14th International Conference on Structural Safety and Reliability - ICOSSAR'25
1-6 June, 2025 - Los Angeles California, USA
R. Ghanem, E. Johnson, S. Masri and A. Olivier (Eds)

PLANNING INSPECTIONS TO MINIMIZE UNCERTAINTY IN THICKNESS MONITORING DATA

CONAL H. BROWN[1, 2] AND SCOTT FERSON[2]

[1]TÜV Rheinland
The Malt Building, Greenall's Avenue, Warrington WA4 6HL, UK
conal.brown@tuv.com www.tuv.com

[2] Institute for Risk and Uncertainty, University of Liverpool
765 Brownlow Hill, Liverpool, L69 7ZX, UK
ferson@liverpool.ac.uk https://www.riskinstitute.org.uk/

Key words: Corrosion Assessment, Kolmogorov-Smirnov Test, Interval Data, Uncertainty.

Abstract. The Kolmogorov-Smirnov test is applied to data containing uncertainty from thickness monitoring inspections. It is shown that the test can be a useful tool for structural corrosion assessment and decision-making under uncertainty. An example then demonstrates how the test can help decide how best to plan future inspections to reduce epistemic uncertainty. The test is applied to empirical thickness data taken from inspections and finds that reducing measurement incertitude is a more effective strategy than collecting additional sample data.

1 INTRODUCTION

Engineers rely on material thickness measurements taken periodically during in-service inspections to monitor steel structures and pressure systems for corrosion. Thickness measurements contain both aleatory and epistemic uncertainty. Aleatory uncertainty exists because the thickness of any structural component will form a distribution rather than being uniformly equal to its nominal design thickness. It arises from two sources: variation inherent in manufacturing processes that affects the initial as-built thickness, and variability in the initiation and growth of corrosion and other deterioration mechanisms across a component's surface once in use. Epistemic uncertainty, which can be represented using interval-valued data, arises because thickness measurements using non-destructive techniques such as ultrasonics have associated measurement uncertainty, are generally only taken from a sample of the surface area, and have a limited probability of detecting defects of any given size.

Any corrosion of a component that is occurring usually progresses slowly, so the associated rate of thickness loss is low. This means that even though inspections are performed infrequently, the change in thickness between inspections taken at different times, which would indicate corrosion, is often comparable to the uncertainty within the thickness measurements taken at any one inspection. If the resulting uncertainty is not acknowledged and managed when planning inspections and analysing thickness monitoring data, subsequent decisions could mitigate risk sub optimally.

This paper analyses how optimal choices can be made in inspection planning to minimize uncertainty in thickness monitoring data. It compares the effect of reducing sample uncertainty and reducing measurement incertitude on the confidence with which corrosion can be confirmed (or discounted). These comparisons are done by applying the imprecise two-sample Kolmogorov-Smirnov test to the problem of detecting a change in thickness with uncertain data. The measurements are represented as p-boxes to characterize the uncertainties, following an imprecise probability paradigm. It briefly discusses the challenges of inspecting the same location at consecutive inspections (for paired sampling) rather than different locations (requiring independent 'population' sampling).

The paper concludes by discussing how best to decide the date and scope of future inspections when there is residual uncertainty and the data does not indicate with a high degree of confidence whether corrosion is active or not.

2 INSPECTION PLANNING AND STUCTURAL INTEGRITY ASSURANCE

Inspection performs an important role in assuring the safety and reliability of steel structures and pressure systems. The material thickness of a structure is an important factor in its ability to withstand the loads for which it is designed. Steel structures are subject to corrosion and potentially subsequent thickness loss over time. It is therefore important to periodically monitor their thickness through inspection. Material thickness data collected through an inspection helps to determine whether or not corrosion is active and the structure is therefore at risk of degradation. Once corrosion is detected, appropriate mitigations can be enacted.

The presence of both aleatory and epistemic uncertainty in thickness data taken from in-service inspections using non-destructive testing methods poses challenges in determining whether thickness loss has occurred.

Aleatory uncertainty is present in thickness data because the thickness across the surface area will form a distribution as a result of variances during manufacturing and fabrication of the steel. This introduces a degree of *irreducible* randomness or stochasticity [1] associated with the thickness at any particular point across the structure. Upon entry into service the distribution is typically normal, due to the large number of factors influencing thickness. Its central tendency is the nominal thickness of the material, and its variance is determined by the manufacturing tolerances of the steel grade (plate steel is generally made to tighter tolerances than tube steel).

Epistemic uncertainty arises because of the limitations of measurement techniques in in-service inspection applications. Firstly, it is often technically and usually economically not possible to measure every part of a structure. A partial coverage inspection is therefore typical [2], which introduces sample uncertainty. Secondly, non-destructive testing methods do not yield completely accurate thickness readings – especially when deployed in field conditions [3,4]. This introduces measurement incertitude. Finally, non-destructive testing methods are unable to reliably detect flaws below a certain size threshold [5]. This potential for non-detects introduces further epistemic uncertainty.

Industry recommended practice is to represent aleatory uncertainty in inspection thickness data using Empirical Cumulative Distribution Functions (ECDFs) to represent the thickness distribution across the sample [6]. A recent contribution has demonstrated that both aleatory and epistemic uncertainty within thickness data can be represented by using probability boxes (p-boxes) [7]. In a p-box, epistemic uncertainty is represented by the interval between the left and right edges of the ECDFs [8]. Figure 1 shows an example of a p-box.

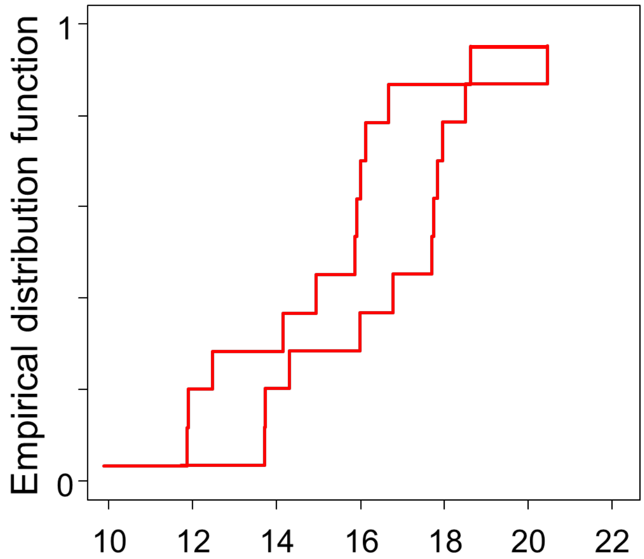


Figure 1: Example of a p-box (probability box) representing aleatory and epistemic uncertainties

Corrosion usually progresses slowly, so the associated rate of thickness loss is low. Given inspections are performed infrequently, the change in average or minimum thickness between inspections taken at different times, which would indicate corrosion, is often comparable to the uncertainty within the thickness measurements taken at any one inspection. This creates the risk of incorrect decision-making if not taken into account. The magnitude of the uncertainty compounds the challenge of managing the combination of the sample uncertainty and measurement incertitude alongside the underlying aleatory uncertainty.

Integrity engineers should take the uncertainty into account when evaluating whether or not corrosion has occurred between two inspections taken some time apart by using statistical tests, rather than simply comparing the average and minimum thickness values found. Changes in measured averages and minima could have occurred by chance, unless shown to be statistically significant. It is clearly beneficial to plan inspections in such a way that uncertainty in the thickness data is minimised because it will allow conclusive decisions about changes in thickness to be reached sooner in the life of the asset being inspected, allowing for more effective management of its structural integrity. At least three aspects of inspection planning can affect uncertainty. One important decision in inspection planning is the trade-off between the number of measurement samples taken and the accuracy of the technique used to obtain each sample. For a given cost it is often possible to take a greater number of lower accuracy measurements, or a smaller number of higher accuracy measurements. Other trade-offs are the amount of measurements to take at each inspection and the interval between inspections. Statistical tests can also be used to evaluate the effect of these inspection planning decisions on the resulting level of uncertainty in the thickness data.

3 APPLYING KOLMOGOROV-SMIRNOV TEST WITH UNCERTAIN DATA

3.1 Kolmogorov-Smirnov test and distance

The Kolmogorov-Smirnov (KS) test can be used to determine the likelihood of two samples being from the same underlying distribution. It is applied to test the null hypothesis that both samples come from the same distribution [9]. The goodness of fit measure used by the test is the maximum vertical distance between the ECDFs of each sample, as shown by the vertical black arrow in Figure 2.

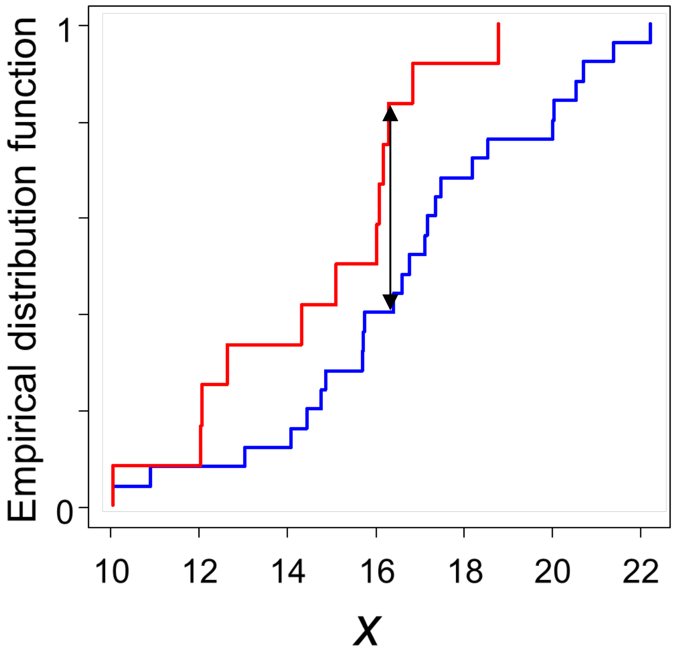


Figure 2: Maximum vertical distance between 2 ECDFs

The ECDF function that generates Figure 2 for n ordered observations in a sample $X_1, \ldots, X_n$ that are independent and identically distributed is defined by [9]

$$\hat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(X_i \leq x), \quad (1)$$

where $\mathbb{1}$ is the indicator function, equal to 1 if $X_i \leq x$ and equal to 0 otherwise.

$\hat{F}_n(x)$ corresponds to the proportion of sample values less than or equal to x. For the two sample KS test we are interested in the similarity between two ECDFs $\hat{F}_n$ and $\hat{G}_m$ based on two samples $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$. The test statistic that calculates the maximum vertical distance between the distributions is [9]

$$D_{n,m} = \sup_{t \in \mathbb{R}} \left| \hat{F}_n(x) - \hat{G}_m(x) \right| \quad (2)$$

$D_{n,m}$ can be referred to as the KS distance. It does not have sense of direction (i.e. in Fig. 1, the distance of the red line above the blue line is not a positive nor a negative distance, and if the maximum distance were at a point where the blue line is above the red line, the distance would similarly neither be positive nor negative).

There is the potential for the vertical distance to change whenever one of the two ECDFs steps upwards, which occurs when there is a change in the magnitude of the ordered values in the sample. The distance at each change depends on the number of values in the sample and the number of data points with the same value. Once the KS distance has been calculated, its significance can be assessed by comparing it with a threshold distance for a particular p-value. The p-values can be obtained from statistical tables or the Kolmogorov distribution [9].

3.3 Kolmogorov-Smirnov test applied to imprecise data

The KS test can be applied to imprecise data characterised by p-boxes [10]. In this situation the vertical distance between the two distributions is now characterised by an interval. The upper and lower boundaries of the interval are defined by [$\overline{\mathrm{F}}$ - $\underline{\mathrm{G}}$] and [$\underline{\mathrm{F}}$ - $\overline{\mathrm{G}}$]. It is possible that the at the lower bound of the interval, the distance is insufficient to reject the null hypothesis that the distributions are not significantly different, while at the upper bound the distance is sufficient to reject the null hypothesis. This means that it is necessary to calculate the maximum 'outer' distance and maximum 'inner' distance and perform the KS test on each of them. Figure 3 shows these two distances, D_{max_outer} and D_{max_inner}.

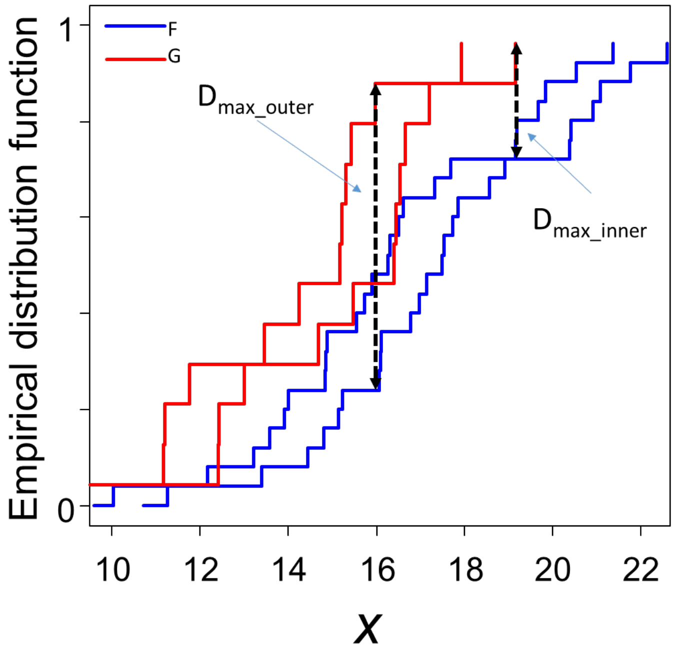


Figure 3: The outer D_{max} and inner D_{max} of two p-boxes

The KS test with the imprecise data has three outcomes [after 10]:

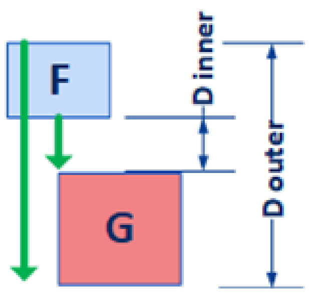
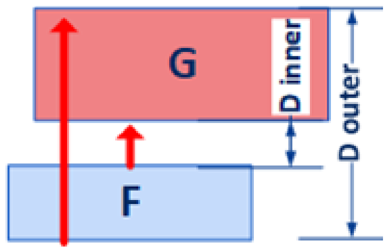
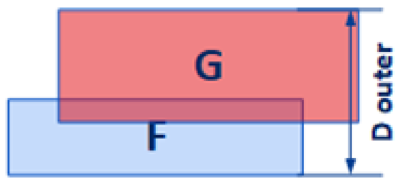
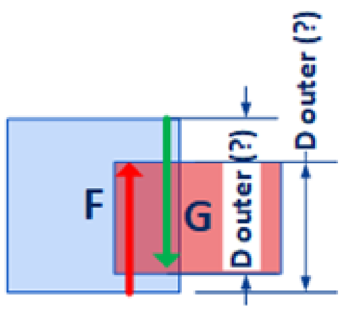
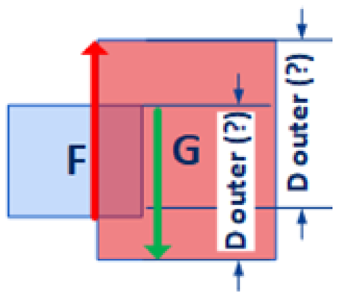
1. Both distances are below the KS test threshold value and the null hypothesis *cannot* be rejected.

2. The outer distance is above the KS test threshold value, but the inner distance is below it. The null hypothesis *can* be rejected (in this outcome, depending on decision criteria, it may be preferable to reject the null hypothesis, accept it, or collect more data).
3. Both distances are above the threshold value for the KS test and the null hypothesis *must* be rejected.

To calculate the maximum outer distance, the overall supremum of the maximum of the distances at each step between [$\overline{F}$ - $\underline{G}$] and [$\underline{F}$ - $\overline{G}$] is taken.

Calculation of the maximum inner distance is more involved because when the two distributions intersect the inner distance is zero. Table 1 gives the six possible scenarios for the relative distances between $\overline{F}$, $\underline{F}$, $\overline{G}$ and $\underline{G}$.

Table 1: The six relative distance scenarios between the extrema of the p-boxes at any given step in the ECDF

Scenarios where F is above G	**Equivalents where G is above F**	**Comment**
Separation between ECDF lines		
		There is a candidate for both the outer D and inner D values in these scenarios.
Partial intersection of ECDF lines		
		There is partial intersection so there is an outer D candidate. The inner D is 0
Full intersection of ECDF lines		
		There is full intersection so there are two possible outer D values, the larger is taken as the candidate.

The green and red arrows on the figures in Table 1 indicate a method for determining whether there are two outer distances being calculated, or an outer and inner distance. When the signs of the distances match (both calculated as either positive or negative) there is no intersection, which means the larger and smaller distances are candidates for the maximum outer and inner distances, respectively. When the signs of the distances differ (one calculated as positive and the other as negative), then there is intersection and therefore both distances can only be potential outer distances. In this situation the larger of the two distances should be taken as the candidate maximum outer distance for that point between the two ECDFs. Desterke and Strauss provide equations for calculating these distances taking into account the situations in Table 1 [10]. The following reproduces their equations but adds the inner and outer distance terminology introduced here to align the visual and mathematical representations.

$$D_{max\_outer} = \sup \max\{|\overline{F}(x) - \underline{G}(x)|, |\underline{F}(x) - \overline{G}(x)|\}, \tag{3}$$

$$D_{max\_inner} = \sup\ D_{F,G}, \tag{4}$$

with

$$D_{\underset{inner}{\max}(x)} = \begin{cases} \{0\ if [\underline{F}(x), \overline{F}(x)] \bigcap [\underline{G}(x), \overline{G}(x)] \neq 0 \\ \min\{|\overline{F}(x) - \underline{G}(x)|, |\underline{F}(x) - \overline{G}(x)|\}\ otherwise \end{cases}$$

4 METHODOLOGY

4.1 Code

An imprecise KS test was developed from the *ksstats* function within the KolmogorovSmirnov.jl source code [11] that is part of the HypothesisTests.jl Julia package [12]. The code can be made available on request. The *ksdistances* function accepts arrays of thickness values representing $\overline{F}$, $\underline{F}$, $\overline{G}$ and $\underline{G}$ as arguments. It returns the number of samples in F, the number of samples in G, the maximum outer KS distance (D_{max_outer}), and the maximum inner KS distance (D_{max_inner}). Returning the number of samples allows the critical KS distance to be calculated and then compared against the two D_{max} values. The code was tested using the interval-valued date provided by Grzegorzewski [Table 1 in 9]. Pluto notebooks [13] were used to perform the analysis.

4.2 Data

The data used in the analysis is piping thickness data from the produced water system of an offshore oil and gas platform. It was obtained through non-destructive testing, and stored in a inspection data management system. The data is company confidential and so cannot be shared.

The measurement uncertainty (measurement error) of the measurements is not stated. Assumed nominal values are used for the measurement incertitude of the thickness readings [14]. To make it possible to include more data in each sample, the data was normalised against the nominal thickness value by dividing the measured thickness by the quoted nominal thickness value. This allows piping of different nominal thicknesses to be included in the same sample. Where no nominal thickness was provided for the thickness monitoring location, that location was excluded from further analysis. To reduce the potential for outliers the data was filtered so that only equipment of type "Pipe" was included in the sample.

4.3 Process to evaluate inspection planning decisions

Both datasets are historical datasets obtained before the research was done. Consequently, inspection 'planning' decisions have to be evaluated by using the data already collected. The process to evaluate inspection planning decisions is as follows.

1. Calculate the outer and inner maximum KS distances between two imprecise ECDFs (each forming a p-box) created from a sample of inspection thickness readings from two different inspections carried out some time apart.
2. Calculate the critical KS distance at the 0.05 significance level.
3. Test whether one, both or none of the D_{max} distances exceed the critical distance and thus the extent to which the null hypothesis, that the distributions are the same, can be rejected.
4. Modify the sample from the second inspection to allow the effect on the KS distances of three aspects of inspection planning to be analysed.
    a. Change the measurement uncertainty associated with the sample uncertainty from the nominal value (±5%) to more accurate (±2.5%) and less accurate (±10%) values. These uncertainties are plausible for thickness measurements taken in the field as opposed to under lab conditions [3].
    b. Change the number of samples. The number of samples is reduced by half by selecting only every other value from the original sample. The number of samples is increased by duplicating the existing samples to create twice as many.
    c. Change the inspection interval (by taking inspection results from later inspections).
5. Record how the change in the inspection strategy affects the test outcome.

5 ANALYSIS AND RESULTS

Table 2 shows the data that was available for analysis once the pre-processing of the data was done as described in Section 4.2. Looking retrospectively at the data, the mean thickness indicates a trend of thickness loss over time. However, when planning the initial inspections of an asset such a trend will not have been established and it is important to understand whether any measured change in thickness between inspections is statistically significant, or whether it could have occurred by chance. The imprecise KS test will help to do this.

Table 2: Number of thickness measurements in each sample

Inspection year	Number of samples	Mean normalised thickness	Min. normalised thickness
2008	43	1.03	0.89
2009	17	0.99	0.91
2011	55	0.96	0.86
2014	72	0.92	0.59

The approximate rejection threshold for the 0.05 probability level of 1.36 is used [10]. This approximate value can be used because the samples sizes are large. Note the 2009 data has a sample size below 30 (but above that available in most tables for *small* samples), which will give slightly worse approximation.

Before the imprecise data was assessed, an approximate two sample KS test was applied to the sample data from 2008 and 2009 using the original scalar values without considering uncertainty. The test failed to reject the null hypothesis, meaning that when measurement uncertainty is not considered there is insufficient distance between the ECDFs created from the two samples to conclude that they are different (and therefore thickness loss due to corrosion could be occurring).

The imprecise data incorporating measurement uncertainty was then analysed. The results in Table 2 were obtained through the process to evaluate inspection planning decisions described in Section 4.3. Eight numbered scenarios were considered. #1 can be considered the base case and #2 - #8 can be considered inspection planning options that vary one or more of measurement uncertainty, sample count and inspection interval.

Table 3: Results of applying imprecise KS test on different inspection planning strategies

#	Sample 1	Sample 2	Measurement Uncertainty	Sample ratio	D_{max} inner	D_{max} outer	Critical Distance	Test result
1	2008	2009	5%	100%	0.1272	0.8714	0.3896	Reject outer
2	2008	2009	2.5%	100%	0.2476	0.7660	0.3896	Reject outer
3	2008	2009	2.5%	200%	0.2120	0.7661	0.3121	Reject outer
4	2008	2011	5%	100%	0.3074	0.8677	0.2768	**Reject both**
5	2008	2011	2.5%	50%	0.7159	0.8123	0.3303	**Reject both**
6	2008	2011	10%	200%	0.1163	0.9412	0.3121	Reject outer
7	2008	2014	10%	100%	0.1860	1.000	0.2621	Reject outer
8	2008	2014	5%	50%	0.8055	0.8837	0.3072	**Reject both**

Scenario #1 shows that, in contrast to the precise data, when a plausible measurement incertitude of 5% is considered (equivalent to ±0.5 mm measurement uncertainty for 10 mm structural thickness) the imprecise KS test finds that the difference between the two distributions is only significant at the outer extreme of the interval. At the inner extreme of the interval, the distance between the two distributions has a greater than 5% likelihood of occurring by chance and so the null hypothesis cannot be rejected at the inner bound.

Scenarios #2 and #3 show that reducing the measurement uncertainty by half does not reduce uncertainty enough to reject the null hypothesis conclusively at both extremes of the interval, even if the number of samples is also doubled.

Scenario #4 shows that increasing the time interval between inspections from 1 to 3 years allows the default measurement accuracy and sample count to establish conclusively that there has been a significant change in the underlying thickness distribution. Scenarios #5 and #6 compare two alternative strategies at the 3-year inspection interval. The strategy to collect fewer, more accurate samples (#5) is shown to be more effective than the one to collect a greater number of less accurate samples (#6) because it leads to both the inner and outer D_{max} values being greater than the critical distance.

Scenarios #7 and #8 again show that a smaller number of more accurate samples leads to a more conclusive test result. By the time there is a six-year interval between the inspections, it is possible to reach a conclusive result rejecting both D_{max} values with the default measurement accuracy and only half the number of thickness measurement readings (#8). However, it is not possible to reject the null hypothesis at both extremes of the distributions when taking the standard number of samples with less accurate measurement (#9).

6 DISCUSSION AND FUTURE WORK

The experiment suggests that when planning inspections, expending resources into more accurate non-destructive testing techniques reduces uncertainty to a greater extent than expending them into increasing inspection coverage. However, without quantification of the relative costs this hypothesis cannot be verified. One area of future work would be to develop a decision-making framework for inspection planning that incorporates cost as well as uncertainty. The best way to reduce uncertainty in inspection data is to increase the interval between subsequent inspections. Of course, ignorance around the condition of equipment only grows during the period between inspections! The risk of failure occurring prior to the next inspection taking place should also be factored into any inspection planning framework.

The research presented has several limitations. The first limitation is the lack of an in-depth engineering review of the data prior to the analysis. It would be worthwhile to do more pre-processing of the data. One step would be to identify and remove outliers, which should lead to more similar tail shapes between samples and less chance of the maximum distances being driven by outliers in the tails. High-value outliers can be caused by measurements of particular components within the “Pipe” equipment type, such as “Tee” sections of pipe being included with measurements of standard straight pipe sections. Low values can be caused by erroneous measurements, but care would be needed not to remove low thickness values without justification. Another way to pre-process the data would be to impute nominal thickness values from similar thickness monitoring locations, which would allow more samples to be used in the test, leading to a reduced critical distance. A final pre-processing step would be to divide the population of thickness monitoring locations into *corrosion circuits* [14] where similarities in materials, environmental conditions and duty mean the presence and rate of corrosion is likely to be similar.

Another limitation of the research is that it used a both-tailed version of the KS test. As corrosion can only ever reduce the thickness of structures, a left-tailed version of the KS test could be used, which would make the test more sensitive to changes in the lower tail of the distribution, which is where any thickness loss due to corrosion would be apparent. Moreover, the KS is only one type of goodness of fit test. Other goodness of fit tests like the Anderson-Darling test and the Chi-squared test could be similarly adapted for uncertain data.

Thickness Monitoring Locations have unique identifiers, which means that a paired sampling approach could be used to test for thickness loss. Paired sampling approaches are generally more powerful than population sampling ones. There are challenges with applying paired sampling to inspections data. These include the lack of methods in the industry for dealing with epistemic uncertainties (measurement imprecision), the difficulty in field-based inspection of inspecting exactly the same location to give a true ‘pair’ of samples, and the need to inspect different areas of an asset over its operating life to ensure a reasonable coverage of the overall surface area. Future work would be to apply the imprecise t-test methodology for interval-valued data proposed by Kreinovich and Servin [15]. The final area of future work is to develop a means of calculating the *rate* of corrosion to complement the test presented in this work. Imprecise regression techniques could be used to calculate an imprecise corrosion rate that would take the epistemic and aleatory uncertainties into account.

It can be concluded that uncertainty should be an important consideration in inspection planning and analysis. Poor decisions with cost and risk implications are possible if uncertainty is not factored into inspection and integrity management. Use of uncertainty-aware statistical tests such as the imprecise KS test has been shown in this paper to be one way of taking uncertainty into account. To reduce uncertainty, inspection planners should prioritize increasing the accuracy of thickness measurements over increasing the number of thickness measurements in future inspection scopes. This is especially the case when locations most at risk of corrosion can reliably be identified, allowing the choice of measurement locations to be targeted at susceptible areas. Where the risk allows, it is also worthwhile extending the inspection interval to increase the likelihood of the ‘signal’ of thickness loss being detectable within the ‘noise’ of the uncertainty in the data, especially where corrosion is suspected but cannot be proven by thickness measurements taken to date.

REFERENCES

[1] Gray, A., Wimbush, A., De Angelis, M., Hristov, P. O., Calleja, D., Miralles-Dolz, E., and Rocchetta, R., 2022, “From Inference to Design: A Comprehensive Framework for Uncertainty Quantification in Engineering with Limited Information,” Mech. Syst. Signal Process., **165**, p. 108210. https://doi.org/10.1016/j.ymssp.2021.108210.

[2] Benstock, D., and Cegla, F., 2017, “Extreme Value Analysis (EVA) of Inspection Data and Its Uncertainties,” NDT E Int., **87**, pp. 68–77. https://doi.org/10.1016/j.ndteint.2017.01.008.

[3] Drury, J. C., 1997, “Corrosion Monitoring and Thickness Measurement – What Are We Doing Wrong?’,” Insight Non-Destr. Test. Cond. Monit., **39**(1), pp. 17–23.

[4] BSI, 2019, “BS EN ISO 16809 Non-Destructive Testing - Ultrasonic Thickness Measurement.”

[5] MIL, 2009, *Nondestructive Evaluation System Reliability Assessment*, MIL-HDBK-1823 A, Department of defense. [Online]. Available: https://statistical-engineering.com/wp-content/uploads/2017/10/MIL-HDBK-1823A2009.pdf. [Accessed: 28-Nov-2024].

[6] Stone, M., 2013, *HOIS Recommended Practice for Statistical Analysis of Inspection Data*, HOIS(12)R8 Issue 1, HOIS / ESR Technology Limited.
[7] Brown, C. H., and Ferson, S., 2025, "Mixed Uncertainty Analysis in Pressure Systems Inspection Applications," ASCE-ASME J. Risk Uncertain. Eng. Syst. Part B Mech. Eng., **11**(2). https://doi.org/10.1115/1.4068152.
[8] Ferson, S., Kreinovick, V., Ginzburg, L., and Sentz, F., 2003, *Constructing Probability Boxes and Dempster-Shafer Structures*, SAND2002-4015, 809606. https://doi.org/10.2172/809606.
[9] Grzegorzewski, P., 2018, "The Kolmogorov–Smirnov Goodness-of-Fit Test for Interval-Valued Data," *The Mathematics of the Uncertain*, E. Gil, E. Gil, J. Gil, and M.Á. Gil, eds., Springer International Publishing, Cham, pp. 615–627. https://doi.org/10.1007/978-3-319-73848-2_57.
[10] Destercke, S., and Strauss, O., 2014, "Kolmogorov-Smirnov Test for Interval Data," *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, A. Laurent, O. Strauss, B. Bouchon-Meunier, and R.R. Yager, eds., Springer International Publishing, Cham, pp. 416–425. https://doi.org/10.1007/978-3-319-08852-5_43.
[11] Sawade, C., 2014, "Kolmogorov_smirnov.Jl." [Online]. Available: https://github.com/JuliaStats/HypothesisTests.jl/blob/master/src/kolmogorov_smirnov.jl. [Accessed: 30-Apr-2025].
[12] Kornblith, S., 2012, "HypothesisTests.Jl." [Online]. Available: https://github.com/JuliaStats/HypothesisTests.jl/tree/master. [Accessed: 30-Apr-2025].
[13] Fons van der Plas, Michiel Dral, Paul Berg, Παναγιώτης Γεωργακόπουλος, Rik Huijzer, Mikołaj Bocheński, Alberto Mengali, Connor Burns, Hirumal Priyashan, Benjamin Lungwitz, Jerry Ling, Eric Zhang, Felipe S. S. Schneider, Ian Weaver, Xiu-zhe (Roger) Luo, Shuhei Kadowaki, Gabriel Wu, Timothy, Luis Müller, Zachary Moon, Supanat, Sergio A. Vargas, Rok Novosel, Jelmar Gerritsen, Vlad Flore, Jeremiah, Ciarán O'Mara, and Michael Hatherly, 2024, "Fonsp/Pluto.Jl: V0.20.4." https://doi.org/10.5281/ZENODO.4792401.
[14] Conder, P., and Stone, M., 2018, *HOIS Guidance for More Effective Pipework Inspection*, HOIS-G-010 Issue 1, HOIS / ESR Technology Limited. [Online]. Available: https://www.bindt.org/shopbindt/hois-documents/HOIS-Guidance-for-More-Effective-Pipework-Inspection-HOIS-G-010-Issue-1.html.
[15] Kreinovich, V., and Servin, C., 2015, *How to Test Hypotheses When Exact Values Are Replaced by Intervals to Protect Privacy: Case of t-Tests*, UTEP-CS-15-09, The University of Texas at El Paso, El Paso. [Online]. Available: https://scholarworks.utep.edu/cs_techrep/892. [Accessed: 30-Apr-2025].