

Automatização da tomada de decisão sobre o Sucateamento de Componentes Eletrônicos da MRS Logística

Uma aplicação de Modelos Reparáveis

Éder Silva de Brito^a, Murilo Sernaglia Fernandes^a, Natally Cunha Redondo^a, Natan Hilario da Silva^a, Rafael Fragoso Marin^a, Renata Navarro Lima^a, Victor Botelho Cardoso^a, Vikson Gianotti Andrade dos Santos^a, Vinicius Loureiro Siqueira^a, Vladimir Morroni Junior^a, Wesley da Silva^a

^aInstituto de Ciências Matemáticas e da Computação, Universidade de São Paulo, São Carlos, Brasil

RESUMO

O trabalho descreve o desenvolvimento de uma modelagem que visa a automatização da tomada de decisão sobre o sucateamento de componentes eletrônicos da MRS Logística, uma empresa responsável por administrar 1.650 km da malha ferroviária brasileira, que está presente nos estados de Minas Gerais, Rio de Janeiro e São Paulo. Partimos da análise da base de dados fornecida pela MRS e, utilizando Análise de Sobrevida e Modelo de Sistemas Reparáveis, desenvolvemos um modelo ARAM para esta tomada de decisão, além de corrigir problemas encontrados na base de dados. Ao final, para trazer uma fácil visualização dos dados pelo usuário, criamos um relatório iterativo automatizado (*Dashboard*) em *Power BI*.

Palavras-chave: Tomada de Decisão, Autimatização, MRS, Sucateamento, Componentes Eletrônicos.

1 Introdução

A MRS¹ é uma operadora logística responsável por administrar cerca de 1.650 km da malha ferroviária brasileira, estando presente nos estados de Minas Gerais, Rio de Janeiro e São Paulo. Ela foi criada em 1996, quando o governo brasileiro decidiu transferir à iniciativa privada a gestão de todo seu sistema ferroviário, e já nos dias atuais, 26 anos depois, ela está entre as maiores companhias ferroviárias de carga do mundo, sendo responsável conectar regiões produtoras de *commodities* minerais e agrícolas e alguns dos principais parques industriais do país aos maiores portos da região Sudeste, carregando quase 20% de tudo o que o Brasil exporta e um terço de toda a carga transportada por trens no país.

Os trens da MRS possuem, entre suas peças, componentes eletrônicos, os quais costumam sofrer desgastes e falhas e tais ocorrências causam avarias às locomotivas, prejudicando o transporte das cargas. No entanto, uma parte destas falhas podem ser evitadas, dado que tais itens já deveriam ter sido sucateados, mas continuaram em uso.

Atualmente, não há parâmetros pré-estabelecidos para a decisão de retirada de circulação ou apenas reparar os itens que falharam, ficando a cargo do técnico de manutenção e sua experiência tal determinação, o que leva a uma falta de padronização em toda a equipe.

Com isto, temos o objetivo de criar um modelo que tome a decisão sobre o sucateamento ou não de um destes componentes eletrônicos, afim de padronizar e automatizar o processo determinatório. Para isto, foram utilizados os códigos dos itens, que identificam cada tipo de componente e, a partir disto, foram analisados a taxa de retorno, as causas do retorno (modo de falha descrito em forma de texto livre) e o reparo realizado. Apesar de o banco de dados possuir um atributo de *status* que indica se uma peça foi "Baixado" (sucateado) ou "Recuperado", notamos que várias peças que foram "baixadas" voltaram a funcionar posteriormente, mostrando uma falha no preenchimento deste atributo, o que nos levou a desconsiderá-lo e utilizar trechos da "Descrição do Reparo", composto por um texto livre, para classificá-lo como sucateado ou não.

Tendo essa visão inicial seguimos com a problemática de como criar um sistema/plataforma que auxilie na tomada de decisão sobre sucatear ou retirar de circulação as peças, para isso utilizamos principalmente a **Análise de Sobrevida e Confiabilidade**, para ser mais específico adentramos na questão de **Modelos de Sistemas Reparáveis**.

Após algumas análises e tentativas de modelagens, o modelo ARAM, variação do ARA, acabou sendo o escolhido e ainda foi desenvolvido um relatório automatizado (*Dashboard*) em *Power BI* para facilitar a visualização dos dados pelo usuário com vários tipos de filtro e disposição das apresentações.

Nas próximas seções, serão apresentadas algumas revisões bibliográficas, abordando como alguns autores tratam este tipo de problema, além dos materiais e métodos utilizados no desenvolvimento da modelagem e, ao final, uma conclusão.

2 Revisão Bibliográfica

2.1 Análise Descritiva

Com intuito de uma visão inicial sobre o banco de dados aplicamos uma **análise exploratória**, onde o principal papel da estatística descritiva de acordo com Bussab e Morettin (2010)² é examinar os dados antes das aplicações de qualquer técnica estatística. Desta forma, conseguimos um entendimento básico dos dados e das relações existentes entre as variáveis analisadas. A análise extrai informações de um conjunto de dados sem o peso das suposições de um modelo probabilístico. Ajudando a ter direções básicas de como lidar com os dados e possíveis problemas para desenvolvimentos futuros.

2.2 Análise de Sobrevivência e Confiabilidade

Utilizando da definição de Colosimo e Giolo (2006)³ em análise de sobrevivência a variável resposta é, geralmente, o tempo até a ocorrência de um evento de interesse. Esse tempo é denominado de **tempo de falha**. A principal característica de dados de sobrevivência é presença de **censura**, ou seja, a observação parcial da resposta, isto refere-se a situações que, por alguma razão, o acompanhamento foi interrompido, em nosso contexto a retirada de operação da peça depois de um número de falhas tenham ocorrido. Dado o entendimento de censuras temos então a principal questões trabalhadas inicialmente no projeto que desenvolve a análise de sobrevivência que são **função sobrevivência** e **função de taxa de falha**

2.2.1 Conceitos fundamentais

Ainda no livro de Colosimo e Giolo (2006), podemos definir esses conceitos mais a fundo. Primeiramente sobre a ideia de **Tempo de Falha**, os indivíduos (em nosso caso peças) devem ser comparáveis desde o início do estudo, com exceção de diferenças medidas pelas covariáveis. É importante, em estudos de sobrevivência, definir de forma clara e precisa o que vem a ser a falha. Em algumas situações ela já é clara, mas em outros pode assumir termos ambíguos. O tempo de falha vai até o tempo inicial de exposição até o produto ficar inapropriado ou inválido.

Definindo também o conceito de **censuras** mais detalhadamente, além da descrição posterior, temos primeiramente uma ressalta que, mesmo censurados, todos os resultados provenientes de um estudo de sobrevivência devem ser usados na análise estatística. Duas razões justificam tal procedimento: (i) mesmo sendo incompletas, as observações censuradas nos fornecem informações sobre o tempo de vida de pacientes; (ii) a omissão das censuras no cálculo das estatísticas de interesse pode acarretar em conclusões viciadas. Ainda temos dois tipos de censura: *Censura Tipo I* é quando o estudo é terminado após um período previamente estabelecido; e a *Censura Tipo II* é quando o estudo será terminado após a ocorrência do evento de interesse em um número previamente estabelecido de indivíduos.

Especificando ainda mais a ideia de tempo de sobrevivência, temos uma variável aleatória não negativa T , que representa o tempo de falha, usualmente especificada em análise de sobrevivência pela sua função de sobrevivência ou pela sua função de taxa de falha (ou risco). Pode também ser especificadas por outras funções relacionadas, essas duas são mais desenvolvidas a seguir.

A **função de sobrevivência** é uma das principais funções probabilísticas para descrever estudos de sobrevivência. Ela é definida como a probabilidade de uma observação não falhar até um certo tempo t , ou seja, a probabilidade de uma observação sobreviver ao tempo t . Em termos probabilísticos, isto é escrito como

$$S(t) = P(T \geq t)$$

Em consequência, a **função de sobrevivência acumulada**, isto é, $F(t) = 1 - S(t)$ é definida como a probabilidade de uma observação não sobreviver ao tempo t . Já a **função de taxa de falha ou risco** é descrita como a probabilidade da falha ocorrer em um intervalo do tempo $[t_1, t_2)$ pode ser expressa em termos da função da sobrevivência como

$$S(t_1) - S(t_2)$$

A taxa de falha no intervalo $[t_1, t_2)$ é definida como a probabilidade de que a falha ocorra nesse intervalo, dado que não ocorreu antes de t_1 , dividida pelo comprimento do intervalo. Assim, a taxa de falha no intervalo $[t_1, t_2)$ é expressa por

$$\frac{S(t_1) - S(t_2)}{(t_2 - t_1) \cdot S(t_1)}$$

De uma forma geral, redefinindo como $[t, t + \Delta t)$, a expressão assume a seguinte forma

$$\lambda(t) = \frac{S(t) - S(t + \Delta t)}{\Delta t \cdot S(t)}$$

Assumindo Δt bem pequeno, $\lambda(t)$ representa a taxa de falha instantânea no tempo t condicional à sobrevivência até o tempo t . Note que as taxas de falha são números positivos, mas sem limite superior. A função de taxa de falha $\lambda(t)$ é bastante útil para descrever a distribuição do tempo de vida. Ela descreve a forma em que a taxa instantânea de falha muda com o tempo. Logo a **função de taxa de falha** T é definida como

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}$$

Assim também temos outra função muito útil em análise de sobrevivência acumulada que fornece a **taxa de falha acumulada** do indivíduo expressa por:

$$\Lambda(t) = \int_0^t \lambda(\mu) d\mu$$

A função de taxa de falha acumulada Λ , não possui uma interpretação direta, mas é útil na avaliação da função de maior interesse que é a taxa de falha. $\lambda(t)$.

Outras duas quantidades de interesse em análise de sobrevivência são: o tempo médio de vida (tvm ou tm) e a vida média residual (vmr). A primeira é obtida pela área sobre a função de sobrevivência, e podemos definir da seguinte forma:

$$tm = \int_0^\infty S(t) dt$$

Agora a vida média residual é definida condicionalmente a um tempo média de vida t , ou seja, indivíduos com idade t tem como medida essa quantidade de tempo médio restante de vida, e podemos definir essa quantidade por:

$$vmr(t) = \frac{\int_t^\infty (\mu - t) f(\mu) d\mu}{S(t)} = \frac{\int_0^\infty S(\mu) d\mu}{S(t)}$$

sendo $f(\cdot)$ a função de densidade de T .

2.2.2 Distribuição Weibull

A princípio na análise encontramos uma distribuição específica que é bem comum em análise de sobrevivência, que é a **Distribuição de Weibull** proposta originalmente por W. Weibull (1954). A popularidade em aplicações práticas dessa distribuição deve-se ao fato de apresentar uma grande variedade de formas, todas com uma propriedade básica: a sua função de taxa de falha é monótona, ou seja, ela é crescente, decrescente ou constante.

Para uma variável aleatória T com distribuição de Weibull se tem a função de densidade de probabilidade dada por:

$$f(t) = \frac{\gamma}{\alpha^\gamma} t^{\gamma-1} \exp\left\{-\frac{t^\gamma}{\alpha^\gamma}\right\}, t \geq 0$$

em que γ é o parâmetro da forma e α o de escala, os quais são ambos positivos. O parâmetro α tem a mesma unidade de medida de t e γ não tem unidade. Assim para essa distribuição, as funções de sobrevivência e risco são, respectivamente:

$$S(t) = \exp\left\{-\left(\frac{t}{\alpha}\right)^\gamma\right\} \quad \text{e} \quad \lambda(t) = \frac{\gamma}{\alpha^\gamma} t^{\gamma-1}, \text{ para } t, \alpha \text{ e } \gamma \geq 0$$

Assim, como mencionado, ela pode tomar três caminhos: estritamente decrescente ($\gamma < 1$), estritamente crescente ($\gamma > 1$) e constante ($\gamma = 1$), neste caso quando constante a distribuição de Weibull se torna uma distribuição exponencial.

As expressões da média e variância dessa distribuição incluem o uso da função gama e podem ser descritas como:

$$E[T] = \alpha \Gamma[1 + (\frac{1}{\gamma})], \quad \text{Var}[T] = \alpha^2 [\Gamma[1 + (\frac{2}{\gamma})] - \Gamma[1 + (\frac{1}{\gamma})]^2], \text{ sendo } \Gamma(r) = (r-1)! \text{ para } r \text{ inteiro}$$

2.3 Modelos de Sistemas Reparáveis

Nossa base de dados tem como intuito indicar quando uma peça é reparada ou baixada (Coluna: **Status**), tendo seus períodos de tempos quando isso ocorre, o que se direciona claramente com a teoria e aplicação de um *sistema reparável*⁴, onde esse um sistema é compreendido como, quando uma falha ocorre, pode-se retornar a uma condição de operação através de algum processo de reparo, evitando que todo sistema seja substituído. Logo, um sistema não reparável é aquele do tipo que é descartado finalizando seu processo com a ocorrência de uma primeira e única falha. Agora partiremos para uma leve introdução aos **conceitos preliminares** desses sistemas.

2.3.1 Processo de Contagem

Suponhamos que o tempo até a falha em um sistema reparável seja representado pela variável aleatória T , sendo que $t = 0$ caracteriza o início do experimento. Denominamos então, para um determinado sistema, que T_j caracteriza o tempo até a ocorrência da j -ésima falha. O total de falhas ocorrentes no intervalo $(0, t)$ pode ser expresso por $N(t)$, sendo considerado um *processo de contagem* de ocorrências de um evento de interesse qualquer.

Formalmente, um processo de contagem $\{N(t), t \geq 0\}$ é um processo estocástico que conta o número de ocorrências de um evento de interesse no intervalo de tempo $(0, t)$ e que satisfaz as seguintes condições:

- $N(t) \geq 0$ e $N(0) = 0$;
- $N(t)$ é um número inteiro;
- $t \rightarrow N(t)$ é contínua a direita;
- $\lim_{h \rightarrow 0} [N(t+h) - N(t)] = 0$

Como consequência direta da definição, temos que, se $s < t$, então $N(s) < N(t)$. Além disso, $N(t) - N(s)$ é o número de eventos que ocorreram no intervalo (s, t) .

Os eventos em um processo de contagem podem ser analisados, equivalentemente, em termos de $\{N(t)\}_{t \geq 0}$ ou $\{T_{N(t)}\}_{t \geq 0}$ ou ainda, por $\{X_j\}$, onde $X_j = T_j - T_{j-1}$, $j = 1, \dots, N(t)$. Assim podemos definir a *função de intensidade* do processo de contagem por:

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P[N(t+\Delta t) - N(t) = 1 | F_t^-]}{\Delta t}$$

onde F_t^- é o histórico de todos os eventos ocorridos antes do tempo t .

2.3.2 Processo de Poisson

O *Processo de Poisson* é um caso particular de processo de contagem que pode ser usado para modelar ocorrências (e contagens) de eventos em um período de tempo, quando estes não são afetados pelas ocorrências passadas. Em particular, este processo é utilizado para descrever e prever falhas de um dado sistema. Além disso, o processo de Poisson é um caso particular de processo de Markov em tempo contínuo, onde a única possibilidade de salto é para o próximo estado da sequência.

Um processo de contagem $N(t)$ é dito Processo de Poisson com intensidade $\lambda(t) > 0$ se

- $N(0) = 0$;
- $N(t)$ tem incrementos independentes;
- $\mathbb{P}[N(t+h) - N(t) = 1] = \lambda(t)h + o(h)$;
- $\mathbb{P}[N(t+h) - N(t) \geq 2] = o(h)$.

Se a função de intensidade de um processo de Poisson for constante, isto é, $\lambda(t) = \lambda, t \geq 0$, dizemos que este é um *Processo de Poisson Homogêneo* (PPH), e podemos reescrever os limites da definição de processo de Poisson como:

- $\mathbb{P}[N(t+h) - N(t) = 1] = \lambda h + o(h)$;
- $\mathbb{P}[N(t+h) - N(t) \geq 2] = o(h)$.

Por outro lado, se a função de intensidade $\lambda(t), t \geq 0$, não for constante, dizemos que o processo é um *Processo de Poisson Não Homogêneo* (PPNH), e podemos reescrever os limites da definição de processo de Poisson como:

- $\mathbb{P}[N(t+h) - N(t) = 1] = \lambda(t)h + o(h);$
- $\mathbb{P}[N(t+h) - N(t) \geq 2] = o(h).$

Um caso paramétrico particular de PPNH, amplamente utilizado em sistemas reparáveis, é o Processo de Lei de Potência (PLP). Um processo $N(t)$ segue um PLP se sua função de intensidade e sua função de intensidade acumulada são dadas, respectivamente, por

$$\lambda(t) = \left(\frac{t}{\eta}\right)^{\beta-1} \quad \text{e} \quad \Lambda(t) = \left(\frac{t}{\eta}\right)^{\beta}$$

onde $\eta > 0$ é o parâmetro de escala e $\beta > 0$ é o parâmetro de forma. O parâmetro η pode ser interpretado como o tempo durante o qual é esperado que exista exatamente uma falha, enquanto β representa a elasticidade do número esperado de falhas em relação ao tempo, então, se $\beta > 1$ significa que o sistema está deteriorando e se $\beta < 1$ o sistema está melhorando ao longo do tempo

2.3.3 Truncamentos

Os truncamentos podem ser descritos como análogos a **censura**, também determinam o fim das observações sobre um determinado conjunto de sistema reparáveis. Podemos denotar os truncamentos em duas formas: *truncamento por tempo* e *truncamento por falhas* para definir melhor ambos precisamos descrever inicialmente um conjunto, então considere um conjunto formado por k sistemas reparáveis, onde $t_{i,j}$ é o j -ésimo tempo de falha observados no sistema no sistema i com $i = 1, \dots, k$ e $j = 1, \dots, n_i$.

- **Por Tempo:** é definido um tempo máximo dos sistemas reparáveis, digamos t_i^* com $i = 1, \dots, k$. Assim sobre esse tipo de truncamento os tempos de falhas são observados para cada sistema i serão dados por $0 < t_{i,1} < t_{i,2} < \dots < t_{i,n_i} < t_i^*$, onde t_{i,n_i} é o último tempo de falha observado.
- **Por Falhas:**
é definido por um número máximo de falhas a serem observadas, digamos n_i^* , com $i = 1, \dots, k$. Logo nesse tipo de truncamento os tempos de falha observadas para cada sistema i serão dados por $0 < t_{i,1} < t_{i,2} < \dots < t_{i,n_i^*}$

2.3.4 Tipos de Reparos

Como descrito no início, sistemas são ditos como reparáveis se ao falharem podem retornar ao funcionamento após um reparo. O tipo de reparo a ser feito depende, essencialmente, do tipo do sistema e do tipo de avaria apresentada. Dessa forma, existem três tipos de reparos quando a qualidade e manutenção desse sistema, que podem causar desde uma decadência no quadro geral do sistema quanto uma melhora. Sendo eles **reparo mínimo**, **reparo perfeito** e **reparo imperfeito**, que serão descritos nos próximos itens:

• Reparo Mínimo

O *reparo mínimo* (RM) concentra-se em corrigir apenas o que originou a falha, deixando assim o sistema na mesma condição em que estava antes, conhecido na literatura como *As Bad as Old* (ABAO), que livremente pode ser traduzido como *tão ruim quanto antes*, esse sistema visa apenas fazer o reparo antes da falha, deixando o sistema com a mesma chance de falha antes de ocorrida, ou seja, a probabilidade de falha naquele sistema continua como se nunca tivesse falhado. Assim no modelo RM, é um processo com função de intensidade sendo a próprio PPHN, isto é se denotarmos por ρ_{RM} a função intensidade do processo de reparação, temos que:

$$\rho_{RM}(t) = \lambda(t) \quad \text{onde } \lambda(t) \text{ é a função de intensidade do PPHN.}$$

Ainda, se considerarmos uma PLP, podemos reescrever $\rho_{RM}(t)$ por

$$\rho_{RM} = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1}$$

• Reparo Perfeito

O *reparo perfeito* (RP) é um tipo de reparo que usualmente substitui-se um componente importante para operação do sistema, assim pode até considerar o sistema como novo, ou seja, o sistema foi reparado de forma que ficou melhor do que antes da falha, na literatura podemos encontrar descrito como a condição *As Good As New* (AGAN), que traduzindo livremente seria *tão bom quanto novo* após o reparo. Neste caso, o processo é modelado a partir de um *processo de renovação*. E esse processo pode ser definido pelas distribuições conjuntas do processo de contagem $N(t)$, é renovável porque a probabilidade de falha começa com cada processo de reparo. A caracterização mais simples é através dos intervalos entre tempos de ocorrência de falha, sendo independente e identicamente distribuídos (i.i.d).

Assim, se todas as variáveis aleatórias X_i são independente e identicamente distribuída (i.i.d.), então $\{X_i; i \geq 1\}$ é um processo de renovação, em que $X_i = T_i - T_{i-1}$ indicam o tempo entre os eventos i e $i - 1$, com $T_0 = 0$ e $i = 1, 2, \dots$

A função de intensidade da falha do modelo RP é dada por

$$\rho_{RP}(t) = \lambda(t - T_{N(t)}),$$

onde $\lambda(t)$ é a função de intensidade inicial do PPNH, $N(t)$ é o número de falhas até o tempo t , $T_{N(t)}$ é o tempo até a $N(t)$ -ésima falha, e portanto, $t - T_{N(t)}$ é o tempo até t desde a última falha estritamente antes de t .

Se considerarmos um PLP, podemos reescrever $\rho_{RP}(t)$ por

$$\rho_{RP}(t) = \frac{\beta}{\eta} \left(\frac{t - T_{N(t)}}{\eta} \right)^{\beta-1}$$

• Reparo Imperfeito

Em muitas situações, as ações de reparo deixam o sistema em um nível intermediário, entre o AGAN e o ABAO, caracterizando uma situação de *reparo imperfeito* (RI). Nesse caso, as ações de reparo são mais direcionadas para a manutenção do sistema como um todo do que para o fator que origina a falha. Esse modelo RI conta com a definição de *idade virtual* do sistema que basicamente seria uma função positiva de sua idade real e seu histórico de falhas, isto é, $V_t = V(t; N(t); T_1, \dots, T_{N(t)})$, onde V_t é a notação para a idade virtual. Esse modelo insere um parâmetro θ tal que $0\theta 1$, que representa o grau de eficiência do reparo. Um sistema com idade virtual $V_t \geq 0$ é assumindo como se comportasse exatamente como um novo sistema que atingiu a idade V_t sem falhar.

Utilizando o artigo como base temos duas classes de modelos para o RI, a *Redução Aritmética de Idade* (ARA - Arithmetic Reduction of Age) e a *Redução Aritmética de Intensidade* (ARI - Arithmetic Reduction of Intensity). Ambas classes de modelos são definidas pela sua memória m , onde m refere-se ao número máximo de falhas anteriores envolvidas no cálculo da função de intensidade. A memória m indica que cada ação de reparo reduz o desgaste do sistema que ocorre após m falhas. Para esse projeto utilizamos uma variação da modelagem ARA a variação ARAm.

2.3.5 Modelagem

Utilizando a fonte de modelagem (ARI, ARA)⁵ podemos definir os modelos da seguinte forma:

• Modelo ARA

O princípio dessa classe de modelos é considerar que o reparo rejuvenesce o sistema de forma que sua intensidade no tempo t seja igual à intensidade inicial no tempo V_t ; onde $V_t < t$. Em outras palavras, o efeito do reparo é expresso por uma redução na idade virtual do sistema.

O modelo ARA tem uma intensidade de falha que é uma função de sua idade virtual $\rho_{ARAm}(t) = \lambda(V_t)$, onde $\lambda(t)$ é a função de intensidade inicial. Omitindo maiores detalhes, podemos definir a função de intensidade de falha para um modelo da classe ARAm por

$$\rho_{ARAm}(t) = \lambda(t - (1 - \theta) \sum_{p=0}^{\min(m-1, N(t)-1)} \theta^p T_{N(t)-p})$$

Dois importantes casos particulares (e extremos) da classe ARAm são os modelos ARA₁ e ARA_∞. Se $m = 1$ temos que a função de intensidade de falha para o modelo ARA₁ é dada por

$$\rho_{ARA_1}(t) = \lambda(t - (1 - \theta) T_{N(t)})$$

Por outro lado, na classe ARA_{∞} assume-se que cada reparo reduz a idade virtual do sistema em um quantidade proporcional à sua idade imediatamente antes do reparo. Nesse caso, a função de intensidade de falha é dada por

$$\rho ARA_{\infty}(t) = \lambda(t - (1 - \theta) \sum_{p=0}^{N(t)-1} \theta^p T_{N(t)-p})$$

• Modelo ARI

A ideia básica da classe de modelos ARI é considerar que cada reparo reduz a intensidade de falha, dependendo do histórico de falhas do processo. No modelo ARI_m assume-se que o reparo reduz o incremento na intensidade da falha desde as últimas m falhas.

Novamente, omitindo maiores detalhes, podemos definir a função de intensidade de falha para um modelo da classe ARI_m por

$$\rho ARI_m(t) = \lambda(t) - (t - \theta) \sum_{p=0}^{\min(m-1, N(t)-1)} \theta^p \lambda(T_{N(t)-p})$$

onde $\lambda(t)$ é a função de intensidade inicial do processo.

Analogamente à classe ARA , na classe ARI_m podemos considerar a existência de dois casos extremos, o ARI_1 e o ARI_{∞} . Se $m = 1$ temos que, pelas mesmas justificativas apresentadas para o modelo ARA_1 , a função de intensidade de falha para o modelo ARI_1 é dada por

$$\rho ARI_1(t) = \lambda(t) - (1 - \theta) \lambda(T_{N(t)})$$

Na classe ARI_{∞} assume-se que cada reparo reduz a intensidade da falha em um valor proporcional à intensidade da falha atual. Nesse caso, a função de intensidade de falha é dada por

$$\rho ARI_{\infty}(t) = \lambda(t) - (1 - \theta) \sum_{p=0}^{N(t)-1} \theta^p \lambda(T_{N(t)-p})$$

3 Materiais e Métodos

Nosso banco de dados possuía 22 variáveis, sendo esses dados descritos na Tabela 1.

Variável	Descrição
Ano	Ano em que foi feito o registro
DataManut	Data da manutenção na máquina
LabRespons	Laboratório
Codigo_Descricao	Descrição (nome) da máquina
Codigo	Código (número) do tipo de máquina
Numero Controle Lab	Código laboratório
Serie	Número de série
SerieLimpo	Número de série com remoção de espaços
DataFabric	Data de fabricação do equipamento
Chave	Chave unificada do equipamento (Código junto com série)
DiasManut	Número de dias desde a fabricação aproximada do item
NumeroOP	
NumeroOS	
Locomotiva	Número da locomotiva da máquina
DescricaoFalha	Variável que descreve a falha da máquina
DescricaoReparo	Descrição do que foi feito no reparo da máquina
FichaMaterialRecuperar	Número da ficha do material a recuperar
FichaMaterialRecuperado	Número da ficha do material que foi recuperado
Subcomponente	
Falha	Tipo de falha da máquina
FalhaLimpo	Variável criada a partir da variável Falha de maneira padronizada
Status	Status do equipamento

Tabela 1. Descrição do Banco de Dados.

Um dos primeiros passos foi fazer uma análise descritiva no conjunto de dados. Com isso, percebemos que precisaríamos criar/mexer em algumas colunas, como a coluna **Status**, que foi mudada através das informações da coluna **DescricaoReparo**, já que alguns itens que tinham como Status "Baixado"(sucateado), continuava aparecendo no nosso banco de dados.

Após essa tratativa e a análise descritiva, tomamos a decisão de fazer uma Análise de Sobrevivência e Confiabilidade, mais especificamente utilizar a teoria dos Modelos de Sistemas Reparáveis. Para isso, foi desenvolvido um código em R utilizando majoritariamente o pacote *survival*. Além disso, nesse mesmo código, foram criadas funções para o cálculo da taxa de falha de uma determinada peça.

Por fim, também foi desenvolvido um relatório automatizado de alguns resultados e do conjunto de dados no Power BI.

4 Resultados e discussões

Para a implementação do modelo as peças foram categorizadas em 6 grupos, sendo cada um composto segundo um tipo de peça. Considerando a presença de dados censurados, casos nos quais a peça apresentou apenas uma única falha e não retornou para a manutenção, adotamos a seguinte função de verossimilhança:

$$L(\beta, \eta, \theta | m, T) = \prod_{i=1}^N \prod_{j=1}^{n_i} f(t_{i,j} | t_{i,j-1}, \dots, t_{i,1})^{\delta} R(t_{i,j} | t_{i,j-1}, \dots, t_{i,1})^{1-\delta}$$

Na qual $\delta = 1$ representa a presença de um evento de falha no tempo $t_{i,j}$ e $\delta = 0$ indica que até o momento não foi identificado nenhum evento de falha. Informação que poder ser sumarizada na Tabela 2 na qual podemos observar os eventos de falhas distribuídos no tempo.

Índice	Nome da Peça	Tempo	Tempos Anteriores	δ
1	FL08080101	342	0	1
1	FL08080101	4947	342	0
2	FL10064357	1365	0	1
2	FL10064357	1596	1365	1
2	FL10064357	1603	1365, 1596	1
2	FL10064357	3125	1365, 1596, 1603	1
2	FL10064357	3678	1365, 1596, 1603, 3125	1
2	FL10064357	4278	1365, 1596, 1603, 3125, 3678	0
...

Tabela 2. Tabela de falhas ao longo do tempo.

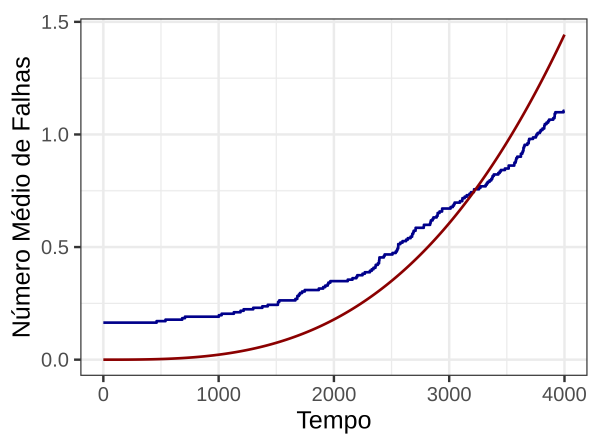
Na Tabela 2 temos a coluna nome da peça, a coluna tempo, que representa a contagem do tempo até o evento registrado seja falha ou não registro, a coluna tempos anteriores, representando o tempo da falha anterior, caso o evento de falha já tenha ocorrido anteriormente. A coluna delta classifica o evento entre zero e um, um representando a presença de falha no tempo decorrido e zero caso contrário. A Tabela 3 apresenta o diagnóstico do modelo estimado e a Figura 1 para avaliarmos o ajuste do modelo aos dados.

Código	m	β	$EP(\beta)$	η	$EP(\eta)$	θ	$EP(\theta)$
R571201097	6	2.0921	0.1038	2931.493	94.8098	1.0000	0.0017
R576810010	2	1.9596	0.0524	2186.388	51.9897	1.0000	0.0003
R576825112	1	1.6057	0.0871	1685.568	77.0812	0.7886	0.1768
R576825093	6	2.3725	0.1207	3758.135	108.7710	1.0000	0.0005
R576825106	1	2.6000	0.2016	2529.857	100.3434	0.5435	0.0720
R571201096	2	3.0215	0.1733	3543.094	87.8251	1.0000	0.0000

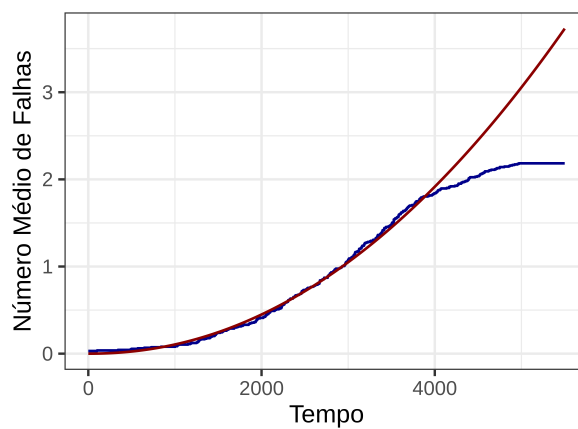
Tabela 3. Diagnóstico do modelo estimado

Para cada tempo t , obtém-se um número médio de falhas das peças que ocorreram até t , estimando a curva empírica dos dados. A construção da curva da média da integral do risco da última falha até o tempo t e o risco acumulado prevê o número de falhas das peças.

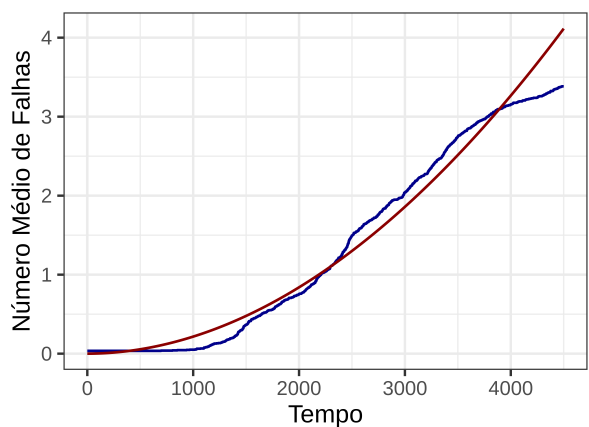
Considerando a Figura 1 identificamos que o ajuste do modelo, curva vermelha, apresenta um número preditos de eventos falhas superior ao número de falhas real, curva azul, ao longo do tempo. A principal hipótese para essa diferença seria o fato de que a base de dados apresenta um número considerável de peças com apenas uma falha ou poucas falhas no período, contribuindo negativamente para o ajuste do modelo considerando que na teoria os valores deveriam se estabilizar para grandes tamanhos amostrais de cada peça.



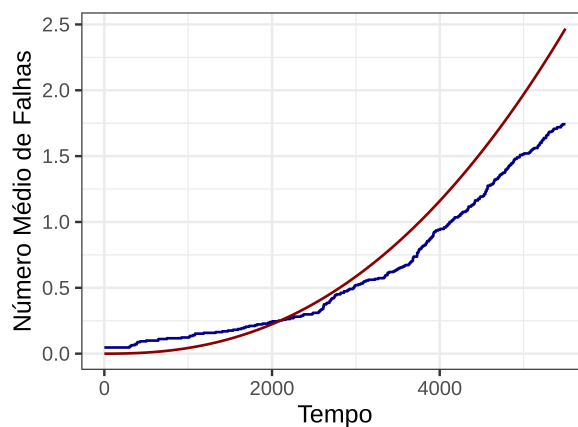
(a) R571201096



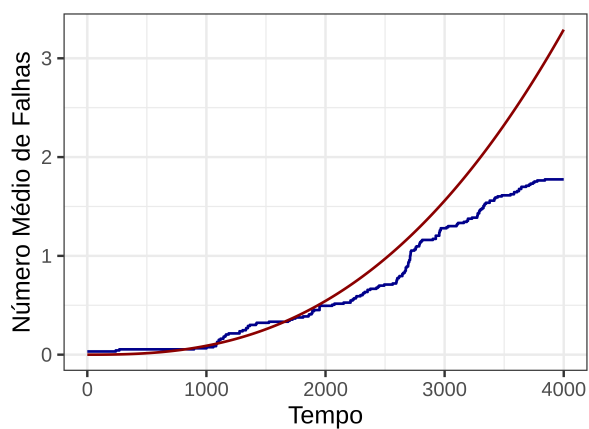
(b) R571201097



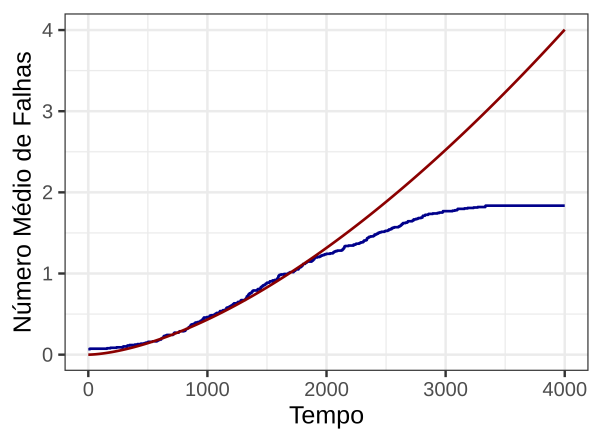
(c) R576810010



(d) R576825093



(e) R576825106



(f) R576825112

Figura 1. Gráficos de número médio de falhas ao longo do tempo

Contudo, podemos observar na Figura 1 que inicialmente a curva predita pelo modelo está próxima à curva empírica indicando que o modelo está consideravelmente ajustado nos valores iniciais. Isso indica que o modelo seria relativamente capaz de prever a distribuição das peças em tempos relativamente baixos, aumentando o grau de incerteza da previsão com o aumento do tempo.

Dado os resultados da modelagem para entrega dos resultados elaboramos um relatório automatizado (*Dashboard*) em *Power BI*, Figura 2, para facilitar a visualização e permitir que os técnicos possam validar a possibilidade de falha da peça considerando as modelagens propostas.

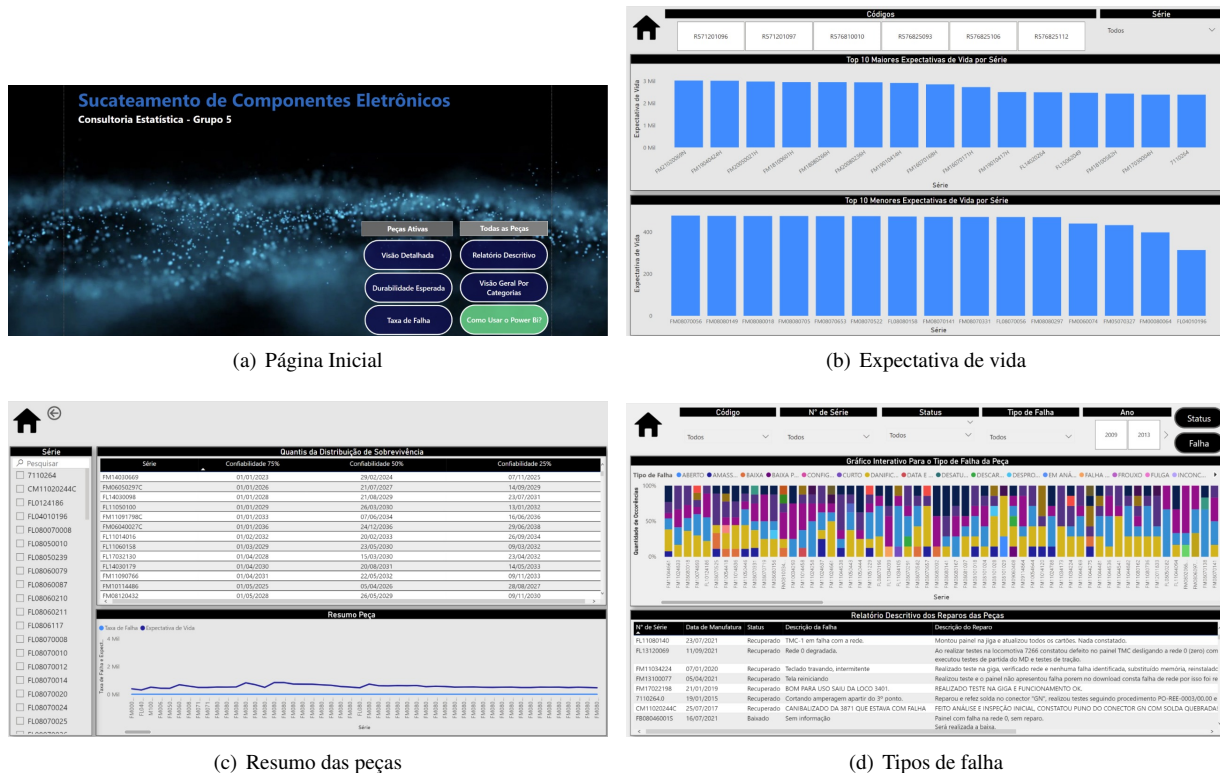


Figura 2. Entrega em *Power BI*

5 Conclusão

Deste modo, podemos concluir que foi necessário utilizar Análise de Sobrevivência e também empregado conceitos Sistemas Reparáveis para uma modelagem que resultou em uma variação do modelo ARA, conhecida por ARAm. Além disso, observamos que não podíamos confiar completamente no atributo de Status do banco de dados, o que nos levou a "corrigi-los", baseado nos valores dos outros atributos desta base de dados.

Apesar de não termos completado a automatização proposta inicialmente, desenvolvemos um relatório automatizado (*Dashboard*) em *Power BI*, que traz uma fácil e iterativa visualização dos dados pelo usuário, que pode analisar a recorrência das falhas e seus tempos, por exemplo.

Referências

1. Logística, M. Quem somos (2014).
2. Morettin, P. A. & de O. Bussab, W. Estatística Básica. *Editora Saraiva* (2010).
3. Colosimo, E. A. & Giolo, S. R. Análise de Sobrevivência Aplicada. *ABE - Projeto Fisher* (2006).
4. Bo Henry Lindqvist. On the Statistical Modeling and Analysis of Repairable Systems. *Statiscian Science* (2006).
5. Maria Luíza Guerra de Toledo, Marta A. Freitas, Enrico A. Colosimo, Gustavo L. Gilardoni. ARA and ARI imperfect repair models: Estimation, goodness-of-fit and reliability prediction. *Elsevier - Reliability Engineering and System Safety* (2015).