Iowa State University
Digital Repository

Creative Components
Iowa State University Capstones, Theses and Dissertations

Spring 2019

Using Degradation Models to Assess Pipeline Life

Qinglong Tian

Follow this and additional works at: https://lib.dr.iastate.edu/creativecomponents

Part of the Applied Statistics Commons

Recommended Citation

Tian, Qinglong, "Using Degradation Models to Assess Pipeline Life" (2019). *Creative Components*. 263.
https://lib.dr.iastate.edu/creativecomponents/263

This Creative Component is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Creative Components by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

Using Degradation Models to Assess Pipeline Life

Qinglong Tian, Shiyao Liu, William Q. Meeker

Department of Statistics

Iowa State University

April 30, 2019

Abstract

Longitudinal inspections of thickness at particular locations along a pipeline provide useful information to assess the remaining life of the pipeline. In applications with different mechanisms of corrosion processes, we have observed various types of general degradation paths. We present two applications of fitting a degradation model to describe the corrosion initiation and growth behavior in a pipeline. We use a Bayesian approach for parameter estimation for the degradation model. The failure-time and remaining lifetime distributions are derived from the degradation model, and we compute Bayesian estimates and credible intervals of the failure-time and remaining lifetime distributions for both individual segments and for the entire pipeline circuit.

Key Words: Bayesian model, degradation model, longitudinal data, pipeline reliability, remaining life.

1 Introduction

1.1 Motivation and Purpose

Repeated measures of wall thickness across time at sampled locations along a pipeline circuit can be used to evaluate the reliability of a pipeline. Degradation models for longitudinal inspections of the pipeline thickness can be used to describe pipeline corrosion behavior, estimate the lifetime distribution of pipeline components, and predict the remaining lifetime of a pipeline circuit. There are two different purposes for such analyses: (1) estimating the life time cumulative distribution function (cdf) of pipeline segments to provide information that can be used to plan the construction of future pipelines and (2) to estimate the remaining life of an existing pipeline circuit. Depending on degradation and corrosion mechanisms, different statistical models and methods are needed to analyze pipeline data. In this paper, we analyze thickness data from two different pipelines and propose degradation models for both applications. In some degradation models, it is computationally challenging to estimate parameters using the traditional likelihood-based method. Bayesian methods with appropriate prior distributions provide an alternative approach for estimating parameters of a complicated degradation model. In addition, estimating complicated functions of the parameters such as quantiles or tail probabilities of the failure time and remaining lifetime distributions is also computationally feasible and computationally efficient when using Bayesian methods.

1.2 Pipeline Data

Figure 1 is a time-series plot of a subset of longitudinal pipeline data from Circuit G in Facility 3. Data were obtained from a sample of thickness measurement locations (TMLs). We show the values for only 15 of 88 of the TMLs so that it is easier to see the nature of the data. For each TML, the thickness was measured at four different quadrants located at the 0, 90, 180, and 270 degree positions (top, right, bottom, and left for a horizontal pipeline). For the first two inspections, only 12 TMLs and quadrant combinations were used. Subsequently, as the perceived risk of failure increased, an additional 76 TML and quadrant combinations were used. Some of these TMLs correspond to elbows and the others correspond to straight pipes. The lines joining the points represent the degradation paths of the different combinations of location and quadrant. The first inspection was performed on February 11, 1995, a number of years after the pipeline had been installed.

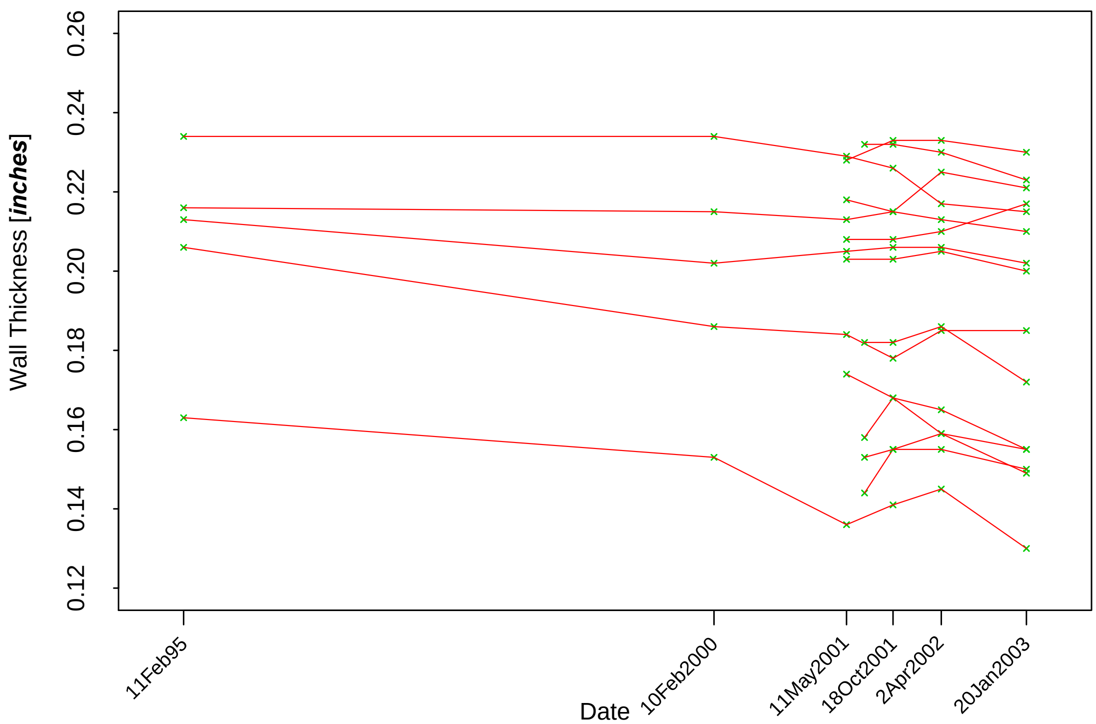


Figure 1: Wall thickness as a function of inspection data for a subset of the TMLs from Circuit G in Facility 3.

The second pipeline data set is from a different facility. Figure 2 displays time plot for the pipeline data from Circuit Q in Facility 1. The data set consists of thickness values at 33 TMLs and each TML was measured at four points in time. Three component types of the pipeline in this data are elbow, straight pipe, and tee. In this facility, the first measurement was taken at the pipeline installation date. The time plot indicates that the original thicknesses vary from TML to TML. Also, the tee pipes are generally thicker than the elbow and straight pipes.

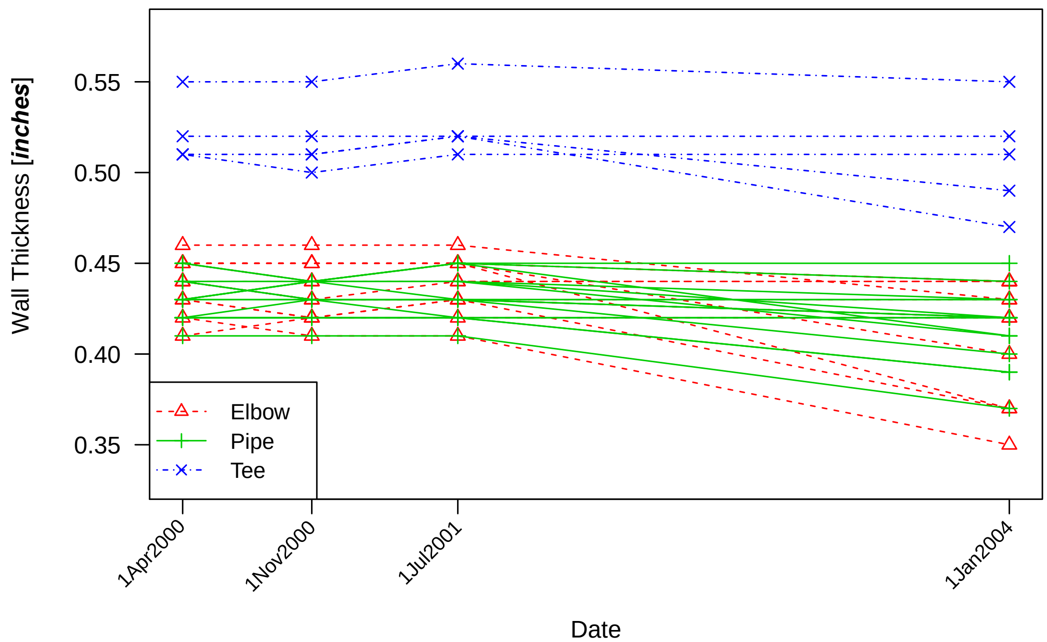


Figure 2: Time plot for pipeline data from Circuit Q in Facility 1.

1.3 Related Work

Degradation models are often used to assess reliability of industrial products. Lu and Meeker (1993) illustrate that under some simple degradation path models, there can be a closed-form expression for the failure time cdf. Chapter 13 of Meeker and Escobar (1998) gives a general introduction to degradation models and describes the relationship between the degradation and failure-time analysis methods of estimating a time-to-failure distribution. Chapter 8 of Hamada et al. (2008) provides an overview of Bayesian degradation models and uses several examples to illustrate how to estimate parameters of a degradation model. Nelson (2009) discusses a model for defect initiation and growth over time and uses maximum likelihood to estimate parameters in the model. Sheikh, Boah, and Hansen (1990) analyze data from water injection pipeline systems and use the Weibull distribution to model the time-to-first-leak. Pandey (1998) uses a probability model to estimate the lifetime distribution of a pipeline before and after failures due to the metal loss.

1.4 Overview

The rest of this paper is organized as follows. Section 2 proposes a degradation model for pipeline data from Circuit G in Facility 3 and uses the Bayesian approach to estimate the parameters in the degradation model. Section 3 derives the failure time and remaining lifetime distributions for the circuit and computes the Bayesian estimates and the corresponding credible intervals. Section 4 analyzes pipeline data from Circuit Q in Facility 1. A degradation model is proposed to describe the corrosion initiation and growth behavior observed in this pipeline. Section 5 evaluates the failure time distribution and predicts the remaining lifetime distribution of Circuit Q in Facility 1. Because there were few inspections in the data from Circuit Q in Facility 1, it was necessary to use some informative prior information about the corrosion rate. In order to study the data needed for estimability without using such prior information, Section 6 analyzes simulated data for a single circuit having more than one inspection after corrosion initiation. Section 7 contains concluding remarks and areas for future research.

2 Modeling Pipeline Data from Circuit G in Facility 3

In this section, we focus on the analysis of the pipeline data from Circuit G in Facility 3 shown in Figure 1. We propose a degradation model and Bayesian estimation with weakly informative prior distributions to estimate the parameters of the degradation model.

2.1 Degradation Model for Pipeline Data from Circuit G in Facility 3

Let $Y_{i_j t_k}$ denote the pipeline thickness at time t_k for TML i ($i = 1, 2, \ldots, 22$; $j = 1, 2, 3, 4$; $k = 1, 2, \ldots, 7$). We assume that the degradation path of Circuit G in Facility 3 is linear with respect to inspection time and has the form

$$Y_{i_j t_k} = y_0 - \beta_{1_{i_j}}(t_k - t_0) + \epsilon_{i_j k} \qquad (1)$$

where $\beta_{1_{i_j}}$ (inches per year) is -1 times the corrosion rate for quadrant j at location i and $\epsilon_{i_j k}$ is the measurement error term. We call this Model 1. Here y_0 is the original thickness at installation time t_0. Specifically, the original thickness y_0 is 0.25 inches and the installation time t_0 is February 12, 1990. The precise dates of installation and beginning-use were not available and this date was obtained by extrapolating backwards in time. Because the corrosion rate, defined as the thickness change per year, varies from location to location and can only be negative, $\beta_{1_{i_j}}$ in the degradation Model 1 in (1) is a positive random variable. To guarantee a positive $\beta_{1_{i_j}}$, We assume that the vector of the logarithms of the slope has a multivariate normal distribution, that is,

$$[\log\left(\beta_{1_{i_1}}\right), \log\left(\beta_{1_{i_2}}\right), \log\left(\beta_{1_{i_3}}\right), \log\left(\beta_{1_{i_4}}\right)]' \sim \mathrm{MVN}([\mu_{\log\left(\beta_{1_1}\right)}, \mu_{\log\left(\beta_{1_1}\right)}, \mu_{\log\left(\beta_{1_1}\right)}, \mu_{\log\left(\beta_{1_1}\right)}]', \mathbf{\Sigma}),$$

$$\mathbf{\Sigma} = \mathbf{LRL},$$

where

$$\mathbf{L} = \mathrm{diag}(\sigma_{\log\left(\beta_{\mathbf{1_1}}\right)}, \sigma_{\log\left(\beta_{\mathbf{1_1}}\right)}, \sigma_{\log\left(\beta_{\mathbf{1_1}}\right)}, \sigma_{\log\left(\beta_{\mathbf{1_1}}\right)}), \mathbf{R} = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{12} & 1 & \rho_{23} & \rho_{24} \\ \rho_{13} & \rho_{23} & 1 & \rho_{34} \\ \rho_{14} & \rho_{24} & \rho_{34} & 1 \end{bmatrix}.$$

And the measurement error is $\epsilon_{i_jk} \sim \mathrm{NORM}\ (0, \sigma_\epsilon)$. Thus the parameters in the Model 1 are: $\boldsymbol{\theta_1} = (\mu_{\log(\beta_1)}, \sigma_{\log(\beta_1)}, \rho_{12}, \rho_{13}, \rho_{14}, \rho_{23}, \rho_{24}, \rho_{34}, \sigma_\epsilon)'$.

2.2 Bayesian Estimation of the Parameters in the Degradation Model

Bayesian estimation with the use of weakly informative prior information is closely related to likelihood estimation (with a flat prior, the Bayesian joint posterior distribution is proportional to the likelihood). Bayesian methods provide a convenient alternative for estimating the parameters in the degradation model, particularly because we need to make inferences on complicated functions of the model parameters.

For the example, we use a normal distribution with mean zero and a large standard deviation [i.e., NORM $(0, 10)$] as the prior distribution for the parameter $\mu_{\log(\beta_1)}$. Gelman (2006) provides general suggestions for choosing proper prior distributions for variance parameters in a hierarchical model. For the lognormal corrosion rate, we use the weakly informative prior distrbution half-Cauchy(0,5) for $\sigma_{\log(\beta_1)}$ and σ_ϵ. The half-Cauchy distribution is defined as follow: If X follows Cauchy$(0, \sigma)$, then $Y = |X|$ follows half-Cauchy$(0, \sigma)$. We obtain a large number of draws from the joint posterior distribution of the degradation model parameters using Hamiltonian Monte Carlo (HMC) implemented in Stan (http://mc-stan.org). Table 1 presents marginal posterior distribution summaries for the parameters in $\boldsymbol{\theta_1}$, including the posterior median and 95% credible intervals. The number of HMC draws was chosen to be 50000 (including 10000 warm-up iterations) so that the endpoints of the credible intervals are accurate to at least ± 1 in the third significant digit. Figure 3 plots the fitted thickness values under Model 1 versus time for Circuit G in Facility 3 with a 10-year extrapolation after the last inspection in January 20, 2003.

Parameters	Posterior Median	Posterior Standard Deviation	95% Credible Interval	
			Lower	Upper
$\mu_{\log(\beta_1)}$	−5.69	0.0920	−5.87	−5.51
$\sigma_{\log(\beta_1)}$	0.593	0.0590	0.497	0.728
ρ_{12}	0.0919	0.226	−0.364	0.501
ρ_{13}	0.302	0.223	−0.211	0.648
ρ_{14}	0.0767	0.162	−0.246	0.387
ρ_{23}	0.568	0.200	0.0369	0.809
ρ_{24}	0.454	0.155	0.0989	0.700
ρ_{34}	0.367	0.165	0.00238	0.639
σ_ϵ	0.00605	0.000255	0.00559	0.00658

Table 1: Marginal posterior distribution summaries of the degradation model parameter estimates for pipeline data from Circuit G in Facility 3 using the degradation Model 1 in (1).

2.3 Statistical Model for Different Quadrants

In Sections 2.1 and 2.2, it was assumed that the corrosion rates of different quadrants from the same location have the same distribution. In non-vertical pipes, however, the corrosion

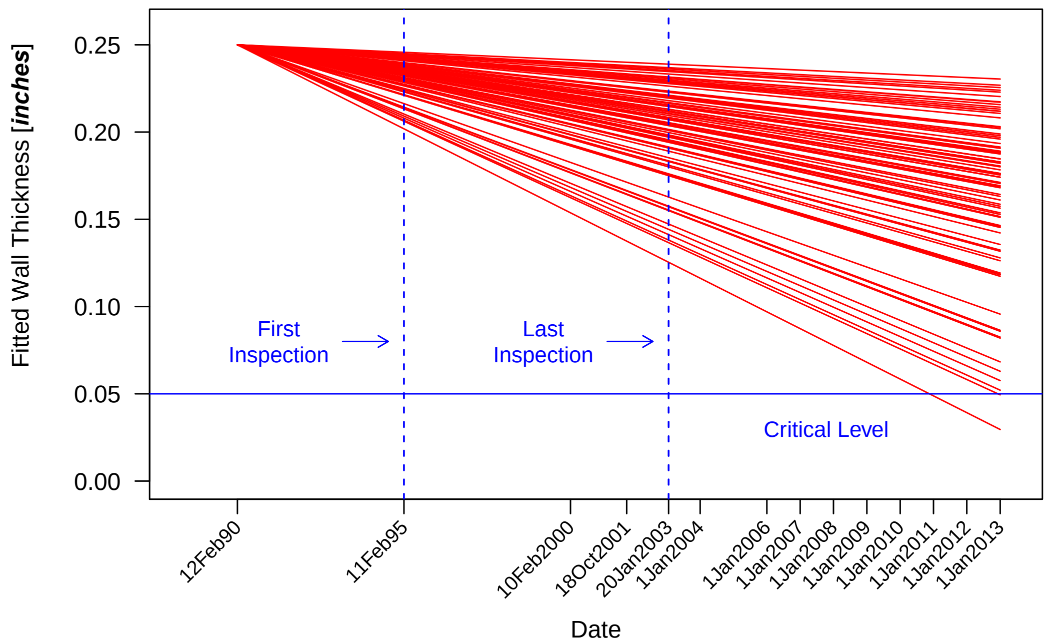


Figure 3: Time plot showing the fitted thickness values for the pipeline data from Circuit G in Facility 3 using the degradation Model 1 in (1).

rate of locations in the upper quadrant might be expected to differ from that in the lower quadrant at the same TML. The degradation model in this section allows the means of the logarithm of the corrosion rates vary from quadrant to quadrant. Assuming that the circuit with initial thickness 0.25 inches was installed on February 12, 1990, the degradation model is

$$Y_{i_j t_k} = y_0 - \beta_{1_{i_j}}(t_k - t_0) + \epsilon_{i_j k} \tag{2}$$

where $\beta_{1_{i_j}}$ (inches per year) is the corrosion rate of quadrant j at TML i ($i = 1, 2, \ldots, 22$; $j = 1, \ldots, 4$) and $\epsilon_{i_j k}$, as before, is the measurement error term at time t_k ($k = 1, \ldots, 7$). Similar to Model 1 in (1), $\beta_{1_{i_j}}$ is also positive in Model 2 in (2). We assume that the vector of the logarithms of the slope has a multivariate normal distribution, that is,

$$[\log\left(\beta_{1_{i_1}}\right), \log\left(\beta_{1_{i_2}}\right), \log\left(\beta_{1_{i_3}}\right), \log\left(\beta_{1_{i_4}}\right)]' \sim MVN([\mu_{\log(\beta_{1_1})}, \mu_{\log(\beta_{1_2})}, \mu_{\log(\beta_{1_3})}, \mu_{\log(\beta_{1_4})}]', \mathbf{\Sigma}),$$

$$\mathbf{\Sigma} = \mathbf{LRL},$$

where

$$\mathbf{L} = \text{diag}(\sigma_{\log(\beta_{\mathbf{1_1}})}, \sigma_{\log(\beta_{\mathbf{1_2}})}, \sigma_{\log(\beta_{\mathbf{1_3}})}, \sigma_{\log(\beta_{\mathbf{1_4}})}), \quad \mathbf{R} = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{12} & 1 & \rho_{23} & \rho_{24} \\ \rho_{13} & \rho_{23} & 1 & \rho_{34} \\ \rho_{14} & \rho_{24} & \rho_{34} & 1 \end{bmatrix}.$$

And we assume that $\epsilon_{i_jk} \sim$ NORM $(0, \sigma_\epsilon)$. The parameters for Model 2 in (2) are:

$$\boldsymbol{\theta_2} = (\mu_{\log(\beta_{1_1})}, \mu_{\log(\beta_{1_2})}, \mu_{\log(\beta_{1_3})}, \mu_{\log(\beta_{1_4})},$$
$$\sigma_{\log(\beta_{1_1})}, \sigma_{\log(\beta_{1_2})}, \sigma_{\log(\beta_{1_3})}, \sigma_{\log(\beta_{1_4})}, \rho_{12}, \rho_{13}, \rho_{14}, \rho_{23}, \rho_{24}, \rho_{34}, \sigma_\epsilon)'.$$

The Bayesian method is again used to estimate $\boldsymbol{\theta_2}$ using weakly informative prior distributions. Lewandowski et al. (2009) introduce the LKJ distribution, which is a weakly informative distribution for a correlation matrix. The parameter η in LKJcorr(η) controls the dependence among the corrosion rates of the four quadrants. When $\eta = 1$, LKJcorr(1) is a uniform distribution over all possible correlation matrices. We use LKJcorr(1) as the prior distribution for $\mathbf{R}$. The prior distributions for σ_ϵ, $\sigma_{\log(\beta_{1_1})}$, $\sigma_{\log(\beta_{1_2})}$, $\sigma_{\log(\beta_{1_3})}$, $\sigma_{\log(\beta_{1_4})}$are half-Cauchy $(0, 5)$. We use a diffuse normal, i.e., NORM(0,10), as the prior distribution for $\mu_{\log(\beta_{1_j})}$, where $j = 1, 2, 3, 4$. Table 2 presents marginal posterior distribution summaries for the parameters in $\boldsymbol{\theta_2}$, including the posterior median and 95% credible intervals. The number of draws was chosen to be 50000 (including 10000 warm-up iterations) and the endpoints of the credible intervals are accurate to at least ± 2 in the second significant digit. Figure 4 shows a time plot of the fitted thickness values versus time for the different quadrants of this circuit.

2.4 The Comparison of Models

The Watanabe-Akaike information criteria (WAIC) was introduced by Watanabe (2010). WAIC is a Bayesian approach for estimating the out-of-sample expectation and it can be used for Bayesian model comparisons. We used the WAIC formula from Vehtari et al. (2017). The WAIC value for Model 1 is -2708.1 with standard error 66.3 and that for Model 2 is -2708.5 with standard error 66.4. Considering the relatively large standard errors, there is no significant difference between the WAIC values for the two models.

Parameters	Posterior Median	Posterior Standard Deviation	95% Credible Interval	
			Lower	Upper
$\mu_{\log(\beta_{1_1})}$	−5.40	0.115	−5.64	−5.18
$\mu_{\log(\beta_{1_2})}$	−5.77	0.116	−6.00	−5.54
$\mu_{\log(\beta_{1_3})}$	−5.69	0.108	−5.90	−5.48
$\mu_{\log(\beta_{1_4})}$	−6.03	0.167	−6.37	−5.71
$\sigma_{\log(\beta_{1_1})}$	0.514	0.0964	0.376	0.751
$\sigma_{\log(\beta_{1_2})}$	0.514	0.0910	0.381	0.735
$\sigma_{\log(\beta_{1_3})}$	0.477	0.0841	0.354	0.682
$\sigma_{\log(\beta_{1_4})}$	0.738	0.131	0.546	1.06
ρ_{12}	0.130	0.207	−0.297	0.504
ρ_{13}	0.284	0.195	−0.136	0.620
ρ_{14}	0.315	0.197	−0.117	0.644
ρ_{23}	0.474	0.174	0.0677	0.739
ρ_{24}	0.479	0.175	0.0731	0.750
ρ_{34}	0.438	0.181	0.0213	0.722
σ_{ϵ}	0.00605	0.000255	0.00557	0.00658

Table 2: Marginal posterior distribution summaries of the degradation model parameter estimates for pipeline data from Circuit G in Facility 3 using the degradation Model 2 in (2).

3 Models Relating Degradation and Failure in Circuit G of Facility 3

3.1 Bayesian Estimation of the Failure Time Distribution

The degradation path over time is $\mathcal{D} = \mathcal{D}(t)$. The failure of an individual segment in a pipeline is said to have happened when the remaining pipeline thickness is less than the critical level $\mathcal{D}_f$ (0.05 inches in this example). Such a failure is known as a "soft failure"

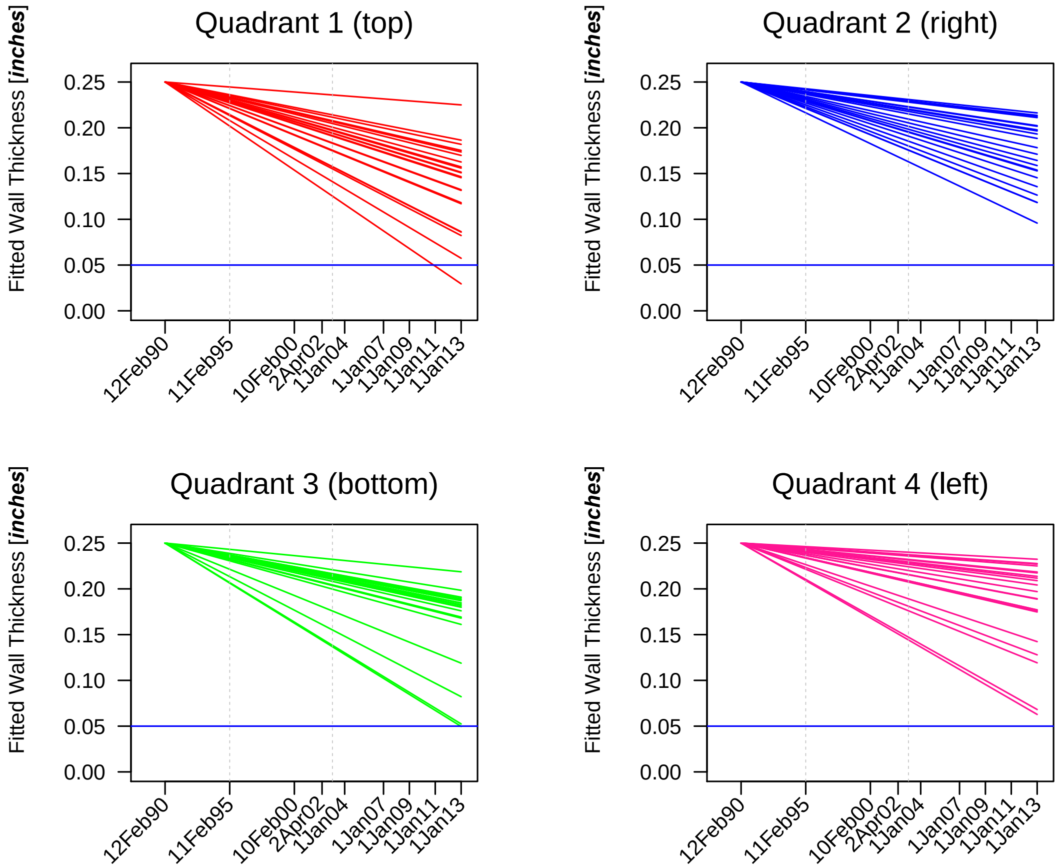


Figure 4: Plots showing the fitted thickness values over time for different quadrants of pipeline data from Circuit G in Facility 3 using the degradation Model 2 in (2).

and such critical levels are determined through engineering judgment as the thickness below which there is risk of a leak. Because $\beta_{1_{i_j}} \sim$ Lognormal $(\mu_{\log(\beta_{1_j})}, \sigma_{\log(\beta_{1_j})}), j = 1, 2, 3, 4$ in Model 2, the failure time cdf $F(t)$ of individual segments in a population of segments of quadrant j in the pipeline can be expressed in a closed form as

$$
\begin{aligned}
F_j(t) &= \Pr(\mathcal{D}(t) \leq \mathcal{D}_f) = \Pr(y_0 - \beta_{1_{i_j}}(t - t_0) \leq 0.05) \\
&= \Pr\left(\beta_{1_{i_j}} \geq \frac{0.20}{t - t_0}\right) = 1 - \Phi_{\text{norm}}\left(\frac{\log(0.20) - \log(t - t_0) - \mu_{\log(\beta_{1_j})}}{\sigma_{\log(\beta_{1_j})}}\right) \\
&= \Phi_{\text{norm}}\left(\frac{\log(t - t_0) - \log(0.20) + \mu_{\log(\beta_{1_j})}}{\sigma_{\log(\beta_{1_j})}}\right). \qquad (3)
\end{aligned}
$$

where Φ_{norm} is the standard normal cdf. The failure time distribution, as a function of the degradation model parameters, can be estimated by using the Bayesian approach. For each draw from the joint posterior distribution, one can, for a particular value of t, evaluate $F_j(t)$ in (3) to obtain a corresponding draw from the marginal posterior distribution of failure time cdf at t. Then approximate quantiles of the empirical distribution of the marginal posterior distribution of $F_j(t)$ gives point estimates and credible intervals for $F_j(t)$. Table 2 and Figure 4 suggest that the corrosion rate of quadrant 1 from the upper quadrant is the largest among these four different quadrants. The upper plot in Figure 5 displays the estimate of the failure time cdf with two-sided 95% and 80% credible intervals for the pipeline data from quadrant 1 of Circuit G in Facility 3. One can also obtain the corresponding failure time cdf plots for other quadrants. But with the largest corrosion rate, the failure time plot for quadrant 1 is the most pessimistic. The estimation of the failure time distribution of individual pipeline segments from the circuit in (3) also allows us to estimate the lifetime of the whole circuit, which consists of many segments.

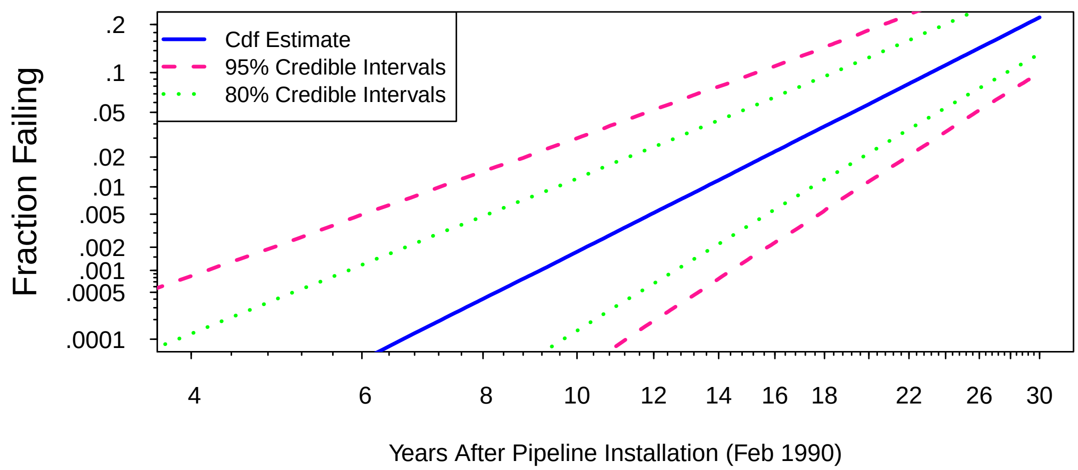
3.2 Prediction of the Remaining Life of the Current Circuit

In the pipeline application, the remaining life of a particular segment of a circuit is an important quantity for assessing the lifetime of the pipeline. The cdf of the remaining lifetime $F_{j_R}(t)$ conditional on surviving until the most recent inspection time (January 2003) for quadrant j is

$$F_{j_R}(t) = \Pr(T \leq t | T > t_c) = \frac{F_j(t) - F_j(t_c)}{1 - F_j(t_c)},\ t \geq t_c \tag{4}$$

where t_c is the most recent inspection time and $F_j(t)$ is the failure time distribution derived in Section 3.1. As before, evaluating (4) at posterior draws provides estimates and the corresponding credible intervals of the remaining lifetime cdf. The plot on the bottom of Figure 5 shows the posterior estimates of the remaining lifetime cdf after the last inspection in January 2003 with 95% and 80% credible intervals.

In pipeline applications, it is of great interest to estimate small quantiles of the minimum remaining lifetime of the population of pipeline segments. To do this, one needs to extrapolate further into the tail of the remaining life distribution estimated for a given segment. A typical pipeline-integrity dataset will contain data from n TML segments, each of which is about one foot long. Recall that our model for the entire pipeline length has M segments



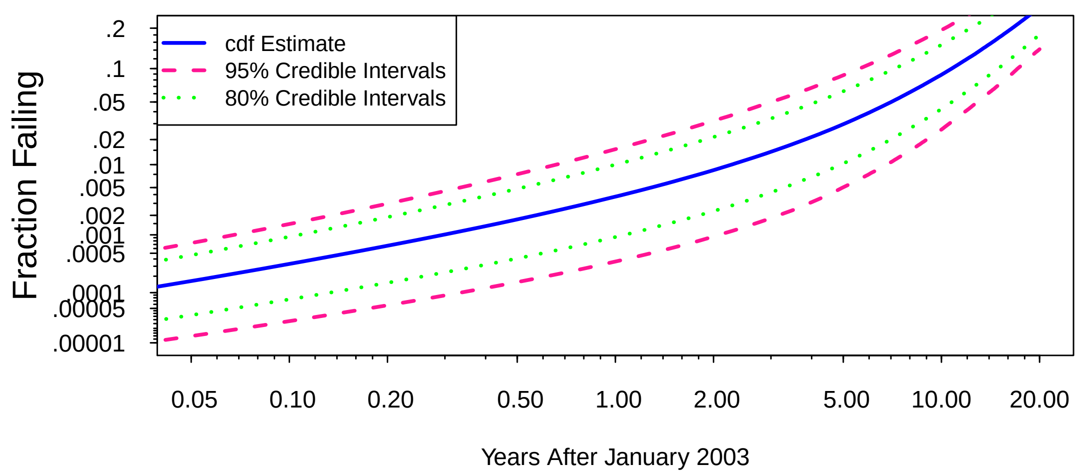


Figure 5: Degradation model estimates of the failure time cdf on lognormal probability paper (top) and the remaining lifetime cdf (bottom) with two-sided 95% and 80% credible intervals for a single randomly chosen segment from quadrant 1 based on the pipeline data from Circuit G in Facility 3.

of this length, where M is typically much larger than n. It would be expected that there is some autocorrelation among the maximum wall-thickness values along the M segments. Because the distance between the n TMLs is large, however, there is no information about the strength of this autocorrelation in the data. Barlow and Proschan (1975) show that when the failure times of a series system's components are positively correlated (the pipeline can be viewed as a series system where each one-foot segment is a component and positive autocorrelation would be expected), assuming independence will provide a conservative prediction of the system's overall reliability. Thus we assume that all M segments are independent. Then the cdf of the minimum remaining life among all of the M segments for quadrant j along the pipeline can be expressed as

$$F_{j_M}(t) = \Pr[T_{\min} \leq t] = 1 - [1 - F_{j_R}(t)]^M \tag{5}$$

where $F_{j_R}(t)$ is the remaining lifetime cdf for a single segment. If one wants to control $F_{j_M}(t)$, such that $F_{j_M}(t) = \Pr[T_{\min} \leq t] = p$, then one would choose the threshold to be $t_{j_p} = F_{j_M}^{-1}(p)$, the p quantile of the distribution of the minimum $T_{\min}$ among the M pipeline segments for quadrant j. The translation to the adjusted quantile in terms of the remaining lifetime cdf $F_{RM}(t)$ is:

$$t_{j_p} = F_{j_M}^{-1}(p) = F_{j_R}^{-1}\left(1 - (1-p)^{\frac{1}{M}}\right). \tag{6}$$

This indicates that the p quantile of the minimum remaining lifetime distribution of the population of M segments for quadrant j corresponds to the $1-(1-p)^{1/M}$ quantile of the remaining lifetime cdf for an individual segment. Figure 6 shows the posterior density of 0.1, 0.2, 0.3, and 0.4 quantiles of the minimum remaining lifetime distribution with the population size $M = 100$ segments using degradation Model 1 and Model 2 (and we only consider the most pessimistic quadrant 1), respectively. Model 2 is more conservative than Model 1 as it generates smaller quantile estimates.

The small quantile estimates suggest that the Circuit G in Facility 3 could have leakage risks within one year after the most recent inspection. One should pay closer attention to this circuit. Careful examination, more frequent inspection at more TMLs, or retirement/replacement of the pipeline would protect against pipeline leakage.

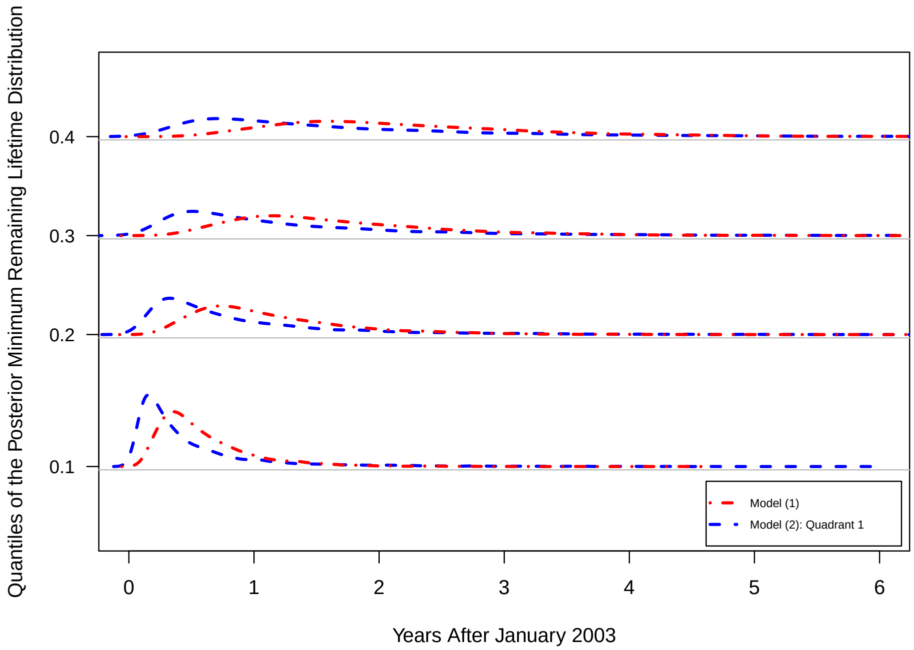


Figure 6: Posterior density of the 0.1, 0.2, 0.3 and 0.4 quantiles of the minimum remaining lifetime distribution (years since the last inspection time t_c: January 2003) with the population size $M = 100$ based on the pipeline data from Circuit G in Facility 3 using the degradation Models 1 and 2.

4 Modeling Pipeline Data from Circuit Q in Facility 1

Figure 7 is a trellis plot for the pipeline data from Circuit Q in Facility 1. Each panel of the trellis plot corresponds to thickness measurements over time for a specific TML. The trellis plot suggests an interesting pipeline corrosion process. For example, at TMLs #1, #2, and #3, there is no detectable thickness loss in the first three inspections. Significant thickness losses, however, were detected at the forth inspection time. This suggests that the corrosion process was likely initiated between the third and forth inspection times. At some TMLs (e.g., TMLs #12, #13, and #33), the corrosion appears unlikely to have initiated before the last inspection time.

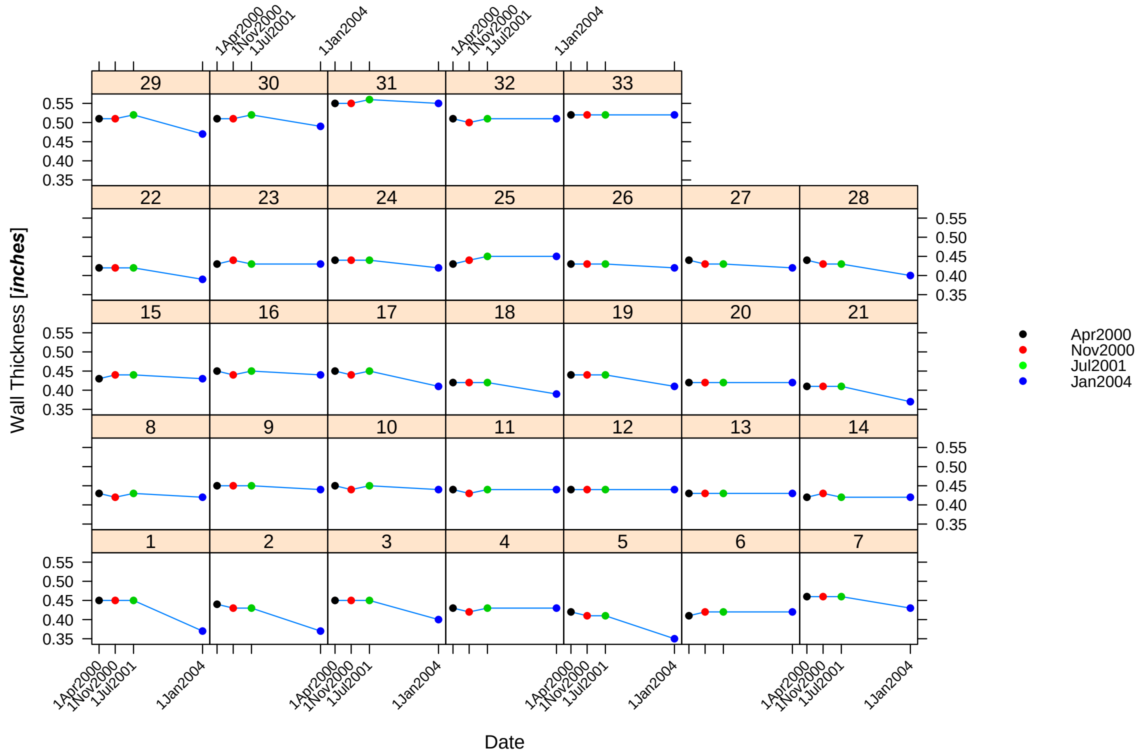


Figure 7: Trellis plot for pipeline data from Circuit Q in Facility 1.

4.1 Degradation Model for Corrosion Initiation and Growth

We assume that after the corrosion initiation, the corrosion rate is constant for a particular location, but may differ from location to location. We propose a degradation model with a random corrosion initiation time and random corrosion rate to describe the overall corrosion initiation and growth process. The degradation model for the pipeline thickness Y_{it_j} at time t_j for the TML i ($i = 1, 2, \ldots, 33$; $j = 1, 2, 3, 4$) is:

$$Y_{it_j} = \begin{cases} Y_{0i} + \epsilon_{ij} & \text{for } t_j < T_{I_i} \\ Y_{0i} - \beta_{1i}(t_j - T_{I_i}) + \epsilon_{ij} & \text{for } t_j \geq T_{I_i}. \end{cases} \tag{7}$$

In this model,

- t_j is the time in years when the measurement j was taken using the installation date as the starting time.

- Y_{it_j} denotes the thickness measurement in inches for TML i at time t_j.

- Y_{0i} is the original thickness (measured in inches) of TML i. Because the distribution of the original thickness depends on the component type of the TML (elbow, tee, or straight pipe), we assume that the initial measurement Y_{0i} has a normal distribution with different means but a common standard deviation:
  - If the TML is an elbow, we assume that $Y_{0i} \sim \text{NORM}\left(\mu_{y0_{\text{elbow}}}, \sigma_{y0}\right)$;
  - If the TML is a pipe, we assume that $Y_{0i} \sim \text{NORM}\left(\mu_{y0_{\text{pipe}}}, \sigma_{y0}\right)$;
  - If the TML is a tee, we assume that $Y_{0i} \sim \text{NORM}\left(\mu_{y0_{\text{tee}}}, \sigma_{y0}\right)$.
- T_{I_i} is the corrosion initiation time (measured in years) at TML i (using the installation date as the starting time) and we assume that $T_{I_i} \sim \text{Lognormal}\left(\mu_{\log(T_I)}, \sigma_{\log(T_I)}\right)$ with density function:

$$f(x|\mu_{\log(T_I)}, \sigma_{\log(T_I)})) = \frac{1}{\sigma_{\log(T_I)}\sqrt{2\pi}}\frac{1}{x}\exp\left[-\frac{1}{2}\left(\frac{\log(x) - \mu_{\log(T_I)}}{\sigma_{\log(T_I)}}\right)^2\right].$$

  By an important property of lognormal distribution, $\mu_{\log(T_I)}$ and $\sigma_{\log(T_I)}$ are, respectively the mean and standard deviation of the distribution of the $\log(T_{I_i})$.
- β_{1i} (inches per year) is the corrosion rate for TML i. First we use $\beta_{1i} \sim \text{Lognormal}\left(\mu_{\log(\beta_1)}, \sigma_{\log(\beta_1)}\right)$, where $\mu_{\log(\beta_1)}$ and $\sigma_{\log(\beta_1)}$ are, respectively the mean (or median) and standard deviation of the distribution of $\log(\beta_{1i})$. Then we use $\beta_{1i} \sim \text{Weibull}\left(\nu_{\beta_1}, \lambda_{\beta_1}\right)$ with density function:

$$f(x|\nu_{\beta_1}, \lambda_{\beta_1}) = \frac{\nu_{\beta_1}}{\lambda_{\beta_1}}\left(\frac{x}{\lambda_{\beta_1}}\right)^{\nu_{\beta_1}-1}\exp\left[-\left(\frac{x}{\lambda_{\beta_1}}\right)^{\nu_{\beta_1}}\right].$$

- ϵ_{ij} is the measurement error and we assume that $\epsilon_{ij} \sim \text{NORM}(0, \sigma_\epsilon)$.

The model parameters are: $\boldsymbol{\theta_3} = (\mu_{y0_{\text{elbow}}}, \mu_{y0_{\text{pipe}}}, \mu_{y0_{\text{tee}}}, \sigma_{y0}, \mu_{\log(\beta_1)}, \sigma_{\log(\beta_1)}, \mu_{\log(T_I)}, \sigma_{\log(T_I)}, \sigma_\epsilon)'$ for the lognormal corrosion rate. When the corrosion rate has a Weibull distribution, the model parameters are: $\boldsymbol{\theta_4} = (\mu_{y0_{\text{elbow}}}, \mu_{y0_{\text{pipe}}}, \mu_{y0_{\text{tee}}}, \sigma_{y0}, \nu_{\beta_1}, \lambda_{\beta_1}, \mu_{\log(T_I)}, \sigma_{\log(T_I)}, \sigma_\epsilon)'$.

4.2 Prior Distributions for The Model Parameters

In addition to the model, we need to specify prior distributions for the parameters in (7). For the lognormal corrosion rate, we also use the weakly informative prior distribution half-Cauchy(0,5) for the standard deviations $\sigma_{y0}, \sigma_{\log(\beta_1)}, \sigma_{\log(T_I)}$, and σ_ϵ.

The fact that pipeline data for the Circuit Q in Facility 1 has no more than one inspection after the corrosion initiation results in difficulty identifying the corrosion rate and initiation times in the degradation model. That is, for a given TML with evidence of an initiation, we cannot distinguish between a) an initiation close to the fourth inspection, followed by a large (in absolute value) corrosion rate and b) an initiation time close to the third inspection time, followed by a smaller corrosion rate. According to the knowledge from experts in the pipeline application, the median corrosion rates for the TMLs should not exceed 0.022 inches per year. Instead of assigning a prior distribution directly to $\mu_{\log(\beta_1)}$, we assign a lognormal distribution to the median of β_1, i.e., $\beta_{1_{0.5}} = \exp\left[\mu_{\log(\beta_1)}\right]$, such that the 0.005 quantile is 10^{-6} inches per year and the 0.995 quantile is 0.022 inches per year. For the prior distributions for the parameters $\mu_{y_{0\text{elbow}}}, \mu_{y_{0\text{pipe}}}$ and $\mu_{y_{0\text{tee}}}$, we use the half-Normal$(0, 10)$ prior distribution. We assign a lognormal prior distribution to the median of T_I, i.e., $\exp(\mu_{\log(T_I)})$, such that the 0.005 quantile is 10^{-6} years and the 0.995 quantile is 200 years. When the corrosion rate has a Weibull distribution, we use the same independent prior distributions for the parameters $\sigma_{y_0}, \sigma_{\log(T_I)}, \sigma_\epsilon, \mu_{y_{0\text{elbow}}}, \mu_{y_{0\text{pipe}}}, \mu_{y_{0\text{tee}}}$, and $\mu_{\log(T_I)}$. The usual parameters of the Weibull and lognormal distributions have different meanings. To make the Weibull and lognormal corrosion rate distributions comparable, for the Weibull distribution, we also put the same prior distributions on the median of the corrosion rate $\beta_{1_{0.5}}$ and the standard deviation of the distribution of the logarithms of the corrosion rate $\sigma_{\log(\beta_1)}$. In the parameterization of the Weibull distribution in Stan, the shape parameter $\lambda_{\beta_1} = \pi/(\sqrt{6}\sigma_{\log(\beta_1)})$ and the scale parameter is $\nu_{\beta_1} = \beta_{1_{0.5}}/\log(2)^{1/\lambda_{\beta_1}}$. Then we can transform the prior distributions on $\beta_{1_{0.5}}$ and $\sigma_{\log \beta_1}$ to the shape and scale parameters of Weibull distribution. Table 3 is a summary of the prior distributions for these two models.

4.3 Bayesian Estimation of the Parameters in the Degradation Model

Tables 4 and 5 present the posterior distribution summaries of the parameters in the degradation model using lognormal and Weibull corrosion rates respectively. Again, the number of HMC draws was chosen to be 50000 (including 10000 warm-up iterations) so that the

Parameter	Model	Prior Distribution
$\mu_{y_{0\text{elbow}}}$, $\mu_{y_{0\text{pipe}}}$, $\mu_{y_{0\text{tee}}}$	Lognormal, Weibull	half-Normal$(0, 10)$
σ_ϵ, σ_{y_0}, $\sigma_{\log(T_I)}$, $\sigma_{\log(\beta_1)}$	Lognormal, Weibull	half-Cauchy$(0, 5)$
$\mu_{\log(T_I)}$	Lognormal, Weibull	Use lognormal distribution for $\beta_{1_{0.5}} = \exp(\mu_{\log(T_I)})$
$\mu_{\log(\beta_1)}$	Lognormal	Use lognormal distribution for $\exp(\mu_{\log(\beta_1)})$
$\beta_{1_{0.5}}$	Weibull	Same as the $\exp(\mu_{\log(\beta_1)})$ in the lognormal model
λ_{β_1}	Weibull	$\lambda_{\beta_1} = \pi/(\sqrt{6}\sigma_{\log(\beta_1)})$
ν_{β_1}	Weibull	$\nu_{\beta_1} = \beta_{1_{0.5}}/\log(2)^{1/\lambda_{\beta_1}}$

Table 3: Summary of prior distributions for degradation models in (7).

endpoints of the credible interval are accurate to at least ± 1 in the third significant digit except for $\beta_{1_{0.5}}$, $\sigma_{\log(\beta_1)}$ and $\sigma_{\log(T_I)}$ from the lognormal corrosion rate as well as $\beta_{1_{0.5}}$, $\sigma_{\log(\beta_1)}$ and $\sigma_{\log(T_I)}$ from the Weibull corrosion rate. These 6 parameters are accurate to ± 2 in the second significant digit. Figures 8 and 9 show the trellis plots of the fitted thickness values for Circuit Q in Facility 1 using lognormal corrosion rates with 10-years of extrapolation after the last inspection on January 1, 2004. Figure 8 presents TMLs that appear to have corrosion that started after the last inspection while Figure 9 is for those TMLs that appear to have corrosion that started before the last inspection. The pink cross in each subplot is the estimated initiation of the corrosion. The x value of the estimated initiation of the corrosion is the median of the marginal posterior distribution of T_{I_i}. The y value of the estimated initiate of the corrosion is the median of the marginal posterior distribution of Y_{0i}.

The WAIC is again used for the Bayesian model comparison. The WAIC value for the model with lognormal corrosion rate is -991.3 and the standard error is 14.9. The WAIC value for the model with Weibull corrosion rate is -990.7 and the standard error is 15.1. Considering the difference is small compared to the standard error, there is no significant difference between these two models in terms of WAIC. Figures 10 and 11 show the box plots of draws from the marginal posterior distributions of the corrosion rates and initiation times for each TML in Circuit Q using the lognormal corrosion rate. These plots indicate that for

Parameters	Posterior Median	Posterior Standard Deviation	95% Credible Interval Lower	95% Credible Interval Upper
$\mu_{y0_{\text{elbow}}}$	0.438	0.00400	0.430	0.446
$\mu_{y0_{\text{pipe}}}$	0.431	0.00348	0.425	0.438
$\mu_{y0_{\text{tee}}}$	0.522	0.00621	0.509	0.534
σ_{y0}	0.0133	0.00192	0.0103	0.0178
$\beta_{1_{0.5}}$	0.0180	0.0113	0.00215	0.04216
$\sigma_{\log(\beta_1)}$	0.555	0.595	0.0684	2.22
$\mu_{\log(T_I)}$	1.13	0.323	0.264	1.55
$\sigma_{\log(T_I)}$	0.547	0.314	0.0515	1.29
σ_{ϵ}	0.00467	0.000399	0.00400	0.00556

Table 4: Marginal posterior distribution summaries of the parameters in the degradation model with a lognormal corrosion rate based on the pipeline data from Circuit Q in Facility 1.

Parameters	Posterior Median	Posterior Standard Deviation	95% Credible Interval Lower	95% Credible Interval Upper
$\mu_{y0_{\text{elbow}}}$	0.438	0.00402	0.430	0.446
$\mu_{y0_{\text{pipe}}}$	0.431	0.00346	0.425	0.438
$\mu_{y0_{\text{tee}}}$	0.522	0.00619	0.510	0.534
σ_{y0}	0.01332	0.00192	0.0104	0.0178
$\beta_{1_{0.5}}$	0.00671	0.00794	0.00161	0.0288
$\sigma_{\log(\beta_1)}$	2.82	1.42	0.171	4.87
$\mu_{\log(T_I)}$	0.748	0.345	0.112	1.35
$\sigma_{\log(T_I)}$	0.328	0.297	0.0110	1.06
σ_{ϵ}	0.00470	0.000402	0.00400	0.00559

Table 5: Marginal posterior distribution summaries of the parameters in the degradation model with Weibull corrosion rate based on the pipeline data from Circuit Q in Facility 1.

the TMLs where pipeline corrosion appears not to have initiated before the last inspection time, the posterior distributions of the initiation times are right skewed. Figure 12 compares plots of the marginal posterior distributions of the initiation times for TMLs with evidence of initiation and without initiation before the last inspections. These plots show that the marginal posterior distributions of the initiation times for the TMLs without initiation are shifted to the right, right skewed, and close to each other.

5 Models Relating Degradation and Failure for Circuit Q in Facility 1

5.1 Bayesian Evaluation of the Failure Time Distribution

As in the analysis of the pipeline data from Circuit G in Facility 3, there are two main purposes for using the degradation model. The first is to assess the lifetime cdf of individual pipeline components or segments. The second is to predict the remaining lifetime of the entire circuit. The degradation path over time is $\mathcal{D} = \mathcal{D}(t)$. A soft failure is defined to be the time at which the remaining pipeline thickness is less than 20% of the mean of the thickness at the installation date.

Suppose that $T_I \sim \text{Lognormal}\,(\mu_{\log(T_I)}, \sigma_{\log(T_I)})$, $Y_0 \sim \text{NORM}\,(\mu_{y_{0_\text{elbow}}}, \sigma_{y_0})$, $Y_0 \sim \text{NORM}\,(\mu_{y_{0_\text{pipe}}}, \sigma_{y_0})$, $Y_0 \sim \text{NORM}\,(\mu_{y_{0_\text{tee}}}, \sigma_{y_0})$, and $\beta_1 \sim \text{Lognormal}\,(\mu_{\log(\beta_1)}, \sigma_{\log(\beta_1)})$. Then the cdf giving the proportion of pipeline segments that have a soft failure as a function of operating time is

$$
\begin{aligned}
F(t) &= \Pr(\mathcal{D}(t) \le \mathcal{D}_f) = \Pr(Y_0 - \beta_1(t - T_I) I(t \ge T_I) \le \mathcal{D}_f) \\
&= \Pr\Big(Y_0 - \beta_1(t - T_I) \le \mathcal{D}_f \bigcap t \ge T_I\Big) + \Pr\Big(Y_0 \le \mathcal{D}_f \bigcap t < T_I\Big) \\
&= \Pr\Big(Y_0 - \beta_1(t - T_I) \le \mathcal{D}_f \bigcap t \ge T_I\Big) + \Pr(Y_0 \le \mathcal{D}_f) \Pr(t < T_I) \\
&= \iiint\limits_{y_0 - \beta_1(t - T_I) \le \mathcal{D}_f,\, t \ge T_I} \frac{1}{\sigma_{y_0}} \phi_\text{NOR}(z_{y_0}) \times \frac{1}{T_I \sigma_{\log(T_I)}} \phi_\text{NOR}(z_{T_I}) \times \frac{1}{\beta_1 \sigma_{\log(\beta_1)}} \phi(z_\beta) \mathrm{d}y_0 \, \mathrm{d}T_I \, \mathrm{d}\beta_1 \\
&\qquad + \Phi_\text{NOR}\left(\frac{\mathcal{D}_f - \mu_{y_0}}{\sigma_{y_0}}\right) \times \left[1 - \Phi_\text{NOR}(z_{T_I})\right],
\end{aligned} \tag{8}
$$

where $z_{y_0} = (y_0 - \mu_{y_0})/\sigma_{y_0}$, $z_{T_I} = (\log(T_I) - \mu_{\log(T_I)})/\sigma_{\log(T_I)}$, and $z_\beta = (\log(\beta_1) - \mu_{\log(\beta_1)})/\sigma_{\log(\beta_1)}$. When the corrosion rate has a lognormal distribution, $\phi(z_\beta) = \phi_\text{NOR}(z_\beta)$ is the standard

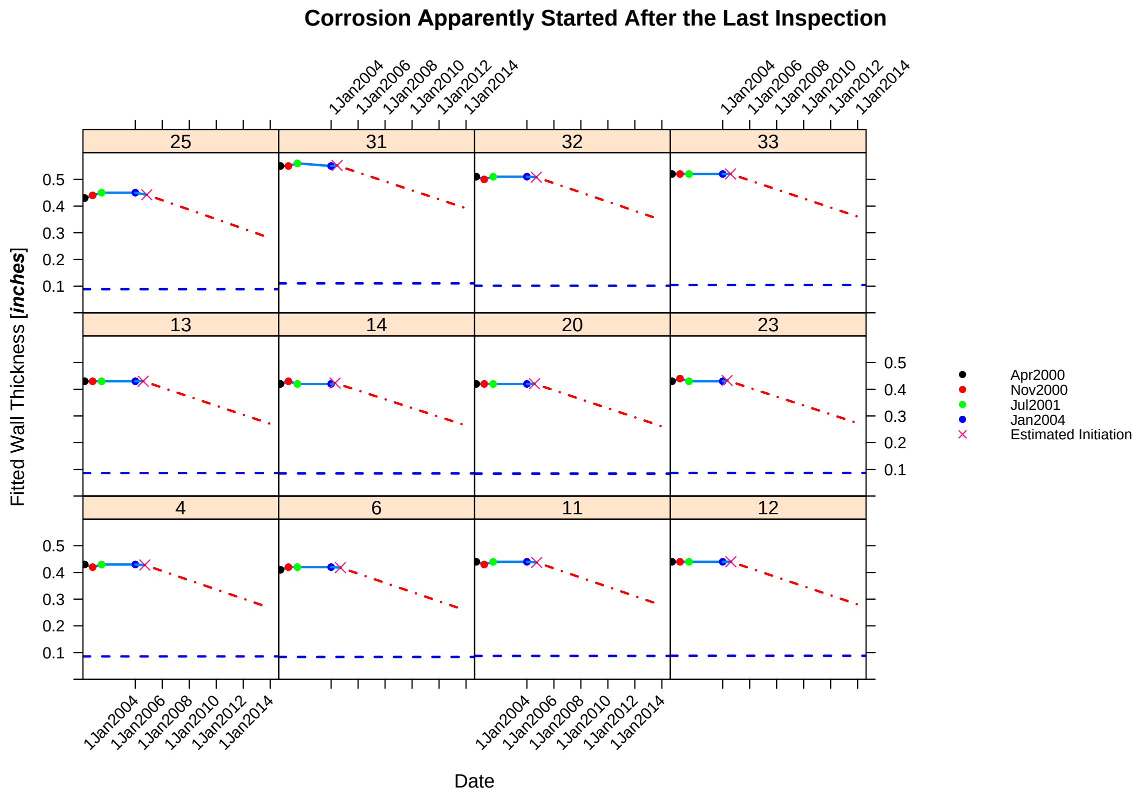


Figure 8: Trellis plot of the fitted thickness values for TMLs from Circuit Q in Facility 1 for which corrosion apparently had not started before the last inspection using the lognormal corrosion rate distribution. The dot-dash lines indicate extrapolation of the predicted degradation path (median of the marginal posterior as a function of time). The horizontal dashed lines indicate the soft failure definition level.

($\mu = 0$, $\sigma = 1$) normal probability density function (pdf). When the corrosion rate has a Weibull distribution, $\phi(z_\beta) = \phi_{\text{SEV}}(z_\beta) = \exp(z_\beta - \exp(z_\beta))$ is the standardized smallest extreme value pdf. Because $F(t)$ in (8) does not have a closed form, the Monte Carlo simulation method described in Section 13.5.3 of Meeker and Escobar (1998) is used to evaluate failure-time cdfs, using 1,000 simulation trials for the evaluation. Figure 13 shows failure-time cdfs for elbows, pipes and tees using lognormal and Weibull disrtibuted corrosion rates. The plots suggest for the Model 3 in (7), the Weibull distributed corrosion rate provides more conservative results compared with the lognormal corrosion rate. Figure 14 shows failure time cdfs for elbow, pipe and tee segments using the lognormal corrosion rate distribution with two-sided 95% and 80% credible intervals.

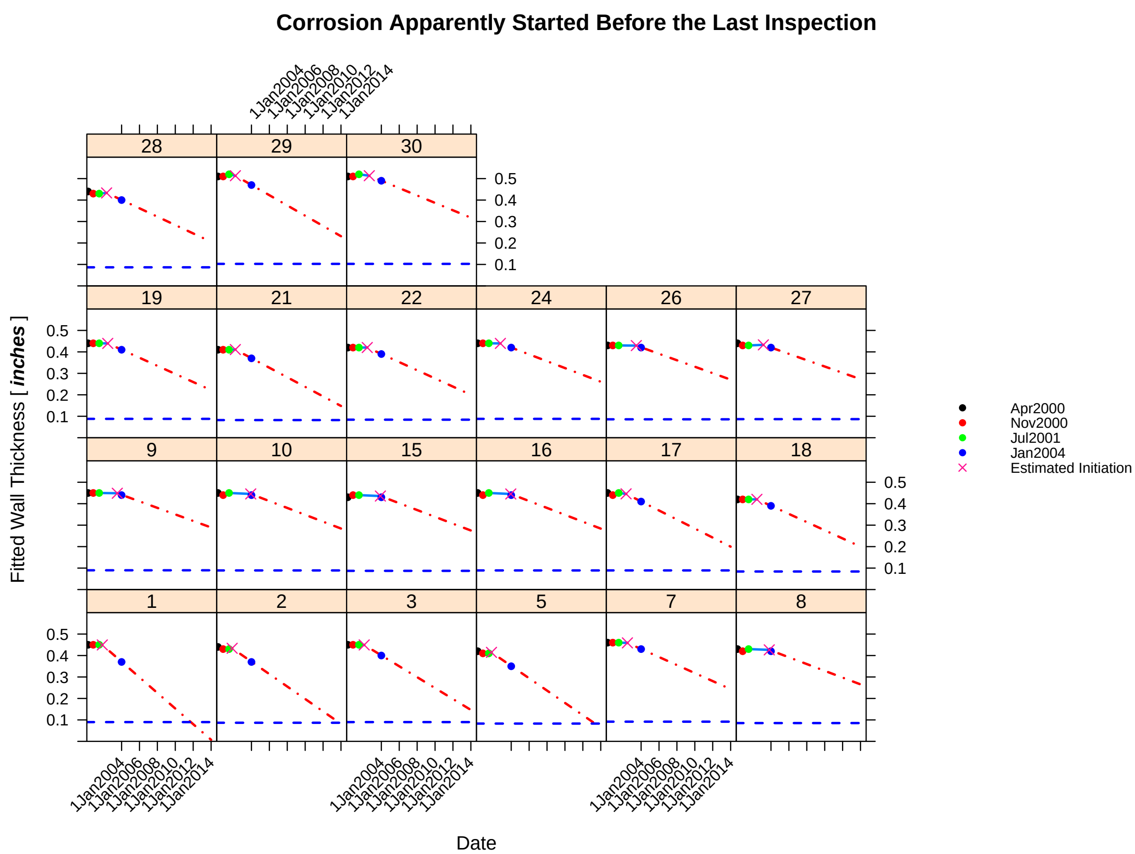


Figure 9: Trellis plot of the fitted thickness values for the TMLs from Circuit Q in Facility 1 for which corrosion had apparently started before the last inspection using the lognormal corrosion rate distribution. The dot-dashed lines indicate extrapolation of the predicted degradation path (median of the marginal posterior as a function of time). The horizontal dashed lines indicate the soft failure definition level.

5.2 Prediction of the Remaining Life of the Current Circuit

Figure 15 compares the remaining lifetime cdfs with lognormal and Weibull corrosion rates for elbows, pipes and tees. The plots suggest that using a Weibull distribution for the corrosion rate in (7) provides a more conservative estimate of remaining life. Figure 16 shows estimates of the remaining lifetime cdfs using the lognormal distribution for the corrosion rate in (7) and the corresponding two-sided 95% and 80% credible intervals.

As in Section 3.2, we are primarily interested in estimating small quantiles of the minimum remaining lifetime cdfs for Circuit Q in Facility 1. Figure 17 shows the posterior

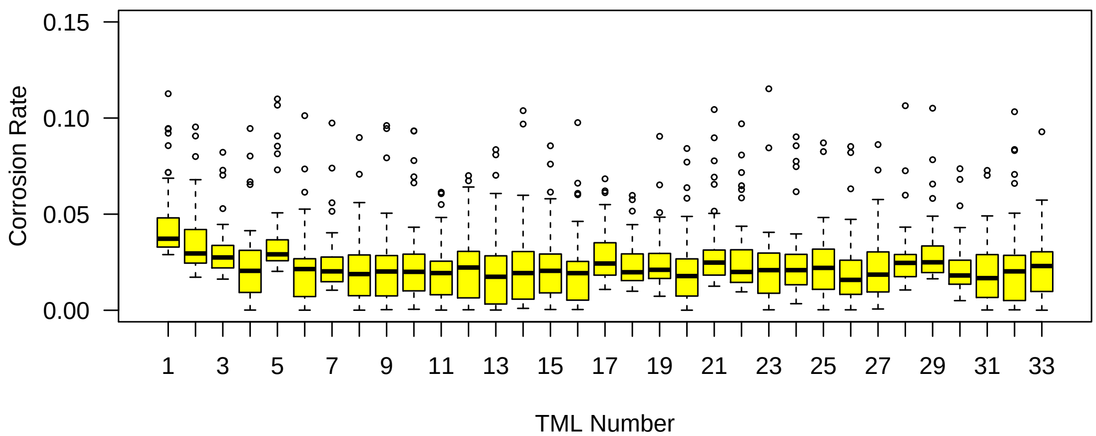


Figure 10: Draws from the marginal posterior distribution of the lognormal corrosion rates for each TML in Circuit Q in Facility 1.

density of the 0.1, 0.2, 0.3, and 0.4 quantiles of the minimum remaining lifetime distribution from the model in (7) with the population size $M = 100$ using the lognormal distribution for corrosion rate. The larger quantile estimates for the tee components indicate that tees have a longer remaining lifetime. The results are consistent with what we observed previously in Figures 14 and 16.

6 Effect of Additional Inspections on Identifiability

In Section 4.2, in the analysis of the pipeline data from the Circuit Q in Facility 1, we used a moderately informative prior distribution to describe knowledge of the median of the corrosion rates, alleviating the identifiability problem that was caused by having no more than one inspection after any of the observed corrosion initiation events. The results of that analysis showed a large amount of uncertainty in predictions of remaining life. In the actual application, the owners of the pipeline would have to wait until after the next inspection time to obtain more precise estimates of remaining life without using informative prior information.

To investigate this identifiability problem, in this section, we simulate data from the

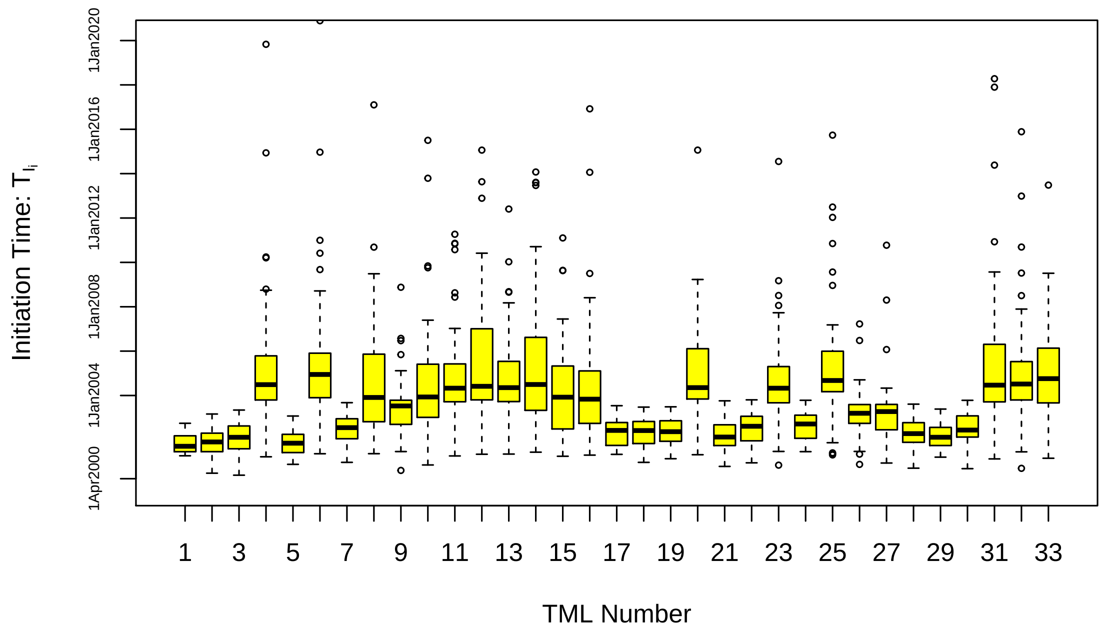


Figure 11: Draws from the marginal posterior distribution of the corrosion initiation times for each TML in Circuit Q in Facility 1 using a lognormal distribution to describe corrosion rates.

Model 3 in (7) such that there is more than one inspection after an initiation (i.e., data that is similar to those from Circuit Q in Facility 1 but with additional inspections at future times). We continue to use a lognormal corrosion rate distribution. Figure 18 displays the time plot for the simulated pipeline data from a single circuit with 33 TMLs and three components: elbow, straight pipe and tee pipe. Corrosion was measured at each TML at 5 times.

We use the same weakly informative prior distributions used in Section 4.2 for all parameters except for the median of the corrosion rates $\beta_{1_{0.5}}$, $i = 1, 2, \ldots, 33$. Because there is more than one inspection after the corrosion initiation in the simulated data, the identifiability problem no longer exists. Therefore, rather than restrict the upper bound of the prior distribution of $\beta_{1_{0.5}}$ to 0.022 inches per year, we can relax the upper bound to 0.10 inches per year providing a weakly informative prior for $\beta_{1_{0.5}}$.

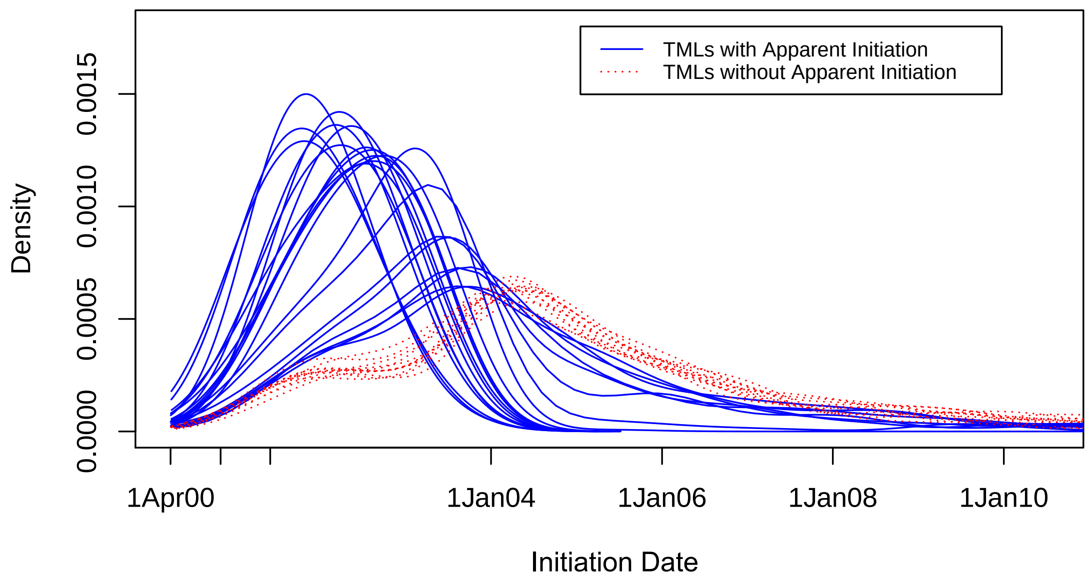


Figure 12: Posterior densities of the initiation times for each TML in Circuit Q in Facility 1 under a lognormal corrosion rate.

For these simulated data, the Bayesian parameter estimates are close to the true parameter values from which the data were simulated. Figure 19 shows the trellis plot of the fitted thickness values for the simulated pipeline data using the weakly informative prior distributions.

As in Section 5, we used the Monte Carlo simulation method to evaluate the marginal posterior distributions of the failure time cdf at chosen points in time. Figure 20 shows the failure time cdfs for the simulated pipeline data of a single circuit, using the weakly informative priors. Compared with the results in Figure 14 for the pipeline data from Circuit Q in Facility 1, the credible intervals in Figure 20 are much narrower. This is because in the simulated data we have more inspections after the corrosion initiation. Thus, the identifiability problem that caused the wide intervals is no longer present. From a practical perspective, having several inspections that occur after an initiation time provides a much more effective estimation of pipeline segment lifetime distributions.

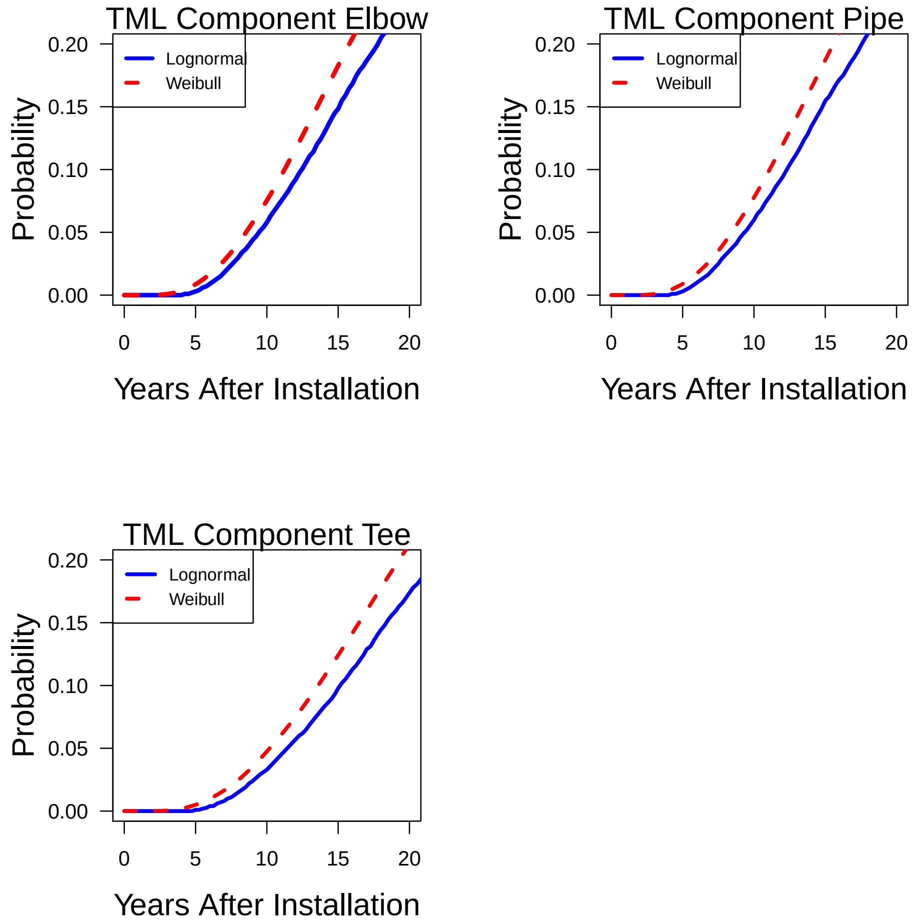


Figure 13: Degradation model estimates of failure-time cdfs for pipeline components from Circuit Q in Facility 1 comparing lognormal and Weibull corrosion rate distributions in the Model 3 in (7).

7 Concluding Remarks and Areas for Future Research

In this paper, we developed degradation models to describe the pipeline corrosion behavior for two particular pipeline data sets. The Bayesian approach with appropriate prior distributions is useful for estimating parameters in the degradation models. The Bayesian method, as an alternative to the likelihood approach, provides a convenient method to estimate and compute credible bounds for functions of the degradation model parameters, even when a

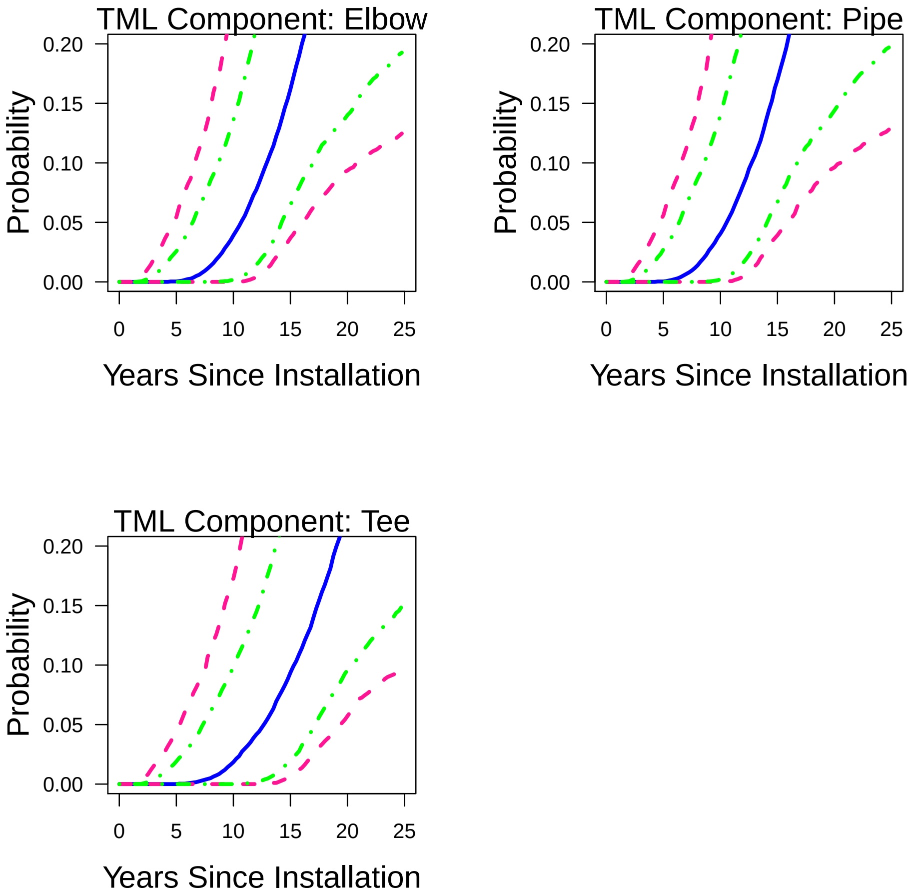


Figure 14: Degradation model estimates (the center lines) of failure time cdfs for pipeline components from Circuit Q in Facility 1 with the lognormal corrosion rate distribution in the Model 3 in (7) and two-sided 95% and 80% credible intervals.

closed-form expression of the function does not exist. A simulation study in Liu, Meeker, and Nordman (2014) shows that these intervals have frequentist coverage probabilities that are close to the nominal credible level. The failure time and the remaining lifetime distributions and small quantile estimates of the minimum remaining lifetime distribution provide useful information to evaluate of the life of a pipeline.

There remains, however, a number of areas for future research. These include:

- In the degradation model for corrosion initiation and growth, test planning methods

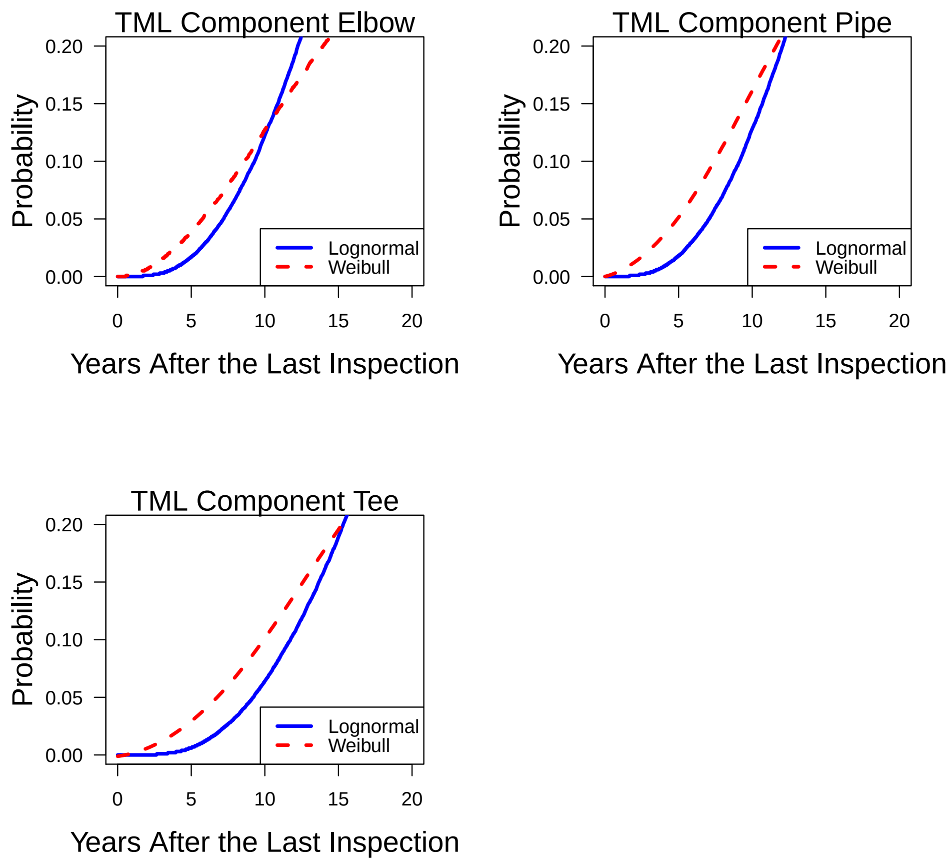


Figure 15: Degradation model estimates of remaining lifetime cdfs for pipeline components from Circuit Q in Facility 1 comparing lognormal and Weibull corrosion rate distributions in the Model 3 in (7).

(see Section 9.6 of Hamada et al. 2008) could be developed to choose an appropriate number of inspections after the corrosion initiation to obtain more precise estimate of the failure time distribution.

- The model with linear degradation paths and the constant corrosion rate can be extended to the models having nonlinear relationships between pipeline thickness and time.

- Each pipeline circuit within a facility, viewed as a series system of many segments, could be considered as a component in a large complex system of circuits. In some applications, the life time of such a pipeline system could be important.

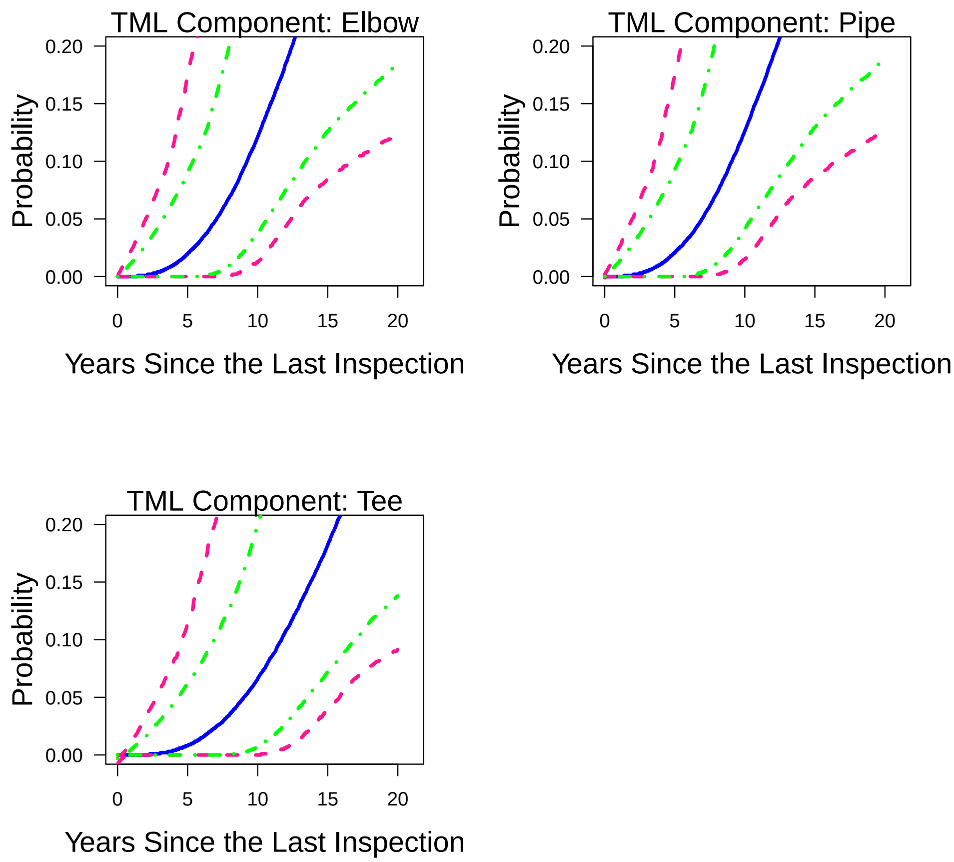


Figure 16: Degradation model estimates (the center lines) of remaining lifetime cdfs for pipeline components from Circuit Q in Facility 1 with the lognormal corrosion rate distribution in the Model 3 in (7) and two-sided 95% and 80% credible intervals.

• In some pipeline applications, it may be possible to obtain dynamic covariate information such as temperature, flow, and type of material. The degradation models could then be generalized by incorporating this dynamic covariate information into the modeling and analysis, in a manner similar to that used in Hong, et al. (2015).

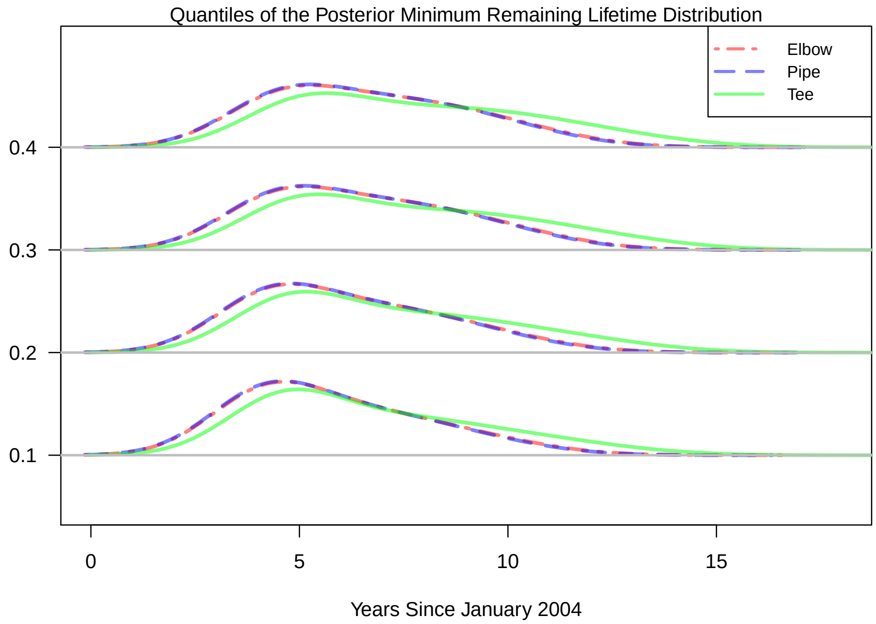


Figure 17: Posterior density of the 0.1, 0.2, 0.3, and 0.4 quantiles of the minimum remaining lifetime distribution (years after the last inspection time t_c: January 2004) with the population size $M = 100$ using the lognormal corrosion rate distribution of pipeline data from Circuit Q in Facility 1.

Acknowledgments

We would like to thank two anonymous referees who made important suggestions to help us improve this paper. We would also like to thank Qianqian Shan for helpful comments on an earlier version of this paper.

References

[1] Barlow, R.E. and Proschan, F. (1975). *Statistical Theory of Reliability and Life Testing.* New York: Holt, Rinehart and Winston.

[2] Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, 1, 515-533.

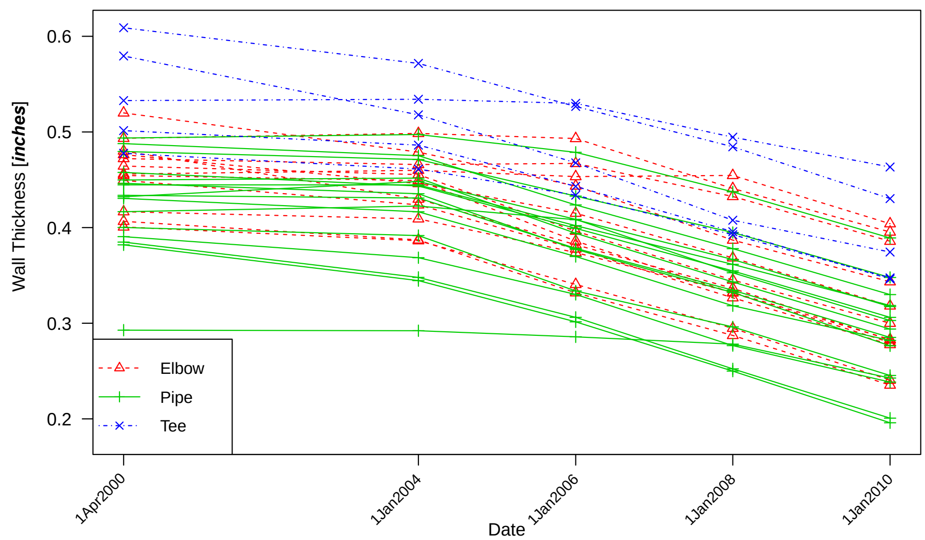


Figure 18: Time plot for the simulated pipeline data from a single circuit with 33 TMLs.

[3] Hamada, M. S., Wilson, A., Reese, C. S. and Martz, H. (2008). *Bayesian Reliability*. New York: Springer.

[4] Hong Y., Duan Y., Meeker, W. Q., Stanley, D. L., Gu, X. (2015). Statistical methods for degradation data with dynamic covariates information and an application to outdoor weathering data. *Technometrics*, 57, 180-193.

[5] Lewandowski, D., Kurowicka D., Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*, 100, 1989-2001.

[6] Liu, J., Nordman, D. J. and Meeker, W. Q., (2016). The number of MCMC iterations needed to compute Bayesian credible intervals. *The American Statistician*, 70, 275-284.

[7] Lu, C. J. and Meeker, W. Q. (1993). Using degradation measures to estimate a time-to-failure distribution. *Technometrics*, 35, 161-174.

[8] Meeker, W. Q. and Escobar, L. A. (1998). *Statistical Methods for Reliability Data*. New York: John Wiley and Sons.

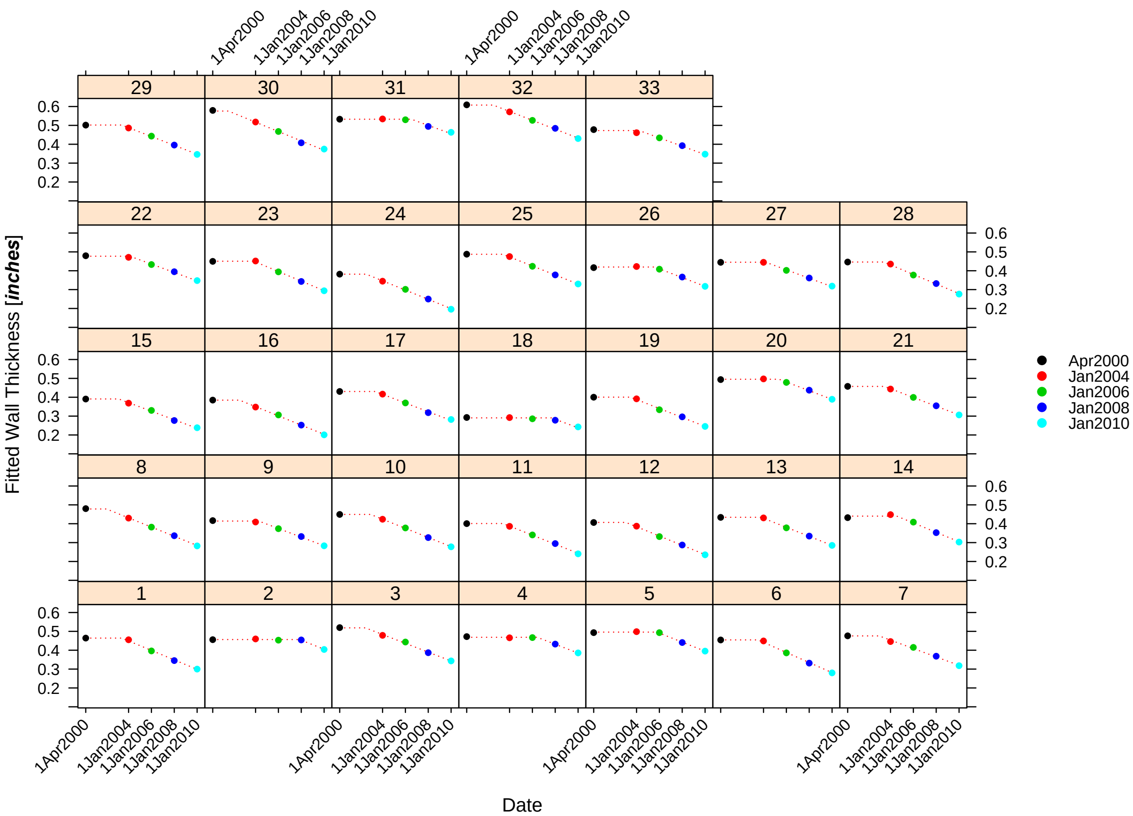


Figure 19: Trellis plot of the fitted thickness values for the simulated pipeline data in a single circuit with 33 TMLs using the weakly informative prior distributions.

[9] Nelson, W. B. (2009). Defect initiation, growth, and failure a general statistical model and data analyses. In M. Nikulin, N. Limnios, N. Balakrishnan, W. Kahle, and C. Huber-Carol (Eds.), *Advances in Degradation Modeling: Applications to Reliability, Survival Analysis, and Finance (Statistics for Industry and Technology)*, 27-37. Boston: Birkhuser.

[10] Pandey, M. D. (1998). Probabilistic models for condition assessment of oil and gas pipelines. *NDT & E International*, 31, 349-358.

[11] Sheikh, A. K., Boah, J. K. and Hansen, D. A. (1990). Statistical modeling of pitting corrosion and pipeline reliability. *Corrosion*, 46, 190-197.

[12] Vehtari, A., Gelman, A. and Gabry, J.(2017). Practical Bayesian model evaluation using

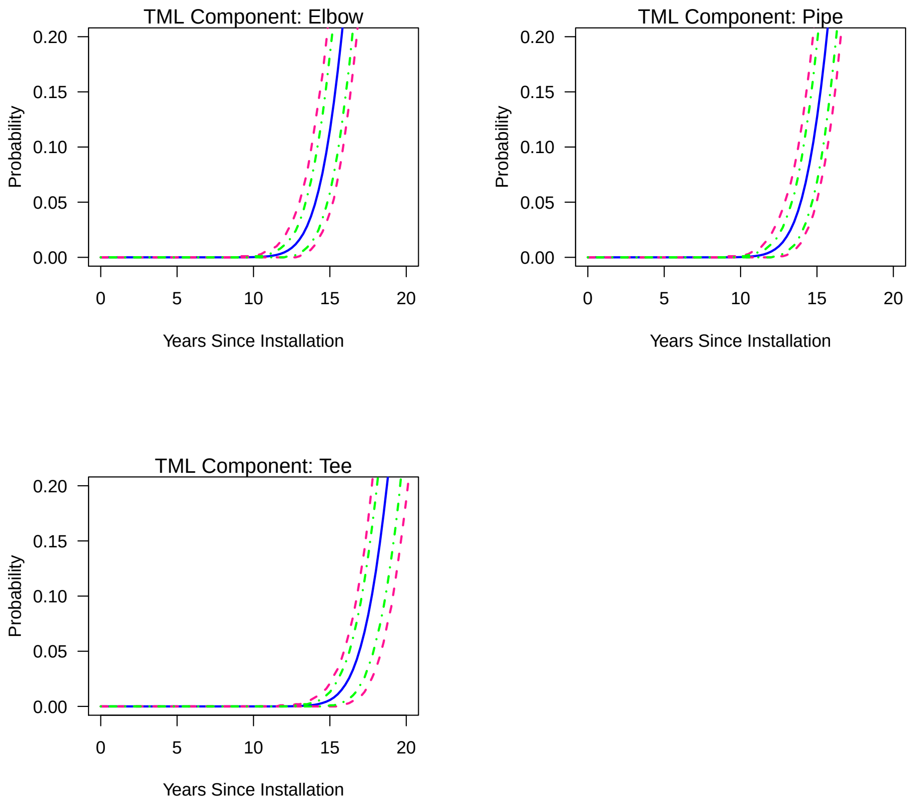


Figure 20: Degradation model estimates of failure time cdfs for the simulated pipeline data in a single circuit with 33 TMLs using the lognormal corrosion rate distribution and weakly informative priors in the Model 3 in (7) and two-sided 95% and 80% credible intervals.

leave-one-out cross-validation and WAIC . *Statistics and Computing*, 27, 1413.

[13] Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and Widely Applicable Information Criterion in singular learning theory. *Journal of Machine Learning Research*, 11, 3571-3594.