Degradation Modelling Under Dynamic Environmental Conditions

Zhe Min Chia
Georgia Institute of Technology
zhemin@gatech.edu

May 6, 2021

Abstract

This paper explores the challenge of degradation modelling under scenarios where the environmental conditions and the corresponding degradation rate are dynamic across time. These environmental conditions will be modelled using Continuous-time Markov chains (CTMCs), and the degradation models under each of these environments would be assumed to be a simple Brownian motion with drift, with different drift parameters. The paper will explore the methodology for simulating such data, and then re-estimating the parameters for the original models used to simulate said data.

1 Introduction and Motivation

Conventional degradation modelling techniques often account for variations in the degradation rate of similar or identical components through slight manufacturing variations or usage environments for each component. The result is that each component would be assumed to have a different degradation rate, but the degradation rate for an individual component would be somewhat consistent and homogeneous.
This paper looks at alternative methodologies for modelling the degradation rate when it appears to be erratic and dynamic over its lifetime. This source of variation would result from environmental conditions (such as humidity, temperature), or operational conditions (output levels, number of hours of daily use) that would change over the course of the component's lifetime, and would be expected to have a significant effect on its degradation rate. Specifically, this paper will look at capturing the dynamism of the degradation rate through "environment profiles".

2 Literature Review

The problem of modelling and analysing degradation rates under dynamic environmental and operational conditions has been examined by Bian, Gebraeel, and Kharoufeh 2015, using continuous-time Markov chains (CTMC) to model the changing environment profiles, and updating the models dynamically using Bayesian frameworks.
Yang 2021 approaches the problem by using deep neural network frameworks instead to model and handle the dynamic nature of the changing degradation rates across different components.

3 Methodology

The project can be broken down in 3 main sections: data generation, model estimation, and remaining useful life (RUL) estimation.

3.1 Data generation

The degradation data used in this project needed to reflect the situation where the degradation rate for a single component was dynamic and changed based on changing environment profiles.

To do so, a 3-state CTMC was used to generate that changing environment profiles, where the 3 states refer to Mild, Moderate and Harsh operating conditions. Under these 3 different conditions, a separate linear degradation model, each with a drift parameter proportional to the degradation under that environmental condition, as captured in the equation below:

$$X(t) = X(0) + \mu_e t + \sigma W(t)$$

where μ_e represents the drift coefficient for the that particular environmental condition.
Using this combined CTMC and 3 separate linear degradation models, a total of 100 different degradation paths were generated (see Figure 2). Assuming a a degradation threshold of $X(t) = 150$, a clipped version of the data was produced in Figure 3.

The final degradation signal for each path can be described using the following equation:

$$S(t) = S(0) + \int_0^t r(\psi(t))dv + \lambda W(t)$$

where $\psi(t)$ is the environment profile of the path at time t, and $r(\psi(t))$ is the rate of degradation corresponding the linear degradation model for the corresponding environment profile at time t.

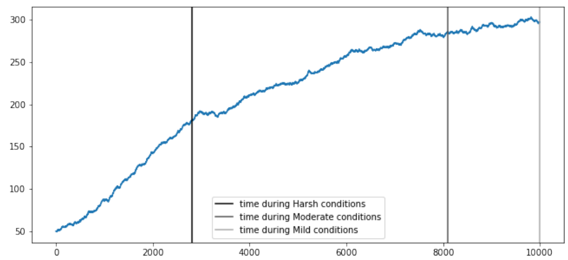

Figure 1: Plot of a single degradation path.

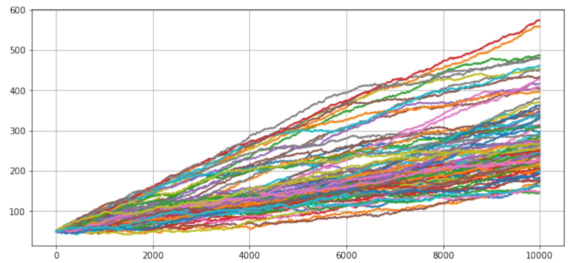

Figure 2: Plot of 100 degradation paths.

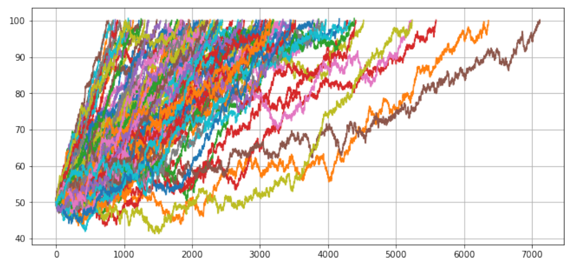

Figure 3: Plot of 100 degradation paths, clipped at $X(t) = 150$.

3.2 Model estimation

After the data has been simulated, various techniques are then applied to re-estimate the parameters in the original model. To simplify the model estimation at this stage, it is assumed that for each of the 100 simulated degradation paths, each path's cycle time the different environment profiles is known. Using the raw data and the cycle time information, the following parameters are estimated:

1. μ_e in the linear degradation model for environment profile
2. Starting distribution for the CTMC
3. Transition matrix for the CTMC
4. λ_e of the exponential distribution for the time spent in each environment profile
5. σ for the linear degradation model

3.2.1 μ_e

Using the information on each path's cycle time through the different environment profiles, we obtain the portions of each degradation path that originate from the linear degradation models belonging to each of the 3 environment profiles in Figure 4. Thereafter, a linear regression was run to determine the respective $\hat{\mu_e}$'s, which were 0.00846, 0.0204, and 0.0490 for Mild, Moderate and Harsh respectively. These estimates are close to the original μ_e's of 0.008, 0.02, and 0.05.

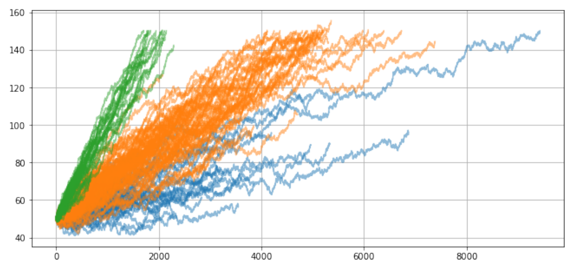

Figure 4: Portions of the degradation path belonging to each environment profile. Mild in blue, Moderate in orange, Harsh in green.

3.2.2 CTMC parameters

By doing simple counting techniques, we can estimate the respective parameters for the starting distribution and the transition matrix for the CTMC.
As for the λ_e of the exponential distribution for the time spent within each environment profile, maximum likelihood estimation was used. In this case, for each environment profile, we have 2 types of data for the time spent in the respective environment: the complete, uncensored time (F), and the censored time (C) (when the experimental data ends). The latter is a Type I Right Censoring, which differs from ordinary experimental circumstances in that the level of censoring is different for each path. Hence, an equation for the MLE estimator was developed as follows:

$$L(\theta_e) = \prod_{i \in F} f(t_i) \prod_{j \in C} R(t_j^*)$$

$$= \lambda_e^r \exp^{-\lambda \sum_{i=1}^r t_i} \exp^{-\lambda \sum_{j=1}^s t_j^*}$$

$$= \lambda_e^r \exp^{-\lambda (\sum_{i=1}^r t_i + \sum_{j=1}^s t_j^*)}$$

where r = # of fully observed sojourn times, s = # of partially observed/censored sojourn times.

$$\ln L(\theta_e) = r \ln \lambda_e - \lambda_e \sum_{i=1}^r t_i - \lambda \sum_{j=1}^s t_j^*$$

let

$$\frac{d \ln(L(\theta_e))}{d\lambda} = 0$$

$$\hat{\lambda}_e = \frac{r}{\sum_{i=1}^{r} t_i + \sum_{j=1}^{s} t_j^*}$$

The respective $\hat{\lambda}_e$ estimates were 1883, 5047, and 2817; which are close to the original λ_e values of 2000, 5000, and 3000.

3.2.3 σ

To obtain estimates for the variance of the linear degradation models (which are constant across the 3 environment profiles), the degradation paths in Figure 4 were detrended by removing the effect of μ_e to produce Figure 5. The difference in consecutive X(t)'s were calculated and fitted using a normal distribution, and the resulting $\hat{\sigma}$ calculated was 0.2504, which is very close to the original σ of 0.25.

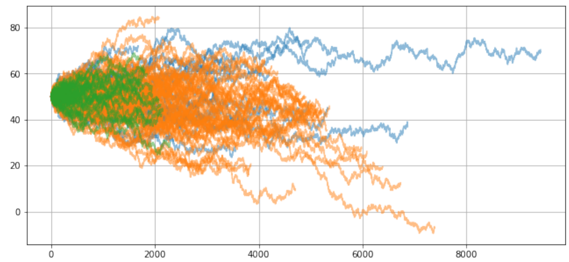

Figure 5: Degradation paths from the respective environment paths, minus the effect of the drift coefficient.

4 Remaining Useful Life (RUL) Estimation

Thereafter, the previous data generated is discarded, and 300 new data points are generated in the same manner for RUL estimation. Out of these data points, 150 were used for estimating all the parameters for the model using the techniques outlined in the previous section, and another 150 were used to test the accuracy of RUL estimation using the estimated model.

4.1 Parameter Estimation

Table 1 below shows the original parameters along with the estimated values, Table 2 shows the true transition matrix and Table 3 shows the estimated transition matrix.

From the tables, we can see that there is a net underestimation of the drift coefficients, and for the transition matrix, the estimated transition probabilities to Mild environment from the other two environments are much higher than the true transition probabilities. Furthermore, the λ estimates experience significant deviation.

	True	Estimated
λ_{mild}	2000	2249.83
$\lambda_{moderate}$	5000	5630.99
λ_{harsh}	3000	2411.58
σ	0.25	0.2501
μ_{mild}	0.008	0.00889
$\mu_{moderate}$	0.02	0.0187
μ_{harsh}	0.05	0.0485

Table 1: True and estimated parameters.

		To		
		Mild	Moderate	Harsh
From	Mild	0	0.7	0.3
	Moderate	0.6	0	0.4
	Harsh	0.3	0.7	0

Table 2: True transition matrix.

		To		
		Mild	Moderate	Harsh
From	Mild	0	0.698	0.302
	Moderate	0.709	0	0.291
	Harsh	0.467	0.533	0

Table 3: Estimated transition matrix.

4.2 RUL Estimates

For the 150 data points used for testing, the simulated degradation paths were clipped at the midpoint of the degradation at 100 (degradation starts at 50, and failure occurs at 150). Thereafter, for each of the test data points, the current degradation signal and the environment profile, along with the estimated parameters from the training data set, are used to simulate the estimated failure time for that given data point.

This simulation is repeated 100 times, the mean estimated failure time, along with the standard deviation and other statistics and aggregated and presented in Table 4 for the first 5 data points. Table 5 shows the aggregated statistics for the standard deviation, estimation error in terms of standard deviation, estimation error, and absolute estimation error.

True Failure Time	Expected Failure Time	S.D.	Estimation error in # of S.D.
5288	5314.660	1022.437425	0.026075
6860	7351.272	1416.623790	0.346791
3462	3901.762	943.465758	0.466113
2758	2928.232	880.337923	0.193371
4373	5235.790	985.914180	0.875117

Table 4: True failure time, estimated failure time, standard deviation, and error/standard deviation for the first 5 data points.

	S.D.	Estimation error in # of S.D.	Estimation error	Absolute estimation error
Mean	982.3	0.6606	-220.15	653.12
S.D.	171.4	0.4394	776.69	471.81
Min	737.4	0.00698	-2609.07	6.73
Max	1629.5	2.333	2391.68	2609.07

Table 5: Aggregated statistics on RUL estimates.

From Table 5, we can see that the RUL estimate generally overestimates the true failure time with a mean of 220.15. This is likely because the estimated parameters for the model have a much larger λ_{mild} and $\lambda_{moderate}$, much smaller λ_{harsh}, and the estimated transition probabilities to the $Mild$ state are much higher than the true transition probabilities. Therefore, the simulated paths used to estimate the RULs spend much more time in the $Mild$ and $Moderate$ states which makes the estimated model degrade at a slower rate than the true model, thus the RUL estimate would be higher than the true RUL.

However, this margin of error in the error is superseded by the much larger standard deviation values for the expected failure times. This larger standard deviations are likely due to the inherent stochasticity in the model from the CTMC itself, and can't be further improved upon while relying on simulation techniques on a CTMC model.

5 Future Extensions

Future work can look at relaxing all the assumptions made in this project and look at:

1. Estimating the number of different states/environment profiles within the dataset
2. Detecting when there is a transition in states/environment profiles within the degradation path
3. Modelling for dynamic exponential degradation models instead of linear degradation models

References

[1] Linkan Bian, Nagi Gebraeel, and Jeffrey P Kharoufeh. "Degradation modeling for real-time estimation of residual lifetimes in dynamic environments". In: *Iie Transactions* 47.5 (2015), pp. 471–486.

[2] Li Yang. "Adaptive Degradation Process with Deep Learning-Driven Trajectory". In: *arXiv preprint arXiv:2103.11598* (2021).