

UM MODELO IDENTIFICÁVEL PARA DADOS DE DEGRADAÇÃO QUANDO UNIDADES OPERAM EM CONDIÇÕES DISTINTAS

Letícia Pereira Pinto

Universidade Federal de Minas Gerais/Departamento de Ciência da Computação
leticiap@dcc.ufmg.br

Cristiano de Carvalho Santos

Universidade Federal de Minas Gerais/Departamento de Estatística
ccsgaus@est.ufmg.br

Rosangela Helena Loschi

Universidade Federal de Minas Gerais/ Departamento de Estatística
loschi@est.ufmg.br

RESUMO

Informações sobre os tempos de falha de produtos que possuem elevada confiabilidade e que dificilmente falham nos testes de vida tradicionais ou testes de vida acelerados podem ser obtidas pelos experimentos de degradação. Ferreira *et al.* (2012) propõem um modelo de degradação onde a taxa de degradação é função de covariáveis categóricas. Neste artigo, aspectos teóricos desse modelo foram estudados e foi demonstrado a sua não identificabilidade. O efeito da não identificabilidade do modelo no processo de amostragem da distribuição *a posteriori* através de métodos Monte Carlo via cadeia de Markov é discutido. Foram propostas modificações na estrutura de tal modelo, obtendo-se um modelo identificável. O modelo proposto foi implementado sob o enfoque bayesiano. Os resultados obtidos, nos estudos de dados simulados realizados, mostraram que o modelo proposto produz estimativas razoáveis para quantidades relacionadas ao tempo até a falha.

Palavra-chave: Confiabilidade; Modelos de degradação; Estatística Bayesiana.

ABSTRACT

Information on the failure times of highly reliability products, that are unlikely to fail traditional life tests or accelerated life tests, can be obtained from degradation experiments. Ferreira et al. (2012) proposed a linear degradation model in which the degradation rate depends on categorical covariates. In this article, theoretical aspects of this model are studied and it is shown that such a model is unidentifiable. The effect of the model non-identifiability in the posterior sampling via Markov chain Monte Carlo methods is discussed. Modifications in such a model structure are proposed, obtaining an identifiable model. It was implemented under the Bayesian approach. The results obtained in the simulated data studies performed showed that the proposed model produces reasonable estimates for time-related quantities until failure.

Keywords: Reliability; Degradation Path Models; Bayesian Statistics.

Como Citar:

Pinto, Letícia P.; Santos, Cristiano C.; Loschi, Rosangela H.. Um modelo identificável para dados de degradação quando unidades operam em condições distintas. *In: SIMPÓSIO DE PESQUISA OPERACIONAL E LOGÍSTICA DA MARINHA*, 19., 2019, Rio de Janeiro, RJ. *Anais* [...]. Rio de Janeiro: Centro de Análises de Sistemas Navais, 2019.

1. INTRODUÇÃO

A melhoria contínua da qualidade e a redução dos custos de produção são imprescindíveis para garantir a sobrevivência das empresas diante de um mercado cada vez mais competitivo e de uma nova postura dos consumidores. Para que um determinado produto seja produzido em alta qualidade, requer-se que a produção seja realizada com alta confiabilidade. Tradicionalmente, testes de vida são utilizados na determinação da confiabilidade de um produto ou de algum de seus componentes. Na prática, no entanto, os testes de vida podem resultar em pouca ou nenhuma falha, inviabilizando a estimativa da confiabilidade.

No cenário atual, onde os prazos de desenvolvimento dos produtos e a verificação de sua confiabilidade precisam ser conduzidos num tempo severamente restrito, medidas de degradação de determinados componentes, tomadas ao longo do tempo, surgem como uma alternativa. Experimentos de degradação são utilizados com o objetivo de extrair informação sobre os tempos de falha de produtos que possuem elevada confiabilidade, os quais, dificilmente falham nos testes de vida tradicionais ou testes de vida acelerados. Nestes casos, informações sobre o tempo de vida do produto são obtidas indiretamente através de características do produto que estão relacionadas à sua falha como, por exemplo, a sua degradação ao longo do tempo. Consequentemente, dados de degradação fornecem informações sobre a confiabilidade do produto.

Nos modelos de degradação, a variável resposta de interesse não é, diretamente, o tempo de falha. Modela-se o comportamento da degradação de alguma característica de qualidade do produto de interesse, tomada ao longo do tempo. Para inferir sobre a confiabilidade, considera-se que uma falha ocorre quando a degradação total, observada durante o período total de duração do estudo, ultrapassa um nível crítico ou limiar de degradação previamente especificado. A justificativa para a utilização deste tipo de abordagem reside no fato de que muitas falhas são resultantes de um mecanismo de degradação ao longo do tempo inerente ao produto. A principal vantagem da utilização de dados de degradação sobre os dados de tempos de falha, oriundos de testes de vida, é que a análise pode ser feita de maneira satisfatória, ainda que nenhuma falha tenha ocorrido, isto é, mesmo que nenhum dos perfis de degradação em observação tenha atingido o limiar crítico de degradação.

n

Em um ensaio de degradação, uma amostra de n unidades é colocada em teste e a degradação desta unidade será medida ao longo do tempo, ou seja, temos dados longitudinalmente observados. A abordagem geral consiste em modelar todos os perfis de degradação das unidades individuais utilizando a mesma forma funcional. As diferenças entre as unidades amostrais são explicadas através da incorporação de efeitos aleatórios no modelo. Os efeitos aleatórios também têm a função de acomodar a correlação existente entre as medidas de degradação oriundas de uma mesma unidade. Alguns exemplos de aplicações deste tipo

de modelo podem ser vistos nos trabalhos de Hamada (2005), Chiao e Hamada (1996, 2001), Lu e Meeker (1993), Lu *et al.* (1996), Tseng *et al.* (1995), Robinson e Crowder (2000) e Freitas *et al.* (2009).

Hamada (2005) utiliza métodos bayesianos em dados de degradação associados a falhas em emissores de lasers. O modelo proposto é baseado em uma relação linear entre degradação e tempo. Um modelo Bayesiano hierárquico é considerado assumindo-se erros normalmente distribuídos e uma distribuição *a priori* Weibull para os efeitos aleatórios. Em Oliveira *et al.* (2010), os dados de degradação de laser são novamente analisados sob enfoque Bayesiano. Estende-se o trabalho de Hamada (2005) por utilizar-se a distribuição preditiva *a posteriori* dos tempos de falha para indicar planos adequados de manutenção e garantia do produto. Modelos bayesianos de degradação capazes de simultaneamente acomodar comportamento assimétrico e caudas pesadas para a taxa de degradação são propostos por Oliveira *et al.* (2018). Nestes modelos, assume-se que o inverso da taxa de degradação tem distribuição na família de distribuições Skew-normal-independente. Abordagens bayesiana não-paramétricas para modelos de degradação são propostas por Yuan e Ji (2015) e Santos e Loschi (2019). Estas classes de modelos, entre outras coisas, são adequadas para tratar dados de degradação oriundos de populações heterogêneas. Para este fim, Yuan e Ji (2015) modelam as taxas de degradação com uma mistura finita de distribuições normais. Santos e Loschi (2019) estendem esta proposta e consideram misturas via processo de Dirichlet para modelarem o comportamento da taxa de degradação.

As características do produto em estudo ou as condições sob as quais tal produto está submetido, muitas vezes podem ajudar a explicar o processo de degradação. A inclusão de covariáveis em modelos de degradação é considerada nos trabalhos de Bagdonavicius e Nikulin (2000) e Ferreira *et al.* (2012). Para avaliar o problema de degradação em rodas de trem, relacionada aos modos de falha dos descarrilamentos de trens, Ferreira *et al.* (2012) assumem modelos de degradação linear log-normal e normal. A taxa de degradação é definida como função de informações disponíveis sobre as posições que as rodas ocupam nos vagões e questões operacionais que exercem influência no desgaste da roda ao longo do tempo.

Neste trabalho, o modelo proposto por Ferreira *et al.* (2012) foi implementado sob o enfoque bayesiano. Estudamos aspectos teóricos deste modelo e demonstramos que é um modelo não-identificável. Assim, não há informação suficiente nos dados de degradação para estimarmos todos os parâmetros e informação adicional é requerida para este fim. Tal problema relacionado à inferência sobre os parâmetros pode ser observado no estudo das cadeias das distribuições *a posteriori* geradas através de métodos Monte Carlo via Cadeia de Markov (MCMC). Diante disto, propusemos (Seção 2) modificações na estrutura deste modelo obtendo-se um modelo identificável. O modelo proposto foi implementado e alguns estudos em dados simulados (Seção 3) foram realizados para avaliar o seu desempenho. Os resultados obtidos são promissores mostrando que o modelo proposto produz estimativas razoáveis para quantidades relacionadas ao tempo até a falha. A Seção 4 fecha o artigo com algumas considerações finais.

2. MODELO DE DEGRADAÇÃO

Considere uma amostra aleatória com n unidades amostrais tal que para cada unidade amostral i observamos m_i medidas de degradação nos tempos t_{ij} predeterminados, em que t_{ij} pode ser o tempo real, o tempo de operação ou alguma outra medida apropriada para uso, como distância ou número de ciclos. Seja Y_{ij} a j -ésima medida de degradação para

unidade amostral i , com $i = 1, \dots, n$ e $j = 1, \dots, m_i$. Assumimos que a medida de degradação Y_{ij} é observada com um erro aleatório ε_{ij} de tal forma que

$$Y_{ij} = D(t_{ij}; \alpha, \beta_i) + \varepsilon_{ij}, \quad (1)$$

em que $\alpha = (\alpha_1, \dots, \alpha_p)^t$ é um vetor $p \times 1$ de efeitos fixos descrevendo características populacionais e $\beta_i = (\beta_{i1}, \dots, \beta_{ik})^t$ é um vetor $k \times 1$ de efeitos aleatórios representando características individuais da unidade i . O termo $D(t; \alpha, \beta)$ representa o perfil teórico de degradação e sua escolha é baseada em uma análise empírica do processo de degradação sob estudo ou, quando possível, no fenômeno físico-químico associado a ele.

Uma suposição comum nestes modelos é que os erros aleatórios ε_{ij} são independentes e identicamente distribuídos (iid) de acordo com uma distribuição Normal (N) com média 0 e variância σ_ε^2 . Além disso, assume-se que os vetores de efeitos aleatórios β_i são independentes dos erros aleatórios σ_ε^2 . Sob estas suposições, segue que as medidas de degradação, dados os efeitos fixos α , os efeitos aleatórios β , e a variância σ_ε^2 dos erros aleatórios, são independentes (*ind*) com distribuição Normal (N) com média $D(t_{ij}; \alpha, \beta_i)$ e variância σ_ε^2 , a

qual é denotada por $Y_{ij} | \alpha, \beta_i, \sigma_\varepsilon^2 \stackrel{ind}{\sim} N(D(t_{ij}; \alpha, \beta_i); \sigma_\varepsilon^2)$. Seja $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{im_i})^t$ o vetor de ordem $m_i \times 1$ representando as medidas de degradação da unidade i e

$\mathbf{y}_i = (y_{i1}, \dots, y_{im_i})^t$ os valores observados para estas medidas. A partir da suposição de independência dos erros aleatórios ε_{ij} , temos que $\mathbf{Y}_i | \alpha, \beta_i, \sigma_\varepsilon^2 \stackrel{ind}{\sim} N(\boldsymbol{\mu}_{\mathbf{Y}_i}; \sigma_\varepsilon^2 \mathbf{I}_{m_i})$, em que

$\boldsymbol{\mu}_{\mathbf{Y}_i} = (D(t_{i1}; \alpha, \beta_i), \dots, D(t_{im_i}; \alpha, \beta_i))^t$ e \mathbf{I}_{m_i} é a matriz identidade de ordem $m_i \times m_i$, para $i = 1, \dots, n$. Sob tais suposições, a função de verossimilhança condicional nos efeitos aleatórios torna-se

$$f(\mathbf{y} | \alpha, \mathbf{B}, \sigma_\varepsilon^2) = \prod_{i=1}^n \left(\frac{1}{2\pi\sigma_\varepsilon^2} \right)^{m_i/2} \exp \left\{ -\frac{1}{2\sigma_\varepsilon^2} (\mathbf{y}_i - \boldsymbol{\mu}_{\mathbf{Y}_i})^t (\mathbf{y}_i - \boldsymbol{\mu}_{\mathbf{Y}_i}) \right\}, \quad (2)$$

em que $\mathbf{Y} = (\mathbf{Y}_1^t, \dots, \mathbf{Y}_n^t)^t$ é o vetor de valores observados para o conjunto completo de medidas de degradação $N \times 1$, com dimensão $N = \sum_{i=1}^n m_i$ e, tal que, $\mathbf{B} = (\beta_1^t, \dots, \beta_n^t)^t$ é um vetor com todos os efeitos aleatórios. Em geral, assume-se também que os β_i são iid com distribuição $f(\beta_i | \theta)$.

A relação entre as medidas de degradação D e o tempo de falha para cada unidade amostral é estabelecida através de um limiar de falha T . Para a unidade amostral i , dizemos que a falha ocorre quando a verdadeira medida de degradação D atinge o limiar T . Isto é, o tempo de falha T_i para a unidade i pode ser definido por

$$T_i = D^{-1}(D_f; \alpha, \beta_i) \quad (3)$$

Uma vez que α é um vetor de parâmetros fixos, a distribuição do tempo de falha depende apenas da distribuição especificada para os efeitos aleatórios.

2.1. ANÁLISE BAYESIANA DO MODELO DE FERREIRA ET AL. (2012)

Visando a acomodar covariáveis no modelo de degradação em (1), Ferreira *et al.* (2012) assume a seguinte relação entre degradação e o tempo:

$$D(t_{ij}; \alpha, \beta_i) = \alpha_0 + \exp \{ \beta_i + \mathbf{X}_{ij}^t \alpha \} t_{ij}, \quad (4)$$

onde α_0 denota o nível inicial de degradação, α o vetor p -dimensional de efeitos fixos associados às covariáveis, β_i o efeito aleatório relacionado à unidade i representando suas características individuais e \mathbf{X}_{ij} o vetor de covariáveis associadas à unidade i .

Assumimos *a priori* que $\alpha_0 \sim \text{LogN}(a, b)$, $\alpha \sim N_p(\mu_\alpha; \Sigma_\alpha)$ e $\beta_i \stackrel{iid}{\sim} N(\mu_\beta; \sigma_\beta^2)$

em que LogN representa a distribuição Log-Normal com média $\exp(a + \frac{b^2}{2})$ e variância $\exp(2a + b^2) (\exp(b^2) - 1)$, e N_p a distribuição Normal p -variada com vetor de médias μ_α e matriz de covariâncias Σ_α . Para a variância dos efeitos aleatórios assumimos

$\sigma_\beta^2 \sim \text{GI}(\lambda_\beta; \psi_\beta)$ em que GI representa a distribuição Gama inversa com média $\frac{\psi_\beta}{(\lambda_\beta - 1)}$ e

variância $\frac{(\psi_\beta)^2}{(\lambda_\beta - 1)^2 (\lambda_\beta - 2)}$, e para a variância do erro do modelo assumimos $\sigma_\varepsilon^2 \sim \text{GI}(\lambda_\varepsilon, \psi_\varepsilon)$.

Para simplificar o estudo de simulação consideramos apenas uma covariável no modelo.

Vários estudos com dados simulados foram realizados. As Figuras 1 e 2 ilustram o comportamento obtido para as cadeias das distribuições *a posteriori* de alguns parâmetros

$(\beta_5, \beta_6, \beta_7, t_5, t_6$ e $t_7)$ no modelo proposto por Ferreira *et al.* (2012). Os efeitos aleatórios

reais foram gerados de uma distribuição Normal com parâmetros de locação $\mu_\beta = -1$ e de

escala $\sigma_\beta^2 = 0,5$. Os seguintes valores reais foram obtidos: $\beta_5 = 0,2015$, $\beta_6 = 0,2849$ e

$\beta_7 = 0,2389$. Consequentemente os tempos de falha reais para as unidades 5, 6 e 7 são:

$t_5 = 22,7948$, $t_6 = 131,3721$ e $t_7 = 229,0339$.

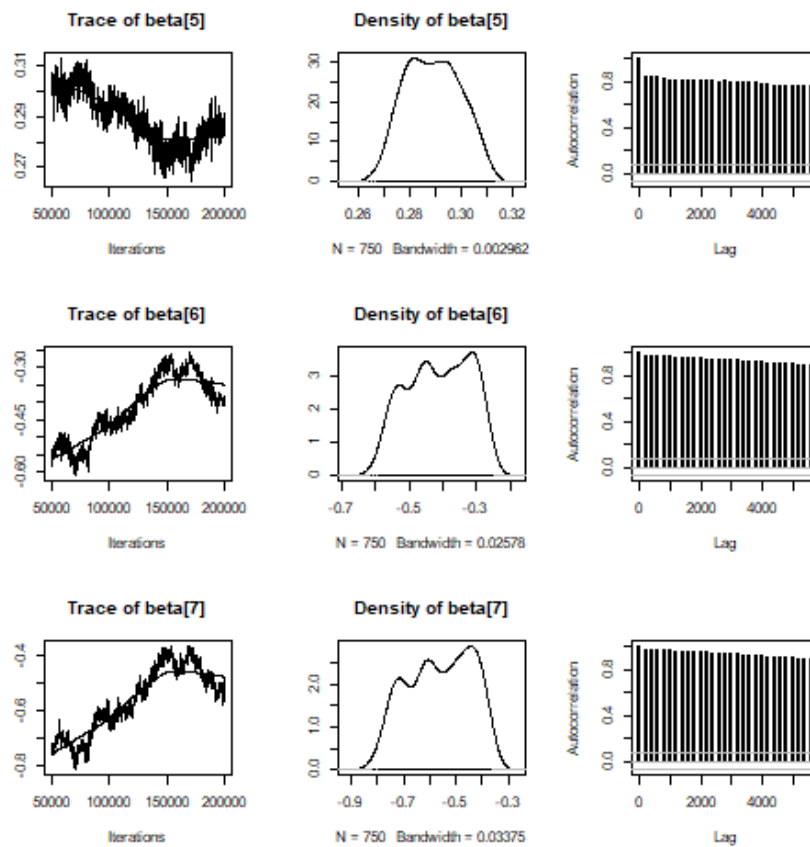


Figura 1 – Gráficos do traço da cadeia, densidade estimada e autocorrelação para os parâmetros β .

Verifica-se que as cadeias formadas para amostrar da distribuição *a posteriori* desses parâmetros não convergiram para o valor real e apresentaram uma alta correlação para todos os parâmetros. O mesmo comportamento foi verificado para os demais efeitos aleatórios e tempos de falha t_i ao utilizarmos este modelo (Figura 2).

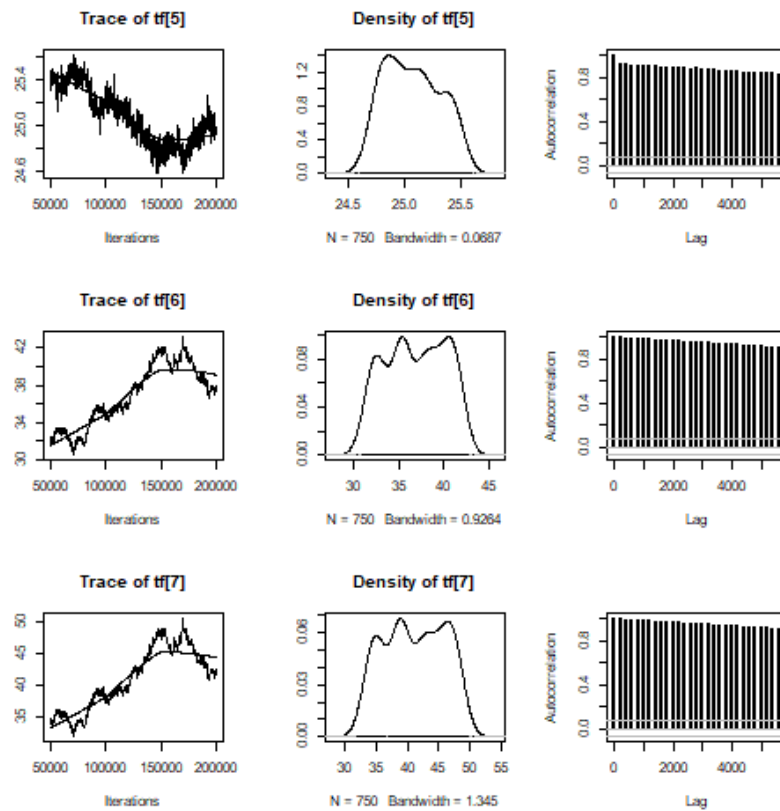


Figura 2 – Gráficos do traço da cadeia, densidade estimada e autocorrelação para os parâmetros t_i .

Investigando teoricamente o problema, constatamos que o modelo proposto por Ferreira *et al.* (2012) é não identificável. Sob as suposições anteriormente mencionadas para modelos de degradação e a estrutura em (4), segue que

$$Y_{ij} \underset{\sim}{ind} N(\alpha_0 + \exp\{\mathbf{X}_{ij}^t \boldsymbol{\alpha} + \beta_i\} t_{ij}; \sigma^2), \quad i = 1, \dots, n; \quad j = 1, \dots, m_i$$

Conseqüentemente, a função de verossimilhança é dada por:

$$f_{\mathbf{Y} | \alpha_0, \boldsymbol{\alpha}, \mathbf{X}_{ij}, \beta, \sigma^2} = \prod_{i=1}^n \prod_{j=1}^{m_i} \left\{ \left(\frac{1}{2\pi\sigma^2} \right)^{1/2} \exp \left[-\frac{1}{2\sigma^2} (y_{ij} - \mu_y)^2 \right] \right\},$$

$$\mu_y = \alpha_0 + \exp\{\mathbf{X}_{ij}^t \boldsymbol{\alpha} + \beta_i\} t_{ij}$$

onde

. Após alguma álgebra, obtemos que

$$f_{\mathbf{Y} | \alpha_0, \boldsymbol{\alpha}, \mathbf{X}_{ij}, \beta, \sigma^2} = \left(\frac{1}{2\pi} \right)^{\sum_{i=1}^n m_i / 2} \left(\frac{1}{\sigma^2} \right)^{\sum_{i=1}^n m_i / 2} \\ \times \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n \sum_{j=1}^{m_i} [\exp\{\alpha_0\} + \exp\{\mathbf{X}_{ij}^t \boldsymbol{\alpha} + \beta_i\} t_{ij}]^2 \right\}$$

$$\begin{aligned} & \times \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n \sum_{j=1}^{m_i} y_{ij}^2\right\} \times \exp\left\{\frac{\exp\{\alpha_0\}}{\sigma^2} \sum_{i=1}^n \sum_{j=1}^{m_i} y_{ij}\right\} \\ & \times \exp\left\{\frac{1}{\sigma^2} \sum_{i=1}^n \sum_{j=1}^{m_i} \exp\{\mathbf{X}_{ij}^t \boldsymbol{\alpha} + \beta_i\} t_{ij} y_{ij}\right\}, \end{aligned}$$

de onde observamos que há estas (α, β) poucas suficientes apenas para estimar os parâmetros α_0 , $\boldsymbol{\alpha}$ e uma função dos parâmetros (β_i) . Isto faz com que o modelo seja não identificável gerando os problemas de convergência observados. Para identificarmos o modelo em (4), devemos (α, β) ou elicitar distribuições *a priori* muito informativas para um dos parâmetros do vetor (α, β) . Isto demandará termos informações adicionais sobre tal parâmetro, provenientes de fontes tais como dados de validação ou outros experimentos. Isto torna o uso do modelo um pouco restritivo. Diante do problema descrito, proporemos, na próxima seção, um novo modelo para o estudo de dados de degradação.

2.2. MODELO PROPOSTO

Visando a contornar o problema de falta de identificabilidade no modelo de degradação em (4), assumimos a seguinte relação entre a degradação e o tempo de medição:

$$D(t_{ij}; \boldsymbol{\alpha}, \beta_i) = \exp\{\alpha_0\} + (\beta_i + \exp\{\mathbf{X}_{ij}^t \boldsymbol{\alpha}\})t_{ij}, \quad (5)$$

onde $\exp\{\alpha_0\}$ é o nível inicial de degradação, $\boldsymbol{\alpha}$ é o vetor p -dimensional de efeitos fixos associados às covariáveis, β_i o efeito aleatório relacionado à unidade i representando suas características individuais e \mathbf{X}_{ij} o vetor de covariáveis associadas à unidade i . Sob esta estrutura e as hipóteses anteriormente mencionadas para os modelos de degradação temos:

$$Y_{ij} | \alpha_0, \boldsymbol{\alpha}, \mathbf{X}_{ij}, \beta_i, \sigma^2 \underset{\sim}{\text{ind}} N(\mu_{y_{ij}}; \sigma^2), i = 1, \dots, n; j = 1, \dots, m_i.$$

Consequentemente, a função de verossimilhança é dada por

$$f_{Y|\alpha_0, \boldsymbol{\alpha}, \mathbf{X}_{ij}, \beta_i, \sigma^2} = \prod_{i=1}^n \prod_{j=1}^{m_i} \left\{ \left(\frac{1}{2\pi\sigma^2}\right)^{1/2} \exp\left[-\frac{1}{2\sigma^2} (y_{ij} - \mu_{y_{ij}})^2\right] \right\},$$

onde $\mu_{y_{ij}} = \exp\{\alpha_0\} + (\beta_i + \exp\{\mathbf{X}_{ij}^t \boldsymbol{\alpha}\})t_{ij}$. Após algumas operações algébricas obtém-se

$$\begin{aligned} f_{Y|\alpha_0, \boldsymbol{\alpha}, \mathbf{X}_{ij}, \beta_i, \sigma^2} &= \left(\frac{1}{2\pi}\right)^{\sum_{i=1}^n m_i / 2} \left(\frac{1}{\sigma^2}\right)^{\sum_{i=1}^n m_i / 2} \\ &\times \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n \sum_{j=1}^{m_i} [\exp\{\alpha_0\} + (\beta_i + \exp\{\mathbf{X}_{ij}^t \boldsymbol{\alpha}\})t_{ij}]^2\right\} \end{aligned}$$

$$\begin{aligned} & \times \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n \sum_{j=1}^{m_i} y_{ij}^2 \right\} \exp \left\{ \frac{\exp\{\alpha_0\}}{\sigma^2} \sum_{i=1}^n \sum_{j=1}^{m_i} y_{ij} \right\} \\ & \times \exp \left\{ \frac{1}{\sigma^2} \sum_{i=1}^n \beta_i \sum_{j=1}^{m_i} t_{ij} y_{ij} \right\} \exp \left\{ \frac{1}{\sigma^2} \sum_{i=1}^n \sum_{j=1}^{m_i} \exp\{\mathbf{X}_{ij}^t \boldsymbol{\alpha}\} t_{ij} y_{ij} \right\} \end{aligned}$$

Neste modelo, há estatísticas suficientes independentes para estimar todos os parâmetros do modelo, mostrando que o modelo proposto é identificável. Para completar a especificação do modelo, assumimos *a priori* que $\alpha_0 \sim N(a, b)$, $\boldsymbol{\alpha} \sim N_p(\boldsymbol{\mu}_\alpha, \boldsymbol{\Sigma}_\alpha)$ e

$\beta_i \stackrel{iid}{\sim} \text{LogN}(\mu_\beta, \sigma_\beta^2)$. Similar ao modelo proposto por Ferreira *et al.* (2012), para a variância do efeito aleatório assumimos $\sigma_\beta^2 \sim GI(\lambda_\beta, \psi_\beta)$ e para a variância do erro do modelo assumimos $\sigma_\varepsilon^2 \sim GI(\lambda_\varepsilon, \psi_\varepsilon)$.

Como consequência da identificação do modelo, ao implementarmos o modelo usando um esquema MCMC, obtivemos cadeias bem mais comportadas, convergindo para o verdadeiro valor do parâmetro, com menor autocorrelação para todos os efeitos aleatórios (β_1, \dots, β_n) e apresentaram um decaimento rápido dos autocorrelogramas para todos os tempos de falha (t_1, \dots, t_n). Isto pode ser visto no estudo de simulação exibido nas Figuras 3 e 4, os efeitos aleatórios reais foram gerados de uma distribuição Log-Normal com parâmetros de locação $\mu_\beta = -1$ e de escala $\sigma_\beta^2 = 0,5$. Os seguintes valores reais foram obtidos: $\beta_5 = 1,4301$, $\beta_6 = 0,3108$ e $\beta_7 = 0,1341$. Consequentemente os tempos de falha reais para as unidades 5, 6 e 7 são: $t_5 = 13,2732$, $t_6 = 26,2153$ e $t_7 = 7,3760$.

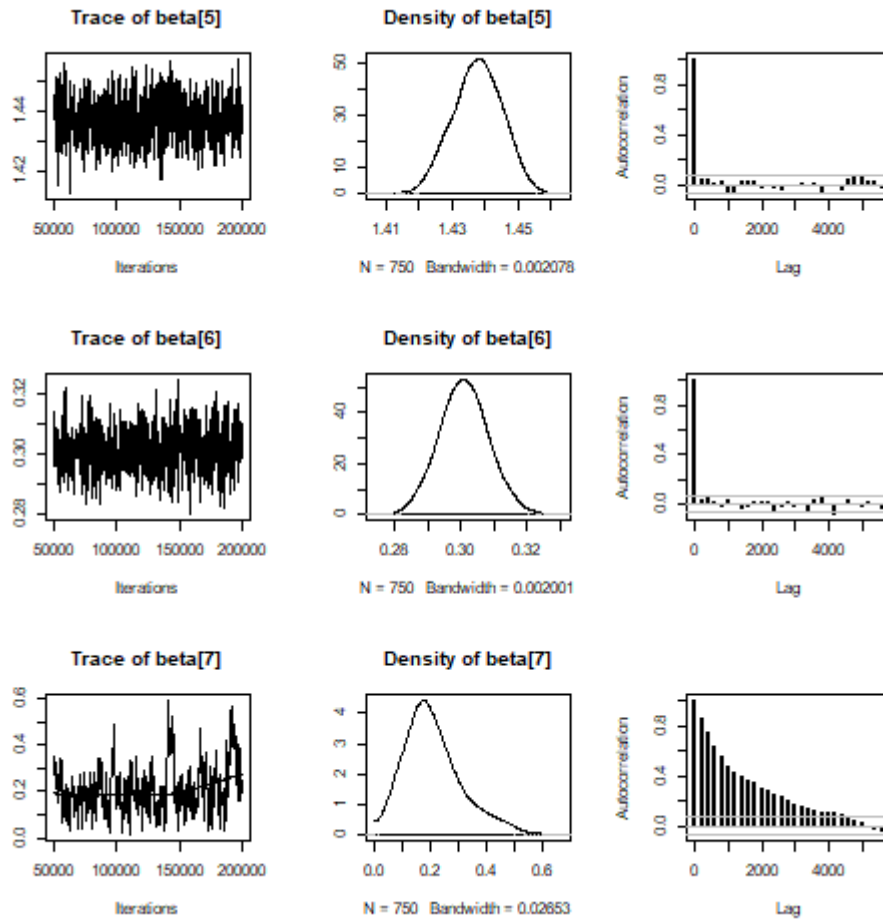


Figura 3 – Gráficos do traço da cadeia, densidade estimada e autocorrelação para os parâmetros β .

Uma possível explicação para o decaimento mais lento da autocorrelação da cadeia relacionada ao efeito aleatório da unidade 7 seria uma maior variabilidade na informação amostral para este perfil de degradação. No entanto, tal comportamento não é observado para a distribuição do tempo até a falha, que é o principal alvo do estudo.

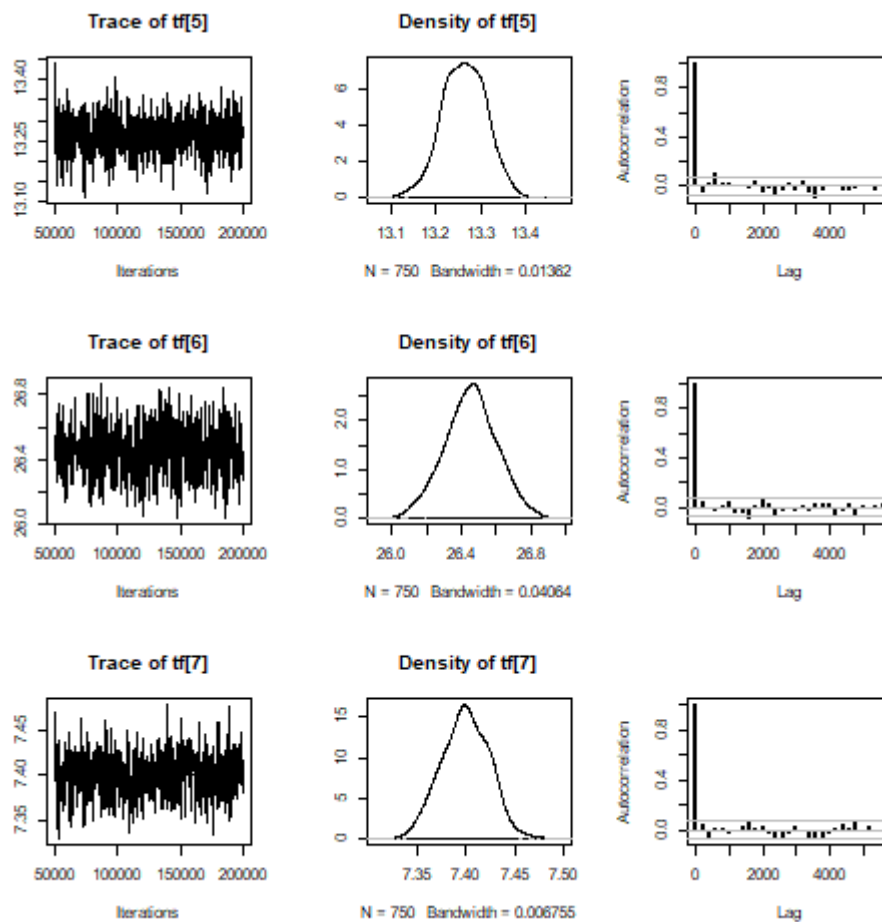


Figura 4 – Gráficos do traço da cadeia, densidade estimada e autocorrelação para os parâmetros t_i .

3. RESULTADOS

Para avaliar a qualidade das estimativas para os parâmetros $\alpha_0, \sigma_\beta^2, \sigma_\epsilon^2, \mu_\beta$ e para os tempos de falha de uma unidade futura sob o modelo proposto, realizamos um estudo de simulação Monte Carlo, em que amostras foram geradas do modelo proposto, assumindo apenas uma covariável. Em cada cenário analisado, foram consideradas 100 réplicas Monte Carlo. Foram geradas amostras de tamanho $n = 15$ e 30, considerando $m_i = 10$ e 20 tempos de medição. Para investigar o efeito da variabilidade do modelo e da incerteza sobre o efeito aleatório de cada unidade, dois valores distintos foram assumidos para cada um destes parâmetros. Fizemos $\sigma_\epsilon^2 = 0,5$ e 20, e $\sigma_\beta^2 = 0,5$ e 20. O modelo proposto foi ajustado assumindo distribuições *a priori* vagas para todos os parâmetros. Foram consideradas as seguintes distribuições *a priori* para média e precisão dos efeitos aleatórios: $\mu_\beta \sim N(m, v\tau_\beta)$, com $m = 0$, $v = 0,01$ e $\tau_\beta \sim Gamma(a_\beta, b_\beta)$, com $a_\beta = 0,01$ e $b_\beta = 0,01$. Para a precisão dos erros assumimos $\tau_\epsilon \sim Gamma(a_\epsilon, b_\epsilon)$, com $a_\epsilon = 0,001$ e $b_\epsilon = 0,001$. A análise dos dados foi efetuada utilizando-se os softwares JAGS 4.3.0 e RStudio (pacote R2jags). A fim de garantir a convergência da cadeia para a distribuição *a posteriori* das estimativas dos parâmetros, assim como minimizar a autocorrelação dessas, foram realizadas

400.000 iterações da simulação MCMC, com um período de aquecimento (*burn-in*) de 100.000, e saltos (*lag*) de 200 para a amostragem, resultando em amostras das distribuições *a posteriori* dos parâmetros $\alpha_0, \alpha, \beta_i, \sigma_\beta^2$ e σ_ε^2 de tamanho 2.000.

As Tabelas 1, 2 e 3 mostram o viés, o erro quadrático médio (EQM), o percentual de cobertura, a maior amplitude (MaHPD) e a menor amplitude (MeHPD) dos intervalos de mais alta densidade *a posteriori* (HPD) com probabilidade de 0,95 para todos os parâmetros do modelo e para o tempo médio até a falha Tf .

Tabela 1: Viés, erro quadrático médio (EQM), percentual de cobertura, maior amplitude (MaHPD) e menor amplitude (MeHPD) dos intervalos HPD com probabilidade de 0,95, para $m_1 = 10$ e 20 em cada unidade

PARÂMETRO	VIÉS	EQM	HPD		% Cobert.
			MeHPD	MaHPD	
$n = 15; m_i = 10; \sigma_\beta^2 = 0,5; \sigma_\varepsilon^2 = 0,5$					
α_0	-0,0082	0,0026	0,2121	0,3123	0,96
α	-0,0087	0,0026	0,0683	0,4977	0,92
β	0,0015	0,0055	0,0640	0,3742	0,93
μ_β	-0,0147	0,0430	0,4400	1,1992	0,94
σ_β^2	0,0525	0,0510	0,2545	2,0279	0,94
σ_ε^2	0,0124	0,0028	0,1838	0,3085	0,97
Tf	0,0795	8,7398	0,8775	19,8519	0,95
$n = 15; m_i = 20; \sigma_\beta^2 = 0,5; \sigma_\varepsilon^2 = 0,5$					
α_0	0,0010	0,0019	0,1669	0,2265	0,94
α	-0,0019	0,0018	0,0773	0,4778	0,94
β	-0,0003	0,0041	0,0621	0,3400	0,94
μ_β	-0,0162	0,0334	0,4532	1,1157	0,96
σ_β^2	0,0354	0,0378	0,2664	1,7366	0,94
σ_ε^2	0,0127	0,0018	0,1336	0,2149	0,96
Tf	-0,0788	5,2554	0,7185	14,5611	0,95

Na Tabela 1 verifica-se que mantendo o número de perfis ($n = 15$) constante, quando aumentamos os intervalos de medição de $m_i = 15$ para $m_i = 20$, o aumento do número de medições em cada unidade reduz o viés e o erro quadrático médio das estimativas. O percentual de cobertura dos intervalos HPD estão próximos de 95% indicando o bom ajuste do modelo considerado.

Na Tabela 2, o número de medições ($m_i = 10$) é mantido fixo, considerando diferentes valores para o número de perfis ($n = 15$ e 20), para σ_β^2 e para σ_ε^2 . Comparando os resultados da Tabela 2 com aqueles mostrados na Tabela 1, observamos que o aumento apenas da variação nos efeitos aleatórios (σ_β^2) leva a um aumento no viés de σ_β^2 , de σ_ε^2 e dos tempos de falha (Tf), e no EQM de todos os parâmetros. Também notamos que havendo um aumento somente na variação do erro relativo à degradação da unidade (σ_ε^2), há um aumento

no viés dos efeitos fixos α , dos efeitos aleatórios (β), de σ_β^2 e de σ_ϵ^2 , e no EQM de todos os parâmetros. Há uma piora substancial nas estimativas do tempo médio até a falha em ambos os casos, o que de certa forma é esperado uma vez que há um aumento substancial na incerteza.

Tabela 2: Viés, erro quadrático médio (EQM), percentual de cobertura, maior amplitude (MaHPD) e menor amplitude (MeHPD) dos intervalos HPD com probabilidade de 0,95, quando mantemos fixos o número de perfis ($n = 15$ e 30) e de medições ($m_i = 10$), e considerando diferentes valores para σ_β^2 e σ_ϵ^2 .

PARÂMETRO	VIÉS	EQM	HPD		
			MeHPD	MaHPD	% Cobert.
$n = 15; m_i = 10; \sigma_\beta^2 = 0,5; \sigma_\epsilon^2 = 20$					
α_0	-0,0094	0,0028	0,9298	6,9766	0,96
α	-0,0103	0,0028	0,0805	0,6322	0,93
β	0,0075	0,0074	0,1476	0,4983	0,94
μ_β	-0,0024	0,0471	0,4853	1,3403	0,92
σ_β^2	0,0546	0,0574	0,2686	2,4849	0,92
σ_ϵ^2	0,5578	4,6677	7,1279	12,5819	0,97
Tf	-1,1444	278,8348	5,4849	105,4141	0,94
$n = 15; m_i = 10; \sigma_\beta^2 = 20; \sigma_\epsilon^2 = 0,5$					
α_0	-0,5002	4,6631	0,1724	16,4898	0,95
α	-0,4945	4,6599	0,0052	12,6349	0,85
β	-2448,788	1,647 e+09	1,518 e-05	8,438 e+02	0,77
μ_β	-16,7803	2257,905	2,6125	394,2598	0,83
σ_β^2	231,594	1,466 e+06	2,3186	3,434 e+04	0,93
σ_ϵ^2	2,059 e+12	1,774 e+26	0,1765	5,336 e+13	0,62
Tf	2,809 e+04	8,572 e+11	3,116 e-04	4,349 e+08	0,92
$n = 30; m_i = 10; \sigma_\beta^2 = 0,5; \sigma_\epsilon^2 = 20$					
α_0	-0,0050	7,323 e-04	0,8577	6,3779	1,00
α	-0,0035	5,201 e-04	0,0618	0,1329	0,98
β	0,0096	2,908 e-03	0,1319	0,2378	0,96
μ_β	0,0012	1,713 e-02	0,4067	0,8078	0,94
σ_β^2	0,0387	2,991 e-02	0,3012	1,2753	0,93
σ_ϵ^2	0,4329	3,271 e+00	5,5443	8,9693	0,94
Tf	-2,5456	4,707 e+02	5,8514	102,8093	0,94
$n = 30; m_i = 10; \sigma_\beta^2 = 20; \sigma_\epsilon^2 = 0,5$					
α_0	-1,728 e-01	4,233 e+00	0,1414	1,627 e+01	0,96
α	-2,057 e-01	4,125 e+00	0,0058	1,214 e+01	0,84
β	-1,317 e+04	3,460 e+11	0	2,008 e+02	0,73
μ_β	-1,910 e+01	2,351 e+03	2,1634	4,586 e+02	0,83
σ_β^2	1,1636 e+02	1,791 e+05	2,3730	1,188 e+04	0,90
σ_ϵ^2	4,4058 e+14	1,910 e+31	0,1486	1,420 e+16	0,47
Tf	8,2246 e+03	1,312 e+11	3 e-04	2,917 e+06	0,92

Na Tabela 3 são mantidos fixos o número de perfis ($n = 30$), a variação nos efeitos aleatórios (σ_{β}^2) e a variação nos erros relativos à degradação da unidade (σ_{ϵ}^2).

Tabela 3: Viés, erro quadrático médio (EQM), percentual de cobertura, maior amplitude (MaHPD) e menor amplitude (MeHPD) dos intervalos HPD com probabilidade de 0,95, para $n = 30$, $m_i = 10$ e 20 e para $\sigma_{\beta}^2 = \sigma_{\epsilon}^2 = 0,5$

PARÂMETRO	VIÉS	EQM	HPD		
			MeHPD	MaHPD	% Cobert.
$n = 30; m_i = 10; \sigma_{\beta}^2 = 0,5; \sigma_{\epsilon}^2 = 0,5$					
α_0	-0,0032	4,4711 e-04	0,1548	0,2296	0,99
α	-0,0032	4,4710 e-04	0,0456	0,1925	0,96
β	0,0017	1,1238 e-03	0,0519	0,1677	0,96
μ_{β}	-0,0240	1,7039 e-02	0,3802	0,7691	0,94
σ_{β}^2	0,0502	2,7597 e-02	0,2709	1,1103	0,93
σ_{ϵ}^2	0,0094	1,9687 e-03	0,1399	0,2275	0,93
Tf	-0,2241	1,4306 e+01	0,9260	21,6647	0,95
$n = 30; m_i = 20; \sigma_{\beta}^2 = 0,5; \sigma_{\epsilon}^2 = 0,5$					
α_0	-0,0033	0,0004	0,1161	0,1464	0,95
α	-0,0031	0,0004	0,0449	0,1357	0,93
β	0,0012	0,0010	0,0467	0,1172	0,95
μ_{β}	-0,0139	0,0157	0,3739	0,7046	0,97
σ_{β}^2	0,0421	0,0227	0,2787	0,9332	0,96
σ_{ϵ}^2	0,0069	0,0009	0,1018	0,1326	0,95
Tf	0,0379	7,8589	0,6095	15,7334	0,94

Comparando-se os resultados apresentados nesta tabela com aqueles mostrados nas tabelas anteriores, percebe-se que aumentando somente o número de perfis (n) há uma redução tanto do viés quanto do EQM de todos os parâmetros. O percentual de cobertura dos intervalos HPD ficam razoavelmente próximos de 0,95 indicando que o modelo apresentou bom ajuste. As amplitudes mínimas do intervalo HPD (MeHPD) para todos os parâmetros não sofreram grandes alterações mantendo-se o número de perfis (n) fixo e aumentando-se o número de medições (m_i) para 20, exceto para o tempo médio até a falha onde se percebe uma redução da MeHPD. Enquanto as amplitudes máximas do intervalo HPD reduziram para todos os parâmetros mostrando que a distribuição *a posteriori* desses torna-se mais concentrada.

4. CONSIDERAÇÕES FINAIS

A obtenção de estimativas de confiabilidade de um produto sem a necessidade de conduzi-lo até a falha representa um avanço, já que, os prazos para o desenvolvimento dos produtos estão cada vez mais reduzidos. Assim, a análise de dados de degradação mostra-se como uma alternativa eficaz em relação aos testes usuais de confiabilidade.

Na tentativa de se implementar o modelo proposto por Ferreira *et al.* (2012) sob o enfoque bayesiano, observamos que as cadeias geradas no MCMC não convergiam. Demonstramos que esse modelo é não identificável, o que nos motivou a propor modificações na sua estrutura.

O modelo de degradação proposto neste trabalho mostrou ser uma abordagem interessante e promissora para a análise de dados de degradação quando há covariáveis que podem explicar seu comportamento. Por ser um modelo identificável, estimativas dos parâmetros são obtidas sem a necessidade de dados de validação ou de informação *a priori* forte. Nos estudos de simulação, o modelo proposto apresentou um bom desempenho, produzindo estimativas razoáveis para quantidades relacionadas ao tempo até a falha, mostrando a sua efetividade.

Agradecimento: Os autores agradecem ao CNPq, projetos no. 301627/2017-7 e 150004/2018-4 (401673/2017-0) (PDJ-Pós-Doutorado Júnior), FAPEMIG e PRPq-UFMG (Projeto 26048-15, [Programa Institucional de Auxílio à Pesquisa de Doutores Recém-contratados da UFMG](#), Edital 11/2017).

5. REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Bagdonavicius, V. and Nikulin, N. (2000). Estimation in degradation models with explanatory variables. *Lifetime Data Analysis*, v.7, 85--103.
- [2] Chiao, C. H. and Hamada, M. (1996). Using Degradation Data from an Experiment to Achieve Robust Reliability for Light Emitting Diodes. *Quality and Reliability Engineering International*, v.12, 89--94.
- [3] Chiao, C. H. and Hamada, M. (2001). Analyzing Experiments with Degradation Data for Improving Reliability and for Achieving Robust Reliability. *Quality and Reliability Engineering International*, v.17, 333--344.
- [4] Ferreira, J. C., Freitas, M. A. e Colosimo, E. A. (2012). Degradation Data Analysis for samples under unequal operating conditions: a case study on train wheels. *Journal of Applied Statistics*, v.39, n. 12, 2721-2739.
- [5] Freitas, M. A., Toledo, M. L. G., Colosimo, E. A., Pires, M. C. (2009). Using Degradation Data to Assess Reliability: A Case Study on Train Wheel Degradation. *Quality and Reliability Engineering International*, v.25, 607--629.
- [6] Hamada, M. (2005). Using Degradation Data to Assess Reliability. *Quality Engineering*, v.17, 615--620.
- [7] Lu, C. J. and Meeker, W. Q. (1993). Using Degradation Measurements to Estimate a Time-to-Failure Distribution. *Technometrics*, v.35: 161--174.
- [8] Lu, C. J., Meeker, W. Q. and Escobar, L. A. (1996). Using Degradation Measurements to Estimate a Time-to-Failure Distribution. *Statistica Sinica*, v.6, 531--546.
- [9] Oliveira, R. P. B., Freitas, M. A., Loschi, R. H. e Colosimo, E. A. (2010). Um modelo de degradação Bayesiano com aplicações para manutenção e determinação de garantia. *Anais do XLII Simpósio Brasileiro de Pesquisa Operacional*, 2010, Bento Gonçalves-RS, pg. 1--11.
- [10] Oliveira, R. P. B., Loschi, R. H. e Freitas, M. A. (2018). Skew-heavy-tailed degradation models: An application to train wheel degradation. *IEEE Transactions on Reliability*, v.67, 129--141.
- [11] Robinson, M. E. and Crowder, M. (2000). Bayesian Methods for a Growth-Curve Degradation Model with Repeated Measures. *Lifetime Data Analysis*, v.6, 357--374.

- [12] Santos, C. C. e Loschi, R. H. (2019). Semi-parametric models for heterogeneous degradation data. Manuscrito submetido à publicação.
- [13] Tseng, S. T., Hamada, M. and Chiao, C. H. (1995). Using degradation data from a factorial experiment to improve fluorescent lamp reliability. *Journal of Quality Technology*, v.27, 363--369.
- [14] Yuan, T. e Ji, Y. (2015). A hierarchical Bayesian degradation model for heterogeneous data, *IEEE Transactions on Reliability*, v. 64, n.1, 63--70.