Analytic Derivation and Validation of Probabilistic Corrosion Growth Model

Yevgeniy Petrov[1], Kregg Philpott[1], Megan Scudder[1]

[1] OneBridge Solutions Inc.

Organized by

Proceedings of the 2024 **Pipeline Technology Conference** (ISSN 2510-6716).

www.pipeline-conference.com/conferences

1 ABSTRACT

Corrosion poses a significant and persistent challenge in the oil and gas industry. Pigging is the most widely used technology used by operators to sample corrosion progress at different times during pipeline operation. Despite improvements in recent years, measuring anomalies with inline inspection (ILI) tools is subject to significant uncertainties. One or several ILI measurements of depths along with their uncertainties and times are required to estimate the future growth of a given anomaly.

In this work, we develop a probabilistic growth model under an assumption that individual depth errors are normally distributed. Its advantage is that it is derived analytically as a natural extension of simpler linear models. As a result, it does not require costly computations to run as opposed to models based on Monte Carlo (MC). The model provides direct predictions of depths at future times without relying on growth. More generally, it outputs probability distributions of both true depth and growth based on historical ILI depths and their tolerances.

We utilize an extensive dataset to validate and analyze the model. 165K miles of ILI data and 50K of repaired metal loss anomalies from 1294 pipelines of 4 large pipeline operators were used. We compare the probabilistic model with other models used in the industry. Various aspects and advantages of its application in pipeline integrity are considered. For example, a depth probability distribution predicted by the model at a specific future time could be used when calculating probability of exceedance for pipeline burst pressures defined with the modified B31G model.

1 INTRODUCTION

Understanding pipeline corrosion process comprehensively is one of the most important challenges presented to operators under their integrity management programs (IMPs). This task depends on a multitude of input factors such as environmental and operational conditions, availability of ILIs and other measured data, and complexity of underlying corrosion models and interpretability of their results.

This paper focuses on a data-driven approach using only ILIs and their measurement uncertainties as the model inputs. It is practical for operators as ILI is widely used data collection technique and the model is straightforward to implement and to interpret. In a previous study [1], we found that regression and probabilistic MC models were preferable for long-term pipe replacement planning while more conservative models might suit better for ongoing excavation planning programs.

Many corrosion models employed by the industry currently ignore uncertainties in ILI depth measurements. Research is ongoing to explore and compare models of different complexities which can take depth and other input uncertainties into account to produce a probabilistic output [2]. While the most complex models are presumably more accurate, there is not enough validation data for different types of corrosion under different environmental scenarios to conclude that any specific model would be the best choice for operators. Some complex models require extensive inputs, computational power, and time, making them impractical and challenging for non-specialists to interpret.

In [1], we examined a simple MC-based probabilistic model utilizing only historical ILI depths and tool tolerances. The model's accuracy matched another high-performing model based on an error-weighted linear regression. However, the model had drawbacks such as inconsistent treatment of corrosion growth rate calculation by using linear regression through the most probable output depths

instead of predicting the most probable depth directly at any future time. It also required extensive computational resources for a scaled implementation due to requirement of running MC for each set of aligned anomalies.

This paper introduces an efficient analytical probabilistic growth model. It processes ILI depth measurements, errors, and times as inputs, generating a probability distribution of future depth. The derivation of the model relies on the assumption of normally distributed depth errors and statistical independence of successive ILI measurements. Extending the PRCI[1] PR-331 model from [3, 4], which used two-point inputs and MC simulations, we provide an analytical solution applicable to any number of input points. This approach has the advantage of utilizing all available data coherently, providing an advantage to operators who perform ILIs frequently. In contrast, the common two-point model may penalize operators by yielding more uncertain results or requiring operators to choose specific ILIs while discarding the rest.

Calculation of corrosion growth rates is an intermediate step towards predicting future depths used in models like the ASME[2] modified B31G and in determining dig conditions. The probabilistic model outputs depth distributions directly and offers a natural way to assign depth uncertainty. For instance, using predicted depth probability distributions improves probability-of-exceedance (PoE) calculations as compared to assuming a normal distribution based on the latest tool run.

2 METHODOLOGY

Following the alignment process outlined in [1], we identify and match anomalies across all ILIs used in the study, forming "*chains*". While one-to-one "*single chain*" matches are common, some scenarios necessitate many-to-one or many-to-many anomaly matching (*multi-matching*). For example, a cluster from one ILI may match several anomalies in the next ILI due to the difference in tool technologies. In this paper, we use single chains as the input for the models. A method described in Section 2.2 provides a way to reduce collections of multi-matched anomalies into single chains and to extend the size of the input dataset.

2.1 DATA COLLECTION AND INITIAL FILTERING

We analyzed ILI data from 1294 pipelines across 4 major operators. The initial dataset is significantly reduced due to the following requirements. An anomaly must be linked to a field repair in the pipe's maintenance history, specifically of the "metal loss" type. The models require a minimum of 3 historical ILI data points, further reducing the dataset size. Most corrosion growth rates are moderate, so there is insufficient time between the latest ILI and a field repair to exhibit a significant difference between model predictions. In order to better distinguish the models, we exclude the most recent point in all chains before analysis.

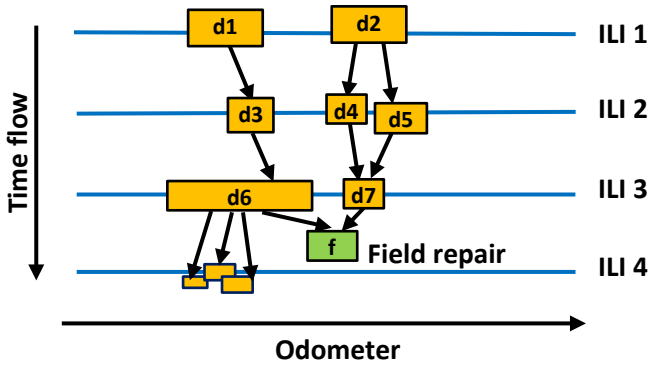
2.2 ANOMALY CHAIN SELECTION

After applying the filters described above, 966 single chains were identified. The dataset size can nearly double with additional 847 collections of multi-matched anomalies. We treat each anomaly collection

[1] Pipeline Research Council International, 15059 Conference Center Drive, #130, Chantilly, VA 20151, USA.

[2] The American Society of Mechanical Engineers, Two Park Avenue, New York, NY 10016-5990, USA.

Figure 1. Collection of multi-matched anomalies across spatially aligned IILs.



as a directional graph, with time flow determining the edge direction between ILI depths represented by graph nodes. We select anomaly chains associated with a field measurement. By backtracking all possible paths from the repair node to the earliest starting nodes, multiple paths may emerge. These paths are merged into a single chain through combining all depths per ILI into a single value per ILI. The combined depth per ILI is distributed normally with the mean and standard deviation given by the Equation 1 where r_i's and σ_i's are individual depths and standard deviations per ILI, and n is the total number of measured anomalies per ILI.

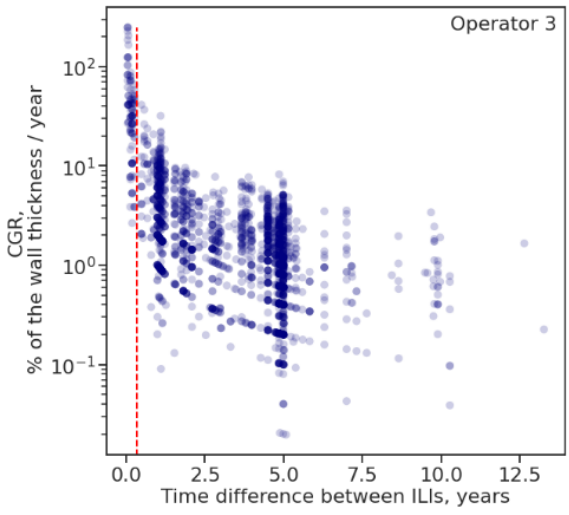
$$mean = \frac{\sum_{i=1}^{n} \frac{r_i}{\sigma_i^2}}{\sum_{i=1}^{n} \sigma_i^{-2}}, \qquad s.t.d. = \sqrt{\frac{1}{\sum_{i=1}^{n} \sigma_i^{-2}}} \tag{1}$$

Assuming independent and normally distributed measurements, this procedure is rigorously proven in [5]. Figure 1 depicts a potential collection of multi-matched anomalies. Paths "d1->d3->d6->f", "d2->d4->d7->f" and "d2->d5->d7->f" are combined into a single chain. The depths per ILI are combined according to the Equation 1 in the following groups: "ILI 1" -> anomalies "d1", "d2"; "ILI 2" -> anomalies "d3", "d4", "d5"; "ILI 3" -> anomalies "d6", "d7". Anomalies in "ILI 4" are a part of the collection but are excluded from possible paths during backtracking from the repair graph node "f" since they are detected after the field repair.

2.3 Averaging of ILI Measurements

Figure 2 illustrates the absolute values of the calculated corrosion growth rates (CGRs) plotted against time differences between successive ILIs across all chains in the data from one operator. Differences

Figure 2. Corrosion growth rates versus times between ILIs

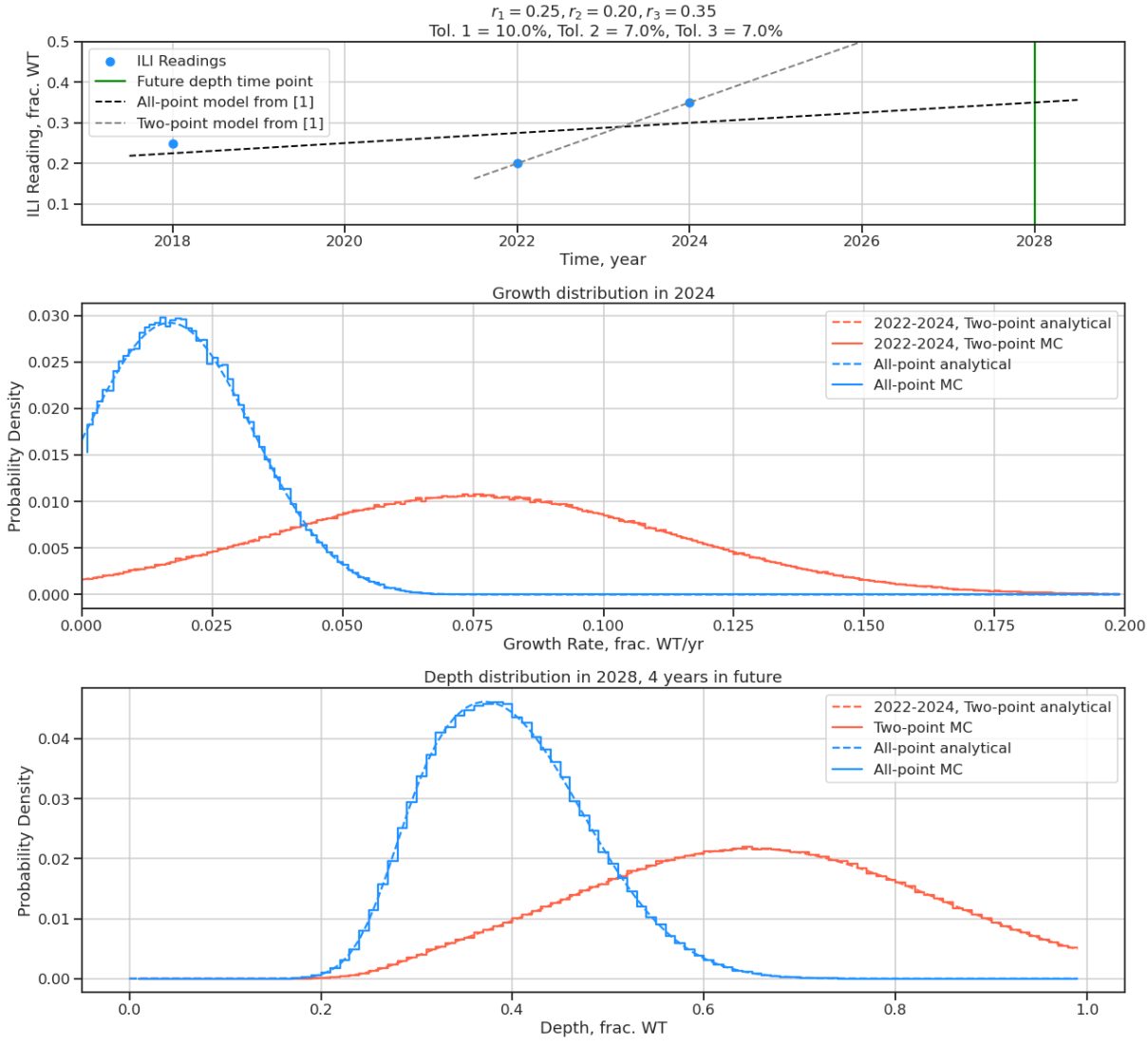


in depth divided by very short time intervals produce a spike of unrealistic CGRs. To address this, we employ an averaging procedure for successive ILI measurements conducted for a specified threshold. The threshold is denoted by a red dashed line at 0.35 years shown in the Figure 2. Similar thresholds, determined from analogous plots, were established for other operators. The averaging is done per chain by assigning a weighted mean and error to the resulting depth value according to Equation 1. New synthetic depths are assigned timestamps equalled to the average of the input times.

2.4 DESCRIPTION OF THE MODELS

We explore several linear models in [1] because linear models are popular in the industry due to their simplicity and interpretability. The simplest among them is the *two-point model*, commonly used for deterministic CGR calculation per anomaly. However, this model can yield overestimated or negative CGRs, limiting its applicability and necessitating corrections. The weighted *all-point model* emerges as a superior choice when more accurate but less conservative estimates are desired. We employ these two models for comparison with the analytical probabilistic model introduced in this paper. A concise description of the two models is provided below, with a comprehensive explanation available in [1].

Figure 3. (Top) A single chain of 3 ILI measurements performed in 2018, 2022, and 2024 with the simple 2-point and all-point trend models from [1]. (Middle) CGR distributions as calculated by the probabilistic model in this work. (Bottom) Predicted depth distributions corresponding to extrapolating the model to 2028.



In the two-point model, a nominal CGR of a matched anomaly is defined as

$$CGR = \frac{\Delta r}{\Delta t} \tag{2}$$

where $\Delta r = r_2 - r_1$ is the depth difference between two ILI readings and Δt is the time elapsed between the two ILI measurements.

The all-point model uses a weighted linear regression where the least-squares minimization assigns weights to measurements based on the reciprocal of their errors [6]. Measurement errors are converted from the tolerances specified by ILI vendors. To prevent unrealistic negative CGRs, the fit is constrained with the requirement that the slope must be non-negative. Any negative values are consequently replaced with zero growth.

2.4.1 Probabilistic Model

Operators commonly rely on ILI measurements as the primary and sole data source for calculating CGRs. The probabilistic model introduced in this section leverages all available ILI measurements and their uncertainties. Typically, CGRs are employed to predict future metal loss for determining excavation conditions or estimating remaining life. In deterministic model approaches, such as those discussed earlier, assigning proper uncertainty to future depth predictions without assumptions about the underlying distribution is challenging or even impossible. For instance, conservative B31G calculations often add the latest ILI tolerance to the predicted depth. The key advantage of the probabilistic approach lies in the model's output – the depth distribution calculated at any future time. With this distribution, operators can naturally identify the most probable future depth and customize the level of uncertainty by selecting desired percentiles from the distribution.

Appendix A has a rigorous derivation of the probabilistic model for two-point and all-point cases. The model is a generalization of the PR-331 described in [3, 4], which uses MC simulations and a fitting procedure to approximate future CGRs. However, an analytical solution is attainable when input depth probability density functions are assumed to be truncated normal distributions – an assumption reasonably applicable to ILI depth measurements.

The simple linear trend models from [1] can exhibit starkly different predicted depths depending on how many measurement points are included (see Figure 3). The top plot shows a single chain of three ILI measurements performed in 2018, 2022 and 2024 with the trend lines showing evolution of corrosion into future using two-point and all-point models described above.

With the probabilistic model, the uncertainty in the CGR distribution is significantly smaller for the all-point case as compared to the two-point case as shown in Figure 3 (middle). The most probable CGRs correspond to the peak values of the distributions and equate to the slope of the trend models in the top plot. The most probable value in this approach is not necessarily the same as the mean since the output distribution is not normal or symmetric in general.

The width of the future depth distribution in 2028 is significantly smaller in the all-point case meaning a smaller uncertainty in the most probable value compared to the two-point case (see bottom of Figure 3). Thus, both depth and growth distributions indicate an advantage of using a larger number of historical ILI measurements for operators.

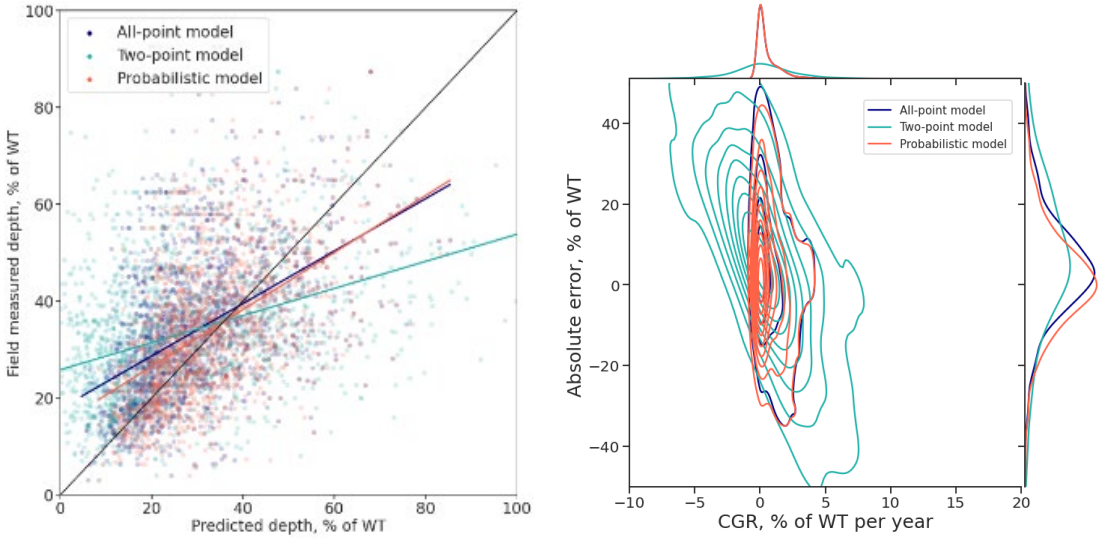
We performed 10^7 MC simulations where growth and depth values are sampled directly from the underlying truncated normal distributions. The results of the simulations are shown by the histograms

in the middle and bottom plots of Figure 3. There is an excellent agreement between the MC and analytical results which serves as an additional validation of the model.

3 RESULTS

We use unity plots, a widely accepted method in the pipeline industry, to evaluate model performance. In the probabilistic model, the predicted depth is defined as the most probable depth derived from the output depth distribution. A comparison of the models in Figure 4 (left) demonstrates that the probabilistic model predictions exhibit a trend line closest to the unity line, while the two-point model deviates the most.

Figure 4. Unity plot for all models (left). Absolute error versus CGRs (right).



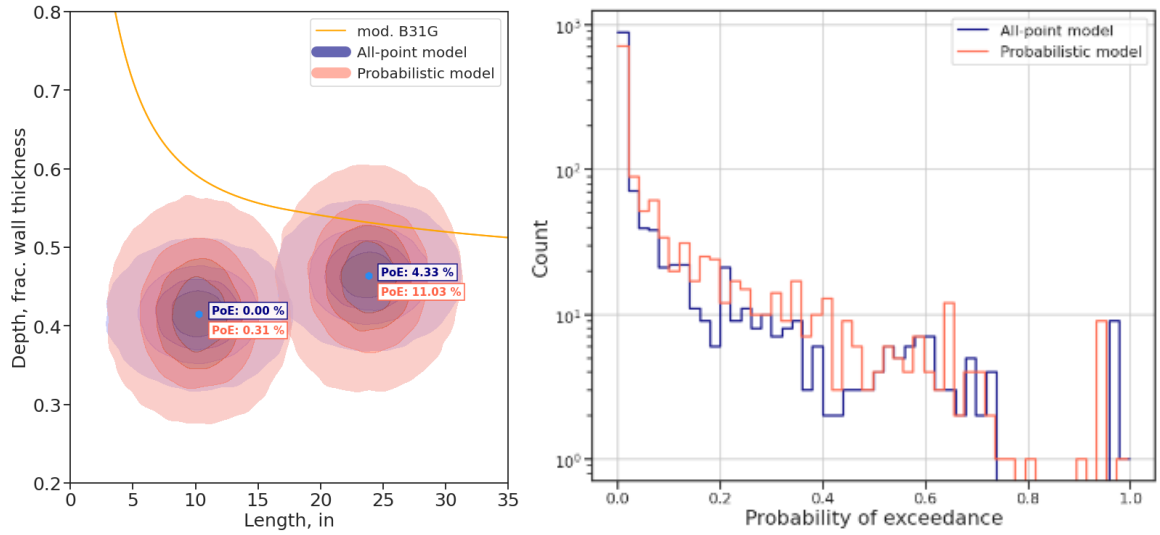
Model performance in terms of its accuracy and conservatism can be represented by a plot of absolute errors against CGRs. An *absolute error* is defined as the difference between the field-measured depth and the predicted depth. The concept of such representation was introduced and extensively described in [1]. Figure 4 (right) shows results for the three models in this study with kernel-density estimations shown on the side plots. The two-point model produces highly over-estimated results in the region of high CGRs. Both all-point and probabilistic models exhibit greater accuracy. The probabilistic model tends to produce slightly over-estimated results.

3.1 MODEL APPLICATION EXAMPLE: PROBABILITY-OF-EXCEEDANCE CALCULATION

Tool tolerances provided to operators by tool vendors can be used to determine the potential state of given anomalies and their proximity to a critical scenario, typically using a burst model. In this section, we consider the ASME modified B31G burst model, which uses the flaw length and depth, pipe properties and maximum operating pressure to predict pipe failure. Using this model integrity engineers perform the probability-of-exceedance (PoE) calculation to have a complete analysis of the possible state of the pipeline in between inspection cycles.

Using the maximum allowable operating pressure in addition to the pipe properties of diameter, wall thickness, and SMYS, the burst curve can be calculated for all depths between zero and the wall

Figure 5. Example calculations (left) and overall distribution of PoE values (right).



thickness using the modified B31G model. In the common scenario, flaw length and future depth are treated as continuous probability distributions according to the tolerances specified by the tool vendor. On the other hand, given the output depth probability distribution from the probabilistic model, one can use it as a direct input to the PoE calculation to achieve improved results.

Figure 5 (left) depicts an example of PoE calculation for two anomalies. Contours show 2D probability density isolines and the PoE is calculated using random samples of 10^5 values distributed around the ILI-reported length and depth calls. The shown examples illustrate that the PoE increases due to the broader depth distribution from the probabilistic model as compared to the normal distribution which is typically used. The right plot shows the distribution of PoE values calculated at repair times using all-point model with the nominal ILI-depth tolerance and the depth distributions taken from the probabilitic model output directly. The PoE values are generally shifted to larger values making the results more conservative.

4 CONCLUSIONS

We introduced and analytically derived a probabilistic model of corrosion which utilizes historical ILI-depth measurements and their uncertainties to the full extent possible. The model is derived under a reasonable assumption that the ILI depths are distributed according to the truncated normal distribution. The equations of the model were validated using MC simulations to produce identical results up to the precision of MC.

Operators can use the future depth probability distribution from the model as a direct input to other calculations to make decisions regarding their integrity management programs. Explicit probability distributions can provide operators with a practical way to assign a realistic uncertainty on the future depth value. For example, it can be defined as borders of the 34% of the distribution to the left and right of the most probable value. Also, one can use the “N-th” percentile to define an upper limit for desired level of conservatism. A PoE calculation provides with an example of a direct usage of the distribution to obtain realistic burst risk values.

APPENDIX A

Let us derive the probability distribution of the growth based on the assumption that it is linear. In the two-point case, the true depth values are related as $d_2 = d_1 + g\Delta t$. Depths d_1 and d_2 are sampled with ILI tools at different times and are assumed to be independent variables. Let us rewrite the sum above as $g\Delta t = d_2 - d_1$. The probability density of the sum of two independent variables is given by a convolution [7] resulting in the following expression in our case.

$$P_G(g) \propto \int_{d_2 \geq d_1} P_{D_2}(d_1 + g\Delta t)\, P_{D_1}(d_1) dd_1 \tag{A.1}$$

The region of integration is restricted to $d_2 \geq d_1$ since corrosion growth cannot be negative. It is reasonable to assume that depths d_1 and d_2 are distributed according to a truncated normal distribution with means corresponding to their ILI readings r_1 and r_2 and standard deviations corresponding to their reported ILI tool errors.

$$P_{D_i}(d_i) \propto \begin{cases} 0 & d_i < 0 \\ \dfrac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(d_i - r_i)^2}{2\sigma_i^2}} & 0 \leq d_i \leq 1\ , \quad i = 1, 2 \\ 0 & d_i > 1 \end{cases} \tag{A.2}$$

The distributions in (A.2) are truncated between 0 and 1 since the depths are expressed as percentage of wall thickness and un-realistic beyond this range. (A.1) becomes the following expression after substitution from (A.2) and completing the square with respect to d_1

$$P_G(g) \propto e^{-\frac{(g\Delta t - \Delta r)^2}{2\sigma^2}} \int_{d_2 \geq d_1} e^{-\frac{\sigma^2}{2\sigma_1^2\sigma_2^2}\left(d_1 - \frac{r_1\sigma_2^2 + (r_2 - g\Delta t)\sigma_1^2}{\sigma^2}\right)^2} dd_1 \tag{A.3}$$

where $\sigma^2 \equiv \sigma_1^2 + \sigma_2^2$. The integral in (A.3) can be rewritten with the following substitutions:

$$a \equiv \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}, \qquad b \equiv \frac{r_1}{\sigma_1^2} + \frac{r_2 - g\Delta t}{\sigma_2^2}, \qquad c \equiv \frac{r_1^2}{\sigma_1^2} + \frac{(r_2 - g\Delta t)^2}{\sigma_2^2} \tag{A.4}$$

Expression (A.3) then becomes:

$$P_G(g) \propto e^{\frac{1}{2}\left(c - \frac{b}{a}\right)} \int_{-\frac{b}{a}}^{1 - g\Delta t - \frac{b}{a}} e^{-\frac{a}{2}\left(d_1 - \frac{b}{a}\right)^2} d\left(d_1 - \frac{b}{a}\right) \tag{A.5}$$

One can notice that the lowest value of d_1 is 0 and the highest value is $1 - g\Delta t$ when $d_2 = 1$. Thus, the integration limits can be expressed explicitly in (A.5) after a variable substitution, $d_1 \to d_1 - \frac{b}{a}$.

The integral in (A.5) can now be written explicitly in terms of error functions as follows:

$$P_G(g) \propto e^{\frac{1}{2}\left(c - \frac{b}{a}\right)} \left(Erf\left\{ \sqrt{\frac{a}{2}} \left(1 - g\Delta t - \frac{b}{a} \right) \right\} + Erf\left\{ \frac{b}{\sqrt{2a}} \right\} \right) \tag{A.5}$$

Let us now extend the above result to the case of N ILI measurements performed at times t_i where $i = 1 \ldots N$. The progression of true depths d_i can be expressed with a set of following equations:

$$\begin{aligned} d_2 &= d_1 + g\Delta t_{1,2}, \\ d_3 &= d_2 + g\Delta t_{2,3}, \\ &\dots \\ d_N &= d_{N-1} + g\Delta t_{N-1,N} \end{aligned} \tag{A.6}$$

One can notice from (A.6) that $d_N = d_1 + g\Delta t_{1,N}$ where $\Delta t_{1,N}$ is the time between the first and the last measurement since the growth, g, is assumed to be constant.

Under assumption of all ILI measurements being independent, the joint probability density of all true depths is given by the product of the individual densities. Each individual probability density can be expressed as a function of d_1 using (A.6) so one can express the probability density of growth as

$$P_G(g) \propto \int_{d_N \geq d_{N-1} \geq \cdots \geq d_2 \geq d_1} P_{D_N}\left(d_1 + g\Delta t_{1,N}\right) P_{D_{N-1}}\left(d_1 + g\Delta t_{1,N-1}\right) \dots P_{D_1}(d_1) dd_1 \tag{A.7}$$

With d_1 being the only variable of integration, the domain is simply $d_N \geq d_1$ i.e. d_1 can have values from 0 to $1 - g\Delta t_{1,N}$ when the last depth value $d_N = 1$.

Let us consider how to obtain explicit expression for $P_G(g)$ in a 3-point case. After substituting explicit probability distributions of the type given in (A.2), we have:

$$P_G(g) \propto \frac{1}{\left(\sqrt{2\pi}\right)^3 \sigma_1 \sigma_2 \sigma_3} \int_{d_3 \geq d_2 \geq d_1} e^{-\frac{\left(d_1 + g\Delta t_{1,3} - r_3\right)^2}{2\sigma_3^2}} e^{-\frac{\left(d_1 + g\Delta t_{1,2} - r_2\right)^2}{2\sigma_2^2}} e^{-\frac{(d_1 - r_1)^2}{2\sigma_1^2}} dd_1 \tag{A.8}$$

Now let us complete the square with respect to d_1 in steps. Also, we will omit constants that do not depend on g. Firstly, we can use the result from (A.5) for completing the square for the two terms corresponding to the first two measurements and then proceed to complete the square for the full expression:

$$\begin{aligned} P_G(g) &\propto \int_{d_3 \geq d_2 \geq d_1} e^{-\frac{\left(d_1 + g\Delta t_{1,3} - r_3\right)^2}{2\sigma_3^2} + \frac{1}{2}\left(c - \frac{b}{a}\right) - \frac{a}{2}\left(d_1 - \frac{b}{a}\right)^2} dd_1 \propto \\ &\propto e^{\frac{1}{2}\left(\tilde{c} - \frac{\tilde{b}^2}{\tilde{a}}\right)} \int_0^{1 - d\Delta t_{1,3} - \frac{\tilde{b}}{\tilde{a}}} e^{-\frac{\tilde{a}}{2}\left(d_1 - \frac{\tilde{b}}{\tilde{a}}\right)^2} d\left(d_1 - \frac{\tilde{b}}{\tilde{a}}\right) = \\ &= e^{\frac{1}{2}\left(\tilde{c} - \frac{\tilde{b}^2}{\tilde{a}}\right)} \sqrt{\frac{2\pi}{\tilde{a}}} \left(Erf\left\{ \sqrt{\frac{\tilde{a}}{2}} \left(1 - g\Delta t_{1,3} - \frac{\tilde{b}}{\tilde{a}}\right) \right\} + Erf\left\{ \frac{\tilde{b}}{\sqrt{2\tilde{a}}} \right\} \right) \end{aligned} \tag{A.9}$$

After laborious algebraic manipulations between the first and second steps in (A.9), it can be shown that the result in (A.9) has the same form as the result for the two-point case, (A.5). The coefficients $\tilde{a}, \tilde{b}$ and $\tilde{c}$ are now defined as follows:

$$\begin{aligned} \tilde{a} \equiv \frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} + \frac{1}{\sigma_3^2}, \qquad \tilde{b} \equiv \frac{r_1}{\sigma_1^2} + \frac{r_2 - g\Delta t_{1,2}}{\sigma_2^2} + \frac{r_3 - g\Delta t_{1,3}}{\sigma_3^2}, \\ \tilde{c} \equiv \frac{r_1^2}{\sigma_1^2} + \frac{\left(r_2 - g\Delta t_{1,2}\right)^2}{\sigma_2^2} + \frac{\left(r_3 - g\Delta t_{1,3}\right)^2}{\sigma_3^2} \end{aligned} \tag{A.10}$$

One can notice that the procedure of completing the square in steps is the same for any number of probability density terms since the previous terms can be wrapped into a single term which is then

used to complete the square for the whole expression in the exponent. The final expression for the growth distribution function in the case of N measurements becomes:

$$P_G(g) \propto e^{\frac{1}{2}\left(c-\frac{b^2}{a}\right)}\left(Erf\left\{\sqrt{\frac{a}{2}}\left(1-g\Delta t_{1,N}-\frac{b}{a}\right)\right\}+Erf\left\{\frac{b}{\sqrt{2a}}\right\}\right) \tag{A.11}$$

where the coefficients a, b and c are re-defined as follows:

$$a \equiv \sum_{i=1}^{N}\frac{1}{\sigma_i^2}, \qquad b \equiv \sum_{i=1}^{N}\frac{r_i-g\Delta t_{1,i}}{\sigma_i^2}, \qquad c \equiv \sum_{i=1}^{N}\frac{\left(r_i-g\Delta t_{1,i}\right)^2}{\sigma_i^2} \tag{A.12}$$

Let us now formulate an expression for the depth probability distribution at a future time t passed since the time of the last measurement t_N. It is based on the previous ILI readings at times $t_1, t_2, \dots, t_N$. The probability of a certain depth occurring at time t is the sum of all the growth paths it could have happened weighted by the probability of each path. Assuming independence of individual depth probabilities $P_{D_1}, P_{D_2}, \dots, P_{D_N}$, the depth probability distribution can be expressed as

$$P_D(d,t) \propto \int_0^{g_{max}} P_{D_N}(d_N)\, P_{D_{N-1}}(d_{N-1}) \dots P_{D_1}(d_1) dg \tag{A.13}$$

where the maximum possible growth, g_{max}, occurs if corrosion starts from depth 0 at time t_1. One can notice that $d = d_1 + g\Delta t_{1,N} + gt$ and therefore $g_{max} = d/(t+\Delta t_{1,N})$.

It is convenient to introduce normalized time $\tau \equiv t/\Delta t_{1,N}$ and normalized growth $\delta \equiv g\Delta t_{1,N}$. Like in the set of equations (A.6), one can express $d_1, d_2, \dots, d_N$ in terms of the normalized variables as follows:

$$\begin{gathered} d_N = d - gt = d - \delta(\tau+\eta_N), \\ d_{N-1} = d - gt - g\Delta t_{N-1,N} = d - \delta(\tau+\eta_{N-1}), \\ \dots \\ d_1 = d - gt - g\Delta t_{1,N} = d - \ \delta(\tau+\eta_1), \end{gathered} \tag{A.14}$$

where $\eta_i \equiv \Delta t_{i,N}/\Delta t_{1,N}$. One can also note that $\eta_N = 0$ and $\eta_1 = 1$ by definition. By substituting (A.14) and normal distributions for $P_{D_1}, P_{D_2}, \dots, P_{D_N}$ into (A.13), we obtain the following expression.

$$\begin{aligned} P_D(d,\tau) &\propto \int_0^{\frac{d}{1+\tau}} e^{-\frac{(d-\delta\tau-r_N)^2}{2\sigma_N^2}} e^{-\frac{(d-\delta(\tau+\eta_{N-1})-r_{N-1})^2}{2\sigma_{N-1}^2}} \dots e^{-\frac{(d-\delta(\tau+\eta_1)-r_1)^2}{2\sigma_1^2}} d\delta = \\ &= \int_0^{\frac{d}{1+\tau}} e^{-\frac{(\delta+(r_N-d)/\tau)^2}{2\sigma_N^2/\tau^2}} e^{-\frac{(\delta+(r_{N-1}-d)/(\tau+\eta_{N-1}))^2}{2\sigma_{N-1}^2/(\tau+\eta_{N-1})^2}} \dots e^{-\frac{(\delta+(r_1-d)/(\tau+1))^2}{2\sigma_1^2/(\tau+1)^2}} d\delta \end{aligned} \tag{A.15}$$

By comparing the final expression in (A.15) with the expression for the probability of growth in (A.8), one can notice that the two have the same form with the following substitutions:

$$\sigma_i \to \frac{\sigma_i}{\tau+\eta_i}, \qquad g\Delta t_{i,N} - r_i \to \frac{r_i-d}{\tau+\eta_i}, \qquad g_{max} \to \delta_{max} = g_{max}\Delta t_{1,N} = \frac{d}{1+\tau} \tag{A.16}$$

We can now define coefficients, in a manner similar to (A.12), as follows.

$$A \equiv \sum_{i=1}^{N}\frac{(\tau+\eta_i)^2}{\sigma_i^2}, \qquad B \equiv \sum_{i=1}^{N}\frac{(\tau+\eta_i)(d-r_i)}{\sigma_i^2}, \qquad C \equiv \sum_{i=1}^{N}\frac{(d-r_i)^2}{\sigma_i^2} \tag{A.17}$$

The final expression for the depth distribution at a normalized time τ in the future since the last measurement is shown below.

$$P_D(d,\tau) \propto e^{\frac{1}{2}\left(\frac{B^2}{A}-C\right)}\sqrt{\frac{2\pi}{A}}\left(Erf\left\{\sqrt{\frac{A}{2}}\left(\frac{d}{1+\tau}-\frac{B}{A}\right)\right\}+Erf\left\{\frac{B}{\sqrt{2A}}\right\}\right) \quad \text{(A.18)}$$

The coefficient A depends on τ, and therefore the term $\sqrt{\frac{2\pi}{A}}$ is left in (A.18) whereas it was skipped in (A.11) previously for simplicity since in that case coefficient a was a constant.

REFERENCES

1. Y. Petrov, M. Scudder “Comparison of Corrosion Growth Models”, Proceedings of the Pipeline Technology Conference 2023.
2. G. Langlois-Rahme, I. Mahmoud, M. Safari, Y. Li, and C. Okoloekwe "A Comprehensive Empirical Evaluation of Probabilistic Corrosion Growth Rate Models." AMPP Annual Conference + Expo, Denver, Colorado, USA, March 2023.
3. J. Dawson, J. Wharf, M. Nessim, 2008, Development of Detailed Procedures for Comparing Successive ILI Runs to Establish Corrosion Growth Rates, PRCI project EC 1-2.
4. M. Nessim, J. Dawson, R. Mora, S. Hassanein “Obtaining Corrosion Growth Rates from Repeat In-Line Inspection Runs and Dealing with The Measurement Uncertainties” IPC2008, paper no. 64378 (Calgary, Alberta, Canada: ASME 2008).
5. T.P. Hill, “Conflations of probability distributions”, Transactions of the American Mathematical Society, 2011, Vol. 363, No. 6, pp. 3368, Theorem 6.1.
6. M. A. Branch, T. F. Coleman, and Y. Li, “A Subspace, Interior, and Conjugate Gradient Method for Large-Scale Bound-Constrained Minimization Problems,” SIAM Journal on Scientific Computing, Vol. 21, Number 1, pp 1-23, 1999.
7. H. Pishro-Nik, "Introduction to Probability, Statistics and Random Processes" (2014). Electrical and Computer Engineering Educational Materials, Example 5.30, https://scholarworks.umass.edu/ece_ed_materials/1