Inspectioneering Journal

ASSET INTEGRITY INTELLIGENCE

Condition Monitoring Optimization Part 2: Making Inference with Uncertain Data

Dr. Andrew Waters, *Director of Data Scientist at Pinnacle*
Ryan Myers, *Product Manager at Pinnacle*

VOLUME 27, ISSUE 6
NOVEMBER | DECEMBER 2021

Condition Monitoring Optimization Part 2: Making Inference with Uncertain Data

Dr. Andrew Waters, *Director of Data Scientist at Pinnacle*
Ryan Myers, *Product Manager at Pinnacle*

Introduction

This article is Part 2 in a series of articles discussing condition monitoring optimization, which aims to provide a framework for quantitatively optimizing inspection scope, techniques, and intervals based on historical inspection data and subject matter expertise, while also dynamically updating the inspection plan to maximize reliability and return on investment (ROI) as new information becomes available. Using this methodology, data collected through inspection can be used to improve confidence in the asset damage state and determine when additional data is required, when inspection adds little or no value, or when corrective maintenance is required. In Part 1 of this series, entitled "Condition Monitoring Optimization: Going Beyond Traditional CML Optimization" and published in the September/October issue of *Inspectioneering Journal*, it was assumed that inspection coverage and techniques were sufficient to capture the true damage state of the asset.

While the above scenario is certainly ideal, it is often not representative of the real-world situations that are encountered in practice. As a specific example, inspection coverage may be limited to a fraction of the total susceptible area of an asset for a variety of reasons, such as inaccessibility. A radiographic testing (RT) scan conducted on a piping elbow, for example, generally captures only a single angle of the potential surface, leaving the inspection professional with imperfect information regarding degradation. Alternatively, even when an inspection does cover the entire area susceptible to a particular damage mechanism, the inspection technique utilized may not provide a 100% probability of detection. Consider the case of using magnetic particle inspection to identify a surface breaking crack—even when a crack is present, this technique may only have a 90% chance of detecting the crack, which leaves the inspection professional with the job of considering that damage may be present even when the inspection technique finds no evidence.

Fundamentally, the industry is faced with the challenge of making statistically meaningful inferences in the presence of limited or potentially erroneous data. The industry can combat this situation by using Bayesian statistical analysis, which combines measured inspection data with prior information derived from subject matter expertise or historical experience. This article will outline this data analysis methodology across a series of practical examples, which focus on local degradation. First, the article will examine the case of thinning on a piping circuit with limited inspection coverage. Second, the article will consider the case of local pitting on heat exchanger tubes with limited inspection data, as well as an imperfect probability of detecting damage where inspections are conducted.

Extreme Value Analysis with Limited Inspection Coverage

While, ideally, one would like to inspect 100% of the susceptible area for any type of damage, such a comprehensive inspection is often either cost-prohibitive (e.g., scanning a large surface area in its entirety) or impossible (e.g., portions of the asset are inaccessible). In such a scenario, one is faced with making an inference about an asset given limited inspection data. In the case where damage is detected using the limited inspection, one can take action to remedy the situation. However, what options exist when no evidence of damage is detected? One can assume that there is no severe damage on the asset, but must also consider the possibility that there is significant damage that has simply not been detected due to the limits in inspection coverage. Ultimately, it's important to quantifiably answer the following question: What is a reasonable estimate of the damage state of the asset given the data that has been collected?

To provide statistically valid answers to this question, one can make use of methods such as extreme value analysis (EVA). A full description of EVA is beyond the scope of this article, however D. Benstock and F. Cegla prepared an introductory treatment of EVA related specifically to non-destructive inspection in their 2017 article "Extreme value analysis (EVA) of inspection data and its uncertainties"[1]. This article will review EVA using a specific example involving a piping circuit. In this example, assume that the piping circuit is inspected with an automated ultrasonic testing (AUT) technique, and the pipe can be divided into condition monitoring locations (CMLs) where each CML corresponds to one continuously scanned "AUT band" of the pipe. Each AUT inspection scans a single CML of the pipe and collects the thickness data at various locations within the CML. As is customary in the industry, however, an inspector typically records only the lowest thickness value for each CML. The minimum thickness data is used to calculate degradation rates to serve as the basis for any inspection program including a risk-based inspection (RBI) program.

Assume that the inspector has conducted an AUT scan over some smaller number of the available CMLs and has recorded the lowest thickness measurements for each of the CMLs. The objective now is to estimate the minimum thickness on the remaining uninspected CMLs given data collected from the inspected CMLs. For example, there are 100 potential AUT CMLs, of which the inspector has sampled only ten CMLs. For this example, the measured data at each CML comes from the bi-modal distribution shown in **Figure 1** (left). Each mode of this distribution corresponds to a location experiencing either general thinning with a relatively low degradation rate (say, 3 mils/year) or an area experiencing

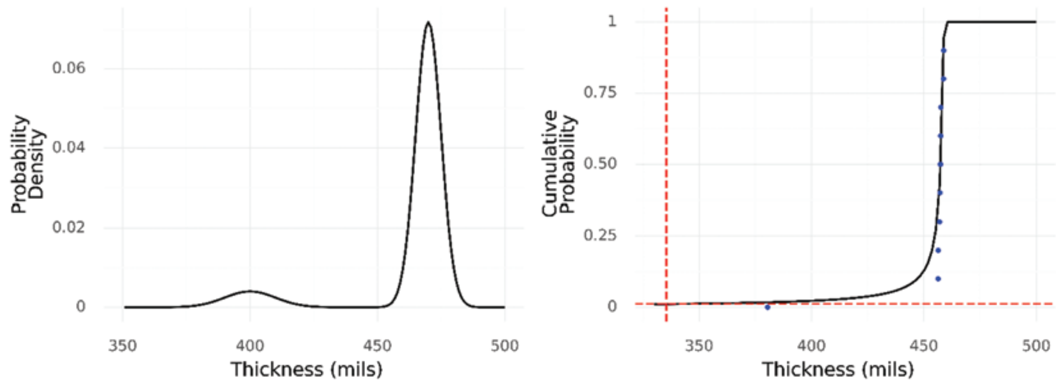


Figure 1. Left: The true thickness distribution minima for each CML of the pipe. **Right:** The cumulative distribution function for the EVA distribution with sample points (blue) and the projection of the minimal thickness for the remaining 90 CMLs on the piping circuit (red dashed lines). The EVA distribution fits reasonably well with the data and the resulting inference regarding the expected minimal thickness is reasonable.

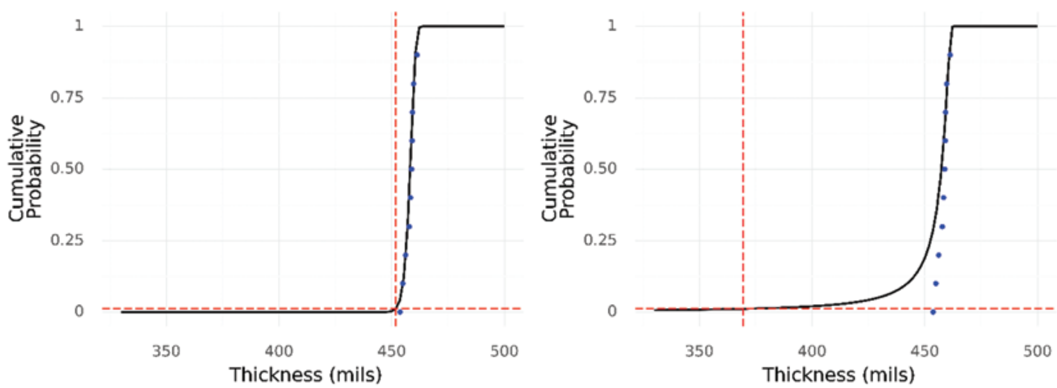


Figure 2. Left: A standard EVA fit on data where no local thinning was observed. The fit to the data is extremely good but ultimately skews the prediction of minimal thickness since no local thinning has been observed. **Right:** Bayesian EVA fit to the data. The fit to the observed data is poorer but the prior information ultimately aids in predicting a much more realistic expectation on the true thickness minima for the circuit.

local degradation at a higher degradation rate (say, 10 mils/year). The AUT scanner collects a total of 100 measurements at each CML and the inspector records the minima for each (ten total minimum thickness readings).

The results of applying a standard EVA analysis are shown in **Figure 1** (right). The resulting distribution provides a reasonable fit to the inspection data (measured data generally lies very close to distribution curve). From this distribution, the expected minimum thickness that would be observed on the remaining 90 CMLs of interest is calculated. In this case, the EVA analysis predicts a minimal thickness of 335 mils (shown by the red dashed lines), which is reasonable and matches what is expected given what is known about the true thickness distribution for this circuit.

The previous example included a CML where local thinning was present, which ultimately enabled the standard EVA analysis to make reasonable predictions about future behavior on the unobserved portion of the piping circuit. However, given the low prevalence of local thinning on this circuit, it is quite possible that the collected inspection data might have only encountered general thinning and no local thinning. In this scenario, a traditional EVA will be overfit to the observed data and ultimately fail to predict the extreme minimal thickness that was observed previously even though our prior expectation is that such a low thickness is indeed very likely to occur. This overfitting scenario is shown in **Figure 2** (left). While the EVA fit to our observed data is extremely good, the resulting prediction over the remainder of the circuit is far too optimistic given our prior expectations about the presence of local thinning. This has been one of the primary challenges in the application and acceptance of EVA in the mechanical integrity industry.

One can combat this problem by exploiting Bayesian statistical methods[2]. Applied to inspection, Bayesian methods combine prior subject matter expert (SME) knowledge with measured inspection data in order to provide a more robust estimate of the damage state of the asset. In our local thinning example, Bayesian EVA assumes a prior distribution on the parameters of the EVA distribution itself. The selection of this prior information is guided by mechanical integrity and corrosion engineer SMEs who would utilize experience along with historical inspection and process data to inform the model of what would be possible for the given asset. An example analysis using Bayesian EVA using the same data as the previous example where no local thinning was observed directly is shown in **Figure 2** (right). With suitable prior information, the EVA still predicts the possibility of local thinning given the very small amount of data that was collected. It is important to note that if SMEs were to continue to collect data using highly effective inspection techniques on a greater

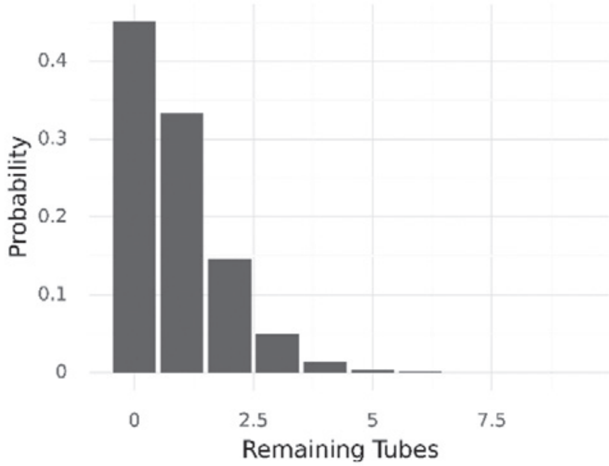


Figure 3. Probability distribution of the remaining pitted tubes for a heat exchanger. The probability that no pitted tubes remain is only 45%, meaning that SMEs would likely find more pitting if they continued to inspect.

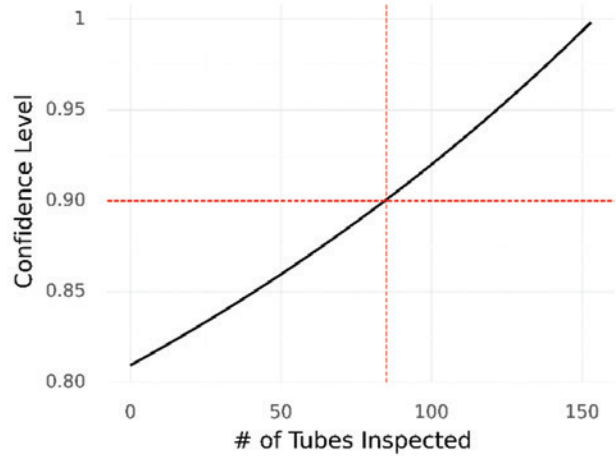


Figure 4. The confidence level that all tubes are free of pitting given an error-free inspection for a subset of tubes. For this case, 90% confidence of no pitting in the population requires inspecting 85 tubes without any evidence of damage. To obtain 95% confidence we must inspect 125 tubes.

portion of the susceptible area and continued to find only general degradation, the Bayesian EVA would ultimately reject the prior and begin to fit more closely to the observed data.

This example highlights both the value and importance of combining SME knowledge with data science to utilize all available information and more accurately estimate the damage state of the asset and improve the assessment of probability of failure and risk. Additionally, this approach can be used to select the inspection techniques, define the scope of the inspection, and determine the inspection intervals to optimize the inspection plan. The optimized inspection plan for a given asset provides the data required to dynamically update the model as new information is available, such as changes in operating conditions.

Dealing with Imperfect Inspection Techniques

In the previous example, inspection coverage was limited to a small percentage of the entire susceptible area of the asset. However, it was assumed that the measurement data itself was accurate and that, for example, local thinning would be detected if it was present in the given inspection CML. This, however, is not always the case in practical situations, where the utilized inspection techniques may have some probability of failing to detect damage even if damage is present.

The next scenario explored in this article uses data collected from a heat exchanger tube bundle. This exchanger consists of 154 tubes, of which 44 have been inspected for evidence of pitting. However, the ability to detect pitting in any given tube is imperfect. Concretely, assume that an inspector examining a tube with pitting using pulsed eddy current (PEC) will only detect the pitting with a probability of 0.7 (70%). Of the 44 tubes inspected, the inspector identifies pitting in four tubes. Given this data, we wish to estimate how many of the remaining 110 tubes in the population are likely to have pitting.

As in the case of the EVA example, it is assumed that a Bayesian analysis has been performed for the tubes affected by pitting. Based on experience and prior inspection history, it is expected that, on average, one out of every 100 tubes will experience pitting.

Figure 3 shows the output of the analysis. The probability that no pitted tubes remain is 0.45. More than likely, there are additional tubes with pitting beyond the four that have been detected.

The above example assumes that an inspection has already been conducted and, after having obtained some data, the problem of interest was in estimating the damage state of the uninspected portion of the asset.

The methodology applied can also be used for future inspection planning. Assume now that, during inspection planning, one wishes to know the number of tubes that must be inspected such that, if no damage is detected, the entire tube population can be declared defect-free. **Figure 4** demonstrates the result where the confidence (probability) of the entire population of tubes is defect free given that SMEs have inspected a given number of tubes and found no evidence of damage.

Depending on the level of confidence that is required, one can determine the optimal number of tubes to inspect. If the actual inspection results are different than the expected damage, additional sampling will be needed to achieve the desired level of confidence that can be calculated, and for inspection to be completed in real time.

Conclusion

This article considers the problem of accurately characterizing the damage state of an asset experiencing local degradation in cases where:

A. inspection is conducted on a subset of the total area, or

B. when the inspection technique has some probability of failing to detect damage when damage is present.

In these situations, statistical inference techniques on the measured data can be utilized to provide reasonable expectations regarding the true extent of damage on the asset. When SME knowledge is available, Bayesian statistical methods can be utilized to provide a more robust estimation, which is especially important when damage is hard to detect. This article demonstrated that a Bayesian framework can overcome previous challenges in applying EVA type analyses, and statistical analysis in general, to local degradation.

It is important to note that while statistical techniques can greatly aid our understanding of asset degradation, they are not meant to be used in isolation from corrosion and mechanical integrity subject matter expertise. Rather, the techniques and methodology proposed here are meant to complement the work of the SMEs and provide deeper insights into the inspection data collected as part of their reliability program. As a concrete example, SMEs will have ideally identified the potential damage mechanisms and susceptible areas for each mechanism prior to performing the EVA methodology described here. By treating each susceptible area in isolation, the resulting goodness-of-fit for the EVA distribution to the inspection data will improve dramatically and provide far more reliable (and stable) inference results.

Future use cases of condition monitoring optimization will explore scenarios where similar principles are applied to damage modes of cracking, an optimized inspection plan is developed based on condition monitoring optimization, and how facilities can use real-time process data and integrity operating windows (IOWs) to update the expected remaining life, risk, and inspection plan of their assets. This series will highlight where and how to apply data analytics to improve inspection programs so that they effectively manage both risk and cost, as well as the importance of integrating SME knowledge and expertise with data analytics to provide a more complete understanding of asset degradation and enable smarter reliability decisions. ■

For more information on this subject or the author, please email us at inquiries@inspectioneering.com.

REFERENCES

1. Benstock, D. and Cegla, F., 2017, "Extreme Value Analysis (EVA) of Inspection Data and its Uncertainties," NDT & E International, 87, pp. 68-77, ISSN 0963-8695, https://doi.org/10.1016/j.ndteint.2017.01.008.

2. Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B., 1995, "Bayesian Data Analysis," Chapman and Hall/CRC.

CONTRIBUTING AUTHORS

Ryan Myers

Ryan Myers, Product Manager at Pinnacle, oversees all new product development activities. He leads multi-disciplinary technical teams across engineering, data science, and software development fields to drive the creation of new products and services focused on increasing customer value through transforming their reliability, integrity, and maintenance programs. Ryan specializes in mechanical integrity and reliability engineering, project management, operational excellence, digital transformation, and product management. Ryan also has extensive experience both developing and applying decision analytics and mathematical models to improve business performance. Ryan obtained his Bachelor of Science in Mechanical Engineering with a minor in business from The University of Texas and is also a certified Lean Six Sigma Black Belt.

Andrew Waters

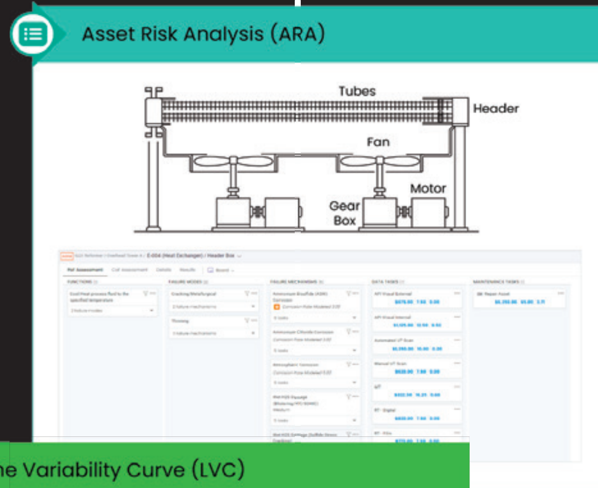
Dr. Andrew Waters is Chief Data Scientist at Pinnacle, focusing on developing data-driven algorithms to enhance a variety of reliability and maintenance applications. Dr. Waters also specializes in utilizing machine learning methods to improve and augment human decision making. He has utilized these skills across a diverse set of industries including finance, communication systems, engineering, signal processing, optimizing student learning outcomes, and hiring and recruitment programs. Dr. Waters holds a doctorate in Electrical and Computer Engineering from Rice University and is the author of over 20 publications in the areas of signal processing, machine learning, and Bayesian statistical methods. His research interests include sparse signal recovery, natural language processing, convex optimization, and non-parametric statistics.

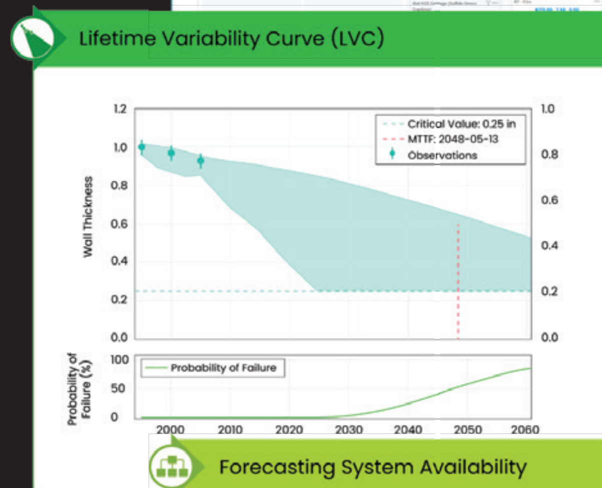
PINNACLE
DATA-DRIVEN RELIABILITY

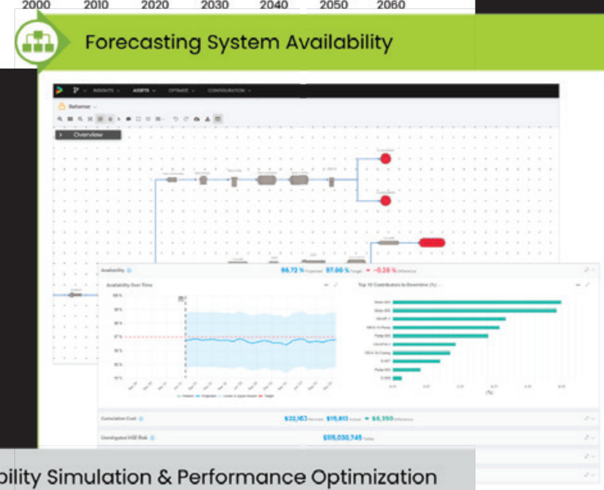
Quantitative Reliability Optimization (QRO) –

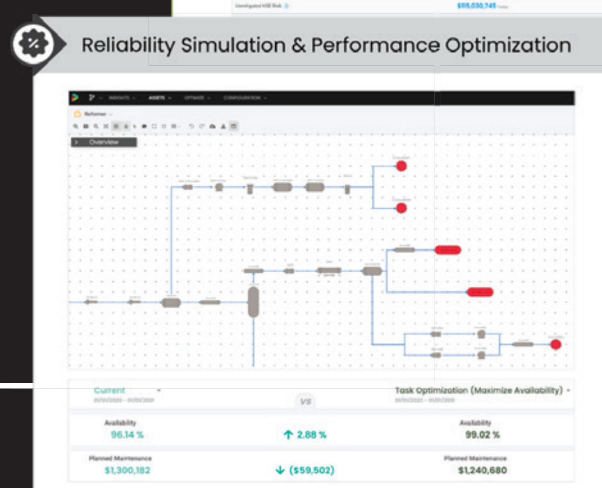
Combining the best of RBI, RCM, FMEA, RAM, PHA, and Multi-Variate Machine Learning

Learn more about the next evolution of reliability modeling









Visit pinnaclereliability.com