

Uma aplicação ao setor de transporte agrícola da análise de confiabilidade

An application to the agricultural transport sector of the reliability analysis

Una aplicación al sector del transporte agrícola del análisis de confiabilidad

Recebido: 08/01/2021 | Revisado: 10/01/2021 | Aceito: 14/01/2021 | Publicado: 18/01/2021

Isaac Pereira Santos

ORCID: <https://orcid.org/0000-0001-9006-3440>
Escola Cidadã Integral e Técnica Estadual Ezequiel Fernandes, Brasil
E-mail: isaac20.santos@gmail.com

Pablo Lourenço Ribeiro de Almeida

ORCID: <https://orcid.org/0000-0002-5945-562X>
Universidade Federal Rural de Pernambuco, Brasil
E-mail: pabloestatistico@gmail.com

Deise Pereira da Silva

ORCID: <https://orcid.org/0000-0002-9263-1234>
Escola Cidadã Integral Teonas da Cunha Cavalcante, Brasil
E-mail: deysmartins@live.com

Cleanderson Romualdo Fidelis

ORCID: <https://orcid.org/0000-0002-3726-5833>
Escola Superior de Agricultura Luiz de Queiroz, Brasil
E-mail: cleanderson@usp.br

Elias Silva Medeiros

ORCID: <https://orcid.org/0000-0002-9694-4019>
Universidade Federal dos Grandes Dourados, Brasil
E-mail: eliasmedeiros@ufgd.edu.br

Tiago Almeida de Oliveira

ORCID: <https://orcid.org/0000-0003-4147-7721>
Universidade Estadual da Paraíba, Brasil
E-mail: tadolive@servidor.uepb.edu.br

Resumo

A análise de confiabilidade é a conversão de técnicas de análise de sobrevivência aplicadas no departamento de produção. Para realizar a análise de confiabilidade, o conjunto de dados em estudo deve atender às condições necessárias para que seja possível fazer estimativas razoáveis das funções de risco e confiabilidade. O objetivo deste trabalho é utilizar a análise de confiabilidade, utilizando as técnicas não paramétricas, semi-paramétricas e paramétricas, para explorar e modelar o tempo de do pneu 11.00R22. Foi realizado um estudo com os 11.00R22, e o tempo em quilômetros rodados foi calculado por meio de um computador de bordo acoplado em caminhões. Os conjuntos de dados têm um total de 552 pneus, os quais apresentaram diferença entre os três tipos de tempo de vida. Na execução deste trabalho, aplicou-se o estimador limite produto de Kaplan-Meier para os três grupos de vida, e o teste log-rank para verificar a existência de diferença significativa entre as curvas de sobrevivência ambos métodos não paramétricos, e em sequência, a abordagem paramétrica com o uso do modelo de regressão para verificar qual distribuição se adequava ao tempo de vida dos pneus e ainda o modelo de Cox para modelar o risco abordagem semi-paramétrica. Na aplicação dos métodos foi utilizado o software R por meio do pacote survival. O modelo Weibull foi mais adequado para modelar o tempo de vida dos pneus 11.00R22.

Palavras-chave: Modelos paramétricos; Modelos de riscos proporcionais; Análise de sobrevivência.

Abstract

Reliability analysis is the conversion of survival analysis techniques applied in the production department. To perform the reliability analysis, the data set under study must meet the necessary conditions so that it is possible to make reasonable estimates of the risk and reliability functions. The objective of this work is to use the reliability analysis, using non-parametric, semi-parametric and parametric techniques, to explore and model the time of the 11.00R22 tire. A study was carried out with the 11.00R22, and the time in kilometers traveled was calculated using an on-board computer attached to trucks. The data sets have a total of 552 tires, which differed between the three types of lifespan. In carrying out this work, the Kaplan-Meier product limit estimator was applied for the three life groups, and the log-rank test to verify the existence of a significant difference between the survival curves, both non-parametric methods, and in sequence, the parametric approach with the use of the regression model to verify which distribution suited the tire life and also the Cox model to model the risk semi-parametric approach. In the application of the methods, software R was used through the survival package. The Weibull model was best suited to model the life span of 11.00R22 tires.

Keywords: Parametric models; Proportional risk models; Survival analysis.

Resumen

El análisis de confiabilidad es la conversión de técnicas de análisis de supervivencia aplicadas en el departamento de producción. Para realizar el análisis de confiabilidad, el conjunto de datos en estudio debe reunir las condiciones necesarias para que sea posible realizar estimaciones razonables de las funciones de riesgo y confiabilidad. El objetivo de este trabajo es utilizar el análisis de confiabilidad, utilizando técnicas no paramétricas, semiparamétricas y paramétricas, para explorar y modelar el tiempo del neumático 11.00R22. Se realizó un estudio con el 11.00R22, y se calculó el tiempo en kilómetros recorridos mediante una computadora de a bordo acoplada a camiones. Los conjuntos de datos tienen un total de 552 neumáticos, que difieren entre los tres tipos de vida útil. Para la realización de este trabajo se aplicó el estimador límite de producto de Kaplan-Meier para los tres grupos de vida, y la prueba de log-rank para verificar la existencia de una diferencia significativa entre las curvas de supervivencia, ambos métodos no paramétricos, y en secuencia, el enfoque paramétrico con el uso del modelo de regresión para verificar qué distribución se adapta a la vida útil del neumático y también el modelo de Cox para modelar el enfoque semiparamétrico de riesgo. En la aplicación de los métodos se utilizó el software R a través del paquete de supervivencia. El modelo Weibull era el más adecuado para modelar la vida útil de los neumáticos 11.00R22.

Palabras clave: Modelos paramétricos; Modelos de riesgo proporcional; Análise de la supervivencia.

1. Introdução

No Brasil, a maior concentração na matriz de transporte encontra-se no modal rodoviário. O Brasil detém a maior participação na matriz do transporte de cargas no Brasil, 61% da movimentação de mercadorias e 95% de passageiros (CNT, 2019). As condições das estradas impactam consideravelmente os custos operacionais, elevando os gastos com o consumo de combustível, manutenção dos veículos, troca de diversas peças, óleo do motor e, em especial, com pneus.

Em relação aos gastos descritos anteriormente, depois dos custos variáveis com combustível, os transportes de carga têm o maior destaque relativo aos custos dos pneus (Oliveira & Melo, 2019). Para a Associação Brasileira do Segmento de Reforma de Pneus (ABR), no setor de transportes, o Brasil é o país que mais reforma pneus no mundo, devido ao seu baixo custo, reduzindo em até 57% o custo/km, podendo reformar cada pneu em média duas vezes, estendendo sua vida útil.

Com os problemas decorrentes, surge a necessidade de estudar métodos que contribuam para a redução na probabilidade de falhas em pneus, o que resulta numa ênfase crescente na confiabilidade. Nessa perspectiva, é importante utilizar ferramentas que descrevem o comportamento de possíveis falhas, aliados a uma metodologia capaz de obter e tratar essas informações, sendo viável adotar uma análise de confiabilidade.

Geralmente, confiabilidade pode ser definida como a operação bem-sucedida de um projeto, unidade ou sistema dentro de um determinado período de tempo sob condições de operação estabelecidas sem qualquer interrupção ou falha (Fogliatto & Ribeiro, 2009). No entanto, Santos (2016) apontou que a análise de confiabilidade precisa ser definida quantitativamente com base na probabilidade. Algumas distribuições estatísticas podem ser utilizadas para indicar a probabilidade de falha, de forma que seja possível determinar quando o desempenho do item, unidade ou sistema analisado está abaixo do nível aceitável indicado, para que a ocorrência de falha possa ser considerada (Haviasas, 2005; Leal, 2018).

Diante do exposto, o presente trabalho tem por objetivo modelar a durabilidade de uma marca comercial de pneus especificação 11.00R22, por meio de uma análise estatística de confiabilidade, se utilizando de técnicas não paramétricas, paramétricas e semi-paramétricas para tanto.

2. Revisão de Literatura

Estudos envolvendo análise de confiabilidade dependem das informações do tempo até a ocorrência do evento, dos tempos censurados, de uma clara definição de origem do estudo (Silva et al., 2015). A comparação entre os indivíduos pode ser realizada apenas a partir do tempo observado até a ocorrência para cada um destes e, em alguns casos, pode haver discrepâncias entre estes tempos, estas medidas a partir de covariáveis que foram mensuradas nestes indivíduos.

A observação do tempo até a ocorrência da falha do objeto é considerada a partir do momento que este começa a ser acompanhado. Logo, o tempo inicial de todo objeto é definido a partir de $t = 0$. Após isso, o tempo transcorrido entre o tempo inicial e o tempo observado até a falha é mensurado. Esses tempos observados são considerados aleatórios e define-se uma variável aleatória T com o objetivo de captar essas variações. Assim, seja uma variável aleatória T , continuamente distribuída no intervalo $[0, t)$, caracterizada por uma função densidade de probabilidade dada por $f(t)$ e função de distribuição acumulada dada por:

$$F(t) = P(T \leq t) = \int_0^t f(u) du, \quad t > 0 \quad (1)$$

A função de distribuição acumulada, apresenta a probabilidade de uma unidade observada falhar em um tempo menor que t . Por definição, temos que

$$f(t) = F'(t) = \frac{d}{dt} F(t) = \frac{F(t + \Delta t) - F(t)}{\Delta t} = \frac{P(t < T \leq t + \Delta t)}{\Delta t}$$

2.1 Função de confiabilidade, $R(t)$

A fim de calcular a função de confiabilidade, vamos supor que n observações foram realizadas em objetos distintos sob condições idênticas. A função de confiabilidade é encontrada a partir da probabilidade acumulada de sucesso, definida por

$$R(t) = \frac{n_s(t)}{n_s(t) + n_f(t)} = \frac{n_s(t)}{n_0} \quad (3)$$

onde n_s é o número de unidades que não falharam, n_f é o número de unidades que falharam, dentre as n unidades observadas, uma vez que $n_s + n_f = n$, estas observadas em um intervalo de tempo $(t - \Delta t, t)$, sendo Δt a variação em relação ao tempo inicial $t = 0$.

Assim, a partir da definição acima, podemos definir a função de confiabilidade em um tempo t , sob a variável aleatória contínua T , da seguinte forma:

$$R(t) = P(T > t). \quad (4)$$

Logo, podemos concluir que a função de confiabilidade (sobrevivência), $R(t)$, remete a probabilidade que uma unidade observada não falha antes do tempo t , considerando o intervalo dado por $(0, t)$.

2.2 Função de risco $h(t)$

A probabilidade do objeto falhar entre $(t; t + \Delta t]$ pode ser mensurada a partir da função de risco. Supondo que o objeto está funcionando no tempo t observado, a função de risco será dada por

$$h(t) = \frac{R(t) - R(t + \Delta t)}{R(t)\Delta t} = \frac{-R'(t)}{R(t)} = \frac{f(t)}{R(t)}, \quad t \geq 0 \quad (5)$$

uma vez que esta pode ser escrita como uma razão entre a função densidade de probabilidade e a função de confiabilidade. De acordo com Ramires (2013), esta função pode ser um indicativo de como o tempo afeta a unidade observada, e também a

quantidade risco sob uma unidade no tempo t . A partir da função de risco, podemos calcular a função de risco acumulada, dada por

$$H(t) = \int_0^t h(u)du, \quad t \geq 0 \quad (6)$$

uma vez que esta função apresenta uma alternativa a utilização da Função dada em (5), mesmo assemelhando-se na sua forma gráfica independentemente da função de probabilidade atribuída para a variável aleatória T (Fogliatto & Ribeiro, 2009).

2.3 Tempo médio até falha, MTTF

O valor esperado para a variável aleatória T , ou seja, o tempo médio até a falha (MTTF) é encontrado a partir

$$MTTF = E(T) = \int_0^{+\infty} tf(t)dt, \quad (7)$$

sendo que uma alternativa a este valor é definida por

$$MTTF = \int_0^{\infty} R(t)dt. \quad (8)$$

onde a Função (8) pode apresentar facilidades no cálculo se comparado a expressão dada em (7), dependendo da distribuição de probabilidade assumida para T .

2.4 Análise de dados censurados

Análises de estudos de confiabilidade espera-se, pelo menos em teoria, que todos os tempos até a falha dos objetos sejam observados. No entanto, por razões diversas alguns desses tempos até a falha não podem ser observados, que neste caso configura-se um conjunto de dados com informações incompletas. Nestes termos, essas informações incompletas dos tempos até a ocorrência da falha são denominadas de censuras.

A censura é caracterizada como uma observação parcial da resposta, ocorrendo por várias razões, dentre as quais destaca-se: perda do acompanhamento da unidade, tempo até a falha muito extenso, falha em equipamentos subsequentes que este objeto depende, etc. Os mecanismos de censuras podem ser do tipo à direita, esquerda e intervalar. Em nosso caso, estamos assumindo um mecanismo de censura à direita, introduzido a partir da variável

$$\delta = \begin{cases} 1, & \text{quando } T \leq S, \\ 0, & \text{quando } T > S. \end{cases} \quad (9)$$

a qual auxilia na identificação se o tempo observado foi um tempo até a censura ou até a falha.

Dentre as possíveis formas que foram desenvolvidas para analisar dados de estudos de confiabilidade, merece destaque a abordagem que trata dos tempos até a ocorrência do evento de maneira probabilística, incluindo ambas as informações na função de verossimilhança e encontrando parâmetros associados a essas distribuições de probabilidade que são capazes de prover uma razoável caracterização desses dados (Wang, Li & Reddy, 2019).

2.5 Distribuições de probabilidade em confiabilidade

Em estudos envolvendo confiabilidade de alguma unidade a medida mais utilizada é escrita em termos da probabilidade

de esta não falhar até um determinado tempo t . Essa probabilidade é dada em função do estudo dos tempos até a falha de n unidades sob estudo. Assim, uma vez que uma distribuição de probabilidade é adequada para modelar este tipo de evento, a partir desta é possível inferir estimando os parâmetros do modelo e conduzindo a conclusões plausíveis sobre os tempos até a falha de cada conjunto de objetos estudados. Por essa causa, as modelagens dos tempos até a falha dos objetos estudados a partir de distribuições de probabilidade adequadas fornecem uma maneira plausível e confiável para explicar resultados obtidos.

Na literatura, pode-se encontrar várias maneiras de abordar a análise dos tempos até a ocorrência da falha neste tipo de estudo. Uma maneira bastante usual é supor que a variável aleatória T adequa-se a uma distribuição contínua de probabilidade. A seguir serão apresentados alguns dos modelos mais utilizados para modelar esses tipos de eventos.

Se uma variável aleatória contínua T , representando o tempo até a falha dos objetos, segue uma distribuição exponencial com parâmetro λ , a sua função densidade de probabilidade é dada por

$$f(t) = \lambda \exp\{-\lambda t\}, \quad (10)$$
$$t \geq 0 \text{ e } \lambda > 0.$$

e sua função de confiabilidade (sobrevivência) é dada por

$$R(t) = \exp\{-\lambda t\}, \quad t \geq 0 \text{ e } \lambda > 0, \quad (11)$$

com função taxa de falha (risco) dada por

$$h(t) = \lambda, \quad t \geq 0 \text{ e } \lambda > 0. \quad (12)$$

Perceba que a distribuição exponencial possui apenas um parâmetro, modelando dados que possuem variabilidade e valor médio iguais, uma vez que estes são encontrados pelo inverso do parâmetro λ , ou seja, são dados por $\frac{1}{\lambda}$.

Supondo que a variável aleatória T se comporta de acordo com uma distribuição de Weibull, temos que a sua função densidade de probabilidade, função de confiabilidade e função taxa de falha são dadas, respectivamente, por:

$$f(t) = \lambda s t^{s-1} \exp\{-\lambda t^s\} \quad t \geq 0 \text{ e } s, \lambda > 0, \quad (13)$$

$$R(t) = \exp\{-\lambda t^s\} \quad t \geq 0 \text{ e } s, \lambda > 0, \quad (14)$$

e

$$h(t) = \lambda s t^{s-1} \quad t \geq 0 \text{ e } s, \lambda > 0. \quad (15)$$

Note que a distribuição Weibull é um caso mais geral da distribuição exponencial, uma vez que se tomarmos o parâmetro $s = 1$ esta distribuição se resume a distribuição exponencial. A função taxa de falha neste caso pode ser de três formas:

- i. $s < 1 \rightarrow$ a função da taxa de falha decresce,
- ii. $s > 1 \rightarrow$ a função da taxa de falha cresce,
- iii. $s = 1 \rightarrow$ a função da taxa de falha é constante.

Outra distribuição comumente utilizada para descrever os tempos para estudos envolvendo análise de confiabilidade é a distribuição log-normal. Supondo que a variável aleatória T se comporta de acordo com uma distribuição log-normal, a função densidade de probabilidade associada a esta é da seguinte forma:

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\log\log(t) - \mu}{\sigma}\right)^2\right\} \quad (16)$$

$$t > 0, \mu \in \mathbf{R}, \sigma > 0,$$

em que, μ , é a média do logaritmo do tempo até falha, assim como σ é o desvio padrão (Colosimo & Giolo, 2006).

A função de confiabilidade do modelo Log-Normal é dada por:

$$R(t) = \Phi\left(\frac{-\log\log(t) + \mu}{\sigma}\right) \quad t > 0, \mu \in \mathbf{R}, \sigma > 0, \quad (17)$$

e a função taxa de falha por:

$$h(t) = \frac{f(t)}{R(t)} \quad t > 0, \mu \in \mathbf{R}, \sigma > 0, \quad (18)$$

em que $\Phi(\cdot)$, é a função de distribuição acumulada de uma distribuição normal padrão (Lee & Wang, 2003). Por causa da dependência da função de distribuição acumulada de uma distribuição normal padrão, as funções de confiabilidade e taxa de falha não possuem uma forma analítica semelhante as distribuições apresentadas anteriormente, possuindo um comportamento não monótono com um crescimento, atingindo o ponto de máximo e depois decrescendo.

A fim de encontrar as estimativas para os parâmetros associadas a cada distribuição de probabilidade utilizada, podemos utilizar o método da máxima verossimilhança. Se assumirmos um conjunto de dados completo (ausente de censuras), sob condições de regularidade, podemos assumir que a função de verossimilhança pode ser escrita como

$$L(\theta) = \prod_{i=1}^n f(t_i, \theta) \quad (20)$$

A função acima remete a informação de as variáveis T_1, \dots, T_n assumirem os valores t_1, \dots, t_n ; estes valores são contabilizados a partir da função densidade de probabilidade do modelo assumido para a variável aleatória T. Note que a função dada em (20) tem como valor desconhecido o valor θ , sendo a estimativa para este encontrada a partir da maximização de $L(\theta)$. Este processo ocorre derivando $L(\theta)$ em relação a θ , igualando-se a zero a equação resultando e, em seguida isola-se θ .

A função apresentada acima é viável quando se conhece todos os tempos até a ocorrência da falha, que na maioria dos problemas não é algo possível. Assim, dispondo de uma amostra contendo tempos observados até a falha e tempos censurados, temos que a função acima pode ser reescrita como

$$L(\theta) = \prod_{i=1}^r f(t_i, \theta) \prod_{i=1}^{n-r} R(t_i^+, \theta) \quad (21)$$

onde θ representa os parâmetros da distribuição de probabilidade assumida com uma função de confiabilidade dada

por $R(t_i^+, \theta)$, sendo esta avaliada em t_i^+ . A função de confiabilidade dada representa a probabilidade do objeto não falhar por pelo menos um tempo t_i^+ . Na maioria dos casos é mais simples encontrar as estimativas para θ utilizando o logaritmo na função de verossimilhança. Assim, a função resultante será dada por

$$l(\theta) = \sum_{i=1}^r \ln f(t_i, \theta) + \sum_{i=1}^{n-r} \ln R(t_i^+, \theta). \quad (22)$$

2.6 Escolha do modelo paramétrico

Quando se ajusta um modelo paramétrico a conjunto de dados várias suposições são realizadas. Entretanto, se essas suposições realizadas não são atendidas, pode-se recair em inferências e conclusões erradas de mais diversas possibilidades (Herrmann, 2011). Dentre as possibilidades de escolha dos modelos ajustados pode ocasionar de escolher-se entre distribuições de probabilidade para a variável aleatória, dentre um conjunto covariáveis, etc.

Existem vários métodos propostos para avaliar as diversas nuances que foram supostas para a realização do ajuste do modelo aos dados. Destaca-se, porém, o teste da razão de verossimilhanças (TRV). Este teste é utilizado quando se dispõe de modelos aninhados, isto é, modelos que possuem maior número de parâmetros (modelo mais geral) em relação a um que possui menor número de parâmetros (modelo particular), buscando avaliar qual possui melhor ajuste aos dados.

O teste da razão de verossimilhanças é realizado a partir das funções de verossimilhança maximizada sob cada modelo, sendo os casos sem restrição o sob hipótese de nulidade (H_0). A estatística do teste é dada por:

$$TRV = -2 \left[\log [L(\tilde{\theta})] - \log [L(\hat{\theta})] \right] \quad (23)$$

uma vez que $\log [L(\tilde{\theta})]$ é o logaritmo da função de verossimilhança do modelo e $\log [L(\hat{\theta})]$ é o logaritmo da função de verossimilhança sob o modelo completo. Assintoticamente, compara-se o valor encontrado com o valor de tabela de uma distribuição χ^2 com número de graus de liberdades igual à diferença entre os modelos, concluindo em favor ou contra a rejeição dependendo do nível de significância assumido (Carvalho, 2011). Uma forma de se concluir em favor ou contra a hipótese nula é a observação do p-valor comparado a distribuição do nível de significância assumido (Herrmann, 2011).

Uma outra maneira de avaliar o modelo ajustado é a partir de uma análise gráfica a partir das distribuições de probabilidades propostas. Entre as principais contribuições deste tipo de avaliação, esta a possibilidade de comparar de maneira exploratória os modelos ajustados. De maneira que as propostas gráficas de linearizar a função de confiabilidade associada ao modelo e a de comparar os ajustes das curvas de confiabilidade de maneira paramétrica com uma estimação por um método não paramétrico de Kaplan-Meier, por exemplo, recebem destaque na literatura (Hermann, 2011).

De acordo com Carvalho (2011), a utilização da função desvio (deviance) como uma medida global de ajuste é uma das possibilidades para avaliação da qualidade do ajuste do modelo. A ideia é de comparar um modelo mais geral (completo) versus um modelo particular (reduzido), podendo o modelo completo ser ou não igual ao modelo saturado. Assim, prioriza-se os modelos que tiverem menor valor da deviance.

2.7 Modelo de riscos proporcionais

Conhecido como modelo de riscos proporcionais, o modelo formulado por Cox (1972) objetiva descrever uma relação existente entre os tempos até a falha e as covariáveis associadas a estes. Este modelo é do tipo semi-paramétrico e recebe o nome de riscos proporcionais devido supor que existe uma razão constante entre as taxas de riscos entre os tempos observados para os

indivíduos e as covariáveis. Vamos assumir um risco basal $\lambda_0(t)$ e um vetor contendo as covariáveis x . Supondo que o tempo de vida T de um objeto no instante t , temos que a função de risco $\lambda(t|x)$ é dada por

$$\lambda(t|x) = \lambda_0(t) \exp(x_1\beta_1 + x_2\beta_2 + \dots + x_p\beta_p) = \lambda_0(t) \exp(x\beta),$$

em que x é o vetor de covariáveis associadas aos tempos até a ocorrência da falha e os parâmetros $\beta = (\beta_1, \beta_2, \dots, \beta_p)$ são os efeitos associados a cada efeito dessas covariáveis.

3. Metodologia

A amostra aqui estudada é composta dos tempos até a interrupção ou recauchutagem de 552 pneus com especificação 11.00R22 de caminhões transportadores de safras, sendo estes avaliados em três vidas úteis. Este tempo até a recauchutagem ou interrupção para a troca foi medido em quilômetros através de um computador de bordo instalado nos caminhões. Sendo assim, trata-se de um estudo quantitativo longitudinal de carácter prospectivo no contexto de análise de confiabilidade, pois há a presença de censura que é a observação parcial da resposta (Pneus que não vem a ficar ruim/falhar). Pereira, Parreira & Shitsuka (2018) apresentam um suporte para este tipo de metodologia de pesquisa e Cox (2018) a metodologia estatística.

Os pneus foram classificados em três grupos, a partir do seu estado de conservação. As classificações ficaram da forma, G1: pneus nunca usados (1ª vida útil); G2: pneus que foram consertados pela primeira vez (2ª vida útil) e G3: pneus que foram consertados pela segunda vez (3ª vida útil). Utilizou-se o estimador de Kaplan-Meier para encontrar uma estimativa não paramétrica da função de sobrevivência considerando todos os tempos e também considerando os tempos por grupo. Após isso, aplicou-se o teste log-rank para testar de maneira não paramétrica se havia diferença ou similaridade entre as curvas para cada grupo de pneus.

Foram utilizados modelos paramétricos exponencial, Weibull e Log-Normal para ajustar os tempos até a falha dos pneus. Os pressupostos e a qualidade do ajuste dos modelos foram avaliados a partir do teste da razão de verossimilhanças, critério de AIC, análises gráficas e análise de resíduo (Cox, 2018).

Também se ajustou o modelo de riscos proporcionais com o objetivo de propor um ajuste semi-paramétrico. Para este modelo foi proposto uma covariável que discrimina cada tempo até a falha do pneu considerando o seu grupo.

O procedimento realizado para analisar a sobrevivência de um total de 552 pneus 11.00R22 foi pelo modelo Cox:

$$\lambda(t|x) = \lambda_0(t) \exp(x_1\beta_1)$$

sendo, x_1 a variável indicadora de grupo_{*i*} (*i*=1,2,3).

A avaliação dos pressupostos do modelo foi realizada por meio do resíduo de Schoenfeld e também uma verificação se o pressuposto de proporcionalidade dos riscos estava sendo atendido no modelo proposto. Para as análises foi utilizada o *software* R 3.2.2 (Team, 2018) com o auxílio da biblioteca *survival* na versão 2.36 (Therneau & Lumley, 2014) e o software ActionEstat.

4. Resultados

Os principais resultados obtidos por meio da análise de confiabilidade para os pneus 11.00R22 são apresentados em ordem: Análise não paramétrica, paramétrica e semi-paramétrica. Inicialmente na Tabela 1 e no gráfico da Figura 1 tem-se os resultados do estimador de Kaplan- Meier.

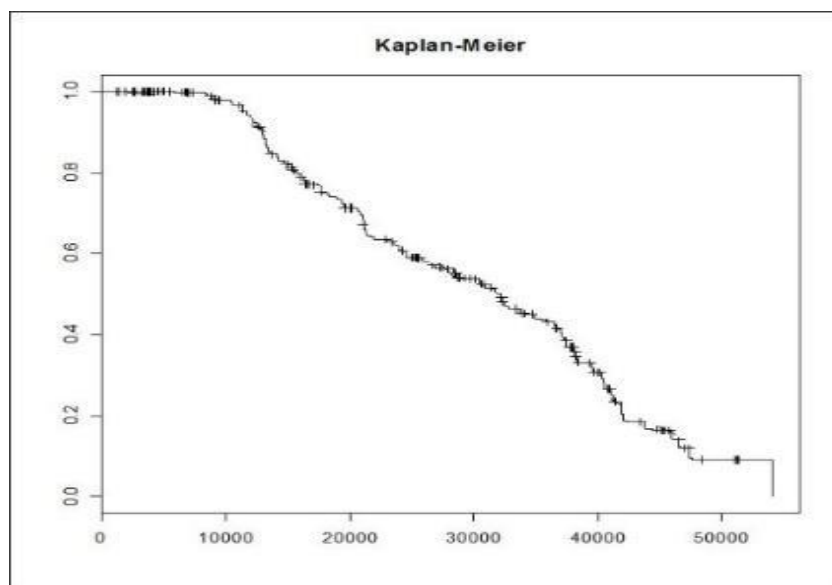
Tabela 1 – Tempo mediano e distribuição dos eventos\falha de acordo com o grupo de vida dos pneus 11.00R22.

	n	Evento\Falha	Tempo mediano
Vida 1	552	449	25953
Vida 2	416	195	36781
Vida 3	138	5	NA
Geral	1106	649	32229

Fonte: Autores.

Com relação ao tempo de rodagem pode se observar que 649 (58,67%) do total (1106) tempos de rodagem tiveram alguma falha e as demais observações censuraram (Tabela 1). A primeira vida útil foi constituída de 552 pneus. Já a vida útil seguinte (segunda) foi constituída de pneus que não falharam até este momento, da mesma forma a terceira vida útil na qual foi constituída apenas de 138 pneus dos quais 5 experimentaram a falha. O número de 32000 quilômetros foi o tempo mediano de vida útil dos pneus.

Figura 1 – Confiabilidade estimada por Kaplan-Meier para os Pneus 11.00R22.



Fonte: Autores.

A função estimada de confiabilidade via estimador produto limite ($R(t)$) pode ser visualizada na Figura 1. Conforme definição, trata-se de uma função representada na forma de escada de valor constante em cada intervalo de tempo. Com valor igual a 1 (um) no primeiro intervalo $[0, t_1)$ e 0 (zero) no último intervalo $[t_k, \infty)$; é zero, se o maior tempo observado for uma falha, e não atingirá o zero se for uma censura. Desta forma, na Figura 1 observa-se o comportamento da curva de confiabilidade com as ocorrências de falhas e censuras, com a observação que no último intervalo existe a representação da censura. Após esta etapa de aplicação do estimador produto limite de Kaplan Meier aos dados para obtenção da curva de confiabilidade estimada, o teste log-rank foi aplicado. Este teste foi utilizado para verificar se há diferença estatística entre os tempos de vida (grupos) em relação a curva de confiabilidade em cada grupo.

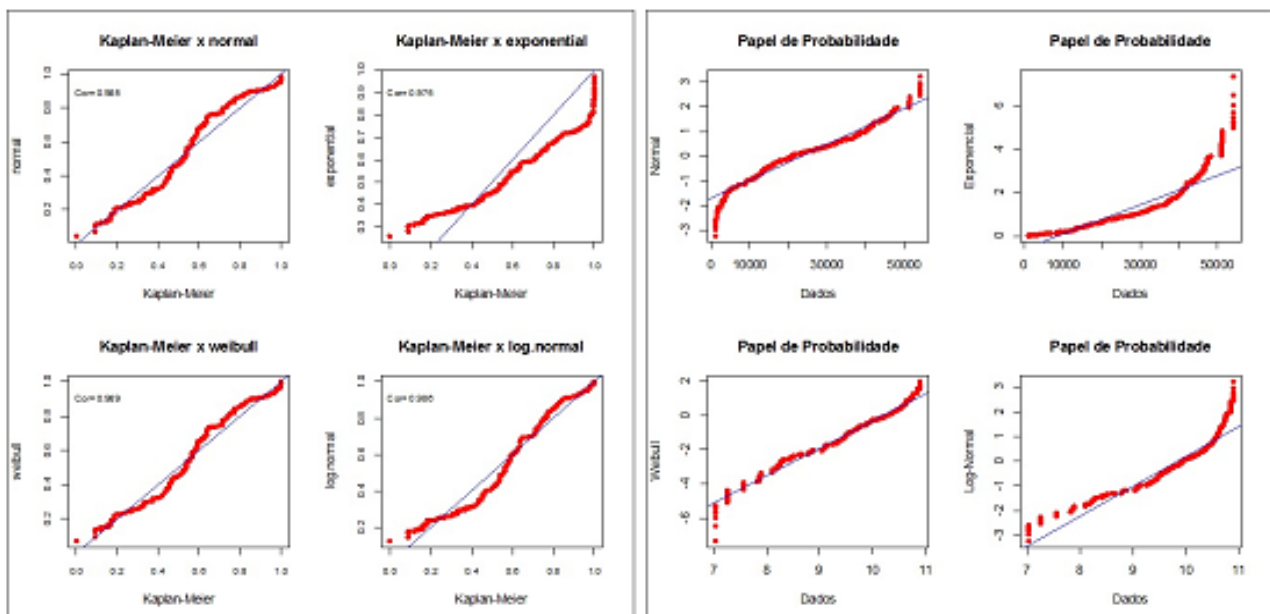
Tabela 2 – Teste de *log-rank* aplicados aos grupos de tempo de vida dos Pneus.

Log-Rank	N	Informação	Valor numérico
grupos = 1	552	Qui-quadrado	73,84
grupos = 2	416	p-valor	0
Log-Rank	N	Informação	Valor numérico
grupos = 2	416	Qui-quadrado	34,79
grupos = 3	138	p-valor	< 0,01
Log-Rank	N	Informação	Valor numérico
grupos = 1	552	Qui-quadrado	72,73
grupos = 3	138	p-valor	0

Fonte: Autores.

O tempo de vida útil dos pneus na primeira vida útil é maior que nas demais vidas, sendo estes estatisticamente diferentes de acordo com o teste *log-rank* (Tabela 2).

Figura 2 – Gráfico do ajuste de Kaplan-Meier versus as distribuições Normal, Exponencial, Weibull e Log-Normal e Representação gráfica dos Papéis de probabilidade para as mesmas distribuições.



Fonte: Autores.

As distribuições Weibull e Log-Normal melhor se ajustaram a reta, tanto na comparação com o estimador produto limite de Kaplan Meier para função de confiabilidade como no papel de probabilidade (Figura 2). Os resultados das estimativas dos parâmetros das distribuições Weibull e Log-Normal, no qual a forma da distribuição é determinada pelo parâmetro de forma (Tabela 3). Este parâmetro define como os dados são distribuídos, sem afetar a distribuição em sua escala. Curvas assimétricas à esquerda são encontradas com valores altos dos parâmetros de forma e curvas assimétricas a direita com valores menores. Neste sentido, os valores estimados caracterizam distribuições aproximadamente assimétricas. A magnitude do erro padrão das estimativas dos parâmetros são muitas vezes menores que as grandezas estimadas, isso é um indício de que o valor do parâmetro

de forma foi bem estimado em ambas as distribuições, além disso, o intervalo de confiança tem comprimento pequeno/estrito.

Tabela 3 – Estimativas dos parâmetros das distribuições Weibull e Log-Normal.

Distr.	Par.	Est. (D.P)	Lim. Inf.	Lim. Sup.
WE	Forma	2,30(0,07)	2,16	2,44
	Escala	35231,23(0,00)	35231,22	35231,23
LN	Forma	10,24 (0,02)	10,20	10,28
	Escala	0,56 (0,01)	0,53	0,59

Fonte: Autores.

O parâmetro de escala de uma distribuição determina a escala da função de distribuição. Um valor maior para a estimativa deste parâmetro amplia o seu pico deixando-a leptocúrtica, possuindo a curva da função de distribuição mais afunilada com um pico mais alto do que a distribuição normal. Neste caso dizemos que essa distribuição possui caudas pesadas.

Tabela 4 – Estimativas dos percentis obtidos por meio das distribuições Weibull e Log-Normal com respectivos desvios padrão e intervalos de confiança.

Distr.	P (%)	Est. (D.P.)	Lim. Inf.	Lim. Sup.
WE	0,01	4783,15 (294,77)	4205,40	5360,90
	0,05	9704,89 (386,17)	8948,01	10461,77
	0,1	13264,53 (399,90)	12480,73	14048,31
	0,5	30049,23 (147,54)	29760,04	30338,41
	0,9	50600,92 (565,39)	49492,79	51709,07
LN	0,01	7585,33 (287,32)	7022,17	8148,48
	0,05	11134,85 (325,28)	10497,30	11772,40
	0,1	13663,34 (345,11)	12986,92	14339,76
	0,5	28122,61 (564,67)	27015,86	29229,36
	0,9	57883,41(1818,04)	54320,11	61446,72

Fonte: Autores.

Os resultados da Tabela 4, no intervalo de 13264,53 à 13663,34 quilômetros rodados, 10% dos pneus falharam. E apenas 10% dos pneus continuaram a ser utilizados após 50600,92 Km e 57883,41 Km, de acordo com os resultados das distribuições Weibull e Log-Normal, respectivamente. O tempo mediano de vida dos pneus estudados é maior pela estimativa da distribuição Weibull (30049,23 Km) em relação a distribuição Log-Normal (28122,61km). Estes resultados, ajudam a balizar o setor produtivo no que tange o tempo de vida útil da banda de rodagem, para poder aplicar parâmetros de garantia aos clientes.

Tabela 5 – Quartis estimados do tempo médio até a falha, mediana e desvio padrão obtidos das distribuições Weibull e Log-Normal.

Distribuições	Índices	Valor
Weibull	MTTF	31212,45
	Desvio Padrão	14370,71
	Mediana	30049,23
	1º Quartil	20514,10
	3º Quartil	40597,98
Log-Normal	MTTF	32957,18
	Desvio Padrão	20138,28
	Mediana	28122,61
	1º Quartil	19233,53
	3º Quartil	41119,91

Fonte: Autores.

De acordo com a Tabela 5, o tempo médio para falha (MTTF) da distribuição Weibull é 31212,45 km, e o tempo médio para falha da distribuição Log-Normal é 32957,18. 25% dos pneus falharam nas distribuições Weibull e Log-Normal antes de 20514,10 e 19233,53, e restavam 25% dos pneus aptos para uso após 40597,98 e 41119,91 km, respectivamente. Na Figura 3, pode-se observar que a distribuição Weibull modelada para esses dados apresenta risco aumentado, a função densidade à direita é levemente assimétrica, e o ajuste ao modelo Weibull e o Kaplan-Meier satisfatório, pois a distribuição sobrevivência pode modelar bem as estimativas obtidas por Kaplan-Meier. O mesmo acontece com a distribuição Log-Normal, com exceção que o risco da distribuição Log-Normal assume inicialmente risco crescente e posteriormente o risco decresce. O teste de razão de verossimilhança e o critério AIC foi utilizado para ajudar na escolha da melhor distribuição dos dados. Os resultados apresentados na Tabela 7 mostram a adequação do modelo Log-Normal por meio dos critérios AIC (menor valor) e TRV (Valor $P > 0,05$). No entanto, devido às técnicas gráficas mostradas nas Figuras 4 e 5, a distribuição Weibull é a distribuição mais adequada.

Tabela 6 - Teste da razão de verossimilhança - critérios do TRV e AIC para distribuição do estudo.

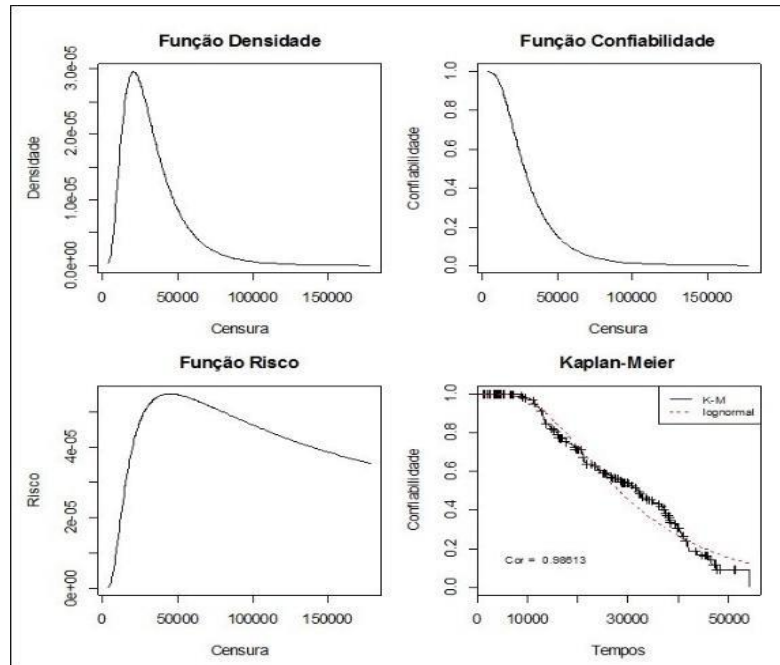
Distr.	Loglik	AIC	Par.	TRV	Valor P
Exp.	-7526,09	15054,18	1	534,7	0
WE	-7267,64	14538,38	2	17,8	<0,01
LN	-7259,70	14523,37	2	1,92	0,16
NO	-7310,02	14624,04	2	102,56	0
GG	-7258,74	14523,48	3	-	-

Fonte: Autores.

Após aplicação dos testes inferenciais aos dados de Pneus 11.00R22, procedeu-se com uma análise de resíduos (Figuras 4 e 5). A diferença entre a Figura 4 e a Figura 5 é o tipo de resíduo utilizado, sendo que na primeira figura é utilizado o resíduo

padrão e, na segunda, resíduo Cox-Snell.

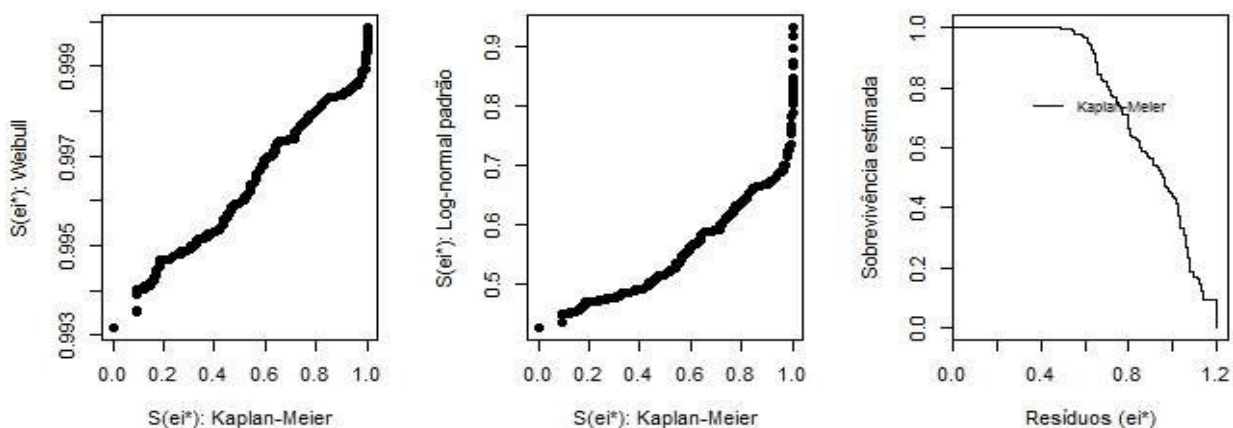
Figura 3 – Gráfico das funções estimadas para o modelo Log-Normal.



Fonte: Autores.

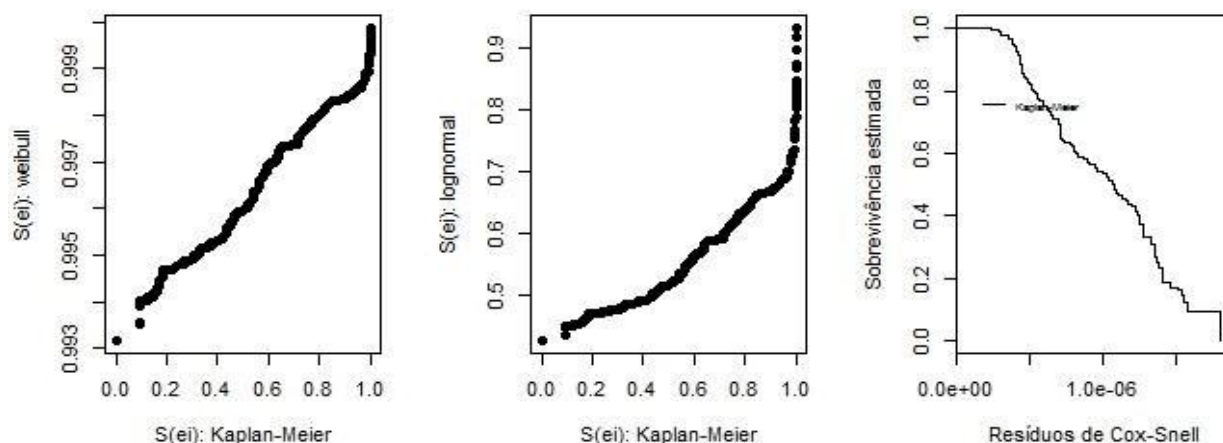
Ambas distribuições candidatas obtidas pela Tabela 6, foram utilizadas para comparação dos resíduos (Weibull e Log-Normal). Na Figura 4 e na Figura 5, pode-se observar que os resíduos padronizados da distribuição Weibull e os resíduos de Cox-Snell se comportam mais próximos de uma linha reta de 45 graus, o que indica que a distribuição possui um melhor ajuste que a distribuição Log-Normal, pois esta última apresenta maior curvatura. Diferentes critérios têm sido utilizados para mensurar a qualidade do ajuste dos modelos aos dados na análise de confiabilidade (sobrevivência). Eles são todos obtidos por meio de diferentes procedimentos e mostram grande discrepância quando calculados simultaneamente, além disso, dependendo da distribuição escolhida sem ter diferentes recomendações a serem dadas quanto ao tempo e estatísticas de interesse para a durabilidade dos Pneus 11.00R22.

Figura 4 – Gráfico dos resíduos padronizados para as distribuições Weibull, Log-Normal e para o estimador de Kaplan-Meier.



Fonte: Autores.

Figura 5 – Gráfico dos resíduos Cox-Snell para as distribuições Weibull, Log-Normal e para o estimador de Kaplan-Meier.



Fonte: Autores.

Referente a análise do modelo de Cox para o modelar os tempos de vida dos Pneus 11.00R22, se aplicou o modelo de Cox sem interação. Na Tabela 7, é checada a suposição de riscos proporcionais e de acordo com o Valor P encontrado ($p < 0,05$), a um nível de 5% de significância, há evidências de violação em todas os níveis da variável vida, como o valor P global é menor que 0,05, o fator vida não foi proporcional, ou seja, é uma variável tempo-dependente.

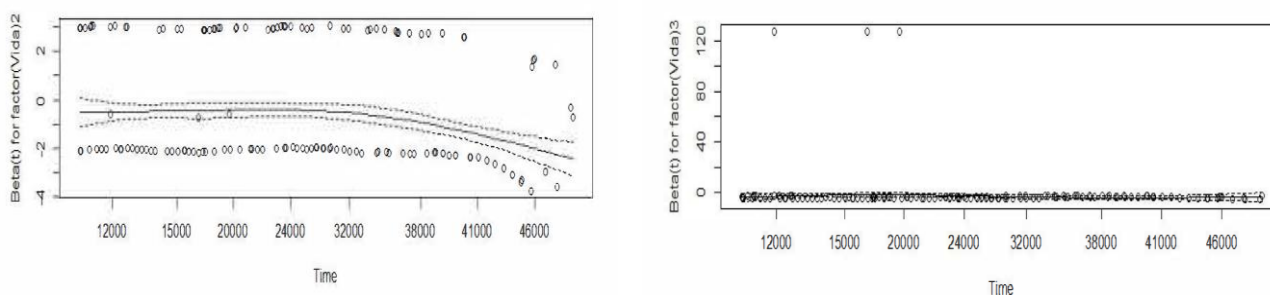
Tabela 7 - Teste de proporcionalidade das funções de risco no modelo ajustado.

Efeitos	ρ	χ^2	Valor P
Vida 2	-0,1754	21,02	<0,01
Vida 3	-0,04921	1,56	0,02
Global	NA	22,00	<0,01

Fonte: Autores.

Para confirmação, pode-se perceber a partir da Figura 6, que os resíduos não têm um padrão aleatório em torno de zero, os resíduos situados aproximadamente a 38.000 km, apresentam um padrão decrescente, sugerindo violação do princípio de proporcionalidade da função de risco.

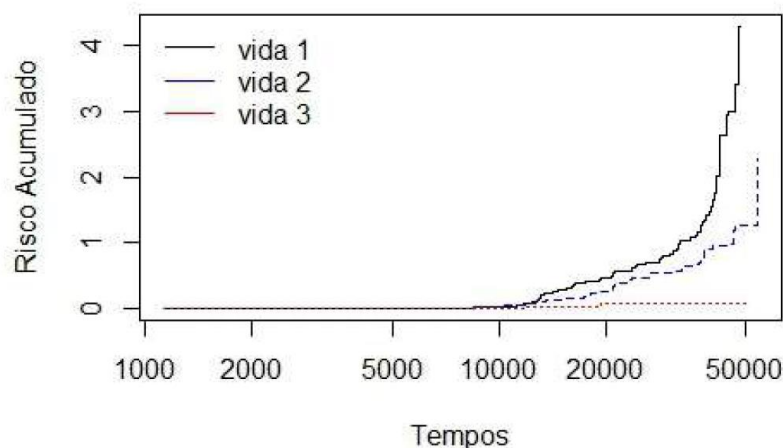
Figura 6 - Suposição de riscos proporcionais fazendo uso dos resíduos padronizados de Schoenfeld.



Fonte: Autores.

Devido a não proporcionalidade dos riscos, o modelo de Cox foi aplicado com a estratificação da função de risco por meio do modelo de Cox estratificado. Como só dispomos de uma covariável e esta é a covariável de estratificação, o modelo de Cox estratificado será apresentado apenas através da função de base em cada grupo de forma gráfica pelos riscos acumulados. Na Figura 7 as curvas obtidas pelo ajuste da função de risco via modelo de Cox.

Figura 7 - Curvas de risco acumulado de base dos estratos obtidas por meio do modelo de Cox estratificado.



Fonte: Autores.

Cada linha na Figura 7 representa um nível do grupo de vida dos pneus. Nota-se que de 10.000 km rodados em diante o risco aumenta consideravelmente com uma maior diferenciação das curvas de partir de 20.000 km rodados, nota-se um crescimento no risco acumulado, onde os estratos com vida inicial (vida 1), evidencia risco de falha maior. Esse aumento após um pequeno intervalo de tempo inicial tende a acelerar ao longo dos quilômetros rodados. A funções de risco acumulado de base obtidas para o estrato se cruzam, reforçando, pois, o indício de violação da suposição de riscos proporcionais. Os resultados mostram que o modelo de Cox estratificado para a análise dos pneus 11.00R22 parece ser o mais indicado, já que os riscos não apresentaram proporcionalidade entre os estratos.

5. Discussão

Após a aplicação da estatística descritiva, observou-se o comportamento das covariáveis em estudo. A tabela de frequência constatou que a vida útil média para troca de pneus era de aproximadamente 32.000 km, para um conjunto com 552 pneus. A curva de Kaplan-Meier foi ajustada aos dados, indicando a existência de diferença entre os grupos. Para provar isso, o teste de log-rank confirmou que havia de fato uma existência significativa entre os três grupos.

As distribuições Log-Normal e Weibull se apresentaram como fortes candidatas para modelar a vida útil dos pneus 11.00R22, ajudando a estabelecer o tempo de troca dos pneus, reduzindo assim o custo com tempo perdido e danos ao transporte. Observou-se que o tempo médio para trocar os pneus 11.00R22 é de aproximadamente 30000 quilômetros quando a distribuição Weibull é usada para descrever os dados e 28122.61km quando a distribuição Log-Normal é usada. Embora as distribuições Log-Normal e Weibull tenham provado ser boas candidatas para modelar a vida de pneus 11.00R22; A distribuição de Weibull apresentou melhor ajuste através da análise gráfica, conforme mostrado nas Figuras 4 e 5. Porém, os resultados apresentados na Tabela 6 indicam a adequação do modelo Log-Normal para os critérios AIC (menor valor) e TRV (valor $P > 0,05$). Segundo Carvalho (2011), os métodos gráficos devem ser usados para descartar modelos claramente inadequados, e não para demonstrar que um modelo paramétrico é melhor, uma vez que os critérios de seleção seriam mais adequados para ele. No modelo de riscos

proporcionais de Cox, a variável de grupo foi significativa ao nível de 5% de significância. Uma vez ajustado o modelo de Cox, os pressupostos de proporcionalidade foram analisados ao longo do tempo e os resíduos de Schoenfeld mostraram a existência de uma violação em todos os grupos.

Ao contrário, o modelo estratificado de Cox foi utilizado e os resultados reforçaram a evidência de violação da suposição de riscos proporcionais. A adequação do padrão estratificado ao longo do tempo de uso dos pneus mostrou que o fator Vida influencia o tempo até o pneu atingir a quilometragem máxima.

6. Conclusão

O modelo paramétrico Weibull foi o mais adequado para modelar o tempo de vida medido em quilômetros rodados. Por meio deste modelo foi possível constatar que o tempo mediano de troca dos Pneus 11.00R22 é de aproximadamente 30.000 quilômetros. O modelo de Cox estratificado foi utilizado e os resultados reforçaram a evidência de violação da suposição dos riscos proporcionais. O ajuste do modelo estratificado no período de uso dos pneus, mostrou que o fator Vida influencia no tempo até que o pneu atinja a quilometragem máxima. Concluímos que o modelo de Cox estratificado é o mais indicado que o modelo de Cox para modelagem no tempo de vida dos pneus 11.00R22, porém a modelagem paramétrica via distribuição Weibull é a mais adequada para modelar a vida útil dos pneus 11.00R22.

Para trabalhos futuros pode se pensar em utilizar outras distribuições com mais parâmetros e verificar se elas descrevem melhor o conjunto de dados que a distribuição Weibull, bem como em situações parecidas com a aqui descrita para um estudo de confiabilidade, caso seja possível viabilizar um banco de dados com as mesmas características mas com mais covariáveis, verificar se extensões do modelo de Cox como modelos de fragilidade podem ser úteis para descrever o comportamento do tempo de vida dos Pneus. Outra abordagem que se pode utilizar é a de modelos de inteligência artificial por meio das técnicas de *machine learning*, baseados em árvores de decisão tais como *Random Survival Forest* e prever o tempo de falha dos pneus.

Referências

- Carvalho, M. S., Andreozzi, V. L., Codeço, C. T., Barbosa, M. T., & Shimakura, S. E. (2005). Análise de sobrevida. *Fiocruz*.
- CNT, Confederação Nacional do Transporte (2019). <http://pesquisarodovias.cnt.org.br/Paginas/relatorio-gerencial>.
- Colosimo, E. A., & Giolo, S. R. (2006). *Análise de sobrevivência aplicada*. Editora Blucher.
- Cox D.R. (2018). *Analysis of Survival Data*. Chapman and Hall/CRC: Routledge.
- Fogliato, F., & Ribeiro, J. L. D. (2009). *Confiabilidade e manutenção industrial*. Elsevier Brasil.
- Haviaras, G. (2005). Metodologia para análise de confiabilidade de pneus radiais em frotas de caminhões de longa distância. 2005. 128p (Doctoral dissertation, Dissertação (Mestrado em Engenharia Automotiva) – Escola Politécnica da Universidade de São Paulo, São Paulo).
- Herrmann, L. (2011). Estimativa de curvas de sobrevivência para estudos de custo-efetividade.
- Leal, V. J., & de Resende Andrade, P. C. (2018). Modelagem dos dados de falha de um caminhão fora de estrada. *ForScience*, 6(3).
- Lee, E. T., & Wang, J. (2003). *Statistical methods for survival data analysis*, v. 476. John Wiley & Sons.
- Oliveira, T. N., & de Melo, J. A. M. (2019). O efeito da infraestrutura rodoviária sobre os custos operacionais das transportadoras de cargas. *Negócios Em Projeção*, 10(2), 107-123.
- Pereira, A. S., Shitsuka, D. M., Parreira, F. J., & Shitsuka, R. (2018). *Metodologia da pesquisa científica*.
- Ramires, T. G. (2013). A distribuição beta semi-normal generalizada geométrica (Doctoral dissertation, Universidade de São Paulo).
- Santos, I. P. (2016). Introdução à análise de confiabilidade: Uma aplicação ao setor de Transportes. Monografia de conclusão da Especialização – UEPB. 46p.
- Silva, J. R. S., Souza, L. A. D., Castro, L. Z., Ferreira, T. A., & Campos, M. S. (2015). Análise Da Confiabilidade: Um Estudo de Caso. XXXV Encontro Nacional De Engenharia De Produção, Perspectivas Globais para a Engenharia de Produção - ENEGEP.
- Stacy, E. (1962). Generalization of the gamma distribution. *Ann. Math. Stat.*, 33, 1187-1192.

Team, R. C. (2018). R: A language and environment for statistical computing.

Therneau, T. M., & Lumley, T. (2014). *Package 'survival'*. Survival analysis Published on CRAN, 2, 3.

Wang, P., Li, Y., & Reddy, C. K (2019). Machine learning for survival analysis: A survey. *ACM Computing Surveys (CSUR)*, 51(6), 1-36.