Comparison of Corrosion Growth Models

Yevgeniy Petrov[1], Megan Scudder[1]

[1] OneBridge Solutions Inc.

18TH PIPELINE TECHNOLOGY CONFERENCE
8–11 MAY 2023, BERLIN

Organized by

Euro Institute for Information and Technology Transfer

Proceedings of the 2023 **Pipeline Technology Conference** (ISSN 2510-6716).

www.pipeline-conference.com/conferences

ABSTRACT

The industry currently has multiple models to calculate corrosion growth between several ILI results. There is also little guidance around selecting a model for young pipelines with only one ILI run. Often, operators are tasked with the decision to select more conservative models that can result in an overestimation of digs or less conservative models which can result in high leak risks.

This research looks at the results of six different models using field-found data to determine the accuracy of the predicted depths. Also, we consider applicability of growth models in young pipelines and in negative growth cases, and where the results land on risk vs conservativism spectrum. Our analysis can help operators understand which growth models can be used to make better decisions in specific contexts. For example, short-term dig planning must be based on the instantaneous corrosion depth and growth rate, determined by the latest one or two depth measurements, but long-term pipe replacement planning may rely on a growth rate trend over several measurements.

1 INTRODUCTION

Metal loss poses a significant threat to pipeline integrity, with corrosion being the most common mechanism responsible. If left unaddressed, metal loss can ultimately lead to pipeline failure due to the burst pressure being lower than the operating pressure and/or the defect depth increasing to the critical threshold of wall thickness (WT). There are numerous models for assessing the remaining lifespan of a pipeline, such as the ASME[1] modified B31G, RSTRENG, and others. One of the primary inputs to these models is the corrosion growth rate (CGR), which can be calculated using a variety of corrosion growth models.

Corrosion growth models can be based on the physics and chemistry of the corrosion process, or they can be data-driven. Mixed approaches that incorporate elements of both are also possible. Data-driven models often utilize inline inspection (ILI) data as input. Another way to categorize growth models is by the practicality of their application. In general, ILI data-driven models are easier to implement and require simpler mathematical techniques such as regression and basic statistical approaches. In contrast, theory-based models often involve more advanced concepts such as Bayesian analysis and various stochastic processes. This paper focuses on corrosion growth models that are data-driven and practical, using field data to evaluate the accuracy of their depth predictions.

The first class of models we study is linear models, which assumes that corrosion growth rate is constant over time. The slope of the regression line through the input ILI data is typically taken to be the desired corrosion growth rate. We begin by examining the simplest two-point model, which estimates growth based on just two ILI measurements. We then consider extensions of this model that use a linear trend through all available ILI measurements, with or without accounting for individual measurement errors. Additionally, we evaluate a model known as PR-331, which is used by some operators in the industry.

The second class of models includes more mathematically sophisticated approaches in which the overall growth behavior is not necessarily linear. An exponential model uses an exponential law to

---

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

describe the evolution of depth, while a probabilistic model uses a Monte Carlo approach to determine the most likely path of corrosion evolution based on multiple ILI measurements.

The performance of the models is evaluated based on their predictions compared to the metal loss field data. The distribution of errors, defined as the difference between field measurements and model predictions, can indicate whether a model tends to overestimate or underestimate metal loss depth. It is important to note that overestimation or underestimation do not necessarily make a particular model "good" or "bad." Rather, we consider different scenarios of model application in the context of pipeline integrity, such as short-term dig planning versus long-term pipe replacement management. In this context, an operator can use the tools developed in this study to compare the models and evaluate the impact on a scale of conservatism versus cost-savings.

Lastly, we examine how to handle special-case measurements, such as negative growth rates and cases of unmatched anomalies in young pipelines where the usual half-life approach leads to unphysically high CGRs.

2 METHODOLOGY

Contemporary computational methods enable the automation of concurrent alignment of multiple ILIs throughout the pipeline's operational history [1]. The weld/joint patterns of the same physical pipe are clearly represented across its history by this alignment, and all modifications to the physical pipe can be tracked by the changes in these patterns. Once the same physical pipe has been identified, it becomes possible to identify and match anomalies across all ILIs. We refer to a set of matched anomalies as a “chain”. There are inevitable scenarios that are not well-described by a set of one-to-one matches, but require many-to-one or many-to-many anomaly matching (*multi-matching*). For instance, a cluster of anomalies detected by one ILI may be matched to several distinct anomalies in the next ILI, depending on the tool technologies employed. The topic of anomaly multi-matching and potential methods for resolving this ambiguity merits a separate study. In this paper, we only consider chains with a single anomaly matched across all ILIs, which we term *single chains*.

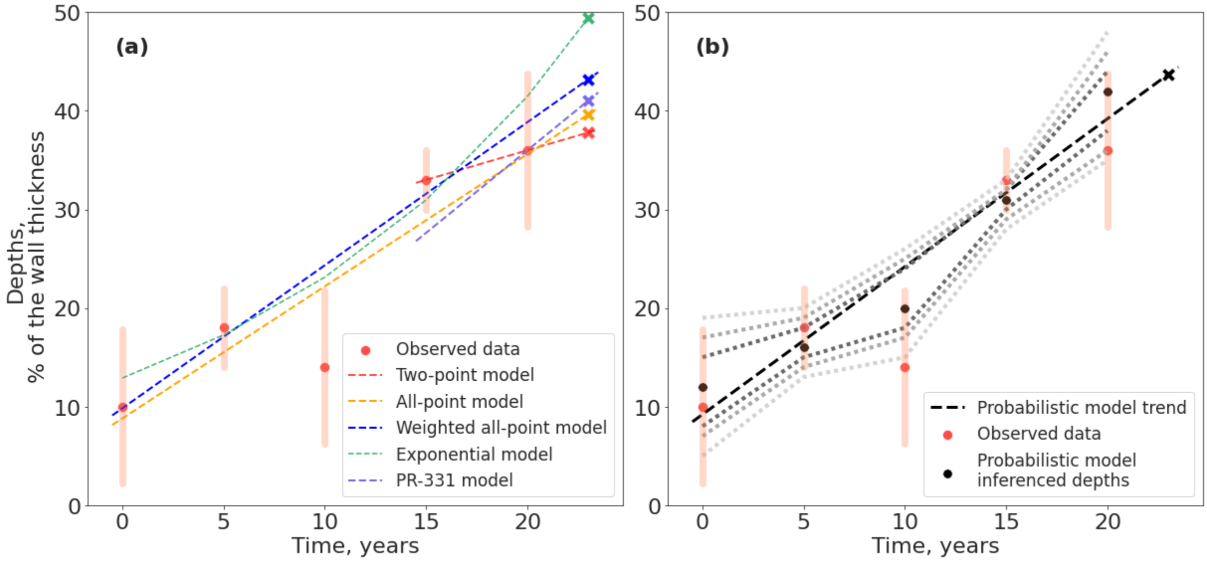
2.1 DATA COLLECTION

We analyze ILI data from 476 pipelines belonging to a single operator. The dataset size was significantly reduced by the two requirements that we consider only single chains of metal loss anomalies and that an anomaly must be associated with a field repair at some point along the pipe's maintenance history. These field data points must also be of the "metal loss" type. Additionally, some of the models under consideration necessitate at least 3 historical ILI data points, further diminishing the dataset size. In total, there are 2600 single chains of matched anomalies selected as a result.

2.2 MODEL DESCRIPTION

Linear models are often used when only two ILI results are available. Mathematical simplicity and tractability make them a popular choice in the pipeline industry. However, it must be noted that systemic and random errors in ILI measurements give rise to substantial deviations from actual corrosion trends, resulting in unrealistic outcomes such as negative or greatly overestimated CGRs.

Figure 1. An example of corrosion growth modelled by different models. Model predictions are shown as crosses. Half-transparent lines at each data point represent standard deviation based on tool tolerances. Gray dotted lines in plot (b) represent top 1st, 5th and 20th percentiles of the simulated chains according to the probabilistic model.



An example of a simulated single chain of a matched anomaly is shown in Figure 1. There are five observed data points at a five-year inspection interval. In this example, we make depth prediction based on the growth rates obtained by different models three years after the final ILI. In the next sections, we will discuss different types of models, commencing with a commonly encountered scenario of two ILIs.

2.2.1 Two-point Linear Model

Corrosion growth rate of a matched anomaly is determined as

$$CGR = \frac{\Delta r}{\Delta t} \tag{1}$$

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. Figure 1a shows an example of a two-point model trendline going through the last two data points with a slope equal to the CGR.

Distribution of CGRs produced by this model can be approximated by a normal distribution (for example, see [2]) with a substantial proportion consisting of unrealistic negative values. In this paper, we use an augmented version of the two-point model where the half-life model is employed to replace all negative CGRs with strictly positive values.

2.2.2 All-point Linear Model

Operators often possess more than two ILIs at their disposal. It is then logical to attempt to approximate the growth rates utilizing all the accessible measurements. Having more measurements should mitigate the impact of individual random errors. The simplest approach is a regression line fitted to all the measurements via least-squares method [3]. This is shown in orange in Figure 1a.

It is unlikely to encounter many single chains where consecutive depth values increase since operators perform field repairs or employ mitigation programs to halt aggressive corrosion growth. Therefore, we anticipate that models with multiple data points will generate less conservative outcomes, i.e. smaller growth rates than two-point models.

2.2.3 Weighted All-point Linear Model

As a more advanced version of the model described in the preceding section, we consider a weighted linear regression fit in which the least-squares minimization assigns a weight to each measurement in accordance with the reciprocal of its error [4]. The measurement errors are obtained from the tolerances specified by ILI vendors. In the absence of a tolerance for a specific ILI measurement, we use an error value corresponding to the commonly employed +/-10% tolerance at a confidence level of 80%. An example of a weighted linear trend is displayed in blue on Figure 1a. In this model, we impose that negative CGRs are unphysical, and thus, the fit is performed under a constraint that the slope must be non-negative and all the negative values are replaced with zero growth.

2.2.4 All-point Exponential Model

This model fits an exponential to all ILI measurements using a nonlinear fit, while considering tolerances on each reading. Negative growth rates are not permitted, and thus, a constrained fitting method is used as described in [5]. The green line in Figure 1a illustrates an exponential fit to the data points. In contrast to the all-point linear models, the exponential model is expected to yield more conservative results, as it tends to generate accelerated growth curves.

2.2.5 PRCI PR-331 Model

PRCI[2] PR-331 Model, employed by some operators in the industry, is based on a PRCI report and a research paper [6, 7]. The model considers probability distributions of the true depths $f_{D_1}(d_1)$ and $f_{D_2}(d_2)$. An assumption is made that the true depth cannot decrease, and it follows a linear growth described by $d_2 = d_1 + g\Delta t$. In this case, the probability distribution of the true growth rate can be expressed as a convolution as shown below.

$$P_G(g) \propto \int_{d_2 \geq d_1} f_{D_2}(d_1 + g\Delta t) f_{D_1}(d_1) dd_1 \tag{2}$$

True physical depths denoted with a symbol d in the above equations are theoretical whereas those denoted with r are measured depths observed with ILIs. The authors of the model utilize Monte Carlo simulations to sample this integral for a given observed rate and find that the true distribution can be approximated by a normal or Weibull distributions, depending on the value of the ratio between the given observed growth and the standard deviation of the growth distribution. Therefore, the ratio is defined as $\alpha = \Delta r/\sigma$. The original version of the model does not specify output for $\alpha < 0.25$ so we augmented the model in that region. The modelled growth rate can then be expressed as follows.

$$g_{PR331} = \begin{cases} 0 & \alpha \leq 0 \\ 4\alpha\mu_{weibull} & 0 < \alpha \leq 0.25 \\ \mu_{weibull} & 0.25 < \alpha \leq 3.0 \\ \dfrac{\Delta r}{\Delta t} & \alpha > 3.0 \end{cases} \tag{3}$$

In the above, $\mu_{weibull}$ represents the mean of the Weibull distribution. The specifics of the calculations can be found in the referenced papers.

2.2.6 Probabilistic Model

In this section, we extend the PR-331 model to be applicable to more than two ILI measurements and eliminate possible constraints by not utilizing any specific functional forms of the true growth probability distribution. The objective here is, given a single chain of ILI measurements with associated

---

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

errors, to determine the most probable chain of depths that returned those readings. There are three constraints on the true depths; true depths must be between 0 and 1 as fraction of wall thickness, the depth at any future time must be always the same or greater than any previous time, and the growth rate over time can only remain constant or increase.

We employ Monte Carlo simulations to sample numerous potential paths across the history of a given single chain. The feasible space of depth values between 0 and 1 is divided into B small bins at each ILI time. For example, for bins spaced at 0.01×WT, there will be 100 bins. For a chain with N data points, the total possible number of paths to sample is B^N. It can be demonstrated numerically that, given the above constraints, the fraction of available paths is 2^{3-2N} or lower for $N > 2$ making it computationally viable to explore this space in reasonable time. We run 10000 simulations per chain.

It can be shown that using Bayes analysis [8] one can estimate the joint probability distribution of the true depths of a chain based on the distributions of individual true depths. The procedure is:

1. Calculate the probability distribution of true depths given a measurement at each ILI in an anomaly chain. Practically, we use a truncated normal distribution above the detection threshold. The normal distribution is truncated at the threshold and in the interval from 0 to the threshold value, we use a uniform distribution. The combined distribution was normalized to have an integral of 1 to properly stitch both distributions at the threshold value. We use detection threshold deduced from the data when possible and a default value of 10% otherwise.
2. Multiply the probabilities together for each of these depths leading to a set of likelihoods.
3. Sort the likelihoods calculated and extract most likely path(s) and the bounds of other likely paths such as 1st or 5th percentiles.

The probabilistic treatment also enables inclusion of undetected anomalies at times where ILI measurements were performed along the timespan of a chain, but the tool was not able to detect or report an anomaly below the threshold. In this case, the model uses a distribution described in step 1 above to randomly draw a depth value below the threshold.

Figure 1b illustrates an example of most probable true depths extracted from the observed data using this approach. Once the true depths are calculated, we fit a linear regression model to obtain the CGR and make a prediction for the future depth.

For all models, we use times of the field data points to make depth predictions. An *absolute error* is defined as the difference between the field-measured depth and the predicted depth. As noted in the introduction, having smaller mean absolute error would not necessarily make a model better as it depends on the goals of an integrity program. Let us illustrate how one can start with the performance of a particular model, and then use our methodology to systematically compare with the other models and improve their integrity program accordingly. Figure 2 shows a representation of the two-point model performance as CGR versus absolute error. The vertical dense band of entries corresponds to the negative CGRs re-calculated using the half-life method. One can see that the coordinate space is divided into three regions. If one wants to improve a model performance, the strategy would be to replace or modify it in a such way that there are no entries in the unrealistic negative CGR part and to reach a suitable balance between conservatism and non-conservatism.

Figure 2. Absolute error versus CGR for the two-point model

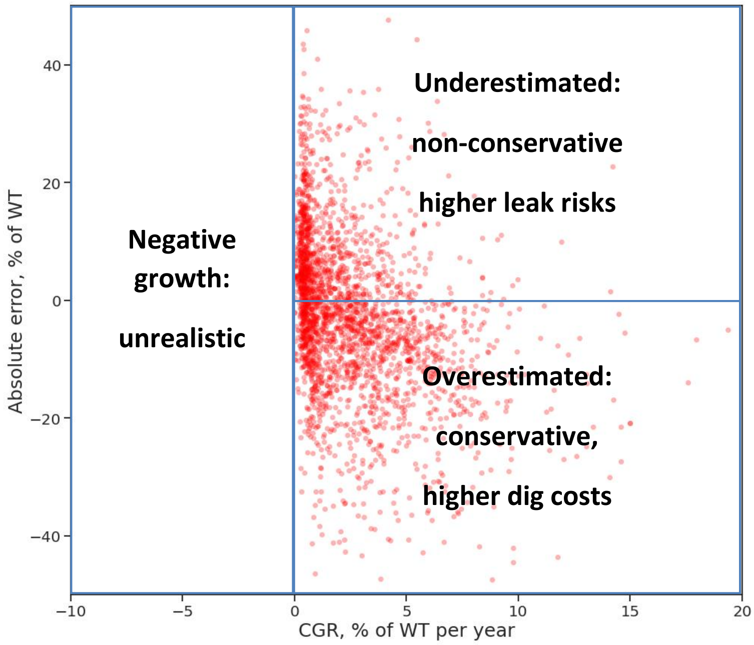


Figure 3 shows the distributions for the other models. The population of negative CGRs re-calculated with the half-life model is also shown, and it reveals an anticipated trend of higher CGRs in younger pipes.

The region of overestimated depths and high growth in Figure 2 is where the two-point model can be improved upon if the goal is to be less conservative and have a model for pipe lifetime estimates and replacement planning. Figure 4 shows results of the weighed all-point model on the same dataset. One can see that the absolute error distribution is centered at zero in the case of the weighted all-point model. By changing the model, we can change the balance between over- and -under prediction. Using

Figure 3. Absolute error versus CGR for other models and for the half-life model

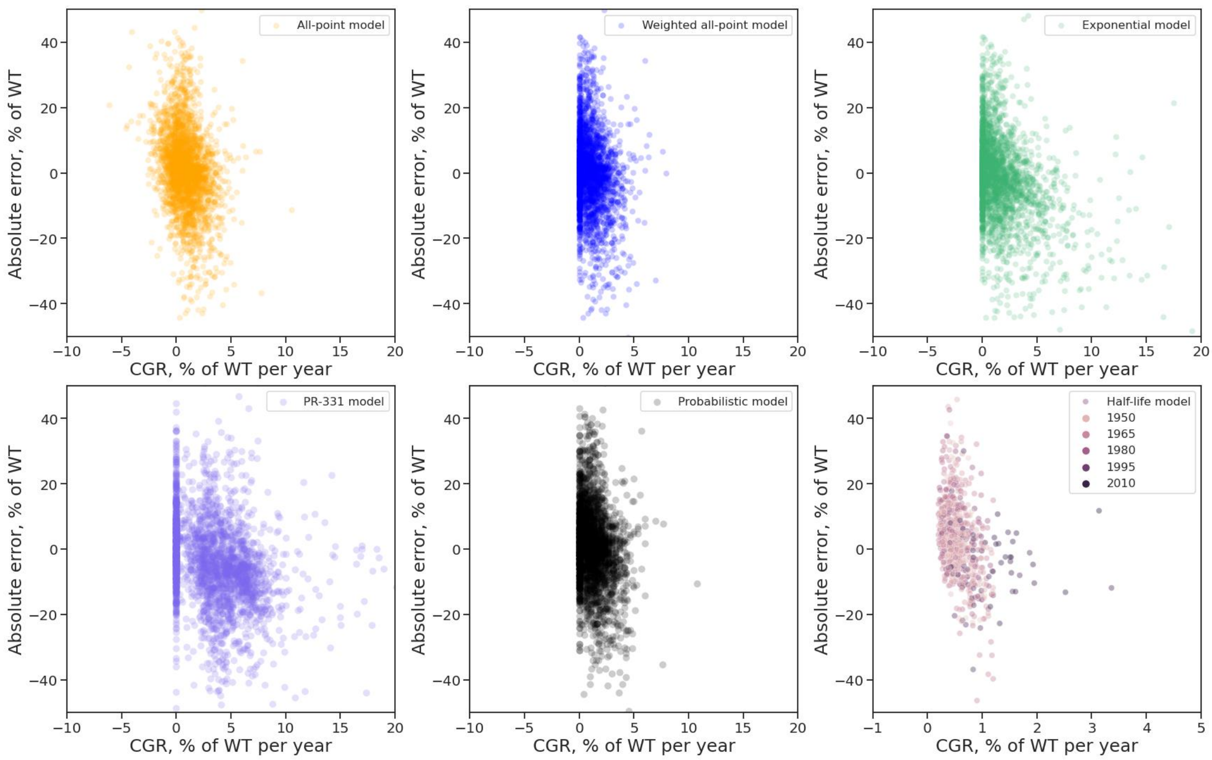
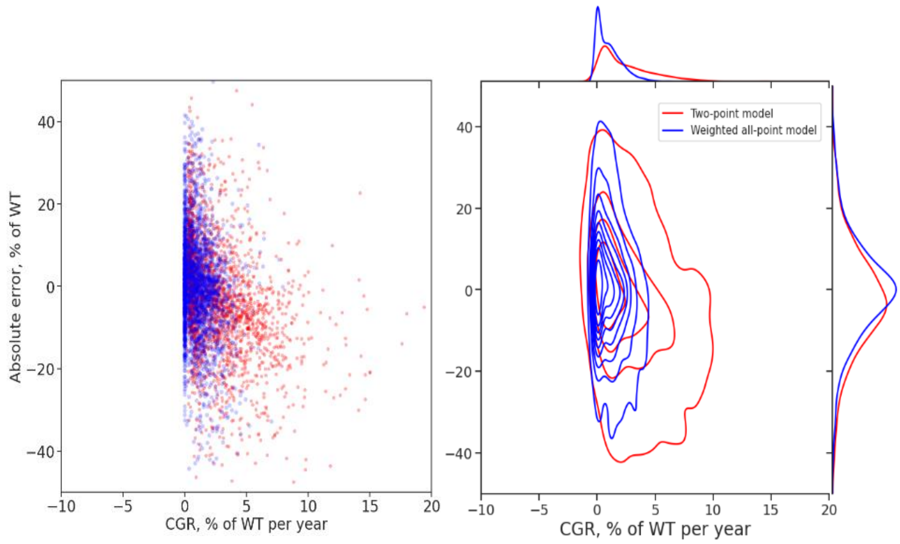


Figure 4. Comparison of the two-point and the weighted all-point models.



the weighted all-point model more evenly balances over- and under-prediction of true CGRs. In the considered scenario, this may be preferable to the behavior of the two-point model, which tends to over-predict CGRs on average.

3 RESULTS

The use of unity plots is a widely accepted method for evaluating model performance within the pipeline industry. As demonstrated in Figure 5, a comparison of all models considered in this study reveals that linear models based on all data points exhibit superior performance as opposed to those based on only two points, with the exponential model occupying a middle ground. In practice, the probabilistic model is computationally intensive and mathematically intricate to implement compared to other simple models. Although the weighted all-point model performs slightly better than the all-point model, tool tolerances are not always readily accessible, making the all-point model the optimal choice for practical application, as evidenced by the extensive analysis of a large ILI dataset.

Figure 5. Combined unity plot for all models

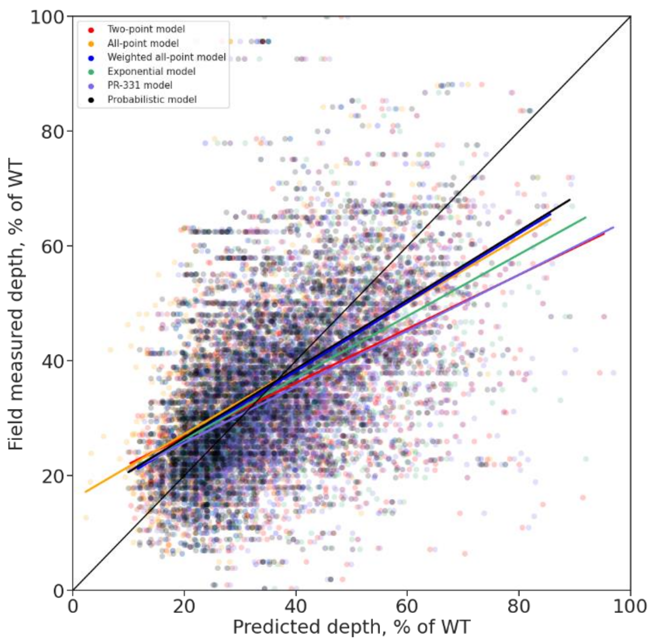
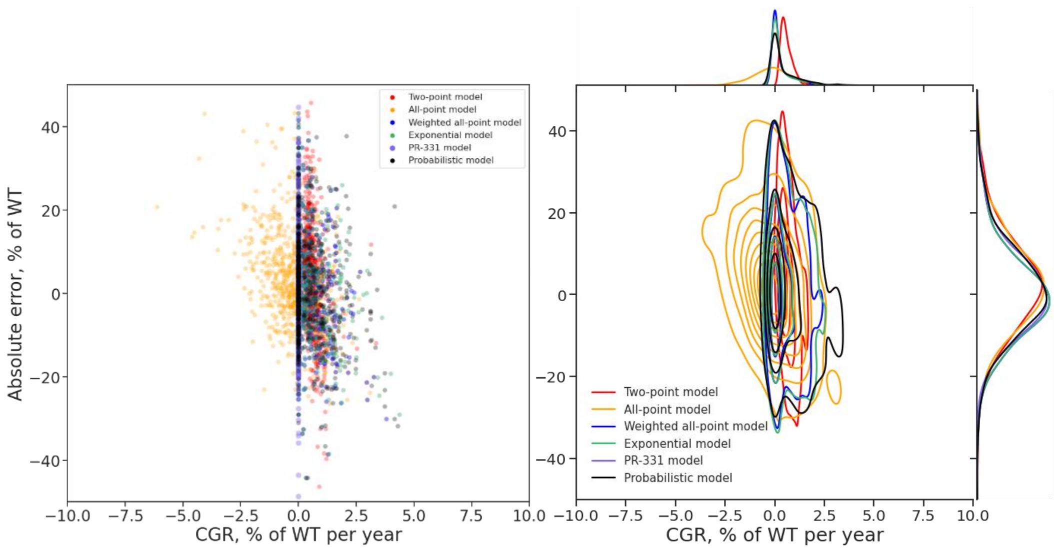


Figure 6. Results for negative growth data



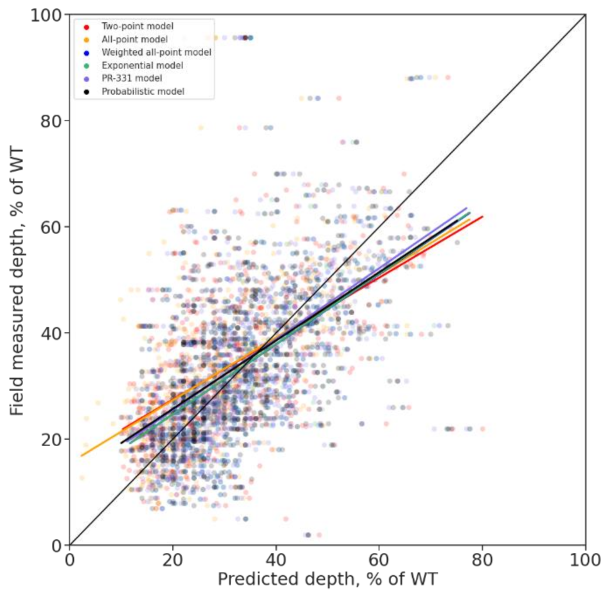
We shall now examine additional scenarios in which models can be applied, such as instances of negative growth in the two-point model and half-life calculations for unmatched anomalies.

3.1 Negative growth

Negative growth is a non-physical phenomenon, resulting from measurement errors and noise. In this section, we evaluate the outcomes of applying all models to negative growth, as observed by the last two measurements.

Figure 6 illustrates the distributions of data points, along with their corresponding kernel density estimated curves, for the negative growth data. The distributions are highly similar. The PR-331, probabilistic, and exponential models generate some outliers in the overestimated region. There is no clear preference for a particular model when dealing with negative growth. The unity plot depicted in Figure 7 displays similar results for all models.

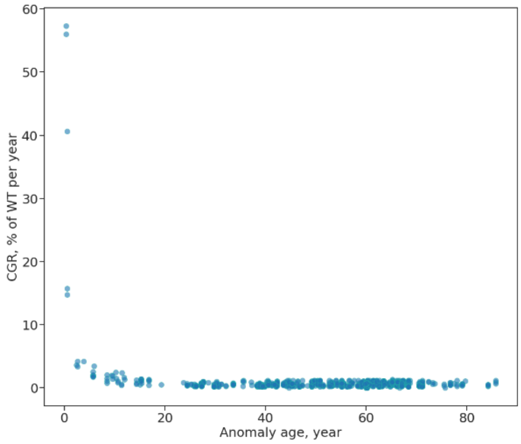
Figure 7. Unity plot for negative growth data



3.2 Half-life model

The half-life model is employed for unmatched anomalies to calculate CGRs, based on the assumption that corrosion growth initiates from zero depth, halfway between the pipe installation and an ILI measurement. This approach yields reasonable results in most cases. However, when pipelines are

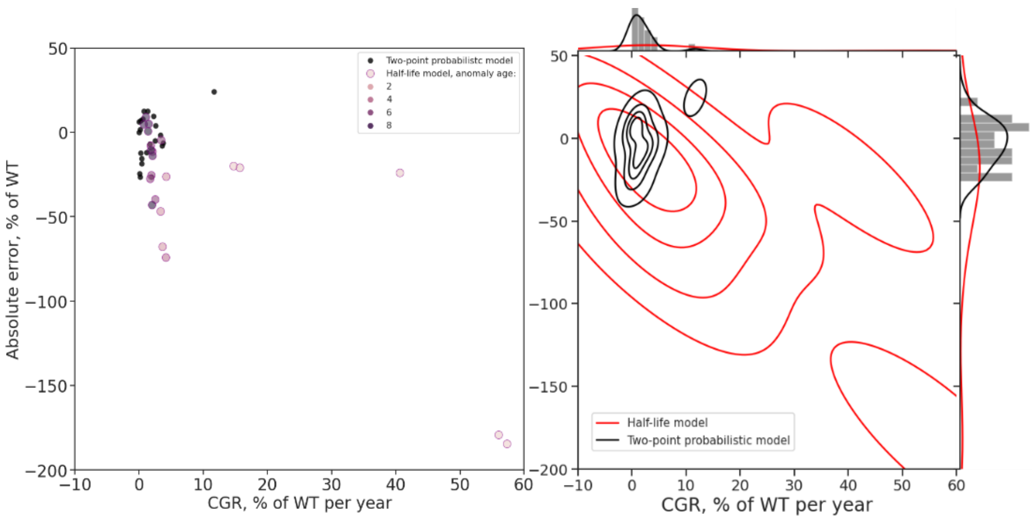
Figure 8. CGRs of un-matched anomalies based on the half-life model.



"young", meaning their installation date is only a few years prior to the onset of detectable corrosion, the half-life approach results in highly overestimated CGRs. Figure 8 illustrates this scenario using our dataset. In this case, "young" pipes could be defined as those whose anomaly age is less than 10 years. Calculated CGRs begin to increase exponentially as the age decreases.

An effective method for treating the population of data from young pipes is to draw a depth and time value from a uniform distribution, with the depth value being below or equal to a threshold. This distribution is integrated as a component of the probabilistic model and was utilized to replace zero depths at half of the anomaly age. Figure 9 illustrates significant enhancements in terms of both accuracy and a reduction in numbers of overestimated CGRs for young pipes using this approach.

Figure 9. Comparison between the half-life model and the two-point probabilistic model



4 Conclusions

We analyzed a large dataset of ILI runs in conjunction with metal loss field data to evaluate the performance of six corrosion growth models. The predicted depths are used to calculate the distribution of absolute error (accuracy) of each model. However, the accuracy alone is not a sufficient metric to determine the applicability of a model. We presented a method to evaluate the positioning of a particular model on the conservatism versus cost-saving spectrum. Decisions can be made by comparing model results in the space of absolute error versus CGRs. As an example, we demonstrated that utilizing the all-point linear regression leads to a balance between over- and under-prediction. This balanced accuracy may be preferable in certain pipeline management scenarios such as long-term pipe replacement planning. Unity plots also serve as a standard tool to compare the performance of models.

In addition to comparing overall performance of the models, the methodology introduced in this paper can be used to evaluate and enhance growth calculations and depth predictions within specific domains of ILI data. In this paper, we analyzed how the treatment of negative growth data can be improved to mitigate the effects of ILI measurement errors. Furthermore, we demonstrated how the use of a probabilistic approach can eliminate unrealistically high growth rates in young pipes.

This general framework also has other practical applications such as the analysis of high-growing chains, which represent outliers in comparison to the main CGR distribution, or an analysis of error reduction in CGR calculation for chains where multiple ILIs are separated by short time intervals of a few months. This can be achieved by averaging the ILI readings between such ILIs.

References

1. "How Automated Data Normalization & Alignment Helps Operators Make Better Decisions", J. Dubuc, C. Scott, https://www.onebridgesolutions.com/wp-content/uploads/2022/12/DPSF-2022-How-Automated-Data-Normalization-Alignment-Helps-Operators-Make-Better-Decisions.pdf (accessed January 2022)
2. G. Coleman, S. Miller "ILI Tool Tolerance and Repeatability Effect on Corrosion Growth Rates" IPC2010, paper no. 31381 (Calgary, Alberta, Canada: ASME 2010)
3. N. Draper, H. Smith "Applied Regression Analysis, 3rd Edition", ISBN: 978-0-471-17082-2, p.15
4. Y. Dodge, "The Concise Encyclopedia of Statistics", ISBN: 978-0-387-31742-7, p.566
5. 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.
6. 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.
7. 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)
8. Y. Dodge, "The Concise Encyclopedia of Statistics", ISBN: 978-0-387-31742-7, p.33