Iowa State University
Digital Repository

Statistics Preprints Statistics

10-2014

Using Degradation Models to Assess Pipeline Life

Shiyao Liu
Iowa State University

William Q. Meeker
Iowa State University, wqmeeker@iastate.edu

Follow this and additional works at: http://lib.dr.iastate.edu/stat_las_preprints

Part of the Statistics and Probability Commons

Recommended Citation

Liu, Shiyao and Meeker, William Q., "Using Degradation Models to Assess Pipeline Life" (2014). *Statistics Preprints*. 127.
http://lib.dr.iastate.edu/stat_las_preprints/127

This Article is brought to you for free and open access by the Statistics at Iowa State University Digital Repository. It has been accepted for inclusion in Statistics Preprints 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

Abstract
Longitudinal inspections of thickness at particular locations along a pipeline provide useful information to assess the lifetime 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 the 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 an entire pipeline circuit.

Keywords
Bayesian, longitudinal data, pipeline reliability, remaining life

Disciplines
Statistics and Probability

Comments
This preprint was published as Shiyao Liu and William Q. Meeker, "Using Degradation Models to Assess Pipeline Life".

This article is available at Iowa State University Digital Repository: http://lib.dr.iastate.edu/stat_las_preprints/127

Using Degradation Models to Assess Pipeline Life

Shiyao Liu, William Q. Meeker

Department of Statistics

Iowa State University

October 8, 2014

Abstract

Longitudinal inspections of thickness at particular locations along a pipeline provide useful information to assess the lifetime 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 the 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 an entire pipeline circuit.

Key Words: Bayesian; 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 curcuit. 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 each application. 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, evaluation of the failure time and remaining lifetime distributions is also computationally feasible and efficient when using the Bayesian method.

1.2 Pipeline Data

Figure 1 shows time plot of longitudinal pipeline data from Circuit G in Facility 3. Data were obtained from a sample thickness measurement locations (TMLs). For the first two inspections, only 12 TMLs were used. Subsequently, as perceived risk of failure increased, an additional 76 TMLs were used. Some of these TMLs correspond to elbows and the others correspond to straight pipes. For each TML, the thickness was measured at four different quadrants located at the 0, 90, 180, and 270 degree position (top, right, bottom, and left for a horizontal pipeline). 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.

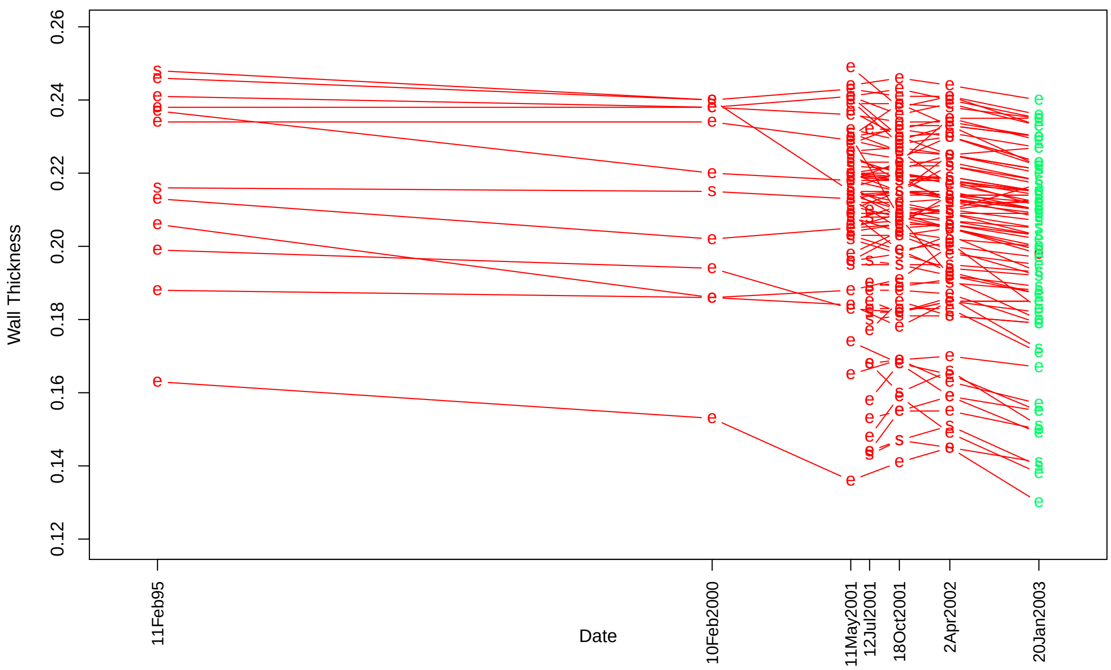

Figure 1: Time plot for pipeline data from Circuit G in Facility 3.

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 4 times. 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.

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

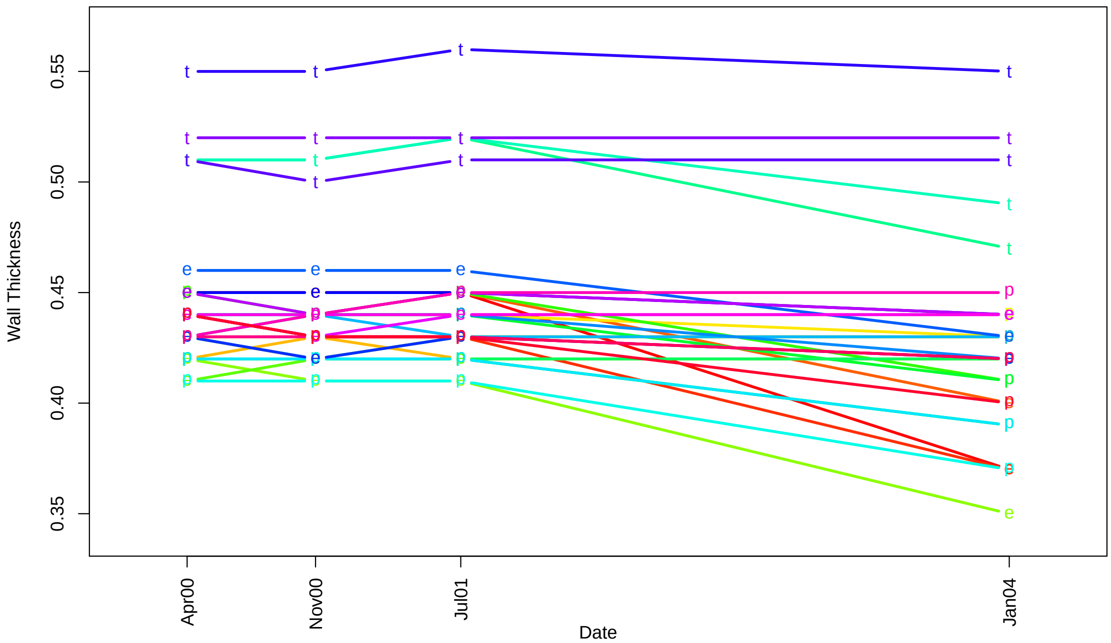

Figure 2: Time plot for pipeline data from Circuit Q in Facility 1.

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 repair 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 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. In order to study the data needed for estimability, Section 6 analyzes simulated data for a single circuit having more than one inspection after corrosion initiation. Section 7 contains the 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 diffuse 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_{it_k} denote the pipeline thickness at time t_k for TML i ($i = 1, 2, \ldots, 88$; $k = 1, 2, \ldots, 7$). We assume that the degradation path of Circuit G in Facility 3 is linear with respect to inspection time and has the form

$$Y_{it_k} = y_0 - \beta_{1_i}(t_k - t_0) + \epsilon_{ik} \tag{1}$$

where β_{1_i} is -1 times the corrosion rate at location i and ϵ_{ik} is the measurement error term. 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 could only be negative, β_{1_i} in the degradation model (1) is a positive random variable. To guarantee a positive β_{1_i}, we assume that β_{1_i} has a lognormal distribution [i.e., $\beta_{1_i} \sim$ Lognormal $(\mu_{\beta_1}, \sigma^2_{\beta_1})$] and that the measurement error is $\epsilon_{ik} \sim$ NOR $(0, \sigma^2_\epsilon)$. Thus the parameters in the degradation model (1) are: $\boldsymbol{\theta} = (\mu_{\beta_1}, \sigma_{\beta_1}, \sigma_\epsilon)'$.

2.2 Bayesian Estimation of the Parameters in the Degradation Model

Bayesian estimation with the use of diffuse 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 variance [i.e., NOR $(0, 10^3)$] as the prior distribution for the parameter μ_{β_1}. The prior distributions for σ_{β_1} and σ_ϵ are Uniform $(0, 5)$. We obtain a large number of draws from the joint posterior distribution of the degradation model parameters using Markov Chain Monte Carlo (MCMC) implemented in OpenBUGS. Table 1 presents marginal posterior distribution summaries for the parameters in $\boldsymbol{\theta}$, including the mean and 95% credible intervals. Figure 3 shows the time plot of the fitted thickness values for Circuit G in Facility 3 with a 10-years extrapolation after the last inspection in January 20, 2003.

			95% Credible Interval	
Parameters	Posterior Mean	Posterior Std. Dev.	Lower	Upper
μ_{β_1}	-11.62	0.06263	-11.75	-11.50
σ_{β_1}	0.5753	0.04743	0.4910	0.6768
σ_ϵ	0.006045	2.537E$-$4	0.005575	0.006574

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).

2.3 Statistical Model for Different Quadrants

In Sections 2.1 and 2.2, we assumed that the corrosion rates of different quadrants from the same location follow the same distribution. In non-vertical pipes, however, the corrosion 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 assumes that means of

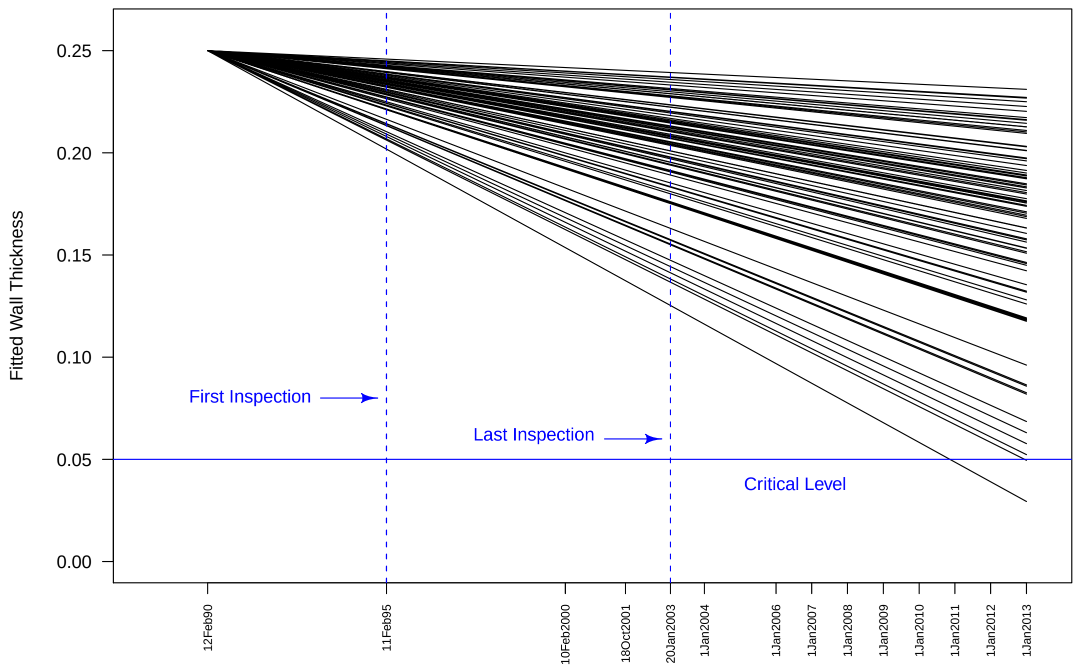

Figure 3: Time plot showing the fitted thickness values for the pipeline data from Circuit G in Facility 3 using the degradation model (1).

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

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

where $\beta_{1_{i_j}}$ is the corrosion rate of quadrant j at TML i ($i = 1, 2, \ldots, 22$; $k = 1, 2, \ldots, 7$; $j = 1, \ldots, 4$) and ϵ_{ijk}, as before, is the measurement error term. Similar to model (1), $\beta_{1_{i_j}}$ is also positive in model (2). We assume that $\beta_{1_{i_j}}$ has a lognormal distribution [i.e., $\beta_{1_{i_j}} \sim$ Lognormal $(\mu_{\beta_{1_j}}, \sigma^2_{\beta_1})$] and $\epsilon_{ijk} \sim$ NOR $(0, \sigma^2_\epsilon)$. The parameters in model (2) are: $\boldsymbol{\theta} = (\mu_{\beta_{1_1}}, \mu_{\beta_{1_2}}, \mu_{\beta_{1_3}}, \mu_{\beta_{1_4}}, \sigma_{\beta_1}, \sigma_\epsilon)'$. The Bayesian method is again used to estimate $\boldsymbol{\theta}$. Table 2 presents marginal posterior distribution summaries for the parameters in $\boldsymbol{\theta}$, including the mean and 95% credible intervals. Figure 4 shows the time plot of the fitted thickness values for different quadrants of this circuit.

			95% Credible Interval	
Parameters	Posterior Mean	Posterior Std. Dev.	Lower	Upper
$\mu_{\beta_{1_1}}$	−11.31	0.1182	−11.54	−11.08
$\mu_{\beta_{1_2}}$	−11.67	0.1179	−11.90	−11.45
$\mu_{\beta_{1_3}}$	−11.59	0.1167	−11.82	−11.36
$\mu_{\beta_{1_4}}$	−11.92	0.1219	−12.16	−11.68
σ_{β_1}	0.5416	0.04551	0.4619	0.6377
σ_{ϵ}	0.006033	2.556E−4	0.005556	0.006551

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).

The deviance information criterion (DIC) (defined in Gelman et al. 2003 on page 182-184), a measure of model goodness-of-fit and complexity, is used for the Bayesian model comparison. The values of DIC for models (1) and (2) are −2574.0 and −2633.0, respectively. Because model (2) has an importantly smaller DIC than model (1), we can conclude that there is a quadrant effect.

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 $\mathcal{D} = \mathcal{D}(t, \boldsymbol{\theta})$. 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 our example). This is known as a "soft failure" definition and such critical levels are determined through engineering judgment as the thickness below which there is risk of a leak. Because $\beta_{1_{ij}} \sim$ Lognormal $(\mu_{\beta_{1_j}}, \sigma^2_{\beta_1})$ in model (2), the failure time cumulative distribution function (cdf) $F(t)$ of individual segments in a population of

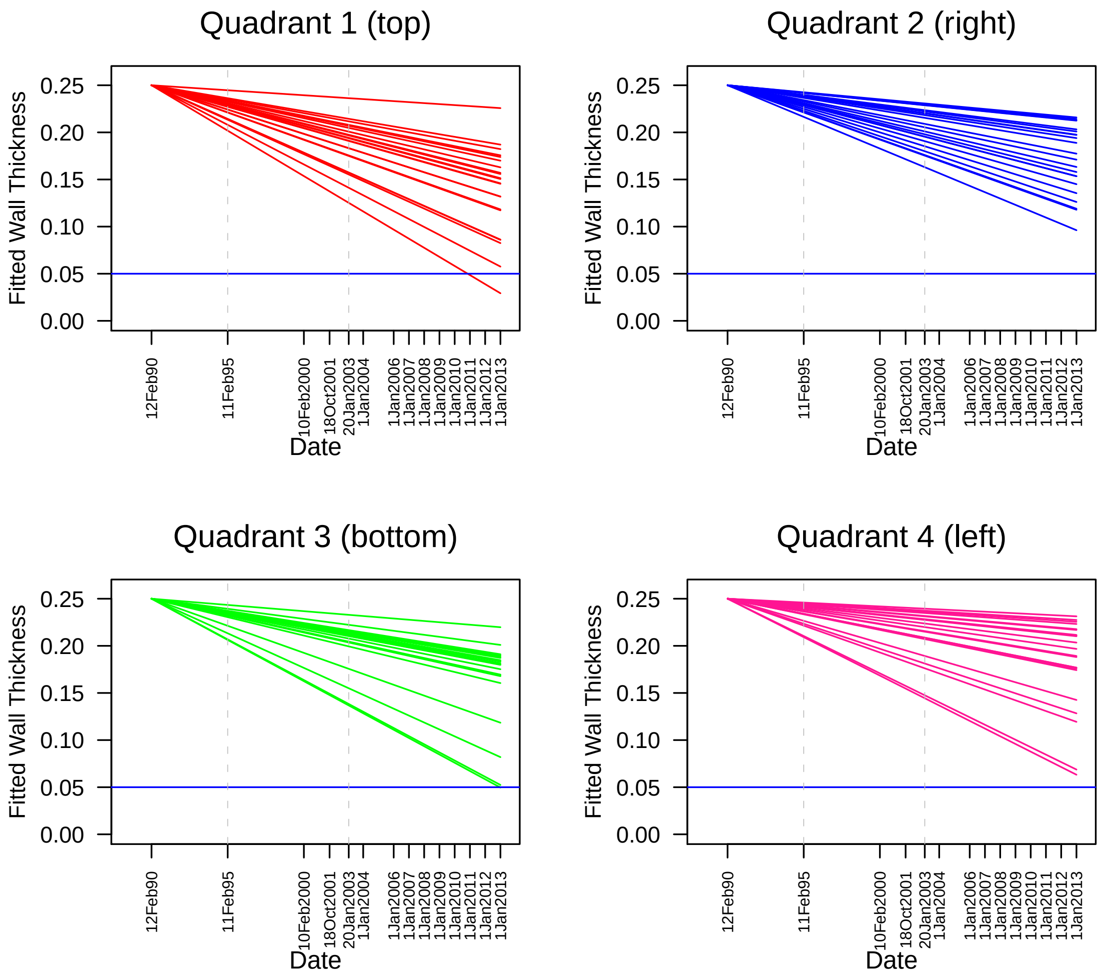

Figure 4: Time plot showing the fitted thickness values for different quadrants of pipeline data from Circuit G in Facility 3 using the degradation model (2).

segments of quadrant j in the pipeline can be expressed in a closed form:

$$\begin{aligned} F(t) &= \Pr(\mathcal{D}(t) \leq \mathcal{D}_f) = \Pr(y_0 - \beta_{1_{i_j}}(t_k - t_0) \leq 0.05) \\ &= \Pr\left(\beta_{1_{i_j}} \geq \frac{0.20}{t_k - t_0}\right) = 1 - \Phi_{\text{nor}}\left(\frac{\log(0.20) - \log(t_k - t_0) - \mu_{\beta_{1_j}}}{\sigma_{\beta_1}}\right) \\ &= \Phi_{\text{nor}}\left(\frac{\log(t_k - t_0) - \log(0.20) + \mu_{\beta_{1_j}}}{\sigma_{\beta_1}}\right). \end{aligned} \tag{3}$$

where Φ_{nor} is the standard normal cdf.

The failure time distribution, as a function of the degradation parameters, can be evaluated simply by using the Bayesian approach. For each draw from the joint posterior distribution, one can evaluate $F(t)$ in (3) to obtain a corresponding draw from the marginal posterior distribution of failure time cdf. 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. Figure 5 (a) 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 failure time cdf in (3) is an estimate of the cdf for an individual pipeline segment. Although the primary interest is to estimate the lifetime of a pipeline viewed as a series system of many segments, the life time cdf of an individual pipeline segment provides useful information to plan the construction of future pipelines.

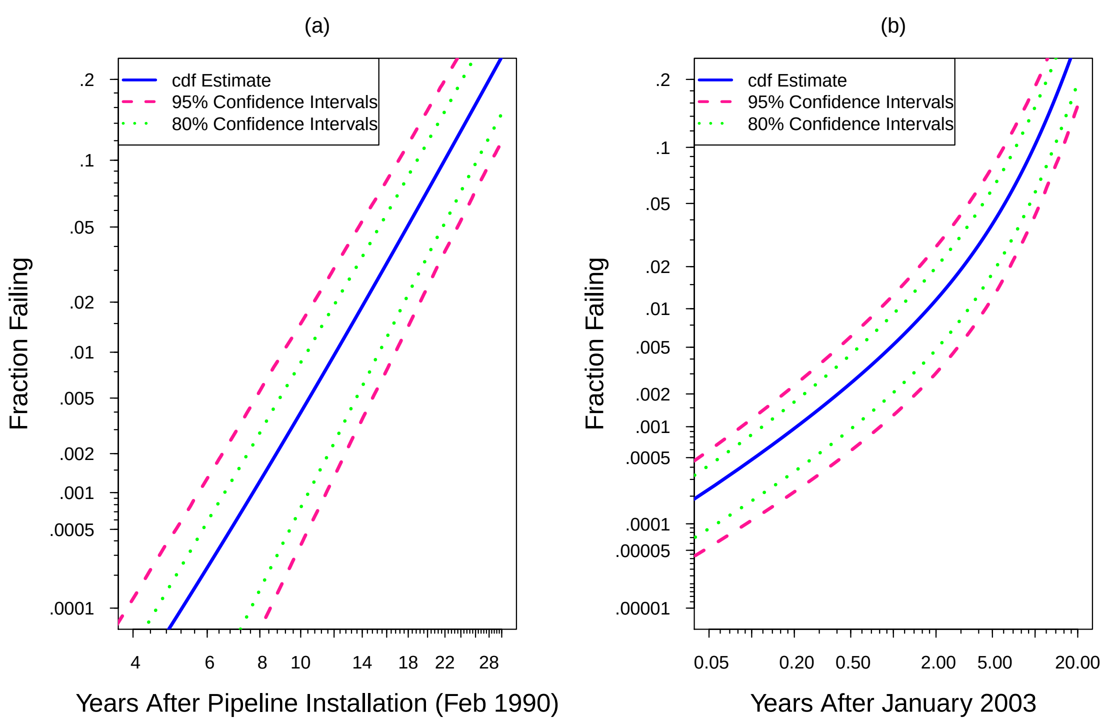

Figure 5: Degradation model estimates of (a) failure time cdf (years after pipeline installation) and (b) remaining lifetime cdf (years after the last inspection t_c) with two-sided 95% and 80% credible intervals on the lognormal paper for pipeline data from quadrant 1 of Circuit G in Facility 3.

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 $F_{RM}(t)$ conditional on surviving until the last inspection time (January 2003) is

$$F_{RM}(t) = \Pr(T \le t | T > t_c) = \frac{F(t; \boldsymbol{\theta}) - F(t_c; \boldsymbol{\theta})}{1 - F(t_c; \boldsymbol{\theta})},\ t \ge t_c \tag{4}$$

where t_c is the last inspection time and $F(t)$ 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. Figure 5 (b) shows the posterior estimates of the remaining lifetime cdf after the last inspection in January 2003 with 95% and 80% credible intervals.

In the pipeline application, it is of great interest to estimate small quantiles of the minimum remaining lifetime of the population. To do this, one needs to extrapolate further into the tail of the remaining life distribution estimated for a given segment. Typically a TML segment is about one foot long. Suppose that the entire pipeline length has M segments of this length. Then the cdf of the minimum remaining life among all of the M segments along the pipeline can be expressed as

$$F_M(t) = \Pr[T_{\min} \le t] = 1 - [1 - F_{RM}(t)]^M \tag{5}$$

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

$$t_p = F_M^{-1}(p) = F_{RM}^{-1}\left(1 - (1-p)^{\frac{1}{M}}\right). \tag{6}$$

This indicates that p quantile of the minimum remaining lifetime distribution of the population of M segments corresponds to the $1 - (1-p)^{1/M}$ quantile of the remaining lifetime cdf for each 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 $M = 100$ using the degradation models (1) and (2) respectively. Model (2) is more conservative than model (1) as it generates the smaller quantile estimates.

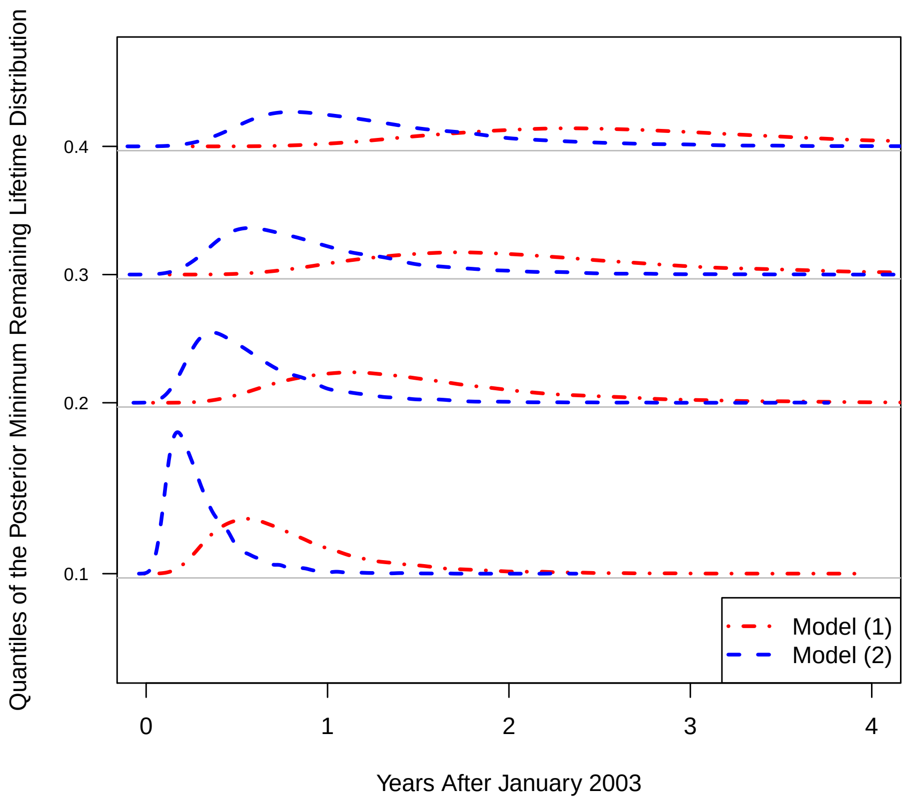

Figure 6: Posterior density of the 0.1, 0.2, 0.3 and 0.4 quantiles of the minimum remaining lifetime distribution (years since the last inspection time t_c: January 2003) with the population size $M = 100$ of pipeline data from Circuit G in Facility 3 using the degradation models (1) and (2).

The small quantile estimates suggest that the Circuit G in Facility 3 could have leakage risks within one year after the last 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 the unexpected pipeline leakage.

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 for a specific TML. The trellis plot suggests an interesting pipeline corrosion process. For example, in 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 initiated between the third and forth inspection times. At some TMLs (e.g., TMLs #12, #13, and #33), the corrosion appears not to have initiated before the last inspection time.

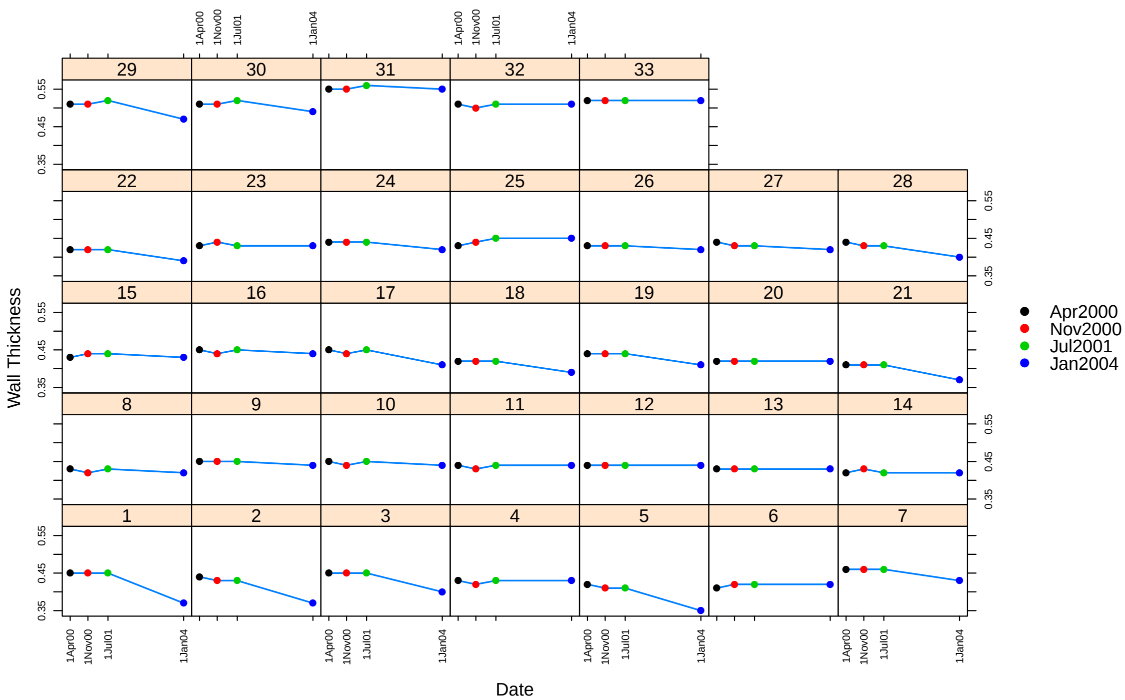

Figure 7: Trellis plot for pipeline data from Circuit Q in Facility 1.

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 ($i = 1, 2, \dots, 33$; $j = 1, 2, \dots, 4$) 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,

- Y_{it_j} denotes the thickness measurement for TML i at time t_j.
- Y_{0i} is the original thickness 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 $Y_{0i} \sim \text{NOR}\ (\mu_{y0_{\text{elbow}}}, \sigma^2_{y0})$;
  - If the TML is a pipe, we assume that $Y_{0i} \sim \text{NOR}\ (\mu_{y0_{\text{pipe}}}, \sigma^2_{y0})$;
  - If the TML is a tee, we assume that $Y_{0i} \sim \text{NOR}\ (\mu_{y0_{\text{tee}}}, \sigma^2_{y0})$.
- β_{1i} is the corrosion rate for TML i and we assume that $\beta_{1i} \sim \text{Lognormal}\ (\mu_{\beta_1}, \sigma^2_{\beta_1})$ or $\beta_{1i} \sim \text{Weibull}\ (\nu_{\beta_1}, \lambda_{\beta_1})$.
- T_{I_i} is the corrosion initiation time at TML i and we assume that $T_{I_i} \sim \text{Lognormal}\ (\mu_{T_I}, \sigma^2_{T_I})$.
- ϵ_{ij} is the measurement error and we assume that $\epsilon_{ij} \sim \text{NOR}\ (0, \sigma^2_{\epsilon})$.
- t_j is the time when the measurement j was taken.

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

4.2 Bayesian Estimation of the Parameters in the Degradation Model

In addition to the model, we need to specify prior distributions for the parameters in the degradation model (7). Gelman (2006) provided general suggestions for choosing proper prior distributions for variance parameters in the hierarchical model. We use the following diffuse prior distributions for the standard deviations $\sigma_{y0}, \sigma_{\beta_1}, \sigma_{T_I}$, and σ_ϵ:

$$\sigma_{y0} \sim \text{Uniform } (10^{-5}, 5),$$

$$\sigma_{T_I} \sim \text{Uniform } (10^{-5}, 10),$$

$$\sigma_{\beta_1} \sim \text{Uniform } (10^{-5}, 5),$$

$$\sigma_\epsilon \sim \text{Uniform } (10^{-5}, 0.25).$$

The fact that pipeline data of 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 an initiation close to the fourth inspection and a large (in absolute value) corrosion rate and an initiation time close to the third inspection time and 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. Thus we specify a somewhat informative prior distribution for the median of corrosion rates for TMLs $\beta_{1i_{0.5}}$ that implies an upper bound on the corrosion rate:

$$\beta_{1i_{0.5}} \sim \text{Uniform } (10^{-6}, 0.022).$$

Regarding the prior distributions for the parameters $\mu_{y0\text{elbow}}, \mu_{y0\text{pipe}}, \mu_{y0\text{tee}}$, and μ_{T_I}, we use the following priors by specifying the lower and upper bounds of the uniform distributions:

$$\mu_{y0_{\text{elbow}}} \sim \text{Uniform } (0.4, 0.47),$$

$$\mu_{y0_{\text{pipe}}} \sim \text{Uniform } (0.4, 0.47),$$

$$\mu_{y0_{\text{tee}}} \sim \text{Uniform } (0.5, 0.62),$$

$$\mu_{T_I} \sim \text{Uniform } (9.31, 10^6).$$

The lower bound of the uniform distribution for μ_{T_I} is determined by the assumption that the corrosion initiation can only occur after the installation date. Similarly, if the corrosion rate has a Weibull distribution, we specify the same independent prior distributions for the parameters $\sigma_{y0}, \sigma_{T_I}, \sigma_\epsilon, \mu_{y0\text{elbow}}, \mu_{y0\text{pipe}}, \mu_{y0\text{tee}}$, and μ_{T_I}. For the Weibull corrosion rate distribution, we specify the prior distribution in terms of ν_{β_1}, the Weibull shape parameter and $\beta_{1i_{0.5}}$, the median of corrosion rates for TMLs. The following are priors for the shape parameter ν_{β_1} and $\beta_{1i_{0.5}}$ in the Weibull distribution:

$$\nu_{\beta_1} \sim \text{Uniform } (1.5, 5),$$

$$\beta_{1i_{0.5}} \sim \text{Uniform } (10^{-6}, 0.022),$$

where $\lambda_{\beta_1} = \log_e(2)/\beta_{1i_{0.5}}^{\nu_{\beta_1}}$ is the alternative second parameter used in the OpenBUGS parameterization of the Weibull distribution. Tables 3 and 4 present the posterior distribution summaries of parameters in the degradation model using lognormal and Weibull corrosion rates respectively. Figures 8 and 9 show the trellis plot of the fitted thickness values for Circuit Q in Facility 1 using lognormal and Weibull corrosion rates with 10-years of extrapolation after the last inspection on January 1, 2004.

			95% Credible Interval	
Parameters	Posterior Mean	Posterior Std. Dev.	Lower	Upper
$\mu_{y0_{\text{elbow}}}$	0.4379	0.004035	0.4301	0.4461
$\mu_{y0_{\text{pipe}}}$	0.4315	0.003448	0.4246	0.4382
$\mu_{y0_{\text{tee}}}$	0.5215	0.006226	0.5093	0.5336
σ_{y0}	0.01342	0.00192	0.01022	0.01767
$\beta_{1i_{0.5}}$	0.01815	0.003171	0.009919	0.02187
σ_{β_1}	0.5902	0.3129	0.1877	1.391
μ_{T_I}	9.411	0.0122	9.389	9.438
σ_{T_I}	0.04408	0.01536	0.01804	0.07968
σ_ϵ	0.004693	4.005E−4	0.003973	0.005552

Table 3: Marginal posterior distribution summaries of the parameters in the degradation model with a lognormal corrosion rate for pipeline data from Circuit Q in Facility 1.

The deviance information criterion (DIC) is again used for the Bayesian model comparison. DIC values for models with lognormal and Weibull corrosion rates are −1513.0 and −1016.0, respectively. The model using the lognormal distribution for the corrosion rate has

			95% Credible Interval	
Parameters	Posterior Mean	Posterior Std. Dev.	Lower	Upper
$\mu_{y0_{\text{elbow}}}$	0.4380	0.004059	0.4299	0.4458
$\mu_{y0_{\text{pipe}}}$	0.4314	0.003478	0.4245	0.4383
$\mu_{y0_{\text{tee}}}$	0.5215	0.006039	0.5096	0.5333
σ_{y0}	0.01347	0.001942	0.01027	0.01797
ν_{β_1}	2.890	0.9342	1.556	4.816
$\beta_{1i_{0.5}}$	0.0197	0.001908	0.01488	0.02192
μ_{T_I}	9.412	0.01221	9.392	9.441
σ_{T_I}	0.04744	0.01431	0.02657	0.08333
σ_{ϵ}	0.004715	4.016E−4	0.004007	0.005566

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

a smaller DIC, indicating a better fit. Figures 10 and 11 show the box plots of samples from the marginal posterior distributions of corrosion rates and initiation times for each TML in Circuit Q using the lognormal corrosion rate. These plots indicate that for 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.

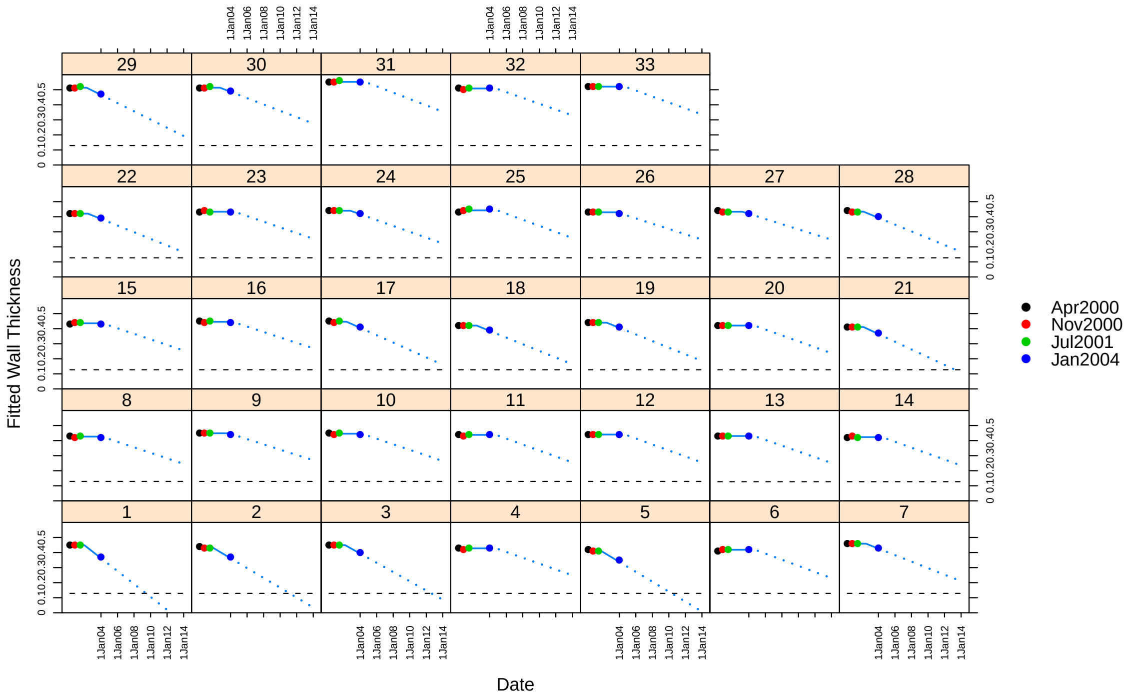

Figure 8: Trellis plot of the fitted thickness values for Circuit Q in Facility 1 using the lognormal corrosion rate distribution. The dotted lines indicate extrapolation.

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 $\mathcal{D} = \mathcal{D}(t, \boldsymbol{\theta})$. 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 $T_I \sim$ Lognormal $(\mu_{T_I}, \sigma^2_{T_I})$, $Y_0 \sim$ NOR $(\mu_{y0_{\text{elbow}}}, \sigma^2_{y0})$, $Y_0 \sim$ NOR $(\mu_{y0_{\text{pipe}}}, \sigma^2_{y0})$, $Y_0 \sim$ NOR $(\mu_{y0_{\text{tee}}}, \sigma^2_{y0})$, and $\beta_1 \sim$ Lognormal $(\mu_{\beta_1}, \sigma^2_{\beta_1})$. Then the cdf giving the proportion of pipeline segments that have a soft failure as a function of operating

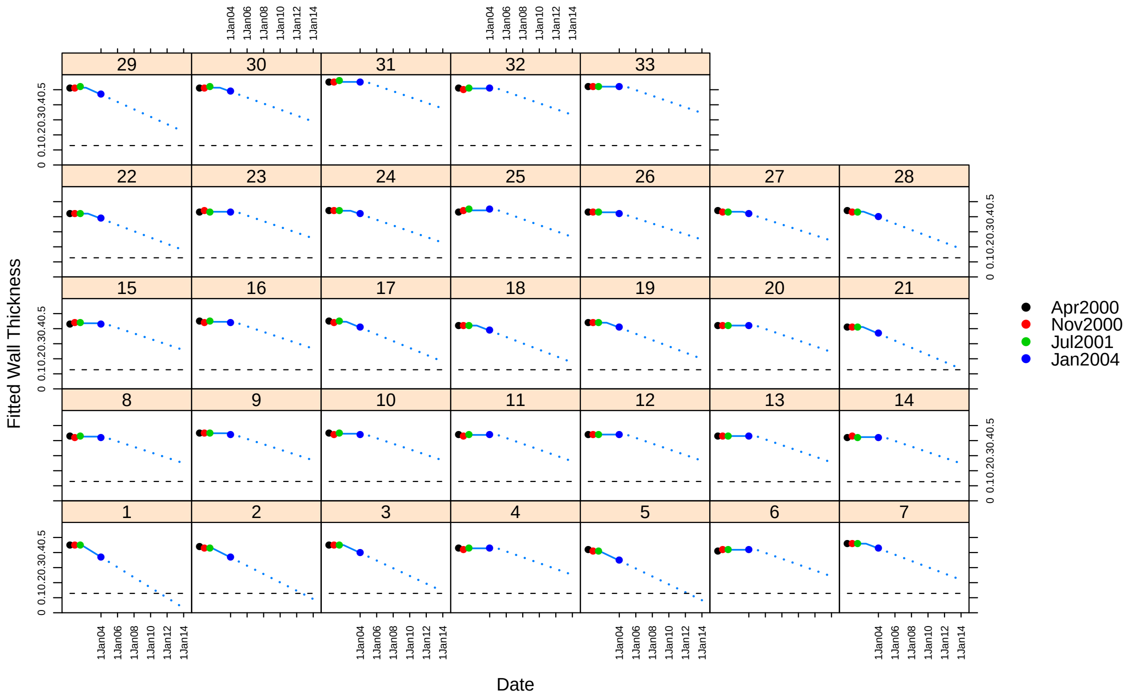

Figure 9: Trellis plot of the fitted thickness values for Circuit Q in Facility 1 using the Weibull corrosion rate distribution. The dotted lines indicate extrapolation.

time is

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

where $z_{y_0} = (y_0 - \mu_{y_0})/\sigma_{y_0}$, $z_{T_I} = (\log(T_I) - \mu_{T_I})/\sigma_{T_I}$, and $z_\beta = (\log(\beta_1) - \mu_{\beta_1})/\sigma_{\beta_1}$. When the corrosion rate has a lognormal distribution, $\phi(z_\beta) = \phi_{\mathrm{NOR}}(z_\beta)$ is the standard ($\mu = 0$, $\sigma = 1$) normal probability density function (pdf). When the corrosion rate has a Weibull distribution, $\phi(z_\beta) = \phi_{\mathrm{SEV}}(z_\beta) = \exp(z_\beta - \exp(z_\beta))$ is the standardized smallest extreme value pdf. Because $F(t)$ in (8) does not have a closed form, the Monte Carlo simulation

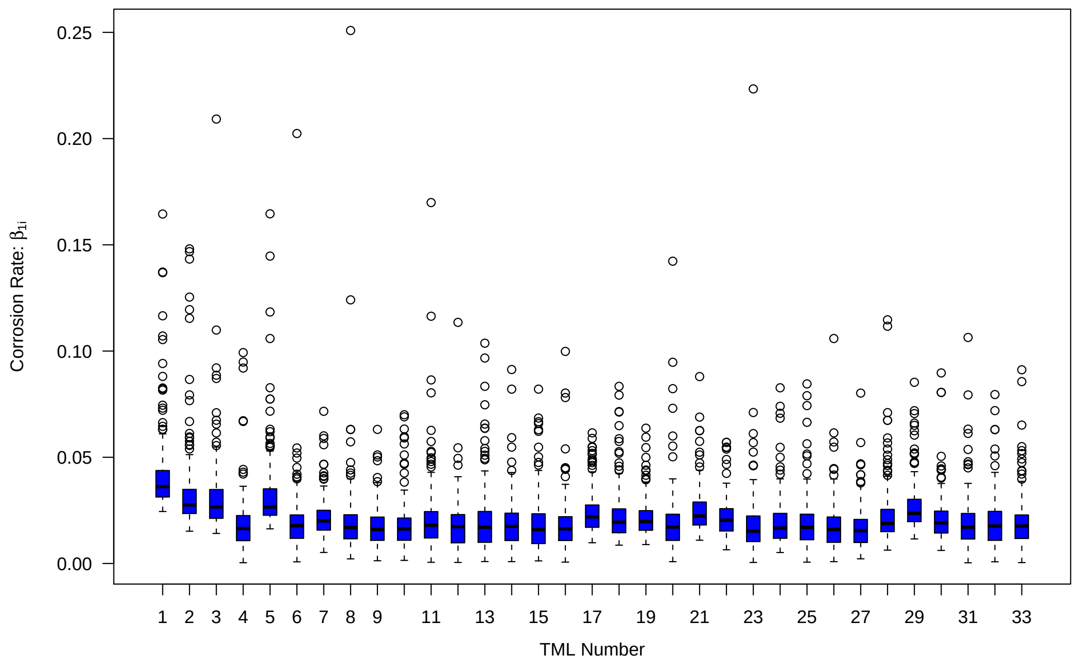

Figure 10: Samples from the marginal posterior distribution of the lognormal corrosion rates for each TML in Circuit Q in Facility 1.

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 normal, lognormal and Weibull corrosion rates. The plots suggest that in the degradation model (7), the lognormal corrosion rate provides the most conservative results compared with the other two model assumptions on the 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.

5.2 Predication of the Remaining Life of the Current Circuit

Figure 15 compares the remaining lifetime cdfs with normal, lognormal and Weibull corrosion rates for elbows, pipes and tees. The plots suggest that a lognormal distribution for the corrosion rate in the degradation model (7) provides the most conservative estimates of remaining life. This is due to the long upper tail of the lognormal distribution. Figure 16

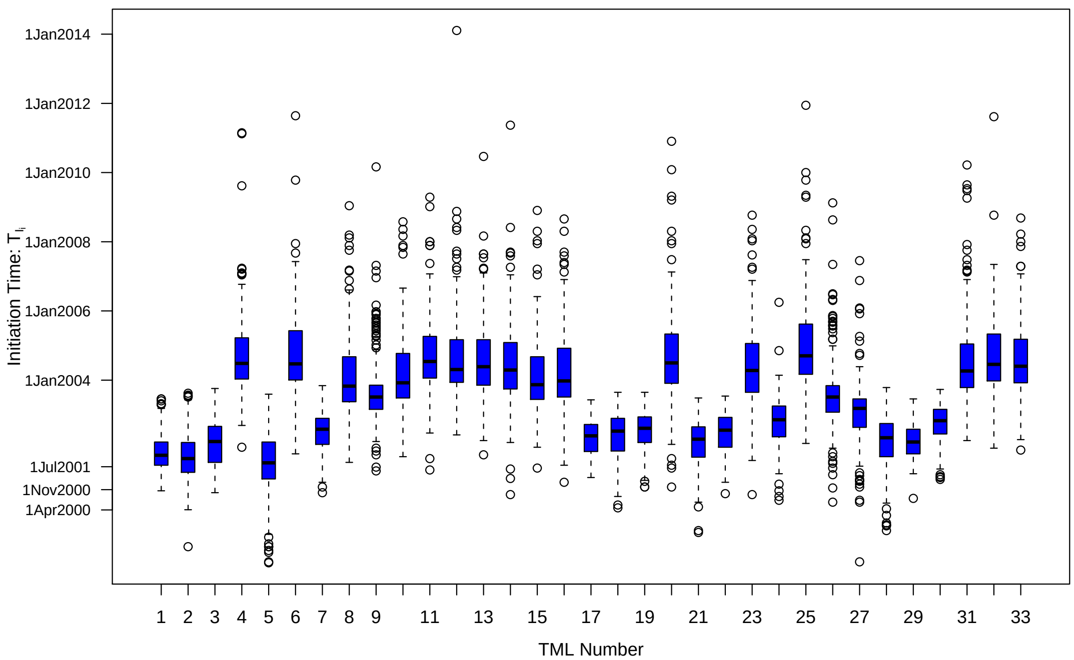

Figure 11: Samples from the marginal posterior distribution of the corrosion initiation times for each TML in Circuit Q in Facility 1 using a lognormal distribution to describe corrosion rates.

shows estimates of the remaining lifetime cdfs using the lognormal corrosion rate in the degradation model (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 density of 0.1, 0.2, 0.3, and 0.4 quantiles of the minimum remaining lifetime distribution from the degradation model (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.

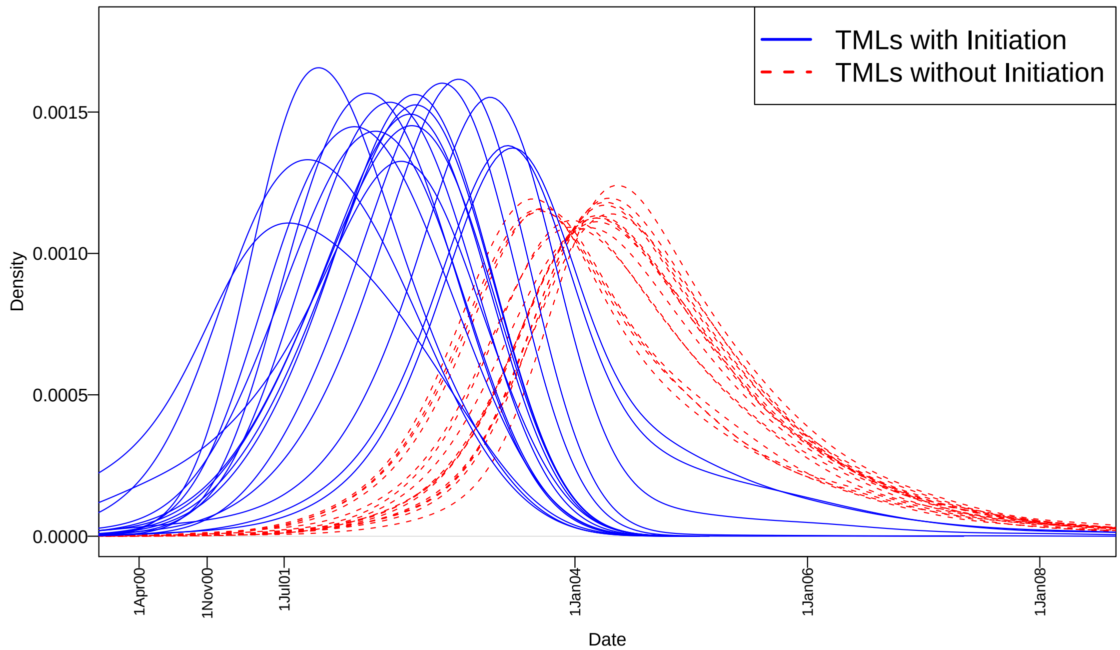

Figure 12: Posterior densities of the initiation times for each TML in Circuit Q in Facility 1.

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 prior knowledge about 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 model (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:

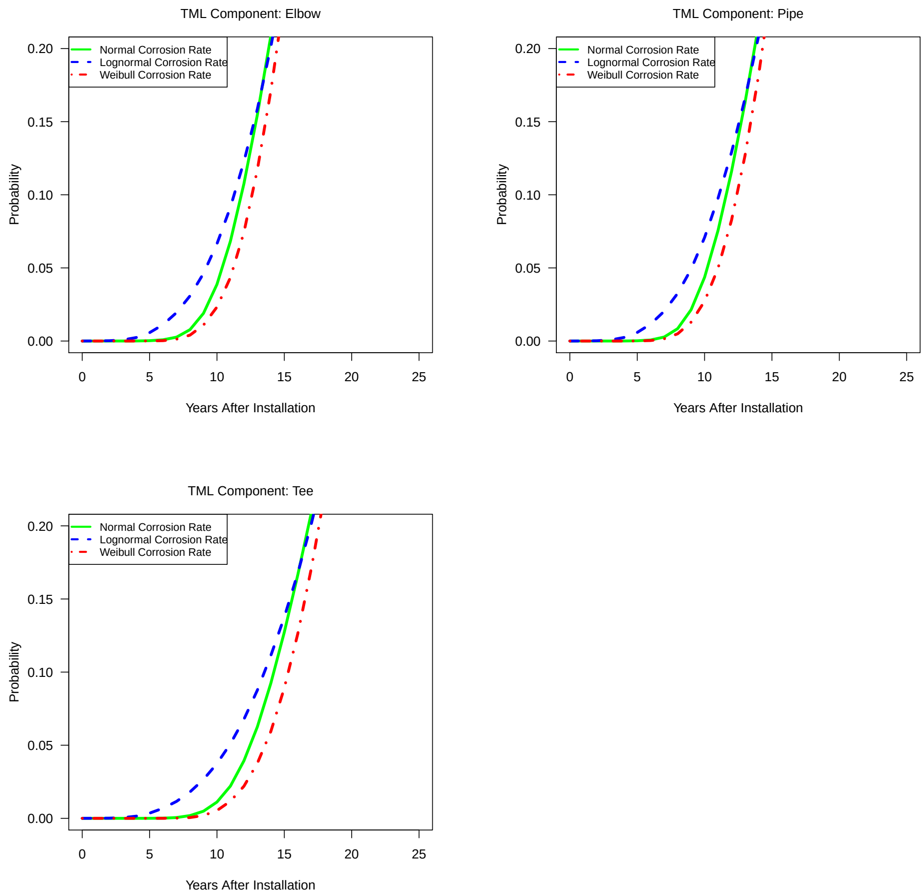

Figure 13: Degradation model estimates of failure time cdfs for pipeline components from Circuit Q in Facility 1 comparing normal, lognormal and Weibull corrosion rate distributions in the degradation model (7).

elbow, straight pipe and tee pipe. Corrosion was measured at each TML at 5 times.

We use the same diffuse prior distributions used in Section 4.2 for all parameters except for the median of the corrosion rates $\beta_{1i_{0.5}}$, $i = 1, 2, \ldots, 33$. 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 $\beta_{1i_{0.5}}$ to 0.022, we can relax the upper bound to 0.10 providing a diffuse prior for $\beta_{1i_{0.5}}$ [i.e. $\beta_{1i_{0.5}} \sim \text{Uniform}\ (10^{-6}, 0.10)$].

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

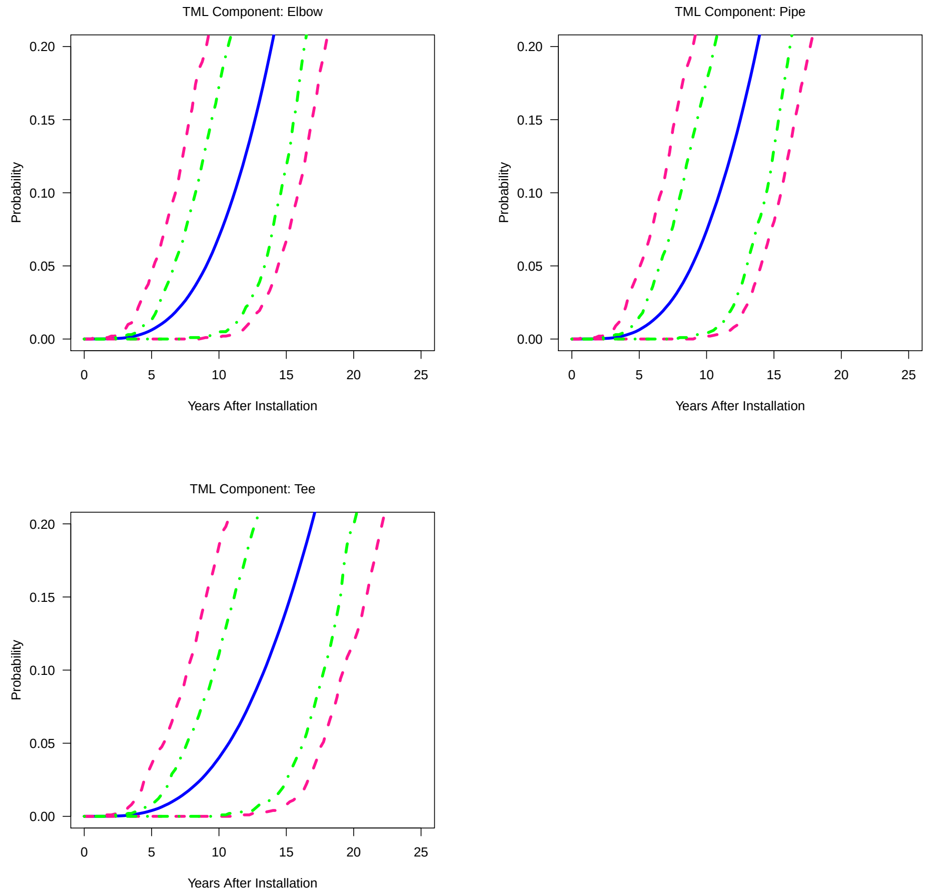

Figure 14: Degradation model estimates (the center lines) of failure time cdfs for pipeline components from Circuit Q in Facility 1 with the lognormal corrosion rate distribution in the degradation model (7) and two-sided 95% and 80% credible intervals.

thickness values for the simulated pipeline data using the diffuse 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 diffuse 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. The reason is that 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

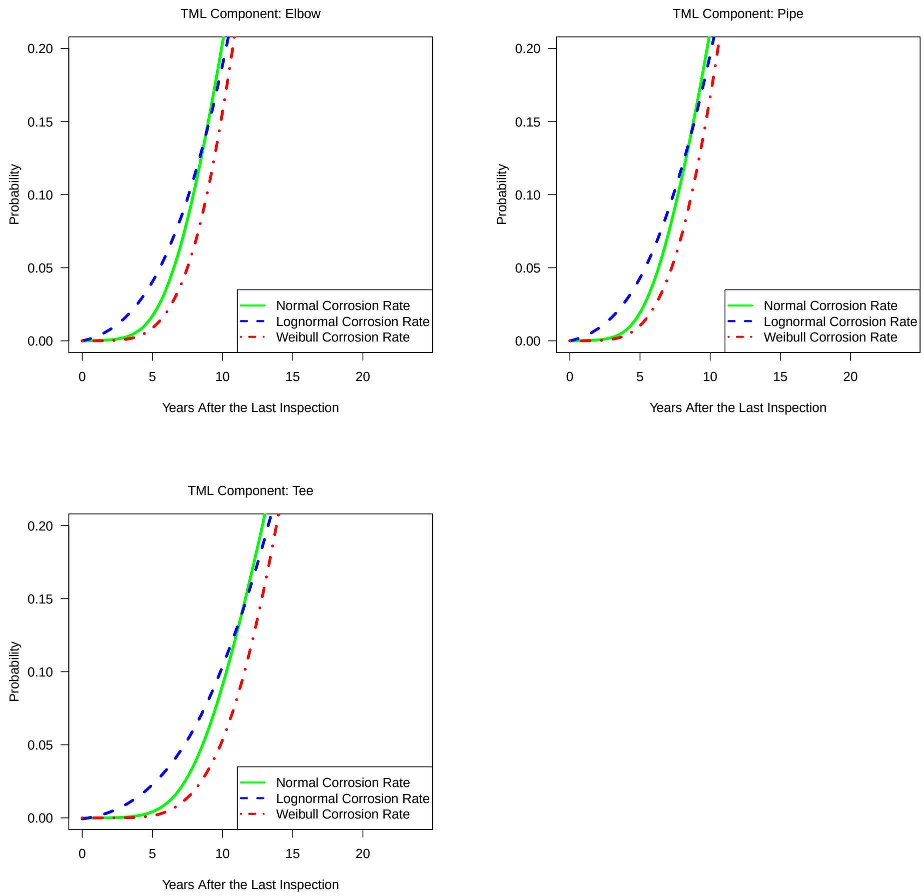

Figure 15: Degradation model estimates of remaining lifetime cdfs for pipeline components from Circuit Q in Facility 1 comparing normal, lognormal and Weibull corrosion rate distributions in the degradation model (7).

pipeline segment lifetime distributions.

7 Concluding Remarks and Areas for Future Research

In this paper, we developed degradation models to describe the pipeline corrosion behaviors 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

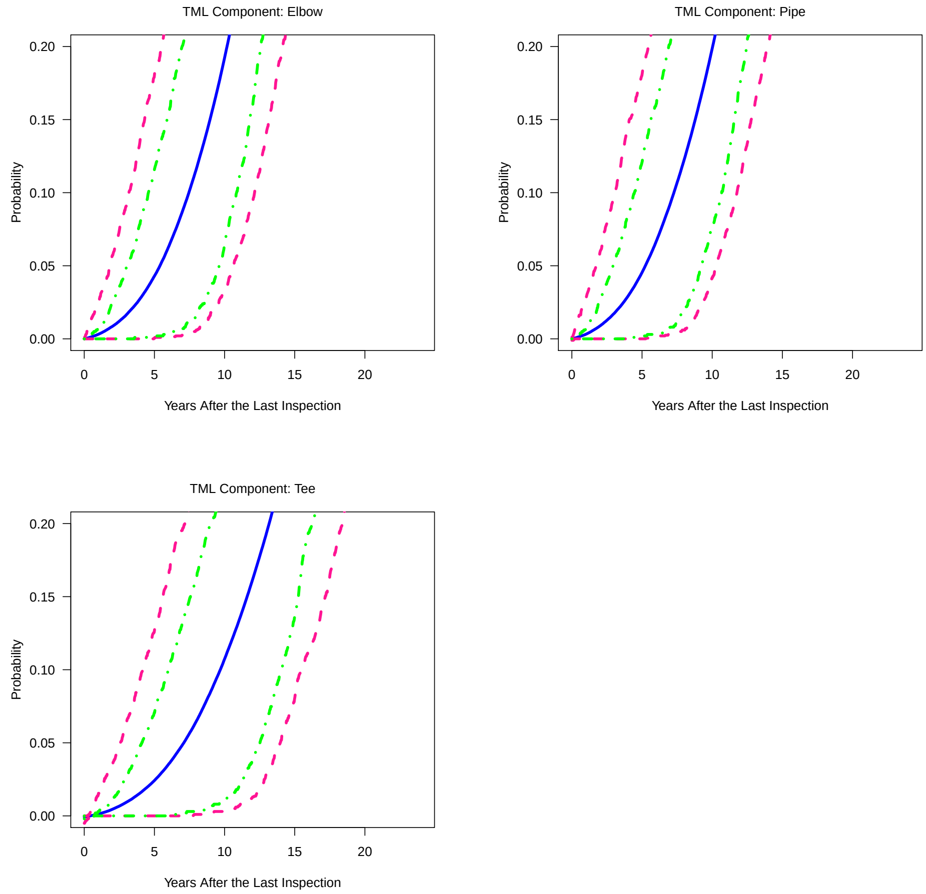

Figure 16: Degradation model estimates (the center lines) of remaining lifetime cdfs for pipeline components from Circuit Q in Facility 1 with the lognormal corrosion rate distribution in the degradation model (7) and two-sided 95% and 80% credible intervals.

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 (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

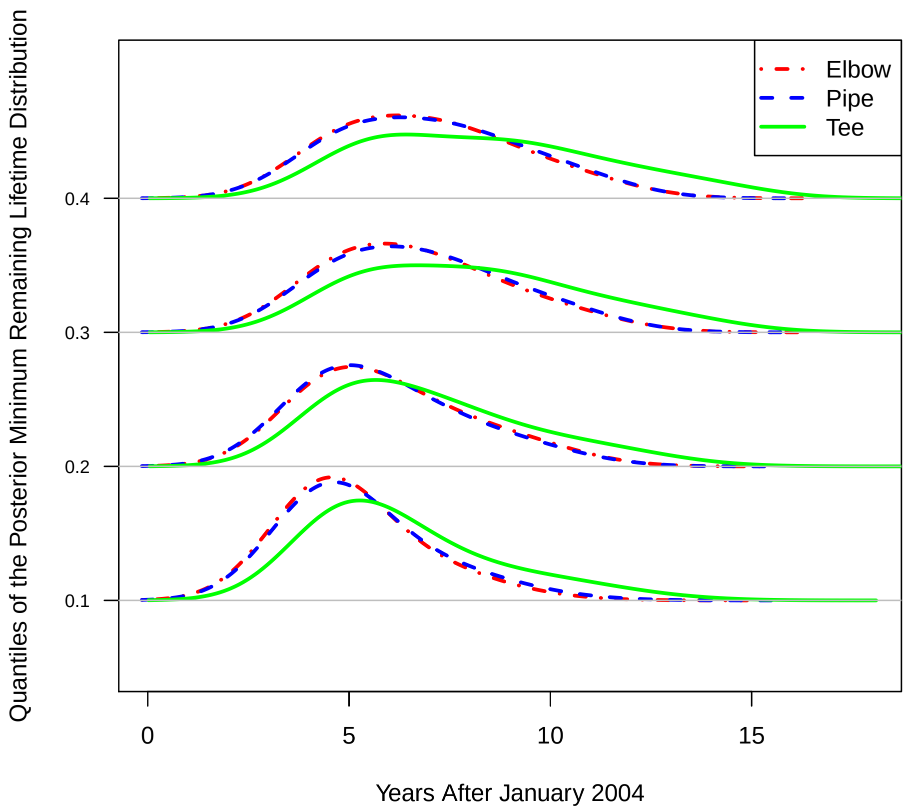

Figure 17: Posterior density of the 0.1, 0.2, 0.3, and 0.4 quantiles of the minimum remaining lifetime distribution (years after the last inspection time t_c: January 2004) with the population size $M = 100$ using the lognormal corrosion rate distribution of pipeline data from Circuit Q in Facility 1.

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.

• 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

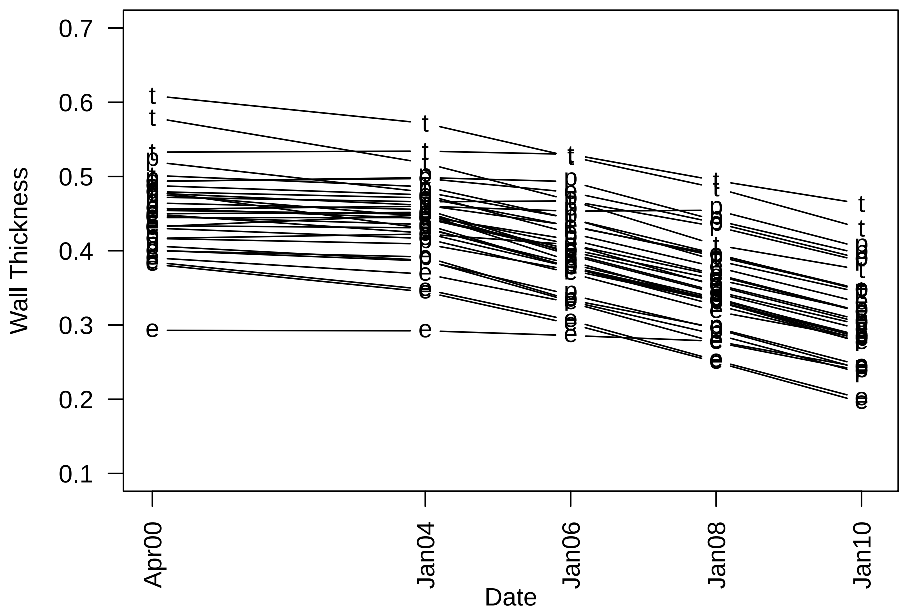

Figure 18: Time plot for the simulated pipeline data from a single circuit with 33 TMLs.

analysis.

References

[1] Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2003). *Bayesian Data Analysis*, second edition. London: Chapman and Hall.

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

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

[4] Liu, J., Meeker, W. Q. and Nordman, D. J. (2014). The number of MCMC iterations needed to compute Bayesian credible intervals. Preprint, Department of Statistics, Iowa State University.

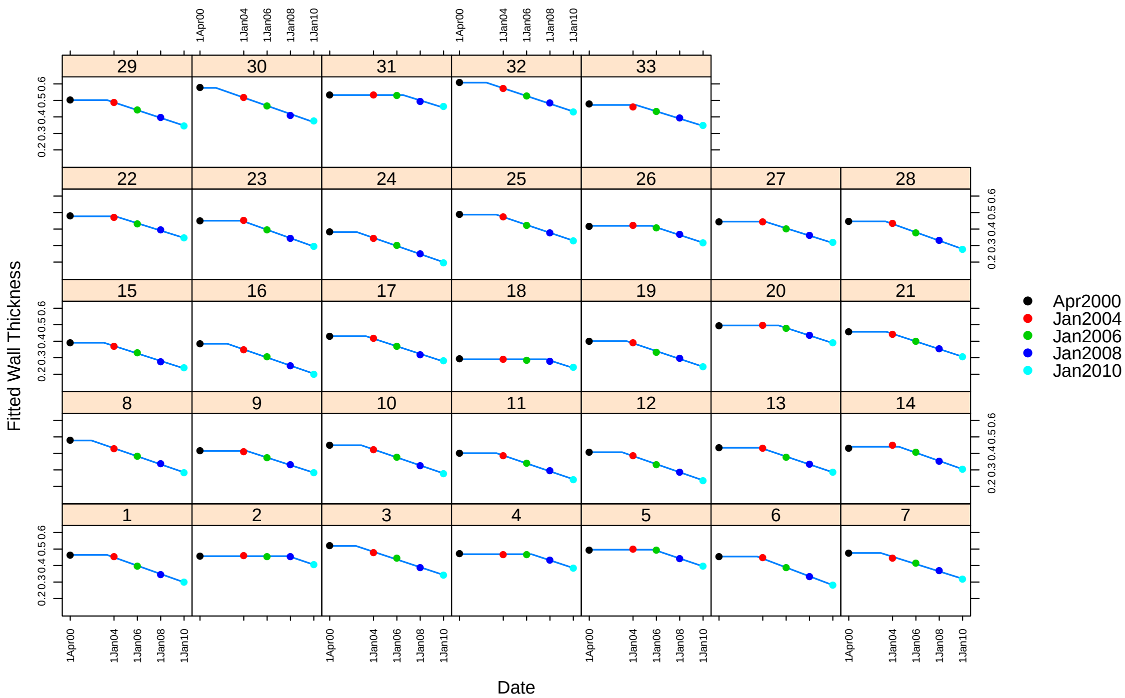

Figure 19: Trellis plot of the fitted thickness values for the simulated pipeline data in a single circuit with 33 TMLs using the diffuse prior distributions.

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

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

[7] 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)*, pp. 27-37. Boston: Birkhuser.

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

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

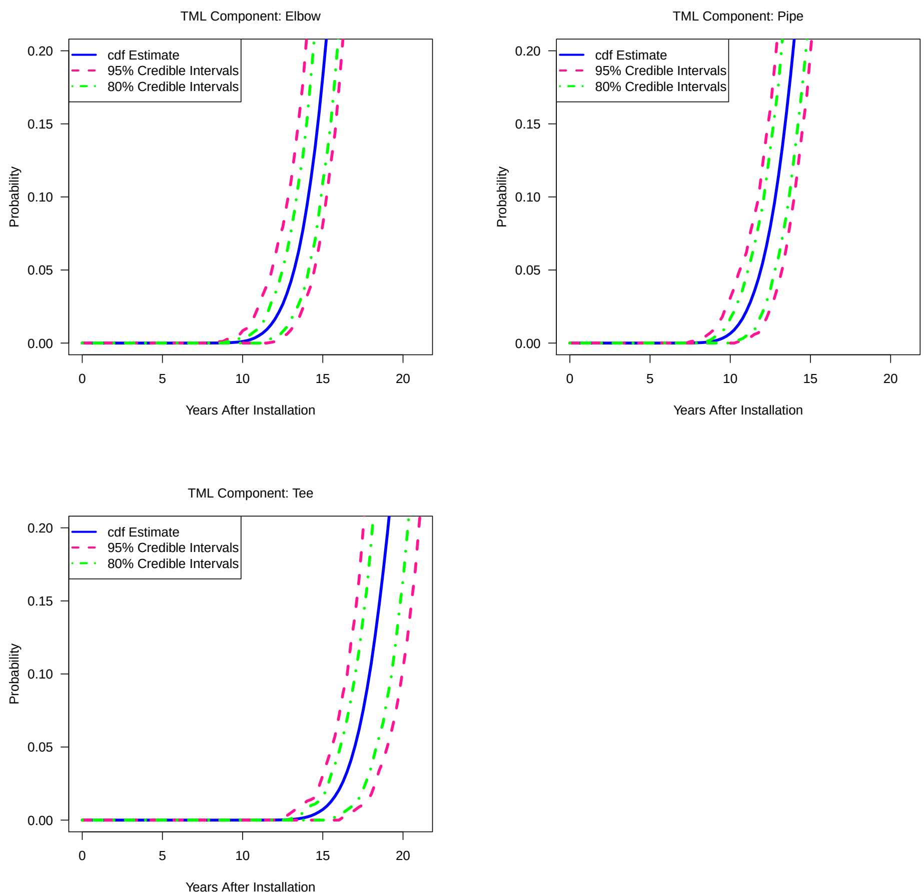

Figure 20: Degradation model estimates of failure time cdfs for the simulated pipeline data in a single circuit with 33 TMLs using the lognormal corrosion rate distribution and diffuse priors in the degradation model (7) and two-sided 95% and 80% credible intervals.