

Produto & Produção, vol. 9, n. 3, p. 55-65, out. 2008



Aplicação de um modelo de riscos concorrentes na análise de confiabilidade de dados de garantia

Gilberto Tavares dos Santos, Dr.

Programa de Pós-Graduação em Engenharia de Produção - UFRGS
gtavares@producao.ufrgs.br

Flávio Sanson Fogliatto, PhD.

Programa de Pós-Graduação em Engenharia de Produção - UFRGS
ffogliatto@producao.ufrgs.br

José Luis Duarte Ribeiro, Dr.

Programa de Pós-Graduação em Engenharia de Produção - UFRGS
ribeiro@producao.ufrgs.br

Marvin Rausand, PhD.

Noewegian Technology University (NTNU - Noruega)
marvin.rausand@ntnu.no

Na análise de confiabilidade, espera-se que dados de vida de equipamentos sigam uma distribuição de probabilidade conhecida; por exemplo, uma distribuição de Weibull ou Lognormal. Entretanto, quando se modelam falhas originadas em campo, essas podem decorrer de causas múltiplas e os tempos até a falha podem, assim, estar associados a diferentes estágios do ciclo de vida de um equipamento, não se ajustando a uma única distribuição de probabilidade. Neste artigo, propõe-se a aplicação de um modelo de confiabilidade para riscos concorrentes em dados de garantia oriundos de duas fases do ciclo de vida de um equipamento: a fase de vida operacional e a fase de envelhecimento (desgaste). Considera-se que falhas surgidas nessas duas fases podem ocorrer simultaneamente e têm início tão logo o equipamento seja colocado em funcionamento. O modelo combina elementos de uma distribuição exponencial e de uma distribuição de Weibull com dois parâmetros. Os parâmetros do modelo proposto são derivados utilizando estimadores de máxima verossimilhança e um teste de ajuste é utilizado para verificar o desempenho do modelo. Equações de confiabilidade são desenvolvidas para ilustrar os desenvolvimentos propostos em um estudo de caso.

Palavras-chave: modelo de confiabilidade; análise de dados de vida; dados de garantia

In reliability analysis, life data are typically expected to follow a known probability distribution such as the Weibull or Lognormal distributions. However when modeling field data, failures may be due to multiple causes and the resulting times-to-failure may be associated with different stages of the product life cycle, not conforming to a single probability distribution. In this paper we propose a competing risk model to analyze reliability warranty data coming from two phases of the product life cycle: the operational life and the wear-out phases. It is considered that failures coming from these two phases can occur simultaneously as soon as the product is put into operation. The model combines elements of the exponential and the two-parameter Weibull distribution. Reliability equations and maximum likelihood estimators of the model parameters are derived for the new model. Real data obtained from warranty claims on air conditioning equipments are used to illustrate the developments proposed and a goodness-of-fit test is used to verify the performance of the proposed model.

Keywords: reliability model; life data analysis; warranty data

1 Introdução

Muitas empresas de manufatura, em particular aquelas que competem em mercados internacionais, alocam uma grande quantidade de recursos no projeto de equipamentos confiáveis (NELSON, 1988; WASSERMAN e

SUDJANTO, 1996). Tais equipamentos, se adequadamente projetados, apresentam um número mínimo de falhas operacionais durante seu período de garantia e, de forma geral, durante sua vida útil. Para obter níveis satisfatórios

de confiabilidade em seus equipamentos, técnicas como FMEA (Failure modes and effect analysis - análise de modos e efeitos de falhas), FTA (Fault tree analysis - Análise da árvore de falhas) e testes acelerados de confiabilidade costumam ser utilizadas em seu desenvolvimento. Entretanto, para o usuário, o desempenho em campo de um equipamento é o melhor indicador de sua confiabilidade.

Inferências a respeito da confiabilidade de equipamentos, relacionadas principalmente a características de garantia e políticas de reposição de peças, dependem da disponibilidade de dados de desempenho desses equipamentos. Na maioria das situações, os dados são coletados na forma de tempos até a falha de unidades do equipamento por simulação do desempenho em laboratório ou por coleta de informações de consumo (em campo). Porém, pressões de mercado relacionadas a custos de desenvolvimento dos equipamentos e restrições de tempo para tomada de decisão estão conduzindo os equipamentores a realizar um número menor de testes, de mais curta duração. Nesse contexto, dados de garantia proporcionam informações valiosas que podem ser utilizadas para alcançar melhorias na modelagem dos equipamentos.

Dados de garantia são obtidos a partir de solicitações de consertos ou reposições por parte de clientes, usualmente motivadas por falhas ou baixo desempenho. O período de garantia pode variar de poucos meses a muitos anos, cobrindo diferentes fases da vida dos equipamentos. Falhas que induzem a utilização da garantia de um equipamento podem ter origem em causas especiais, comuns ou a uma conjugação de ambas. Em qualquer caso, dado o tipo de falha, é possível identificar em um gráfico da função de risco o intervalo de tempo em que a falha provavelmente tenha ocorrido. Um conjunto de dados de garantia pode ser constituído de registros de falhas causadas por diferentes mecanismos. Assim, os tempos-até-falha associados a esses dados podem não seguir uma distribuição de probabilidade específica, mas uma combinação de distribuições. Além disso, dados de garantia são altamente censurados, já que poucas unidades do equipamento devem falhar durante o período de garantia.

Neste artigo, propõe-se um modelo de riscos concorrentes para a análise estatística de dados de garantia. O interesse recai, particularmente, na aplicação deste modelo em dados amostrais oriundos de duas populações com taxas de falha constante e crescente, respectivamente, tipicamente modelados por uma mistura de distribuições de Weibull, com modelo de mistura dado pela eq.(1) (LEEMIS, 1995; NAIR et al., 2001):

$$f(t) @ \hat{\sum}_{l @ 1}^2 p_l f_l \# | \mathbf{t}_l , \quad (1)$$

onde $f_l \# | \mathbf{t}_l ,$ representa a função de densidade de probabilidade da l -ésima população com vetor de parâmetros $\mathbf{t}_l ,$ p_l e $\hat{\sum}_{l @ 1}^2 p_l @ 1$ e $p_l , 0 ,$

A eq. (1) representa uma abordagem clássica para modelar dados de garantia ou dados de vida coletados em campo, em que falhas devidas a múltiplas causas ocorrem simultaneamente. Tal modelo, entretanto, pode oferecer ajuste deficiente aos dados já que o parâmetro de mistura p_l não é função do tempo, ao contrário do que ocorre em grande parte das situações práticas. Em outras palavras, a proporção de falhas devidas a causas comuns (isto é, aleatórias) e a causas especiais varia com o tempo: causas comuns prevalecem em fases iniciais da vida do equipamento (mortalidade infantil); causas especiais, usualmente associadas ao uso e degradação do equipamento, são predominantes em fases de maturidade e finais da vida do equipamento (RAUSAND; HØYLAND, 2004).

O modelo de confiabilidade apresentado neste artigo está fundamentado em uma combinação de distribuições de probabilidade: a distribuição exponencial, adequada para modelar tempos-até-falha na fase de vida operacional de um equipamento, associados a falhas aleatórias (constantes) e uma distribuição de Weibull com dois parâmetros, adequada para modelar tempos-até-falha, associados às falhas de desgaste (crescentes), apesar de uma distribuição de Weibull com três parâmetros ser a mais ajustável para interpretar falhas dessa natureza. Porém, a aplicação do modelo proposto pressupõe que: (i) a ocorrência simultânea desses dois tipos de falha pode se dar tão logo o equipamento em estudo seja colocado em funcionamento, e (ii) as possíveis falhas por mortalidade infantil, com origem nas fases de produção e montagem do equipamento, são corrigidas em testes de *burn-in* e, por isso, não são consideradas no modelo. A partir do pressuposto em (i), a utilização de uma distribuição de Weibull com dois parâmetros pode ser considerada como limitação do modelo proposto. Estudos que utilizem uma distribuição de Weibull com três parâmetros podem ser realizados futuramente.

Dessa forma, a função de risco da modelagem proposta inicia apresentando uma taxa de falha c em $t=0$, correspondendo ao parâmetro da distribuição exponencial, crescendo a uma taxa de $d/e \# | e ,^{d @ 1}$, onde d e e correspondem aos parâmetros da distribuição de Weibull. A principal suposição associada ao novo modelo é que causas comuns e especiais de falhas ocorrem simultaneamente e, por isso, nenhum parâmetro de mistura é necessário no modelo.

A análise aqui proposta é ilustrada através de um exemplo oriundo da indústria de aparelhos elétricos. No estudo de caso, são analisados dados provenientes do uso de garantia por clientes de um novo modelo de aparelho de ar condicionado produzido por um fabricante do ramo eletro-eletrônico.

As justificativas para realização do trabalho podem ser divididas em práticas e teóricas. O enfoque prático fundamenta-se no fato de que estudos dessa natureza são escassos no Brasil, a despeito do parque industrial brasileiro apresentar um porte considerável e estar em expansão. Estudos desenvolvidos em empresas relacionados à análise de confiabilidade limitam-se, em muitos casos, a identificar as falhas ocorridas e realizar alguma previsão de reposição de peças com monitoramento de entrada e saída de estoques. Considere-se, nesse caso, que a elevada dose de empirismo na tomada de decisão pode submeter empresas a riscos que se transformam em recrudescimento dos custos de produção, manutenção, reposição de peças e perda de competitividade com possível aumento do preço do equipamento final ao comprador. Com relação à justificativa teórica, o estudo dos modos de falha e a possibilidade de estabelecer modelos que prevejam ocorrência de falhas durante o período de garantia é um desafio constante para empresas e pesquisadores.

Numa visão mais tradicional, quando se realiza uma análise estatística para avaliar a confiabilidade de um sistema, costuma-se ajustar os dados a apenas uma distribuição de probabilidade, o que, em muitas situações, não corresponde à realidade, dado que modos de falhas distintos podem se ajustar a distribuições de probabilidade específicas. O uso de uma única distribuição de probabilidade, nestes casos, permite o ajuste de parte dos dados observados e tornar qualquer inferência menos precisa. O modelo proposto neste artigo apresenta um perfil matemático que busca incorporar a análise de modos de falhas predominantes em um modelo que assimile a ocorrência simultânea (concorrência) de dois tipos de distribuições de probabilidade durante o período de garantia. A modelagem de dados por tipo de falha permite conhecer melhor o desempenho do equipamento e gerar um diferencial de mercado para o fabricante o qual poderá planejar adequadamente o período de garantia dos itens comercializados.

O restante do artigo está assim organizado: na Seção 2, apresenta-se uma revisão de literatura sobre modelagem de dados de garantia, com ênfase na apresentação de modelos de probabilidade em que mais de uma distribuição de probabilidade é aplicada. Na Seção 3, são apresentados o modelo proposto, os estimadores de máxima verossimilhança dos seus parâmetros e uma breve discussão sobre a aplicabilidade do modelo. A Seção 4 está dedicada ao estudo de caso em que dados de uso da garantia de aparelhos domésticos de ar-condicionado são analisados. A Seção 5 apresenta a conclusão do artigo.

2 Revisão de literatura

A análise de Confiabilidade é, em sua essência, uma atividade de modelagem estatística. Assim, um pressuposto básico em estudos de Confiabilidade é a existência de dados acerca do desempenho do equipamento de interesse que permitam essa modelagem. Em grande parte das aplicações, esses dados são coletados na forma de tempos-até-falha de unidades do equipamento, podendo ser obtidos simulando a operação das unidades em laboratório ou coletando informações acerca de seu desempenho em campo.

O alto custo frequentemente incorrido na coleta de dados de confiabilidade obtidos em laboratório sob condições controladas costuma limitar a utilização prática desta ferramenta (BLANKS, 1998). Como alternativa aos dados de laboratório, a modelagem de confiabilidade pode ser conduzida a partir de dados de utilização da garantia do equipamento. Tais dados refletem o desempenho em campo das unidades estudadas. Esse tipo de dados pode ser utilizado (i) na revisão ou, mesmo, na determinação precisa do período de garantia a ser oferecido ao equipamento, tendo em vista um custo-alvo; (ii) em projetos de melhoria de equipamentos ou (iii) no balizamento do desenvolvimento de equipamentos novos similares aos já existentes, acerca dos quais dispõe-se de dados de utilização da garantia.

A demanda por métodos estatísticos para analisar dados de garantia surgiu com a crescente capacidade de armazenamento de dados a baixo custo, sendo inicialmente identificada por Hahn e Meeker (1982) e Lawless (1982). Desde então, a modelagem de dados de garantia tem sido abordada em um número relativamente restrito de publicações acadêmicas. Ao mesmo tempo em que tais métodos surgem, cresce a necessidade por métodos simplificados para coletar e analisar dados de garantia, que comecem a ser apontados por Moltoft (1994) e Blanks (1998). Em geral, associa-se o desenvolvimento de tais métodos a um crescente uso de técnicas quantitativas na indústria.

As referências tratando de modelagem de dados de garantia podem ser classificadas em dois grupos. O primeiro relaciona-se (i) à estimação das funções de confiabilidade de dados censurados (tais como dados de garantia), utilizando métodos paramétricos e não-paramétricos, e (ii) à modelagem de dados descritos por modelos mistos (fundamentados em misturas de distribuições de probabilidade). O segundo grupo trata da análise de dados de garantia utilizando modelos paramétricos, mais notoriamente o processo de Poisson. Somente o primeiro grupo de referências será abordado nesta seção, visto que está diretamente relacionado à proposta deste artigo. Como referências características do segundo grupo, destacam-se os trabalhos de Suzuki (1985a, b), Kalbfleisch e Lawless

(1988), Kalbfleisch et al. (1991), Nelson (1988), Wasserman e Sudjianto (1996) e Hu e Lawless (1996a, b).

O tratamento de dados truncados na análise de Confiabilidade foi introduzido por Kaplan e Meier (1958) no contexto da análise de dados oriundos de experimentos médicos. A estimação não-paramétrica da função de confiabilidade a partir de dados truncados foi também explorada por Turnbull (1976), Woodroffe (1985) e Wang (1989); os trabalhos diferem, essencialmente, nos algoritmos propostos para obtenção de estimativas de máxima verossimilhança da função de Confiabilidade.

Robinson (1983), por sua vez, concebeu um algoritmo para estimar a função de confiabilidade a partir de dados com truncamento progressivo, utilizando *bootstrapping* (EFRON e TIBSHIRANI, 1998) e supondo tempos-até-falha seguindo uma distribuição de Weibull com dois parâmetros. Seu procedimento desenvolve-se a partir do trabalho de Lawless (1982), sendo apropriado na modelagem de grandes massas de dados de Confiabilidade. Os desenvolvimentos propostos, todavia, aplicam-se somente a dados completos ou com censura (truncamento) do tipo II (isto é, o tempo até a censura é conhecido e único). Tal pressuposto pode restringir a aplicação do procedimento usando dados de garantia, pois estes apresentam, tipicamente, censura aleatória, já que a utilização da garantia estende-se por vários lotes de produção, com datas de fabricação e comercialização distintas.

Dados descritos por modelos de probabilidade mistos (ou misturados) são apresentados por Leemis (1995) em um contexto onde tempos-até-falha são observados a partir da incidência de diversos modos de falhas, atuando isoladamente ou em conjunto, sobre o equipamento (componente ou sistema) em estudo. Em sua forma mais simplificada, o modelo misto descreve uma situação onde os tempos até falha observados para o equipamento podem ser divididos em m populações baseados em alguma característica do item (por exemplo, o modo de falha dominante); tal modelo é apresentado na eq. (1). Em sua exposição, Leemis (1995) pressupõe ser possível estimar τ_i a partir de amostras de cada população; ou seja, $f(t)$ é determinado por aplicação direta da eq. (1). A determinação de $f(t)$ a partir de dados mistos, como no caso de dados de garantia, não é abordada.

Cacciari et al. (1995) trataram do problema de dados descritos por modelos mistos de Weibull, no contexto de sistemas de isolamento elétrico, onde esse tipo de dados costuma ocorrer com frequência. Sua abordagem utiliza um procedimento numérico iterativo para a determinação dos parâmetros do modelo misto a partir de dados mistos.

Majeske e Herrin (1995) abordaram o tratamento de modelos mistos no contexto da modelagem de dados de garan-

tia. As técnicas utilizadas incluíram análise gráfica e análise da função empírica de risco. Os dados analisados provinham de misturas das distribuições de Weibull e uniforme (contínuas). O procedimento compreendia a modelagem individual das distribuições singulares, pressupondo dados de garantia separados conforme sua distribuição de procedência com base no registro dos modos de falha. Uma vez modeladas as distribuições singulares, procedia-se com a mistura das distribuições usando o modelo na eq. (1). A contribuição dos autores centrava-se em uma série de técnicas gráficas para verificar o ajuste do conjunto completo de dados ao modelo misto obtido. Os autores também propunham um estimador do número futuro de utilizações da garantia baseado em um modelo de Poisson, seguindo desenvolvimentos consolidados em diversos trabalhos compreendidos no segundo grupo de referências mencionadas no início desta seção.

Chan e Meeker (1999) desenvolveram um modelo que combinou duas distribuições de tempos-até-falha relacionadas a falhas prematuras e por desgaste. O modelo foi definido como um GLFP (modelo geral de falhas limitadas - *general limited failure population*). O GLFP pode ser visto como um caso especial de um modelo de riscos concorrentes com duas causas de falhas associadas. O modelo simula a separação dentre os dois tipos de falhas em diferentes tempos, a fim de avaliar o ajuste de dados mais adequado.

Nair et al. (2001) propõem uma abordagem bayesiana para a modelagem de dados mistos no contexto da análise de confiabilidade. O objetivo é obter um modelo que permita descrever dados mistos provenientes das duas fases iniciais, de mortalidade infantil e de vida útil, da curva da banheira. Os autores, todavia, consideraram ser possível obter tal modelo a partir da mistura de duas distribuições exponenciais, apesar de uma mistura de duas distribuições de Weibull ser notoriamente a opção mais adequada. Tal consideração deve-se ao fato de que uma distribuição de Weibull oferece um melhor ajuste a dados oriundos da fase de mortalidade infantil (taxa de falha decrescente) e também pode ser utilizada para ajustar os dados da fase de vida útil (taxa de falha constante), já que a distribuição exponencial é um caso especial da distribuição de Weibull. O curso de ação adotado na análise em Nair et al. (2001) pode ser justificado ao considerar-se que a abordagem bayesiana para a estimação de $f(t)$, na eq. (1), requer a definição de uma distribuição *a priori* para os parâmetros em τ_i . Em casos onde $f_i(t|\tau_i)$ é definida por dois parâmetros, como no caso da distribuição de Weibull, tal distribuição *a priori* é uma distribuição bivariada, o que pode implicar em esforço computacional considerável na determinação da distribuição *a posteriori* de $f(t)$. É importante salientar que a abordagem em Nair et al. (2001) permite a análise de dados mistos nos quais exista uma variável indexadora associada a cada ponto amostral identificando a região da curva da banheira à qual o dado pertenc-

ce. Tal variável, em aplicações práticas, dificilmente estaria disponível a partir de observação empírica, devendo ser estimada a partir de considerações subjetivas posteriores.

Similar à abordagem de Nair et al. (2001), em que duas distribuições exponenciais estão misturadas, Park & Kulasekera (2004) desenvolveram um método inferencial paramétrico para dados mistos. Os autores calcularam estimadores de máxima verossimilhança em um modelo exponencial. Eles utilizaram dados estratificados em grupos, de acordo com o modo de falha dominante, e propuseram, posteriormente, procedimentos de ajuste gráficos.

Jiang e Murthy (1995) desenvolveram estudos com associações de duas distribuições de Weibull com dois e três parâmetros utilizando modelos definidos como seccionais, multiplicativos e de riscos concorrentes, que consideram a ocorrência de falhas em período de tempo simultâneo ou seqüencial. O ponto forte desses estudos está na elaboração de gráficos de papéis de probabilidade, os WPP (*Weibull Probability Plots*), como possibilidade para analisar o ajuste de dados e definir parâmetros das distribuições utilizadas. Zhang e Ren (2002) complementaram a análise de Jiang e Murthy (1995) e estabeleceram WPP para modelos em que três distribuições de Weibull são combinadas. Essas análises gráficas permitem decidir sobre um ajuste exequível de dados, porém tais estudos reconhecem a necessidade de se fazer análises similares em modelos cujas distribuições não são exclusivamente de Weibull, assim como haver a necessidade de realizar estudos com métodos estatísticos mais refinados para ratificar os valores dos parâmetros obtidos graficamente.

Mais recentemente, Majeske (2003) gerou um modelo misto Weibull-uniforme com o intuito de separar a ocorrência de modos de falhas relacionados a problemas de produção e montagem daqueles decorrentes de uso comum do equipamento. Chukova et al. (2004) analisaram dados de garantia originados de equipamentos submetidos a reparos imperfeitos. O conjunto de dados analisado era composto por tempos-até-falha de equipamentos (i) não submetidos a qualquer reparo e (ii) submetidos a reparos imperfeitos. Finalmente, Attardi et al. (2005) criaram um modelo misto de regressão para distribuição de Weibull, que foi utilizado na modelagem de dados de garantia de automóveis.

A solução dos modelos apresentados anteriormente está vinculada a análises matemáticas e/ou gráficas. Essas duas formas de análise propiciam obter valores para os parâmetros das distribuições de falhas em estudo, sejam elas individuais ou mistas. Leemis (1995), por exemplo, utiliza um modelo matemático de mistura em que os modos de falhas necessitam de estar bem caracterizados no contexto global de modos de falhas, sob o risco do modelo gerar previsões de confiabilidade pouco realistas. A

mesma estratégia de análise é utilizada por Majeske (1995; 2003). Nesses trabalhos, como complemento aos estudos de Leemis, métodos gráficos servem de subsídio para adequação da expressão matemática de cada modo de falha identificado. Já o modelo matemático de Chan e Meeker (1999) busca captar uma situação real de ocorrência de falhas concorrentes em que, apesar de não ficar expressa a melhor alternativa de separação dos modos de falhas considerados nos modelos, há possibilidade de se realizar simulação quanto a pontos limítrofes entre os dois modos de falhas considerados.

As análises gráficas de Jiang e Murthy (1995) e Zhang e Ren (2002) oferecem subsídios para a obtenção de parâmetros em situações onde a solução analítica apresenta um alto nível de complexidade e requeira a realização de vários métodos iterativos de cálculo; ou seja, a análise gráfica é recomendada em situações onde houver necessidade de apurar vários parâmetros independentes em um mesmo modelo. Nesse sentido, a solução gráfica auxilia na obtenção de possíveis valores de parâmetros, mesmo que uma margem de erro esteja implícita na obtenção de solução.

3 Modelo proposto

O modelo apresentado neste artigo segue o perfil de riscos (falhas) concorrentes. Em um modelo de riscos concorrentes considera-se que um equipamento pode falhar devido a uma de várias falhas. A falha no equipamento pode ocorrer devido a um de dois modos de falhas concorrentes. O tempo-até-falha (T_1) devido à causa 1 é distribuído de acordo com $F_1(t)$. Analogamente, o tempo-até-falha (T_2) devido à causa 2 é distribuído de acordo com $F_2(t)$. Em ambos os casos, $F_i(t) \in [0, 1]$. A falha do equipamento é apontada pelo mínimo de T_1, T_2 e, como resultado, $R(t)$ é dada por (JIANG; MURTHY, 1995):

$$R(t) = 1 - F_1(t) - F_2(t) + F_1(t)F_2(t) \quad (2)$$

O desenvolvimento do modelo foi motivado pela necessidade de análise de um conjunto de dados reais originados de solicitações de uso de garantia de aparelhos elétricos domésticos. Os períodos de garantia do equipamento estudado cobrem tipicamente de 10 a 20% do seu tempo médio até a falha e a maioria das falhas está relacionada às últimas duas fases da curva da banheira, isto é, a fase de vida operacional e a fase do desgaste (envelhecimento do equipamento), considerando que tais aparelhos são expostos a procedimentos de *burn-in* (mortalidade infantil) antes de alcançar os clientes finais. A partir de um modelo fundamentado em uma combinação das distribuições de Weibull e exponencial, e a suposição de modos de falhas concorrentes, torna-se possível realizar o ajuste para dados de garantia com as características acima.

A função de confiabilidade do modelo aqui proposto é dada por:

$$R(t) = e^{-\alpha (t/e)^d} \tag{3}$$

onde $R(t)$ é a função de confiabilidade, t é o tempo-até-falha ($t > 0$), e , d e α são parâmetros a serem estimados.

As demais funções associadas ao modelo de confiabilidade da eq. (3) são apresentadas na seqüência. A função de distribuição $F(t)$ é dada por:

$$F(t) = 1 - R(t) = 1 - e^{-\alpha (t/e)^d} \tag{4}$$

A função densidade de probabilidade é dada por:

$$f(t) = \frac{d}{e} \alpha (t/e)^{d-1} e^{-\alpha (t/e)^d} \tag{5}$$

É facilmente demonstrável que a eq. (5) satisfaz as seguintes propriedades de uma função de densidade de probabilidade (MOOD et al., 1974): (i) $\int_0^{\infty} f(x) dx = 1$, e (iii) $\int_0^{\infty} x f(x) dx = \int_0^{\infty} x^2 f(x) dx = \dots$

A função de risco $h(t)$ é dada por:

$$h(t) = \frac{f(t)}{R(t)} = \frac{d}{e} \alpha (t/e)^{d-1} e^{\alpha (t/e)^d} \tag{6}$$

A eq. (6) é apresentada na Figura 1 para $\alpha = 0,05$, $d = 1,7$ e $e = 100$. Observa-se que para este conjunto de parâmetros, $h(t)$ é uma IFR (*increasing failure rate* - taxa de risco crescente), começando em α quando $t = 0$ e aumentando à taxa de $d/e \cdot t^{d-1} e^{\alpha (t/e)^d}$. Além disso, há uma clara associação entre α e o parâmetro exponencial, e entre d e e e os parâmetros de uma distribuição de Weibull.

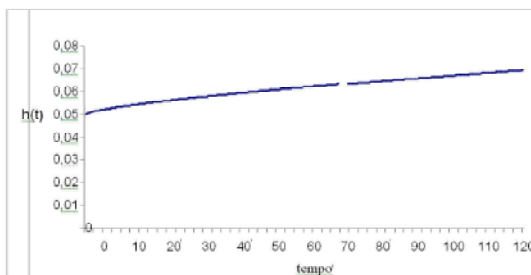


Figura 1 - Gráfico da função de risco quando $\alpha = 0,05$, $d = 1,7$, $e = 100$

3.1 Estimação dos parâmetros

O desenvolvimento a seguir apresenta o cálculo dos estimadores de máxima verossimilhança dos parâmetros da eq. (5), considerando um conjunto de dados censurados. Sejam $T_{1:5}, T_n$ variáveis aleatórias seguindo uma distribuição de probabilidade com densidade $f(t, t)$, onde t é um parâmetro desconhecido. Na exposição que segue, consideram-se r tempos-até-falha observados e indicados por $t_{1:5}, t_r$, e $(n-r)$ tempos-até-falha censurados, indicados por $t_{1:5}^*, t_{n0r}^*$. As expressões são desenvolvidas considerando censura à direita, isto é, $t_i \leq t_{n0r}^* \leq t_0$. A amostra é constituída de n unidades (na ausência de censura, $r = n$). A função de verossimilhança associada à amostra de dados é dada por:

$$L(t) = \prod_{i=1}^r f(t_i, t) \prod_{i=1}^{n0r} R(t_i^*, t) \tag{7}$$

onde $R(t_i^*, t)$ designa a função confiabilidade com o parâmetro t avaliado no tempo t_i^* . O logaritmo da eq. (7) é dado por:

$$l(t) = \sum_{i=1}^r \ln f(t_i, t) + \sum_{i=1}^{n0r} \ln R(t_i^*, t) \tag{8}$$

sendo que essa expressão é utilizada para deduzir as demais expressões apresentadas a seguir. Substituindo as equações (5) e (3) na equação (8), obtém-se:

$$l(\alpha, d, e) = r \ln \alpha + \sum_{i=1}^r \ln \left[\frac{d}{e} \alpha (t_i/e)^{d-1} e^{-\alpha (t_i/e)^d} \right] + \sum_{i=1}^{n0r} \ln \left[e^{-\alpha (t_i^*/e)^d} \right] \tag{9}$$

Os estimadores de α , d e e dados nas equações (10), (11) e (12), respectivamente, são obtidos através do cálculo das derivadas da eq. (9) com relação a cada parâmetro e igualando o resultado de cada uma delas a zero. Exceto por α , não há uma forma direta de cálculo dos parâmetros, e as estimativas de d e e são determinadas utilizando métodos numéricos.

$$\hat{\alpha} = \frac{r}{\sum_{i=1}^r t_i^d + \sum_{i=1}^{n0r} t_i^{*d}} \tag{10}$$

$$\frac{\partial l}{\partial d} = \sum_{i=1}^r \frac{t_i^{d-1}}{t_i^d + e^d} + \sum_{i=1}^{n0r} \frac{t_i^{*d-1}}{t_i^{*d} + e^d} \ln e \tag{11}$$

$$\frac{\partial l}{\partial e} = \frac{r}{e} \ln \left[\frac{1}{\sum_{i=1}^r t_i^d + \sum_{i=1}^{n0r} t_i^{*d}} \right] \tag{12}$$

Os componentes podem ser agrupados em subsistemas. Um exemplo é o subsistema evaporador, apresentado na Figura 2, constituído pelo evaporador, caracol e concha. Esse subsistema trabalha como uma unidade integrada. A falha de um componente reflete imediatamente na funcionalidade dos demais.

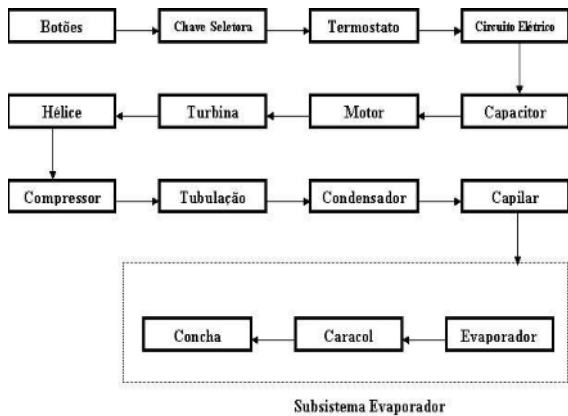


Figura 2 - Diagrama de blocos do aparelho de AC

As principais causas de falhas do AC correspondem às seguintes situações: (i) condições climáticas desfavoráveis, como elevado nível de umidade no ar, o que pode gerar acumulação de água dentro do equipamento, (ii) variabilidade da tensão na rede elétrica domiciliar da região onde o AC está instalado, (iii) uso intensivo do AC, acelerando o uso dos componentes e sua respectiva degradação, e (iv) instalação e manuseio incorreto do AC.

Os modos de falha apontados pelo fabricante correspondem a falhas críticas, isto é, falhas que exigem imediata correção para que a função essencial do AC seja recuperada. Para este estudo, o modelo de AC teve produção e ocorrência de falhas acompanhadas durante vários anos. No período, 94.039 unidades de AC foram produzidas e 3.759 falhas registradas (4%). Isto é, os dados observados são altamente censurados e os aparelhos são postos em operação em diferentes pontos do tempo.

No AC foram identificados 93 modos de falhas nos 15 componentes. Os modos de falha da turbina, objeto de análise neste artigo, são apresentados na Figura 4. A turbina representa 4 % do total de falhas do AC (Figura 3). Os dados analisados neste artigo (Tabela 2) referem-se a um lote onde o componente turbina estava sob suspeita de desempenho inferior. Nesses dados de falha da turbina foram captadas 70% de falhas (144 dados) e 30% de censura (66 dados).

Os modos de falhas são acompanhados pelo Departamento de pós-vendas do fabricante. A utilização da garantia dos equipamentos é monitorada através da reposição de componentes falhos. Junto à informação de reposição anotam-se, ainda, o número de série, semana de produção, código do modo de falha, mês e semana de

ocorrência de cada falha (Tabela 1). Para facilitar a análise dos dados coletados e estabelecer um padrão de avaliação, todas as falhas foram consideradas como ocorrendo na segunda semana do respectivo mês em que o componente apresentou defeito. Com isto, a unidade de tempo considerada neste artigo pôde ser unificada para semana. Para obter o TTF (tempo-até-falha) de cada componente, subtraíram-se os valores na Figura 5 referentes à "semana de falha" e "semana". Não há distinção para o tempo de prateleira. Sabe-se o dia da produção, mas não se identifica quando o equipamento começa a funcionar após a respectiva venda. Porém, é sabido que o tempo de prateleira não é superior a 4 semanas após a produção.

De acordo com o fabricante, alguns componentes do AC, tais como motor, turbina e compressor, começam a se degradar tão logo a sua vida operacional inicie. Por essa razão, falhas devidas à degradação e causas aleatórias competem durante a vida operacional do componente.

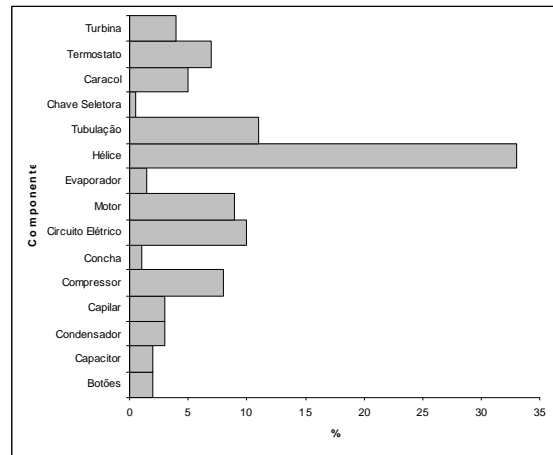


Figura 3 - Percentual de falhas por componente do AC

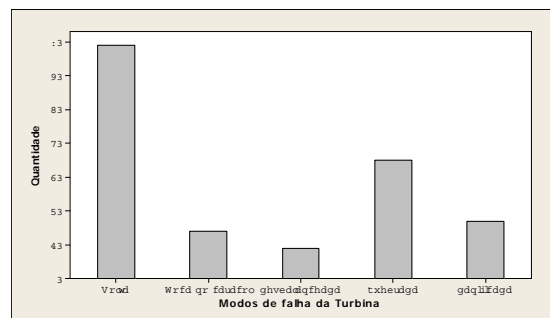


Figura 4 - Modos de falha da Turbina

Turbina (ordem)	Item (nº de série)	Semana de produção	Semana	Modo de falha (código)	Mês de falha	Semana de falha	Tempo-até-falha (TTF)
1	901	01/98	1	444	abr/98	16	15
2	981	01/98	1	449	abr/99	68	67
3	911	01/98	1	446	mai/99	72	71
4	900	01/98	1	446	Jun/99	77	76
5	959	01/98	1	449	Jun/99	77	76

Figura 5 - Banco de dados (exemplo de registros)

O conjunto de dados de tempos-até-falha apresentado na Figura 6 foi coletado de um lote de produção de 205 itens montados utilizando uma turbina defeituosa e deram suporte à análise inferencial subsequente. Há 61 dados censurados não apresentados na tabela; o tempo de censura é de 79 semanas, correspondente ao final do período de garantia de cada AC.

4	9	11	12	13	14	17	18	19	22	23	26	28	35	42	49	57	66
6	10	11	12	13	14	17	18	19	22	23	26	30	35	43	51	58	67
7	10	11	12	13	14	17	18	19	22	23	26	31	35	44	52	58	67
8	10	11	12	13	15	17	18	19	22	23	26	32	36	44	53	60	69
8	10	12	12	13	15	17	18	20	22	23	26	33	37	45	53	60	69
9	10	12	12	13	16	17	18	21	22	24	27	33	37	45	54	60	71
9	11	12	13	14	16	17	18	21	23	25	28	34	38	47	54	64	72
9	11	12	13	14	17	17	19	22	23	25	28	34	40	47	54	66	74

Figura 6 - Tempos-até-falha da turbina

Uma primeira apreciação do desempenho da turbina é apresentada na Figura 7 através de um histograma de frequências de tempos-até-falha.

Considerando a informação do comportamento de degradação do componente fornecida pelo fabricante, testou-se o ajuste do modelo apresentado na Seção 3 aos dados analisados. Os resultados são apresentados a seguir.

Utilizando as equações de máxima verossimilhança [eqs. (10) a (12)] e um algoritmo comercial de otimização não-linear, foram obtidas as seguintes estimativas para os parâmetros do modelo de confiabilidade proposto: $c = 0,004109$, $d = 1,145$ e $e = 66,03$. O algoritmo elaborado demandou a realização de cálculos numéricos iterativos para otimizar uma função objetivo que minimizasse as diferenças entre as parcelas das eqs.(10) a (12), atendendo à condição $\frac{\partial l}{\partial c} \Big|_{c=e} = 0$ mencionada na Seção 4.

A Figura 7 apresenta o histograma (dados observados) de tempos-até-falha da turbina, enquanto que a Figura 8 apresenta a função densidade de probabilidade ajustada para as falhas da turbina utilizando o modelo proposto na Seção 3. Observa-se que o histograma apresenta um pico quando $t > 79$ semanas, correspondendo aos dados censurados, que foram considerados para obter as estimativas dos parâmetros apresentados anteriormente.

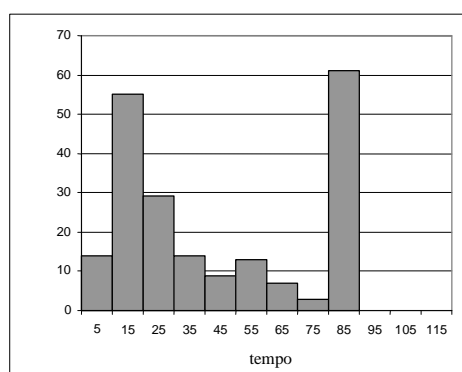


Figura 7 - Histograma de frequências dos tempos-até-falha da turbina

O teste de Kolmogorov-Smirnov (teste K-S) é útil para comparar uma função de distribuição empírica a uma função de distribuição teórica. Trata-se de um dos métodos não-paramétricos mais úteis e genéricos para comparar duas distribuições acumuladas por ser sensível a diferenças de localização e forma entre as distribuições teórica e empírica. O teste K-S foi utilizado neste artigo para testar o ajuste dos dados na Figura 6 às distribuições de Weibull e lognormal, bem como para o modelo proposto. Para verificar o ajuste dos dados às distribuições hipotetizadas, o teste K-S utiliza a máxima distância (D) entre as distribuições acumuladas empírica e teórica. Apresentando um D , o modelo proposto ajusta-se melhor aos dados que os modelos das distribuições de Weibull e Lognormal, para os quais D foi igual a 0,1255 e 0,1236, respectivamente.

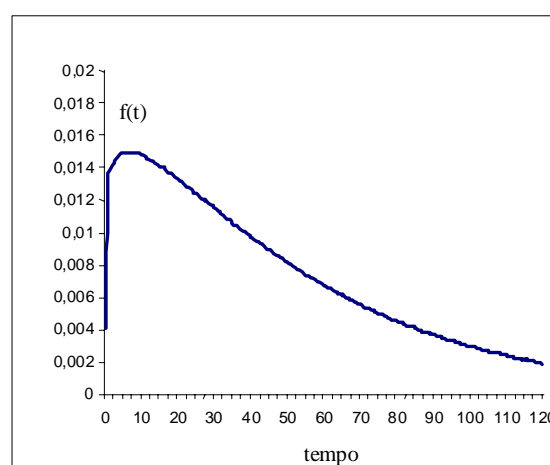


Figura 8 - Função densidade de probabilidade dos tempos-até-falha da turbina

5 Conclusão

Falhas incorridas durante o período do equipamento implicam custos indesejáveis. Por essa razão, a incidência de falhas durante a garantia, bem como a sua natureza, são normalmente registradas pelas empresas, na busca de informações que permitam aprimorar o projeto de equipamentos. Entretanto, pesquisas acerca da modelagem de dados de garantia são recentes e abordam um conjunto relativamente limitado de características desses dados.

Este artigo propõe a aplicação de um modelo de riscos concorrentes para ajustar dados de garantia fundamentando-se na utilização das distribuições exponencial e de Weibull. O modelo é aplicado para um conjunto de dados censurados de um sistema doméstico de ar condicionado. Falhas no equipamento são devidas a causas aleatórias e desgaste dos seus componentes. O modelo proposto apresentou um bom ajuste aos dados como pôde ser verificado gráfica e analiticamente.

Futuras pesquisas sobre o modelo misto de confiabilidade apresentado devem detalhar a utilização de testes de ajustamento.

tes gráficos e analíticos, assim como viabilizar a inclusão de um parâmetro de localização no modelo, no sentido de responder por situações em que causas aleatórias e de desgaste de componentes não se sobreponham no tempo.

Referências

- ATTARDI, L.; GUIDA, M. & PULCINI, G. A mixed-Weibull regression model for the analysis of automotive warranty data. *Reliability Engineering & System Safety*, Vol. 87, p. 265-273, 2005.
- BLANKS, H.S. The challenge of quantitative reliability. *Quality & Reliability Engineering International*, Vol. 14, p.167-176, 1998.
- CACCIARI, M.; CONTIN, A. & MONTANARI, G.C. Use of mixed-weibull distribution for the identification of PD phenomena. *IEEE Transactions on Dielect. Electr. Insul.*, Vol. 2, p.1166-1179, 1995.
- CHAN, V.; MEEKER, W. A failure-time model for infant-mortality and wearout failure modes. *IEEE Transactions on Reliability*, Vol. 48, nº 4, p.377-387, 1999.
- CHUKOVA, S.; ARNOLD, R. & WANG, D. Warranty analysis: an approach to modelling imperfect repair. *International Journal of Production Economics*, Vol. 89, p.57-68, 2004.
- EFRON, B.; TIBSHIRANI, R.J. *An introduction to the bootstrap*. New York: Chapman-Hall, 1998, 436p.
- HAHN, G.J.; MEEKER, W.Q. Pitfalls and practical considerations in product life analysis, Part I. *Journal of Quality Technology*, 14 (3), 144-185, 1982.
- HU, X.J. & LAWLESS, J.F. Estimation from truncated lifetime data with supplementary information on covariates and censoring times. *Biometrika*, Vol.83, p.747-761, 1996b.
- HU, X.J. & LAWLESS, J.F. Estimation of rate and mean functions from truncated recurrent event data. *Journal of the American Statistical Association*, Vol. 91, p.300-310, 1996a.
- JIANG R.; MURTHY, D. N. P. Reliability modeling involving two Weibull distributions. *Reliability Engineering and System Safety*, Vol. 47, p. 187-198, 1995.
- KALBFLEISCH, J.D. & LAWLESS, J.F. Estimation of reliability in field-performance studies. *Technometrics*, Vol. 30, p.365-388, 1988.
- KALBFLEISCH, J.D.; LAWLESS, J.F. & ROBINSON, J.A. Methods for the analysis and prediction of warranty claims. *Technometrics*, Vol. 33, p.273-285, 1991.
- KAPLAN, E.L. & MEIER, P. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, p.459-481, June 1958.
- LAWLESS, J.F. *Statistical models and methods for lifetime data*. New York: John Wiley, 1982.
- LEEMIS, L.M. *Reliability - Probabilistic models and statistical methods*. Englewood Cliffs: Prentice-Hall, 1995.
- MAJESKE, K. D. A mixture model for automobile warranty data. *Reliability Engineering and System Safety*, Vol. 81, p. 71-77, 2003.
- MAJESKE, K.D. & HERRIN, G.D. Assessing mixture-model goodness-of-fit with an application to automobile warranty data. *Proceedings of the Reliability and Maintainability Symposium*, 378-383, 1995.
- MOLTOFT, J. Reliability engineering based on field information - the way ahead. *Quality & Reliability Engineering International*, Vol. 10, p.399-409, 1994.
- MOOD, A.M.; GRAYBILL, F.A. & BOES, D.C. *Introduction to the theory of statistics*. New York: McGraw-Hill, 1974.
- NAIR, V.N.; TANG, B. & XU, L. Bayesian inference for some mixture problems in quality and reliability. *Journal of Quality Technology*, Vol.33, p.16-28, 2001.
- NELSON, W. Graphical analysis of system repair data. *Journal of Quality Technology*, Vol. 20, p.24-35, 1988.
- PARK, C. & KULASEKERA, K. B. Parametric inference of incomplete data with competing risks among several groups. *IEEE Transactions on Reliability*, Vol. 53, p.11-21, 2004.
- RAUSAND, M.; HØYLAND, A. *System reliability theory: models, statistical methods, and applications*. 2ª Ed., New York: John Wiley & Sons, 2004.
- ROBINSON, J.A. Bootstrap confidence intervals in location-scale models with progressive censoring. *Technometrics*, Vol. 25, p.179-187, 1983.
- SUZUKI, K. Estimation of lifetime parameters from incomplete field data. *Technometrics*, Vol. 27, p.263-271, 1985b.
- SUZUKI, K. Nonparametric estimation of lifetime distributions from a record of failures and follow-ups. *Journal of the American Statistical Association*, Vol. 80, p.68-72, 1985a.

TURNBULL, B.W. The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of the Royal Statistical Society, Ser.B.*, Vol. 38, p.290-295, 1976.

WANG, M. A semiparametric model for randomly truncated data. *Journal of the American Statistical Association*, Vol. 84, p.742-748, 1989.

WASSERMAN, G.S. & SUDJIANTO, A. A comparison of three strategies for forecasting warranty claims. *IIE Transactions*, Vol. 28, p.967-977, 1996.

WOODROOFE, M. Estimating a distribution function with truncated data. *The Annals of Statistics*, Vol. 13, p.163-177, 1985.

ZHANG, T.; REN, Y. Failure data analysis by models involving 3 Weibull distributions. *Proceedings Annual Reliability and Maintainability Symposium*, p. 44-50, 2002.