A DATA-DRIVEN PROGNOSTIC METHOD FOR CRUDE-OIL PIPELINE SYSTEMS

Ying Liao, Yisha Xiang, Mario Beruvides, and Lloyd Heinze
Texas Tech University

ABSTRACT
Crude-oil pipelines are the foundation of the liquid energy supply. Reliability of the pipelines are critical to the safe movement of the crude and pipeline maintenance plans. The degradation process of pipelines is complex, because (a) its deterioration is non-stationary and (b) operational environments vary dramatically by location. In this paper, we use a data-driven prognostic approach to model the degradation process of pipeline systems. We model the degradation process, a non-stationary Gamma process, with covariates and use the Expectation-Maximization (EM) algorithm for parameter estimation. The proposed approach is illustrated with a real case study by analyzing the annual wall thickness decrement of in-service pipelines. The results indicates that the proposed model is suitable for pipeline operators to identify the thickness deteriorating condition and the model is also flexible for stationary case without covariates.

KEYWORDS: Crude-oil; Pipeline Wall thickness; Non-stationary Gamma process; Covariates; EM

1. INTRODUCTION

In oil and gas industry, pipelines function as blood vessels and their failures could lead to unacceptably low safety levels for the public, environmental damage and monetary losses. Pipelines in the system are easily affected by the surrounding environment, construction errors, natural disasters and human activities. Different kinds of defects, such as corrosion, crack and mechanical damage may result in reduced strength in pipeline segments. For these defects, the nature of the growth mechanisms are time-dependent. With the use of suitable degradation model, the probability of failure can be estimated for pipelines with particular types of defect. Corrosion is a major integrity threat to oil and gas pipelines. Degradation analysis for metal loss corrosion defect is a vital part of pipeline integrity management [1]. The major concern of pipelines operators is the need to evaluate the pipelines' current reliability and also the time-dependent change in reliability.

The reliability-based pipeline corrosion management typically includes periodic inspections to detect and size corrosion, building the corrosion growth model based on the inspection results and mitigation of the defects [2]. The corrosion growth modeling plays an important role in the pipeline corrosion management in that it is crucial to the development of staged maintenance strategies that meet the safety and resource constraints. For a given operation lifetime, corrosion is of critical interest in maintenance optimization. For this reason, analyzing pipeline wall thickness data to develop the thickness degradation model can identify the deteriorating process. Given the failure wall thickness threshold, we can then optimize inspection interval, plan the replacement strategy to ensure safe operation and minimize the cost.

Degradation model are usually developed based on degradation data or prior understandings of physics behind degradation processes of a specific system. Two common categories of degradation models are data-driven approach and physics-based modelling approach [3]. A power law is applied to model the loss of wall thickness with the time of exposure in [4]. Since it is hard to capture the failure mechanisms or physical phenomena of a complex system, the data-driven methods are becoming increasingly popular in applications. In Data-driven models, stochastic processes models and artificial intelligence models are widely used in characterizing degradation in different industrial applications. Stochastic processes models can take into account the temporal variability of a degradation process, and therefore are very suitable and realistic methods. Two stochastic processes, namely Gamma process and Markov process, are commonly used to characterize the growth of corrosion defect on pipelines [2, 5-7]. For example, the homogeneous Gamma process was presented [2, 5] to model the growth of defect depth. The homogeneous and non-homogeneous Markov processes were used [6, 7] to model the growth of pitting corrosion. Note that for simplicity the corrosion growth was assumed to be a time-independent random variable as opposed to a stochastic process.

In this paper, we apply the non-stationary Gamma process with covariates to model the wall thickness of corroding pipelines. In real-world application, non-stationary Gamma process with covariates is more suitable to deal with the heterogeneity problem. It assumed a monotone increasing shape function and random rate parameter. Expectation-Maximization (EM) algorithm is efficient for parameter estimation and is also the natural choice for missing data problem. For a special case, when the wall thickness degradation path has independent decrement in each time interval among different units, the proposed approach will be simplified to a stationary case and the basic maximum likelihood method can be used to estimate the parameters.

The remainder of this paper is organized as follows. Section 2 will briefly introduce the non-stationary Gamma process with covariates and the implementation of parameter estimation. Computational study is presented in section 3. Section 4 concludes the paper and states the future work.

2. MODEL DEVELOPMENT

2.1 Non-stationary Gamma Process With Covariates

Degradation of a system often processes a monotone, nondecreasing degradation path and independent increments [8]. Gamma process has both properties so that it is considered to be suitable for degradation modeling. Suppose $\eta(\cdot)$ is a monotone increasing function with $\eta(0) = 0$. The process $\{Y_t; t \geq 0\}$ is called a Gamma process when it has independent and Gamma distributed increments, i.e. $\Delta Y_t = Y_{t+s} - Y_s$ is independent of Y_s, and ΔY_t follows $Gamma(\eta(t+s) - \eta(s), \gamma)$. In our proposed model, in order to capture heterogeneities within a population, a random effects model [9] is introduced by assuming rate parameter γ follows $Gamma(k, \lambda)$. Besides, in reality, the exact form for shape function is hard to determine without any prior knowledge, so a semi-parametric method is used in [9] to estimate the shape parameters instead of assuming a parametric form of the shape function, like an exponential law or power law. First, estimate every $\Delta\eta_j$ nonparametrically, and then fit the estimates to get a plausible parametric form of $\eta(\cdot)$.

We consider n pipeline segments and the cumulative wall thickness decrement of these pipelines is recorded periodically at time point $\boldsymbol{t} = \{t_0, t_1, \cdots, t_m\}$. $\boldsymbol{Y}_i = \{Y_{i,t_0}, Y_{i,t_1}, \cdots, Y_{i,t_m}\}$ is the cumulative decrement of wall thickness for the i^{th} unit, where we suppose $t_0 = 0$ and $Y_{i,t_0} = 0$. Throughout the paper, we shall denote $\eta_j = \eta(t_j)$, $\Delta\eta_j = \eta(t_j) - \eta(t_{j-1})$, $\Delta Y_{i,j} = Y_{i,t_j} - Y_{i,t_{j-1}}$ for $t_j, t_{j-1} \in \boldsymbol{t}$ and $\boldsymbol{Y} = \{\boldsymbol{Y}_i : i = 1, \cdots, n\}$. In this random effects model, the log-likelihood function [10] for the n units is

$$\begin{aligned} ln\, L(\boldsymbol{\theta}|\boldsymbol{Y}) = \textstyle\sum_{i=1}^{n} \big[k\, ln\, \lambda + ln\, \Gamma(\eta_m + k) - ln\, \Gamma(k) - (\eta_m + k)\, ln\big(\lambda + Y_{i,t_m}\big)\big] \\ + \textstyle\sum_{i=1}^{n} \sum_{j=1}^{m} \big[\big(\Delta\eta_j - 1\big)\, ln\, \Delta Y_{i,j} - ln\, \Gamma\big(\Delta\eta_j\big)\big], \end{aligned} \tag{1}$$

where $\boldsymbol{\theta} = \big(\eta_j, k, \lambda; j = 0, \cdots, m\big)$ is the parameter vector. Since analytically maximizing (1) is almost impossible, EM algorithm will be applied to estimate the parameters, presented in the next section. For a special case, assuming the shape function has a linear form with time t, which means $\eta(t) = at + b$ and the rate parameter γ is unknown number, the model will be simplified to a stationary case which means every $\Delta Y_{i,j} \sim Gamma\big(a(t_j - t_{j-1}), \gamma\big)$. Basic MLE method can solve the parameter estimation problem.

2.2 Parameter Estimation Using EM Algorithm

For random effects model, the γ_i's are considered as missing data[9]. The EM algorithm is a natural choice for missing data problem. The algorithm iteratively applies two steps, that is, the expectation step(E-step) and the maximization step(M-step). Denote $D_{miss} = \{\gamma_i; i = 1, \cdots, \mathrm{n}\}$ as missing dataset, $D = \boldsymbol{Y} \cup D_{miss}$ as complete dataset. Given the complete dataset, the log-likelihood function, up to a constant, can be shown as

$$\begin{aligned} ln\, L(\boldsymbol{\theta}|D) = \sum_{i=1}^{n} \sum_{j=1}^{m} \big[\Delta\eta_j \big(ln\, \Delta Y_{i,j} + ln\, \gamma_i\big) - ln\, \Gamma\big(\Delta\eta_j\big)\big] \\ + \textstyle\sum_{i=1}^{n} \big[k\, ln\, \lambda + (k-1)\, ln\, \gamma_i - ln\big(\Gamma(k)\big) - \lambda\, \gamma_i\big]. \end{aligned} \tag{2}$$

Further denote $\boldsymbol{\theta}^{(N)} = \big(\eta_j^{(N)}, k^{(N)}, \lambda^{(N)}\big)$ the estimated parameters at the N^{th} EM algorithm, at $(N+1)^{th}$ iteration, by considering the missing data as random variables, the EM algorithm evolves as follows [9].

(1) E-step: Compute the Q-function

$$Q\left(\boldsymbol{\theta}\middle|\boldsymbol{\theta}^{(N)}\right) = E\left[\ln L(\boldsymbol{\theta}|D)\,\middle|\boldsymbol{\theta}^{(N)}, \boldsymbol{Y}\right] = \sum_{i=1}^{n}\sum_{j=1}^{m}\left[\Delta\eta_j\left(\ln \Delta Y_{i,j} + \nu_i^{(N)}\right) - \ln \Gamma\left(\Delta\eta_j\right)\right] + \sum_{i=1}^{n}\left[k \ln\lambda + (k-1)\nu_i^{(N)} - \ln\left(\Gamma(k)\right) - \lambda\,\omega_i^{(N)}\right] \tag{3}$$

where $\nu_i^{(N)} = E\left(\ln\gamma_i\,\middle|\boldsymbol{\theta}^{(N)}, \boldsymbol{Y}\right) = \psi\left(\eta_m^{(N)} + k^{(N)}\right) - \ln\left(\lambda^{(N)} + Y_{i,t_m}\right)$, $\psi(\cdot)$ is the digamma function; and $\omega_i^{(N)} = E\left(\gamma_i\middle|\boldsymbol{\theta}^{(N)}, \boldsymbol{Y}\right) = \frac{\eta_m^{(N)} + k^{(N)}}{\lambda^{(N)} + Y_{i,t_m}}$.

(2) M-step: Update $\boldsymbol{\theta}^{(N+1)} = \left(\eta_j^{(N+1)}, k^{(N+1)}, \lambda^{(N+1)}\right)$ by maximizing the Q-function over $\boldsymbol{\theta}$, we have

$$\ln k - \psi(k) = \ln\bar{\omega}^{(N)} - \bar{\nu}^{(N)} \tag{4}$$

$$\lambda^{(N+1)} = \frac{k^{(N+1)}}{\bar{\omega}^{(N)}} \tag{5}$$

$$\Delta\eta_j^{(N+1)} = \psi^{-1}\left(\bar{\nu}^{(N)} + \frac{1}{n}\sum_{i=1}^{n}\ln\Delta Y_{i,j}\right) \tag{6}$$

where $k^{(N+1)}$ is the solution for (4), $\bar{\omega}^{(N)}$ and $\bar{\nu}^{(N)}$ are the means of $\omega_1^{(N)}, \cdots, \omega_n^{(N)}$ and $\nu_1^{(N)}, \cdots, \nu_n^{(N)}$, respectively and $\psi^{-1}(\cdot)$ is the inverse digamma function. Therefore,

$$\eta_j^{(N+1)} = \sum_{h=1}^{j}\Delta\eta_h^{(N+1)} \tag{7}$$

Finally, the EM algorithm terminates when the increment of the log-likelihood value is smaller than a given criterion α.

3. COMPUTATIONAL STUDY

We applied the presented approach to analyze the pipeline wall thickness data from a pipeline operation in Montana. There are 6 pipeline segments distributed in Montana. The wall thickness of each pipeline segment is measured daily from December 2015 to February 2019. Since the pipeline degrades slowly, we use the yearly average wall thickness to represent the degradation level in year 2016-2018. Table 1 shows the yearly thickness data from 2016 to 2018. We calculated the annual degradation decrement from year 2016 and converted the unit from millimeter per year into milli-inch per year(mpy) (shown in Table 2).

First, we apply the random effects model. We get $\widehat{\Delta\eta}_1 = 1.138, \widehat{\Delta\eta}_2 = 0.999$ and $\hat{k} = 8.556, \hat{\lambda} = 7.654$. Using the estimated parameters, we can obtain the 95% confidence variance bound paths and simulated realizations of estimated model compared with the six sample paths as shown in Figure 1. Real data has been mostly covered by the variability bounds. Thus, the estimated degradation model has reasonable goodness-of-fit. Next, we relax the Gamma process to stationary case, the numeric result shows that the annual thickness decrement follows $Gamma(1.024, 0.993)$. By applying a Kolmogorov-Smirnov test, the p-value is 0.925, which means the obtained distribution can fit the sample data well as shown in Figure 2. From this real case study, we can see the non-stationary Gamma process with covariates can not only characterize the heterogeneities between different units in different time intervals, but also be flexible for simple stationary case.

4. CONCLUSION

This paper provides a data-driven prognostic method for analyzing the pipeline wall thickness data. The non-stationary Gamma process with covariates is shown to be reasonably satisfying to characterize the degradation behavior of the pipelines corrosion. EM algorithm is efficient to estimate the parameters for random effects model by assuming the rate parameters as missing data. In addition, the proposed method is flexible to model the stationary case. Therefore, we can optimize maintenance strategies to ensure safe operation and minimize the cost based on the estimated model, which is the direction for future exploration.

REFERENCES

1. Xie, M. and Z. Tian, *Risk-based pipeline re-assessment optimization considering corrosion defects.* Sustainable Cities and Society, 2018. **38**: p. 746-757.
2. Zhang, S. and W. Zhou, *System reliability of corroding pipelines considering stochastic process-based models for defect growth and internal pressure.* International Journal of Pressure Vessels and Piping, 2013. **111**: p. 120-130.
3. Shahraki, A.F., O.P. Yadav, and H. Liao, *A review on degradation modelling and its engineering applications.* International Journal of Performability Engineering, 2017. **13**(3): p. 299.
4. Amirat, A., A. Mohamed-Chateauneuf, and K. Chaoui, *Reliability assessment of underground pipelines under the combined effect of active corrosion and residual stress.* International Journal of Pressure Vessels and Piping, 2006. **83**(2): p. 107-117.
5. Zhou, W., H. Hong, and S. Zhang, *Impact of dependent stochastic defect growth on system reliability of corroding pipelines.* International Journal of Pressure Vessels and Piping, 2012. **96**: p. 68-77.
6. Hong, H.-P., *Application of the stochastic process to pitting corrosion.* Corrosion, 1999. **55**(1): p. 10-16.
7. Hong, H.P., *Inspection and maintenance planning of pipeline under external corrosion considering generation of new defects.* Structural safety, 1999. **21**(3): p. 203-222.
8. Singpurwalla, N.D., *Survival in dynamic environments.* Statistical science, 1995: p. 86-103.
9. Ye, Z.-S., et al., *Semiparametric estimation of gamma processes for deteriorating products.* Technometrics, 2014. **56**(4): p. 504-513.
10. Lawless, J. and M. Crowder, *Covariates and random effects in a gamma process model with application to degradation and failure.* Lifetime Data Analysis, 2004. **10**(3): p. 213-227.

Table 1. Yearly average thickness data from 2016 to 2018 (mm)

Units	2016	2017	2018
1	12.930651	12.834149	12.806557
2	9.0725522	9.0319358	9.0219283
3	10.105685	10.10539	10.074024
4	9.2760896	9.2423857	9.2335435
5	12.893054	12.865728	12.853616
6	11.788882	11.769629	11.763282

Table 2. Annual degradation decrement data since 2016 (mpy)

Units	2016	2017	2018
1	0	3.80219850	4.889317
2	0	1.60028372	1.994581
3	0	0.01165216	1.247451
4	0	1.32793311	1.676317
5	0	1.07666410	1.553869
6	0	0.75854495	1.008631

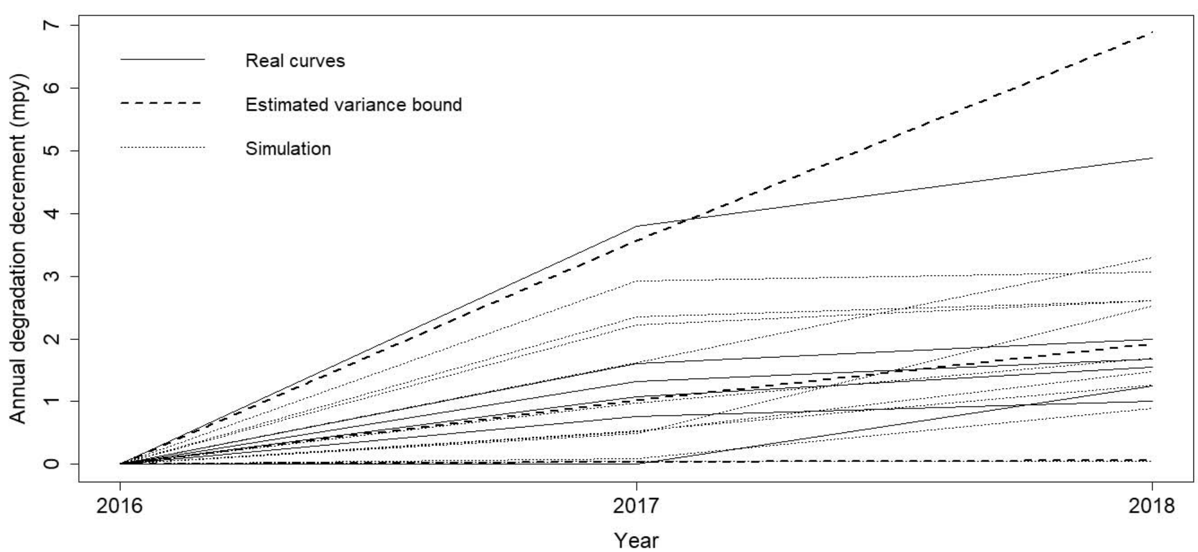

Figure 1. Annual degradation decrement and simulated Gamma process plot

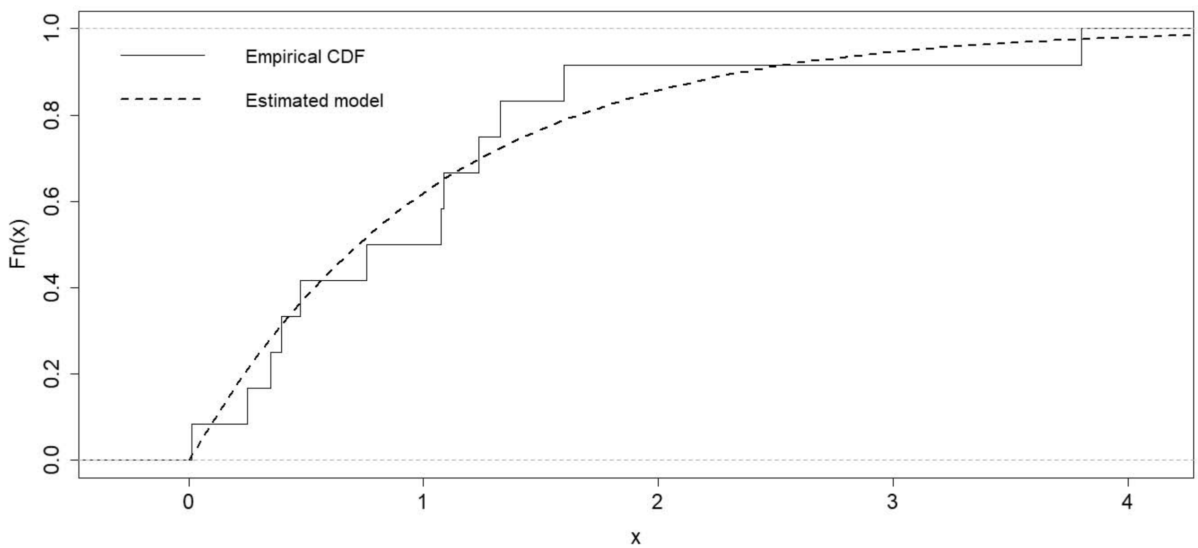

Figure 2. Empirical CDF and estimated Gamma distribution CDF