investigate and identify areas towards creating an effective asset strategy based on age-failure patterns. This top down approach can assist in identifying what assets and what failure modes may benefit performance-wise from a PHM initiative. To protect proprietary information, all variables have been anonymized and age has been scaled.

After a preliminary prioritization assessment based on a benchmarking exercise, a particular type of rotating asset in a particular industrial application was identified as a candidate for prioritization based on high repair cost and counts relative to peer values. The dataset identified for the case study consists of $N = 180$ assets observed over a period of 5 units of time (such as a year) at 3 different sites. There were about 4,000 repairs observed total over that period of time, but not all of the repairs were failures (the breakdown indicator was not used).

The first step was to classify which of the repairs were failure events. We used the classifier in the GE Digital APM commercial software package which classifies a repair event as a failure or not a failure based on free text data. Classifying the repair data resulted in about 650 failure events. We were able to visually inspect the results to evaluate the failure classifications and they were satisfactory. A couple examples demonstrating the differentiation of repair events are shown below in **Table 1**:

Table 1. Examples demonstrating failure classification of repair events

Work description	Is A Failure?
Seal is leaking	True
Shaft seized	True
Motor to be decoupled	False
Check trip condition	False

It was also identified that for each population grouping, we wanted to conduct the analysis for different populations based on the different sites. We assumed that after each failure, the repair was good as new. We converted the observed maintenance data, which was in calendar time, to lifetime data through calculating the time between events.

The Kaplan-Meier survival curves between the 3 sites are shown in **Figure 4**:
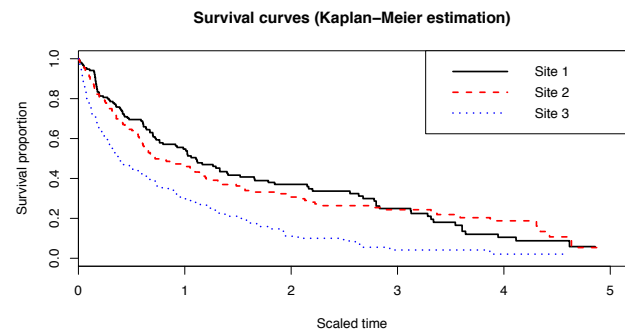

Figure 4. Survival curves for the case study as calculated by Kaplan-Meier estimation.

We calculated the non-parametric smoothed hazard function for each asset population which were grouped by site and show the curves in **Figure 5**. We used the 'muhaz' R package for the kernel smoothing method, which first estimates the cumulative hazard function with the Nelson-Aalan estimator, then smoothed using an Epanechinikov kernel. The hazard function is returned as the first order difference of the smoothed function. In the method, to avoid boundary effects that occur from bias problems near the left endpoint, we used the (default) maximum time for the hazard function estimator to when 10 assets remained at risk. From the survival curves in **Figure 4**, we can see how most of the assets at Site 3 failed before 3 time units.
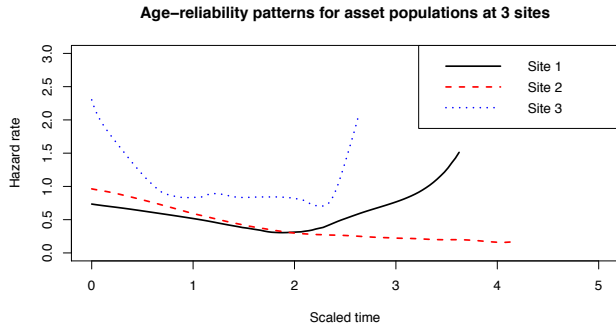
Figure 5. Non-parametrically estimated smoothed hazard functions across population of rotating assets. The end time for the smoothing was determined by the sample size of remaining assets at risk. Since nearly all but 10 failures for Site 3 were observed in the first 3 time units, the smoothed estimate does not account for the few remaining.

Applying the shape assessment to the smoothed hazard functions, the age-failure patterns for the rotating asset populations at Sites 1 and 3 have the characteristic bathtub shape, meaning that there are failure mode patterns which exhibit wear-out characteristics, and performance could potentially be improved through imposing an age-limit or through condition or health monitoring strategies. For Site 2, we observe that there is likely a complex failure mode structure and unless one of those failure modes dominates, imposing an age-limit or applying PHM will do little to improve the overall reliability.

The next step is to drill down into the failure mode structures and identify strategy opportunities. Due to data quality issues, we mine the failure mode information by identifying the maintainable item and failure mechanism in the work orders using the classification algorithms from the GE Digital APM commercial software package.

Investigation of failures at Site 2 revealed that the dominant failure modes were seal leaks (25%), bearing leaks (18%), and lube oil leaks (12%). The percentages were calculated by dividing the number of observed failures for a particular failure mode by the total number of failures. Of the 650 classified failures, 193 were at Site 2. A parametric Weibull analysis showed all of these failure modes as having an infant mortality trend ($\beta < 1$), which is consistent with the monotonically decreasing non-parametric hazard rate estimate.

We drilled down to individual failure event patterns for different individual assets and observed repetitive repair events. **Figure 6** shows a representative example of the event history for one asset. Observe that the same failure mode (seal leak) was observed and recorded repeatedly in a short period of time. What could possibly be happening is that the repairs are not effective, so investigation is needed to

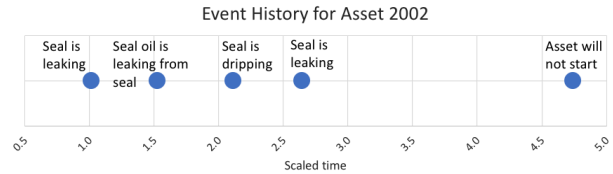determine why repeat repairs are happening, and address the root cause.

Figure 6. Failure event occurrences for an asset at Site 2. Many repetitive seal repairs over a short period of time, implying that possibly value in a strategy to reduce repeat failures and improve fix effectiveness.

At Site 1, failures were driven by the dominant failure mode of seals leaking (40% of failures). The next most common failure mode were general failures (15% of failures) (General failures refers to general failures that are challenging to characterize such as "Asset is not working" or commonly "Pump is not pumping"). No other failures modes were observed with significant frequency (<10%). Site 3 failures had a similar pattern, except most failures were due to a dirty strainer (33%), then seal leaks (11%). Mechanistically, any of these failure modes would be assumed to be due to some sort of wear (seal degradation, strainer accumulating gunk, wearing out of mechanical parts, etc.). In order to identify which of these failures contributed to the observed wear out portion of the bathtub curve observed in the smoothed hazard estimators, we performed a Weibull analysis. The parametric estimated hazard functions for the Weibull analysis for Site 1 are shown in **Figure 7**.
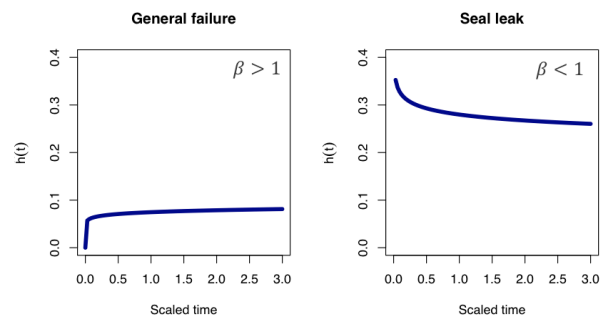

Figure 7. Parametric hazard rates (results of Weibull analysis) for dominant failure modes at Site 1.

The Weibull analysis for both Sites 1 and 3 revealed that seal leaks and dirty strainers had infant mortality patterns similar to Site 2. Further, the wear out contributions were due to general failures. In this particular analysis, if the seal leaks or dirty strainers were associated with wear out, they would be candidates for a cost-effectiveness analysis for a PHM initiative. However, at these sites the engineer should probably recommend a study of the maintenance

effectiveness. Additionally, the analyst should also start investigating and asking questions as to what happened behind the general failure events observed in the work history data to determine the root cause. While the need for the reliability engineer to ask questions to determine what happened is nothing new, a work process utilizing data analytics has been proposed which identifies where to ask the questions and which questions to ask.

5. CONCLUSION

Analytics embedded in work processes have the potential to assist in making maintenance strategy decisions which can have the largest impact on business goals in an organization. However, using data sources common to most industrial facilities as inputs to data-driven analytics in a large-scale fashion introduces many new challenges. It is important to recognize that data-driven analytics alone are far away from replacing human knowledge. Rather, there is need to develop workflows which reduces tedious manual labor to help the human execute work processes in a more informed fashion. What is ideal for incorporating data and analytics are "man-in-the-loop" workflows which supplement the domain expert's experience. In the context of prioritizing a PHM initiative, analytics can be embedded in workflows which help in measuring and improving data quality, identifying areas of opportunity to focus on, and for identifying systematically where opportunities lie for the optimal payback on a PHM investment.

It is important to note that none of these proposed methodologies address the single hardest limitation which is actual successful implementation of the determined optimal strategies. In the FMEA approach, only the challenges associated with identifying which risks to mitigate in the most effective way were addressed in this study. Once a maintenance strategy is defined, the challenge of actually eliminating failure modes remains which is less straightforward and the most difficult part of the process. However, once a strategy is designed and implemented, there is also the possibility and challenge of monitoring for dynamic risk to ensure that the strategy is both successfully executed and effective. Risks may change over time, and asset strategies may need to be re-assessed. While the analytics-driven workflows proposed here probably will not help the operator and maintainer optimally maintain the assets themselves, there is opportunity to modify the proposed methods and workflows to monitor and measure the effectiveness of a strategy. In this fashion, alerts can warn the reliability engineer which assets to continuously check up with respect to strategy implementation.

The population-level methods in this study calculating the hazard rate and Weibull analysis are mathematically the same to methods in the PHM literature referred to as "type I prognostics". The term "type I prognostics" has been coined

to describe traditional reliability analysis in a prognostics perspective (Coble & Hines, 2011) (Coble & Hines, 2009). Reliability analytics models characterize the expected lifetime of an average system operating under average operating conditions. From a prognostics perspective, shortcomings of this approach are that actual remaining useful life (RUL) estimates are not accurate for an individual asset due to lack of information specific to an individual asset's operating environment. However, in our specific use-case, we wish to characterize failure patterns for a group of assets to understand their general behavior as a first step to identifying priority groupings, and to make general assertions about to approach developing a strategy.

REFERENCES

Casto, P. (2010, April/May). Creating a New Partner with Reliability Centered Operations. *Uptime Magazine*, pp. 9-17.

Coble, J., & Hines, J.W. (2011). Applying the general path model to estimation of remaining useful life. *International Journal of Prognostics and Health Management*, 2(1), 71-82.

Coble, J., Ramuhalli, P., Bond, L. J., Hines, J. W., & Ipadhyaya, B. (2015). A review of prognostics and health management applications in nuclear power plants. *International Journal of Prognostics and Health Management*, 6.

Digital, GE. (2017). *APM Software, Version 4.3*. San Ramon, CA.

Goebel, K., Daigle, M., Saxena, A., Sankararaman, S., Roychoudhury, I. & Celaya, J.R. (2017). *Prognostics: The science of making predictions*. Publisher: Author.

Gulati, R. (2009). *Maintenance and reliability best practices*. New York: Industrial Press Inc.

He, B. (2016). *A Machine Learning Approach for Data Unification and Its Application in Asset Performance Management*. Thesis, Virginia Polytechnic Institute and State University, Blacksburg.

Hess, K. & Gentleman, R. (2010). muhaz: Hazard Function Estimation in Survival Analysis. *R package version 1.5*

Hodkiewicz, M., Kelly, P., Sikorska, J., & Gouws, L. (2006). A framework to assess data quality for reliability variables. *Engineering Asset Management*, 137-147.

Hodkiewicz, M., and Ho, M.T.W. (2016). Cleaning historical maintenance work order data for reliability analysis. *Journal of Quality in Maintenance Engineering*, 22(2), 146-163.

International Standards Organization (ISO) (2004). Petroleum, petrochemical and natural gas industries—Collection and exchange of reliability and maintenance data for equipment. In *ISO 14224*. Genève, Switzerland: International Standards Organization.

Koronios, A., Lin, S., & Gao, J. (2005). A data quality model for asset management in engineering organisations. In *IQ*.

Kumar, U., Galar, D., Parida, A., Stenström, C., & Berges, L. (2013). Maintenance performance metrics: a state-of-the-art review. *Journal of Quality in Maintenance Engineering*, 19(3), 233-277.

Lawless, J. (2011). *Statistical models and methods for lifetime data*. (Vol. 362). John Wiley & Sons.

Lee, J., Liao, L., Lapira, E., Ni, J., & Li, L. (2009). Informatics platform for designing and deploying e-manufacturing systems. In *Collaborative Design and Planning for Digital Manufacturing* (pp. 1-35). London: Sprinter.

Lee, J., Wu, F., Zhao, W., Ghaffari, M., Liao, L., & Siegel, D. (2014). Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications. *Mechanical systems and signal processing*, 42.1, 314-334.

Lin, S., Gao, J., Koronios, A., & Chanana, V. S. G. (2007). Developing a data quality framework for asset management in engineering organisations. *International Journal of Information Quality*, 1(1), 100-126.

López-de-Ullibarri, I. & Jácome Pumar, M.A. (2013). survPresmooth: Presmoothed Estimation in Survival Analysis. American Statistical Association. *R package version 1.1-10*.

Lukens, S., Naik, M., Hu, X., Doan, D., & Abado, S. (2017). The role of transactional data in prognostics and health management work processes. *Proceedings of the Annual Conference of the Prognostics and Health Management Society*, (pp. 517-528). St. Petersburg, FL.

Meeker, W. Q., & Escobar, L. A (1998). *Statistical models for reliability data*. New York: John Wiley & Sons, Inc.

Meeker, W. Q., & Hong, Y. (2014). Reliability meets big data: opportunities and challenges. *Quality Engineering*, 26(1), 102-116.

Moubray, J. (1997). *Reliability-centered maintenance*. New York, New York: Industrial Press Inc.

Muller, H. G. (1994). Hazard rate estimation under random censoring with varying kernels and bandwidths. *Biometrics*, 61-76.

Naik, M. (2016, March 17). How Good Does Data Need to Be for Maintenance Practices? *Automation.com*.

Naik, M. & Saetia, K. (2018). Improving data quality using best practices and cognitive analytics. *SMRP Conference Proceedings*. Orlando, FL.

Nowlan, F., & Heap, S. (1978). *Reliability-centered maintenance*. San Francisco CA: United Air Lines Inc.

Rajamani, R., & Bird, J. (2016). Introduction to Prognostics and Health Management. In PHM Society Short Course: PHM Fundamentals and Case Studies (pp. 11-27). Denver.

Saxena, A., Roychoudhury, I., Celaya, J., Saha, S., Saha, B., & Goebel, K. (2010). Requirements Specification for Prognostics Performance-An Overview. *AIAA Infotech@ Aerospace 2010*, (p. 3398).

Sikorska, J., Hammond, L., & Kelly, P. (2007). Identifying failure modes retrospectively using RCM data. In *ICOMS Asset Management Conference*. Melbourne, Australia.

SMRP Best Practices (2017). Atlanta, GA: Society for Maintenance & Reliability Professionals (SMRP). Retrieved from https://smrp.org/

Wang, J. L. (2005). Smoothing hazard rates. In *Encyclopedia of Biostatistics*.

Whitt, R.E. (2009). Improving asset performance and process safety. *Petroleum Technology Quarterly*, Q3, 125-131.

Sarah Lukens lives in Roanoke, Virginia and is a Data Scientist at GE Digital. Her interests are focused on data-driven modeling for reliability applications by combining modern data science techniques with current industry performance data. This work involves analyzing asset maintenance data and creating statistical models that support asset performance management (APM) work processes using components from natural language processing, machine learning, and reliability engineering. Dr. Lukens completed her Ph.D. in mathematics in 2010 from Tulane University with focus on scientific computing and numerical analysis with applications in fluid-structure interaction problems in mathematical biology. Prior to joining Meridium in 2014, she conducted post-doctoral research at the University of Pittsburgh and the University of Notre Dame building data-driven computational models forecasting infectious disease spread and control.

Matt Markham is a Program Manager with GE Digital who works with industrial organizations to design and deploy digital solutions. Matt believes that creating value from digital technology requires careful attention to business process alignment, enablement and change management. Matt has supported many asset performance management projects with organizations in a variety of industries including oil and gas, power, chemical and mining. Matt holds a variety of certifications related to asset management and reliability including Certified Lubrication Specialist (CLS) from the Society of Tribologists and Lubrication Engineers, Certified Maintenance & Reliability Professional (CMRP) from the Society of Maintenance and Reliability Professionals and Certified Reliability Engineer (CRE) from the American Society of Quality. Matt holds a Bachelor's Degree in Mechanical Engineering from the University of Missouri-Rolla and a Master's in Business Administration from Virginia Polytechnic & State University.