

METODOLOGIA DE ANÁLISE INTEGRADA PARA PREDIÇÃO DE INTERVENÇÕES EM JUNTAS ISOLANTES COLADAS NA ESTRADA DE FERRO CARAJÁS

Allan Maciel

Resumo

A Estrada de Ferro Carajás (EFC) destaca-se entre as ferrovias brasileiras pelo volume de transporte e a moderna tecnologia aplicada em seus ativos e processos, projetada em bitola larga, é parte integrante da Companhia Vale, interligando a província mineral de Carajás, no Pará, com o porto de Ponta da Madeira, em São Luís, Maranhão. Nas vias férreas da EFC, circulam um dos maiores trens com mundo, com 330 vagões e 3,3 km de extensão, e estima-se que sejam transportados 120 milhões de toneladas de carga por ano, sendo o minério de ferro, o principal produto. Além disso, são movimentadas cerca de 350 mil pessoas anualmente pela via.

Um dos grandes desafios de toda malha ferroviária, é sem dúvidas, a minimização de ocorrências que geram perdas de capacidade produtiva, desde pequenas interrupções até interrupções catastróficas. Atualmente, a maior parte dos investimentos em ativos e processos são direcionados para redução de riscos de impactos aos negócios e comunidade. Partindo deste contexto, é necessário que cada vez mais, sejam aplicadas ferramentas de análises avançadas para antecipar eventos e impactos no sistema produtivo. O presente artigo descreve um modelo de análise baseado na integração e correlação de defeitos de via permanente com o intuito de prever os serviços necessários em juntas isolantes coladas, tais como esmerilhamento, correção geométrica e a substituição do ativo. Para a construção deste modelo, foram utilizadas ferramentas de big data analytics aplicadas aos dados de inspeção de geometria de via e de trilhos, fundamentos de machine learning, validações e medições em campo e dados históricos de manutenção. Após as validações do modelo, foram construídos painéis de gestão para acompanhamento das condições dos ativos com integração ao sistema de manutenção da via permanente dando suporte para o planejamento da manutenção.

Palavras-chave: manutenção preditiva, juntas isolantes coladas, ferrovia, machine learning, big data analytics

Abstract

The Carajás Railroad (EFC) stands out among Brazilian railways for its transport volume and the modern technology applied to its assets and processes, designed in broad gauge, it is an integral part of Vale company, interconnecting the mineral province of Carajás, in the Pará, with the port of Ponta da Madeira, in São Luís, Maranhão. One of the largest trains in the world runs on the EFC railroads, with 330 wagons and 3.3 km in length, and it is estimated that 120 million tons of cargo are transported per year, with iron ore being the main product. In addition, around 350,000 people move annually along the route.

One of the great challenges of any railway is undoubtedly the minimization of occurrences that generated losses in production capacity, from small to continuous catastrophic ones. Currently, most investments in assets and processes are directed towards reducing the risk of impacts on business and the community. Based on this context, it is increasingly necessary to apply advanced analysis tools to anticipate events and impacts on the production system. This article describes an analysis model based on the integration and correlation of defects in order to predict the services required in insulating joints, such as grinding, tamping, and asset replacement. To build this model, big data analytics tools applied to track geometry and rail inspection data, machine learning fundamentals, field validations and measurements, and historical maintenance data were used. After model validation, management panels were built to monitor asset conditions with integration with the railway maintenance system, supporting maintenance planning.

Keywords: predictive maintenance, insulated joints, railway, machine learning, big data analytics

1 INTRODUÇÃO

A Estrada de Ferro Carajás (EFC) representa um dos mais relevantes corredores logísticos do Brasil, destacando-se pelo elevado volume de transporte de minério de ferro e pela adoção de tecnologias avançadas na operação de seus ativos ferroviários. Projetada em bitola larga e operada pela Companhia Vale, a ferrovia conecta a província mineral de Carajás, no estado do Pará, ao porto de Ponta da Madeira, em São Luís, Maranhão. Com composições ferroviárias que atingem 3,3 km de extensão e movimentam cerca de 120 milhões de toneladas de carga por ano, a EFC impõe demandas severas à infraestrutura da via permanente, exigindo elevados padrões de confiabilidade, segurança e disponibilidade.

As juntas isolantes são componentes instalados entre trilhos com a função de segmentar eletricamente os circuitos de via, permitindo a detecção de ocupação da linha por trens e viabilizando o funcionamento dos sistemas de sinalização e controle de tráfego. Além de sua importância operacional, as JIC estão sujeitas a intensos esforços mecânicos e ciclos de carga dinâmicos, o que acelera sua degradação e torna sua manutenção especialmente desafiadora. A confiabilidade das juntas isolantes coladas (JIC) em ferrovias sinalizadas tem sido objeto de crescente atenção na literatura técnica, dada sua relevância para a segurança e eficiência operacional. Falhas em juntas isolantes podem comprometer não apenas a segurança da circulação, como também ocasionar paralisações operacionais e aumento dos custos de intervenção emergencial. Tressia et al. (2022) realizaram uma análise detalhada dos trilhos utilizados em regiões de JIC na Estrada de Ferro Carajás, identificando defeitos recorrentes como *spalling* e *lipping*, e propondo melhorias na geometria e nos materiais para mitigar falhas elétricas e

mecânicas. A conformidade normativa e os cuidados na instalação também foram destacados como fatores críticos para o desempenho das JIC, conforme discutido por Revista AdNormas (2022).

Diante da complexidade e da criticidade desses ativos, estratégias tradicionais de manutenção corretiva ou preventiva não têm se mostrado suficientes para garantir a continuidade operacional em ambientes de alta demanda. É nesse cenário que desponta a manutenção baseada em condição, fundamentada na análise contínua de dados operacionais, registros históricos e sinais de deterioração física. Essa abordagem visa

antecipar falhas potenciais e priorizar intervenções conforme o risco e a condição real dos componentes.

A manutenção baseada em condição (*Condition-Based Maintenance – CBM*), utiliza dados operacionais e históricos para antecipar falhas e otimizar intervenções e segundo Garcia et al. (2022), técnicas de Inteligência Artificial como Redes Neurais Artificiais (ANN) e Máquinas de Vetores de Suporte (SVM) vêm sendo aplicadas com sucesso em sistemas ferroviários para esse fim.

A aplicação de Big Data Analytics e Machine Learning permite processar grandes volumes de dados heterogêneos, como medições ultrassônicas, inspeções geométricas e

registros de manutenção. Queiroz & Santos (2024) demonstram que essas tecnologias são capazes de construir modelos preditivos robustos, contribuindo para a redução de falhas e aumento da disponibilidade da via permanente. No caso das JIC, a integração dos dados de monitoramento e históricos de manutenção permite não apenas monitorar sua condição, mas também prever com antecedência a necessidade de esmerilhamento, ajuste geométrico ou substituição do ativo.

Assim, o presente trabalho propõe uma metodologia de análise integrada voltada à predição de intervenções em juntas isolantes coladas da EFC, utilizando como base os fundamentos de Big Data Analytics e Machine Learning, com foco na otimização da estratégia de manutenção baseada em condição e na maximização da disponibilidade da via permanente.

2 METODOLOGIA

A metodologia proposta neste estudo visa desenvolver uma abordagem integrada para análise e predição de falhas em juntas isolantes coladas (JIC) em ferrovias sinalizadas, com foco na manutenção baseada em condição (CBM). A seguir, são descritos os fundamentos teóricos e operacionais que sustentam cada etapa da metodologia.

2.1. Caracterização de Defeitos em Juntas Isolantes Coladas (JICs)

As juntas isolantes coladas (JIC) são componentes críticos em ferrovias sinalizadas, responsáveis por garantir a separação elétrica entre seções da via, permitindo o funcionamento dos circuitos de sinalização e a detecção da presença de trens. Devido à sua localização estratégica e à exposição contínua a esforços dinâmicos, as JIC estão sujeitas a diversos mecanismos de degradação que comprometem sua integridade mecânica e elétrica.

A caracterização dos defeitos em JIC envolve a identificação, classificação e análise dos principais tipos de falhas observadas em campo e em laboratório. Segundo Tressia et al. (2022), os defeitos mais recorrentes incluem:

Spalling: perda de material na superfície do trilho, geralmente nas extremidades próximas à junta, causada por tensões de contato excessivas e fadiga superficial. Pode evoluir para trincas profundas e comprometer a colagem do isolante.

Lipping: formação de bordas deformadas nas extremidades dos trilhos, resultante do escoamento plástico do metal sob carga repetitiva. Esse defeito aumenta a concentração de tensões e pode gerar falhas elétricas por contato indevido com o isolante.

Escoamento metálico: deformação progressiva das extremidades dos trilhos, levando à perda de paralelismo e à redução da eficácia da colagem. Está associado à baixa dureza do trilho e à geometria inadequada da região de colagem.

Trincas longitudinais e transversais: fissuras que se propagam a partir de pontos de concentração de tensão, podendo atingir o isolante e comprometer a separação elétrica. São detectadas por inspeções ultrassônicas e análise metalográfica.

Descolamento do isolante: falha na adesão entre o perfil isolante e o trilho, causada por contaminação na superfície, aplicação incorreta do adesivo ou envelhecimento do material. Pode resultar em curto-circuito entre seções da via.

A caracterização desses defeitos é realizada por meio de inspeções visuais, medições de dureza, microscopia eletrônica de varredura (MEV), ensaios ultrassônicos e análise de histórico de intervenções. A adoção de trilhos de maior dureza, ajustes na geometria das extremidades e melhorias no processo de colagem são medidas recomendadas para mitigar a ocorrência dessas falhas.

Essa etapa é fundamental para alimentar os modelos preditivos com dados confiáveis e representativos, permitindo que algoritmos de Machine Learning

identifiquem padrões de degradação e antecipem a necessidade de intervenção com maior precisão.

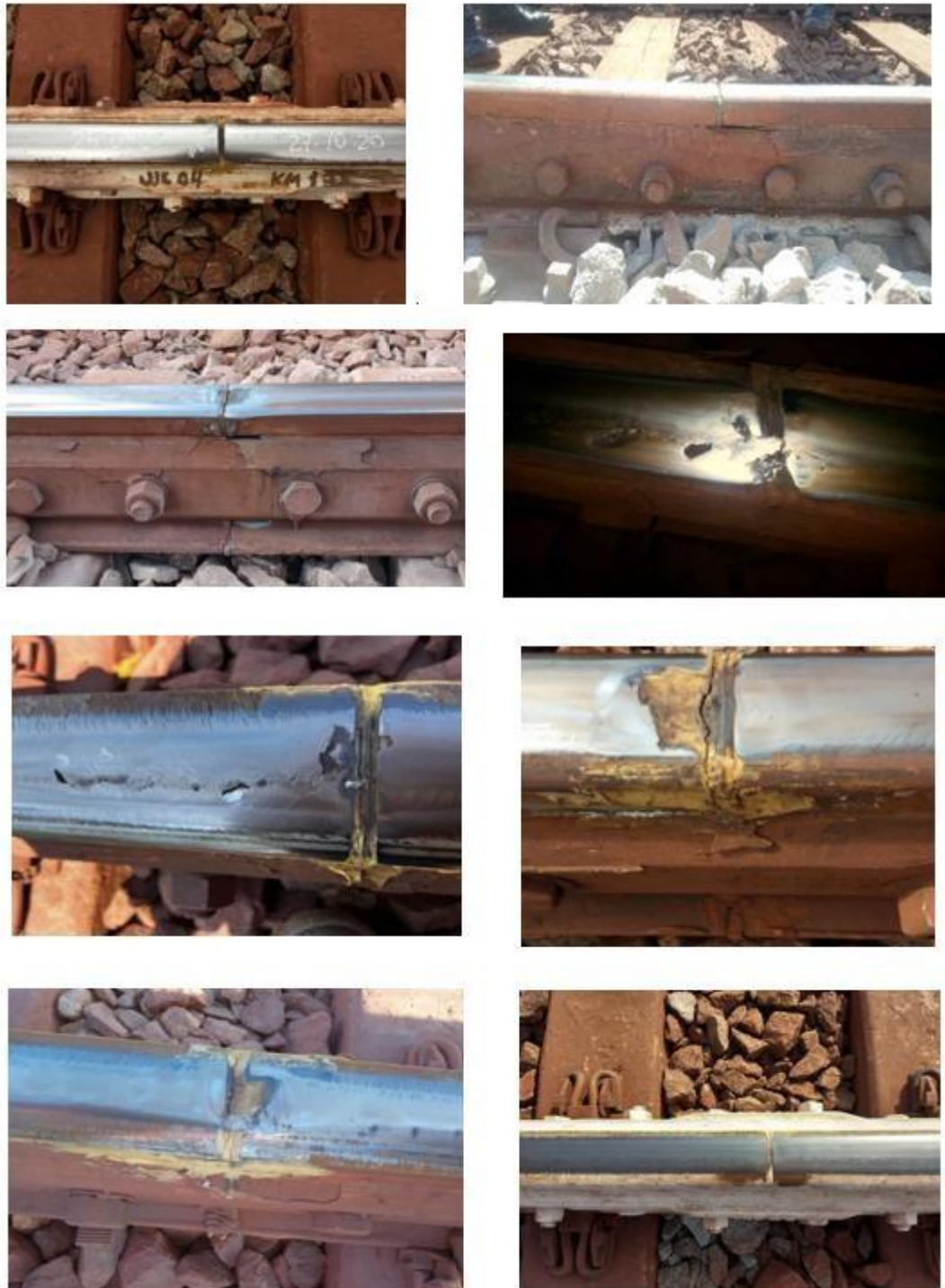


Figura 1 - Defeitos em Juntas Isolantes Coladas

2.2 Manutenção Baseada em Condição (CBM) aplicada às Juntas Isolantes

Coladas

A manutenção baseada em condição (CBM) é uma estratégia que visa otimizar a confiabilidade de ativos por meio do monitoramento contínuo de parâmetros operacionais, permitindo intervenções apenas quando há evidências de degradação funcional. No contexto das juntas isolantes coladas (JIC), essa abordagem é especialmente relevante devido à sua função crítica na separação elétrica entre seções de trilhos e à exposição constante a esforços mecânicos e ambientais.

A CBM para JIC envolve a coleta de dados como variações de resistência elétrica, vibração de passagem, deslocamentos verticais e temperatura. Esses dados são processados por algoritmos de diagnóstico e prognóstico, como redes neurais convolucionais (CNN) e modelos de sobrevivência de *Weibull*, que estimam a vida útil remanescente da junta com base em padrões históricos e condições reais de operação. Essa abordagem permite reduzir falhas inesperadas, otimizar recursos de manutenção e aumentar a disponibilidade da via.

Além disso, a CBM se integra a estratégias como a Manutenção Centrada na Confiabilidade (RCM), que identifica funções críticas e modos de falha, e à norma ISO 13374, que define requisitos para processamento e apresentação de dados de monitoramento. Estudos como os de Carretero et al. (2014) demonstram a eficácia da CBM na detecção precoce de falhas em JIC por meio da análise de resposta vibracional, enquanto Esveld (2001) destaca a importância da manutenção inteligente em trilhos modernos.

2.3 Aquisição e Integração de Dados

A aquisição e integração de dados representam o processo de coleta estruturada e consolidação de informações provenientes de diferentes sistemas, sensores e registros operacionais. Esse processo é essencial para garantir a confiabilidade, rastreabilidade e profundidade das análises realizadas sobre o estado dos ativos ferroviários. A junção coerente e temporal dessas fontes permite compor uma visão holística sobre o desempenho e degradação das juntas isolantes coladas, viabilizando modelos preditivos mais robustos e decisões de manutenção mais assertivas.

Fonte	Equipamento / Sistema	Periodicidade	Principais Variáveis	Volume Médio
Inspeção geométrica, desgaste e vídeo analytics	Carro-controle EM100 (Plasser & Theurer)	Mensal	perfil transversal, alinhamento, nivelamento, batente, desgaste de trilho	≈ 200 GB/mês
Inspeção Sensitiva	Aplicativo Próprio (PowerApps)	Mensal	condição física das juntas	≈ 1 GB/mês
Inspeção ultrassônica	Veículo rodoferroviário CATER UltraWave	Mensal	eco de reflexão, perda de atenuação, profundidade de trinca	≈ 3 GB/mês
Dados operacionais	Plataforma proprietária + módulo GPV (Gestão e Performance de Viagens)	Em tempo real	carga por eixo, velocidade, temperatura de roda, contagem de passagens	≈ 5 GB/mês
Histórico de manutenção	SAP PM + formulários internos (app móvel)	Contínua	tipo de intervenção, material aplicado, condição pós-serviço, data/km	10 anos (≈ 120 k registros)

Tabela 1 – Características das fontes de dados

2.3.1 Inspeções geométricas, desgaste e vídeo analytics

O carro controle EM100, fabricado pela Plasser & Theurer, percorre toda a via a 80 km/h, utilizando sensores inerciais, laser e perfis de contato para medir, em resolução de 0,25 m, parâmetros de geometria (nivelamento, alinhamento, bitola) e perda de seção de trilho. Os arquivos brutos são gerados em formato .dat (proprietário) e convertidos para .csv via software *Offboard* fornecido pelo próprio fabricante preservando metadados (quilômetro, horário, código das linhas). O equipamento possui sistema de vídeo com análise de dados da grade para detecção de anomalias na região dos trilhos (defeitos superficiais) e fixações.

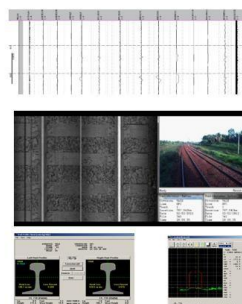


Figura 2 - Carro Controle Plasser & Theurer EM100

2.3.2 Inspeções ultrassônicas

Os veículos CATER US realizam varredura por ondas de contato direto, registrando ecos refletidos por descontinuidades internas. Cada registro contém posição GPS, profundidade da indicação e severidade. A exportação ocorre em .csv, que é normalizado para o modelo de dados relacional para integração com os demais dados monitorados.



Figura 3 - Equipamento Ultrassom rodoferroviário CATER

2.3.3 Inspeções sensitivas

Além dos métodos instrumentais, são realizadas inspeções sensitivas por técnicos capacitados, conforme protocolos da área de via permanente. Essas observações visuais e táteis são fundamentais para identificar:

- Descolamento parcial do isolante ou falhas na colagem
- Presença de lipping, spalling e escoamento metálico não capturados por sensores
- Trincas superficiais e deteriorações por impacto, abrasão ou desgaste lateral
- Condição do torque e fixação dos elementos metálicos

Essas inspeções ocorrem mensalmente em trechos críticos e são registradas em aplicativos móveis, alimentando a base de dados com evidências qualitativas e contexto operacional. A sensibilidade humana, aliada ao registro fotográfico e checklist técnico, complementa os dados automatizados com sinais precoces e nuances nem sempre capturadas por equipamentos.

2.3.3 Dados operacionais (GPV)

O sistema GPV registra os trens com os pesos de cada composição (vazias e carregadas) e esta informação diária é integrada aos dados de via como forma de MTBT

(milhões de toneladas bruta transportada) onde representa a utilização do ativo em cada CDV (circuito de via).

2.3.4 Histórico de manutenção (SAP + formulários)

Ordens de serviço originadas no SAP PM são extraídas via relatório SAP-BO, contendo código funcional, tipo de intervenção, mão de obra e materiais.

Os técnicos de via complementam as informações em formulários móveis (Power Apps), registrando condição visual da JIC antes e depois da intervenção (fotos e checklist de torque). Um processo de ETL consolida ambas as fontes num único fato “Intervenção_JIC” em banco SQL. Essa base retroalimenta os algoritmos com o histórico completo de intervenções por junta.

2.3.5 Integração e Governança

A etapa de integração e governança dos dados representa o elo fundamental entre os diversos sistemas de aquisição e o modelo analítico preditivo proposto neste estudo. Após a coleta das informações provenientes de inspeções geométricas, ultrassônicas, dados operacionais, registros sensitivos e históricos de manutenção, é necessário consolidar, organizar e qualificar esses dados em uma base única e confiável, acessível para análises técnicas e tomadas de decisão.

Esse processo é conduzido sob supervisão do Centro de Monitoramento de Ativos (CMA), área responsável por garantir a governança dos dados de infraestrutura ferroviária na Estrada de Ferro Carajás. O CMA atua na padronização de nomenclaturas, controle de integridade, tratamento de falhas de leitura, rastreabilidade dos registros e preservação histórica das intervenções. Além disso, é responsável pela implementação das políticas de segurança e conformidade digital, em alinhamento com a norma ISO 27001 e diretrizes corporativas de gestão de ativos.

Os dados coletados são armazenados em um ambiente de data lake distribuído, com particionamento por data, quilômetro e ativo específico. A padronização é feita por

metamodelos de referência, catalogados em ambiente em nuvem, permitindo consultas estruturadas e análises comparativas. A integração entre as fontes ocorre por meio de chave espacial-temporal, geralmente baseada na coordenada de quilômetro acrescida de um buffer de localização de até 25 metros. Isso viabiliza a junção precisa entre os dados de inspeção, registros operacionais e histórico de intervenções por junta.

Para garantir qualidade, são aplicadas regras de validação automatizadas — como verificação de formato, intervalos esperados, controle de registros nulos e duplicados. Periodicamente, o CMA realiza auditorias internas para certificar que os dados alimentam corretamente os painéis operacionais e os modelos de aprendizado de máquina, evitando viés estatístico e distorções na análise.

Além do uso técnico, a governança assegura que os dados possam ser utilizados como fonte confiável para auditorias regulatórias, planejamento de recursos e estudos de engenharia de longo prazo. A atuação do CMA é, portanto, um pilar estratégico para que os dados se tornem ativos corporativos de alto valor para a manutenção preditiva e a excelência operacional da ferrovia.

2.4. Aplicação de *Big Data Analytics* e *Machine Learning*

A aplicação de *Big Data Analytics* e técnicas de *Machine Learning* (ML) no contexto da ferrovia apresenta um salto qualitativo na forma como os dados operacionais e de inspeção são tratados, interpretados e convertidos em valor para a gestão de ativos. Considerando a volumetria, velocidade e variabilidade dos dados provenientes de sensores ultrassônicos, inspeções geométricas, históricos de manutenção e operação, é fundamental o uso de soluções analíticas escaláveis e inteligentes para gerar conhecimento útil em tempo hábil. Tradicionalmente, a análise desses dados era realizada de forma isolada, com baixa capacidade preditiva e alta dependência de conhecimento especializado. O uso de *Big Data* e ML surge como resposta para integrar, correlacionar e prever comportamentos anômalos com maior acurácia e rapidez.

As estratégias mais empregadas nesta abordagem estão alinhadas ao tipo de problema que se pretende resolver, e incluem modelos supervisionados, não supervisionados, híbridos e de séries temporais:

Modelos supervisionados, como *Random Forest*, *XGBoost* e *Support Vector Machines* (SVM), vêm sendo utilizados para a classificação preditiva de JIC com alto risco de falha, considerando variáveis como intensidade de carga por eixo, número de passagens, temperatura da roda e histórico de intervenções (GHOFRANI et al., 2018). Em paralelo, técnicas não supervisionadas como K-Means, DBSCAN e agrupamento hierárquico demonstram eficácia na detecção de padrões anômalos em regiões com degradação acelerada, mesmo sem registros prévios de falhas (AGGARWAL, 2016).

Além disso, estratégias baseadas em séries temporais, como os modelos *Long Short-Term Memory* (LSTM), ARIMA e Prophet, são empregadas para prever o deslocamento de isolantes, o aumento de rugosidade e a perda de resistência elétrica ao longo do tempo, a partir da evolução dos parâmetros monitorados (ZHANG; WANG, 2020). Modelos híbridos têm ganhado destaque ao integrar abordagens estatísticas, como a distribuição de Weibull, com redes neurais convolucionais (CNN) aplicadas a imagens térmicas, oferecendo maior robustez na classificação e prognóstico simultâneo de juntas com risco de falha interna (KUMAR; SINGH, 2019).

Essas estratégias são implementadas por meio de um pipeline analítico operado pelo Centro de Monitoramento de Ativos (CMA) e definido pela Confiabilidade de Via Permanente, assegurando rastreabilidade dos dados, atualização e desenvolvimento dos modelos e revalidação contínua das métricas de desempenho, conforme práticas de segurança definidas pela companhia.

2.5. Estratégia de Intervenção

A definição de estratégias de intervenção para juntas isolantes coladas na Estrada de Ferro Carajás é fundamentada em uma abordagem analítica orientada por dados, combinando informações operacionais, de inspeção e históricos de manutenção. Com base nos modelos preditivos desenvolvidos a partir de técnicas de *Big Data Analytics* e Machine Learning, são atribuídas classificações de risco a cada junta, considerando sua condição estrutural, histórico de degradação e projeção de vida útil remanescente.

A partir dessa avaliação técnica, cada JIC é enquadrada em um regime de intervenção otimizado, visando maximizar a disponibilidade da via permanente e

minimizar impactos sobre o cronograma operacional. Juntas com risco crítico e vida residual inferior à janela de planejamento são priorizadas para substituição programada, enquanto aquelas com sinais moderados de deterioração são encaminhadas para ações de manutenção preventiva, como esmerilhamento, ajuste geométrico ou ajuste na região do endpost. Juntas classificadas como estáveis seguem em regime de monitoramento contínuo.

As estratégias de intervenção são integradas ao sistema de gestão de ativos da ferrovia, permitindo que as ordens de serviço, alocação de recursos e definição de materiais sejam automatizadas com base em critérios técnicos e econômicos. Essa lógica decisória, gerenciada pelo Centro de Monitoramento de Ativos (CMA), assegura que cada ação de campo seja guiada por evidências analíticas, elevando a eficiência da manutenção e contribuindo diretamente para a segurança operacional da via.

3 DESENVOLVIMENTO E RESULTADOS

O desenvolvimento metodológico adotado neste estudo baseia-se em uma abordagem analítica robusta que contempla desde a coleta e arquitetura de dados até a modelagem preditiva e a implantação operacional dos resultados. Inicialmente, são reunidas diversas fontes de dados relacionadas à ferrovia, como inspeções de geometria de via realizadas com carro controle, inspeções ultrassônicas dos trilhos, inspeções sensitivas e registros históricos de ordens de serviço acumuladas ao longo de vários anos, totalizando cerca de 120 mil entradas. Esses dados, que apresentam uma taxa média mensal de geração de aproximadamente 15 GB (específicos das regiões da JICs), são armazenados em um Data Lake estruturado com particionamento por data-hora, quilômetro e via férrea. Os metadados são devidamente catalogados garantindo organização e acessibilidade.

A etapa de preparação e engenharia de atributos envolve o agrupamento espacial dos dados no entorno das juntas isolantes, com uma janela de 25 metros. Isso reduz ruídos e melhora a qualidade das análises. A partir desse agrupamento, são gerados 42 atributos que descrevem de forma abrangente cada conjunto “quadro-junta”, incluindo parâmetros como deflexão, a frequência de passagem, a contagem de falhas ultrassônicas nas categorias críticas, a idade das juntas e o histórico recente de intervenções. Para lidar com

o desbalanceamento entre classes de dados, é adotada a análise de refinamento por vizinhança

Na modelagem preditiva, o processo contempla a experimentação com múltiplos algoritmos de aprendizado de máquina, começando pela Random Forest como modelo de referência, seguido por XGBoost e redes neurais recorrentes (LSTM) voltadas à análise de séries temporais. A validação dos modelos é feita com divisão temporal (70% para treino e 30% para teste), garantindo que não haja “*look-ahead*”, ou seja, quando informações do futuro, que não estariam disponíveis no momento da previsão, são indevidamente utilizadas para treinar ou testar o modelo. Além disso, emprega-se a validação cruzada *k-fold* com divisão anual para reforçar a robustez estatística. Para a experimentação e elaboração dos modelos, foi utilizado o software KNIME Analytics Platform.

As métricas adotadas para aferição do desempenho incluem acurácia, precisão, recall, área sob a curva ROC-AUC. Por fim, o sistema preditivo é disponibilizado por meio de uma dashboard desenvolvida em Power BI para a visualização dos resultados e capaz de apresentar rankings de risco por junta, mapas de calor quilométricos e comparativos mensais entre falhas previstas e ocorrências reais.

O Diagrama abaixo mostra os passos do desenvolvimento do modelo:

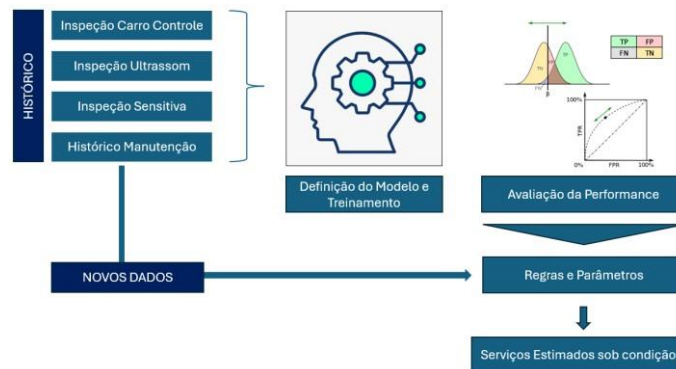


Figura 4 - Fluxograma elaboração do modelo preditor de serviços em juntas isolantes

Os resultados obtidos na avaliação comparativa dos modelos preditivos desenvolvidos para prognóstico de intervenções em juntas isolantes coladas (JIC) na Estrada de Ferro Carajás. Foram considerados três algoritmos amplamente reconhecidos na literatura especializada: *Random Forest*, *Gradient Boosting* (XGBoost) e Redes Neurais Recorrentes (LSTM).

A performance dos modelos foi aferida por meio de métricas tradicionais de classificação: acurácia global, recall específico da classe “Substituir”, área sob a curva ROC (ROC-AUC) e estimativa de economia anual associada à aplicação da metodologia preditiva. Os resultados são sintetizados na Tabela 2.

Modelo	Acurácia	Recall (classe “Substituir”)	ROC-AUC	Custo evitado anual (R\$ milhões)
Random Forest	0,79	0,62	0,85	8,3
XGBoost	0,87	0,74	0,92	11,5
LSTM	0,85	0,72	0,9	10,9

Tabela 2 - Desempenho comparativo dos modelos preditivos

O algoritmo XGBoost apresentou desempenho superior nas quatro métricas avaliadas, destacando-se como a abordagem mais eficiente tanto sob o ponto de vista técnico quanto econômico. Sua acurácia global de 87 % indica elevada confiabilidade na classificação dos eventos de manutenção. O recall da classe “Substituir” (0,74) revela sensibilidade expressiva na identificação de juntas com necessidade crítica de substituição. A métrica ROC-AUC, superior a 0,90, confirma a robustez do modelo na separação entre classes, minimizando falsos positivos e negativos, ou seja, o modelo elaborado possui excelente capacidade de discriminação entre as classes analisadas.

Adicionalmente, estimou-se que a adoção do XGBoost resultaria em uma economia anual de R\$ 11,5 milhões, considerando custos evitados com falhas não planejadas e intervenções corretivas prematuras. As regras definidas pelo modelo para cada tipo de serviço considerando parâmetros medidos pelos processos/sistemas de inspeção na EFC são apresentadas abaixo:

Serviços	Parâmetros
Esmerilhamento	RCF_RH > 500 mm Desgaste < 10 %
Correção Geométrica	Desvio Padrão Nvl > 1,6 mm Média Desvio Padrão Nvl (12M) > 1,6 mm Desgaste < 10 %

Substituição (Defeito)	RCF_RH > 1000 mm Desgaste > 10 %
Substituição (Fadiga)	Desvio Padrão Nvl > 2,0 mm Desgaste > 10 %
Substituição (Fadiga) - Precoce	Média Desvio Padrão Nvl (12M) > 1,6 mm Desgaste > 6%

Tabela 3 - serviços e parâmetros definidos para modelo preditor de manutenção em juntas isolantes.

Após a implantação do modelo e integração à rotina das manutenções condicionais, em um horizonte de 12 meses foram verificados os seguintes ganhos quantitativos e qualitativos:

- Redução de 28 % nas falhas inesperadas de JIC;
- Diminuição de 15 % nos custos totais de manutenção de juntas;
- Antecipação média de 90 dias nas ordens de serviço classificadas como críticas.
- Direcionamento assertivos de manutenção para aumento de vida útil dos ativos;

Esses resultados demonstram que, ao integrar dados heterogêneos e aplicar técnicas avançadas de aprendizado de máquina, é possível aumentar significativamente a eficiência da manutenção ferroviária, otimizando recursos e preservando a confiabilidade operacional da via permanente.

4 CONCLUSÃO

A aplicação coordenada de técnicas de *Big Data Analytics* e *Machine Learning* revelou elevada eficácia na predição de intervenções em juntas isolantes coladas (JIC) ao

longo da malha da Estrada de Ferro Carajás (EFC). A metodologia proposta promoveu a unificação de fontes de dados até então segmentadas — incluindo registros operacionais, inspeções físicas e históricos de manutenção — e possibilitou a geração de atributos de alto valor analítico, fundamentais para o desempenho preditivo dos modelos. A arquitetura analítica permitiu o desenvolvimento e implantação de um modelo supervisionado com acurácia de 87%, validado por métricas como ROC-AUC, configurando uma ferramenta confiável para apoiar decisões de manutenção.

Os ganhos práticos evidenciaram-se na redução de intervenções corretivas não planejadas, otimização da alocação de recursos e melhoria da disponibilidade da via, o que se traduziu em impactos econômicos tangíveis e fortalecimento da segurança operacional. A abordagem baseada em dados mostrou-se uma alternativa tecnicamente superior frente aos modelos tradicionais baseados em calendários fixos ou inspeções periódicas isoladas.

Para avanços futuros, deve-se incorporar variáveis meteorológicas, expansão da instrumentação com sensores embarcados em vagões (viabilizando monitoramento contínuo), nova geração de veículos de inspeção e abrangência da metodologia preditiva aos demais ativos de via permanente, potencializando a abrangência e a eficiência da manutenção ferroviária baseada em condição.

REFERÊNCIAS

AGGARWAL, Charu C. *Outlier Analysis*. 2. ed. Cham: Springer International Publishing, 2017. ISBN 978-3-319-47577-6. DOI: 10.1007/978-3-319-47578-3.

Disponível em: <https://www.sadbhavanapublications.org/research-enrichment-material/2Statistical-Books/Outlier-Analysis.pdf>. Acesso em: 04 ago. 2025.

CARRETERO, D.; WESTON, P.; ROBERTS, C. Condition monitoring of insulated rail joints using vibration response parameters. Proceedings of the IMechE, Part F, 2014. • Esveld, C. Modern Railway Track. 2ª ed., MRT-Productions, 2001. • Mobley, R. K. Maintenance Engineering Handbook. 9ª ed., McGraw-Hill, 2020. • Moubray, J. Reliability-Centered Maintenance. 2ª ed., Industrial Press, 1997. • ISO 13374-1:2019. Condition monitoring and diagnostics of machines — Data processing, communication and presentation.

ESVELD, C. Modern Railway Track. MRT-Productions, 2nd ed., 2001.

GARCIA, Pedro; VISMARI, Lucio Flavio; CUGNASCA, Paulo. Inteligência artificial aplicada à manutenção baseada em condição: uma revisão sistemática da literatura para sistemas ferroviários. In: Anais do V Simpósio de Engenharia Ferroviária, 2022. Campinas: Galoá. Disponível em: <https://proceedings.science/sef/sef-2022/trabalhos/inteligencia-artificial-aplicada-a-manutencao-baseada-em-condicaoouma-revisao-si>. Acesso em: 04 ago. 2025.

GHFRANI, Faeze; YOUSEFIANMOGHADAM, Seyedsina; HE, Qing; STAVRIDIS, Andreas. Rail service failure prediction: an integrated approach using fatigue modeling and data analytics. In: TAIM 2019 – Transportation Analytics and Infrastructure Management. State College, PA, outubro 2019. Disponível em: <https://www.taim.psu.edu/assets/docs/rail-service-failure.pdf>. Acesso em: 04 ago. 2025.

ISO 13374-1:2019. “Condition monitoring and diagnostics of machines—Data processing, communication and presentation.”

MOBLEY, R. K. Maintenance Engineering Handbook, 9th ed. McGraw-Hill, 2020.

MOUBRAY, J. Reliability-Centered Maintenance, 2nd ed. Industrial Press, 1997.

QUEIROZ, Eduardo do Nascimento; SANTOS, Victor Otto dos. Aplicação da inteligência artificial na manutenção preditiva: revisão bibliográfica. Pindamonhangaba: Faculdade de Tecnologia de Pindamonhangaba, 2024. Disponível em: https://ric.cps.sp.gov.br/bitstream/123456789/29427/1/manutencaoindustrial_2024_1_e_duardonascimentoqueiroz_victorottodossantos_aplicacaodainteligenciaartificialna.pdf. Acesso em: 04 ago. 2025.

REVISTA ADNORMAS. A conformidade das juntas isolantes coladas em vias férreas. São Paulo: Revista AdNormas, 2022. Disponível em: <https://revistaadnormas.com.br/2022/12/06/a-conformidade-das-juntas-isolantescoladas-em-vias-ferreas>. Acesso em: 04 ago. 2025.

SANTOS NETO, Manoel Ferreira; LEITE, Denisson Santana; NASCIMENTO, Willem Vieira. Revisão bibliográfica da manutenção preditiva e seus conceitos de tecnologia atrelados à Indústria 4.0. Aracaju: Universidade Tiradentes, 2023. Disponível em: <https://ri.ufs.br/bitstream/riufs/10428/2/RevisaoBibliograficaManutencao.pdf>. Acesso em: 04 ago. 2025.

SINGH, Dippu Kumar. Railway Revolution – AI-Driven Network Asset Change Detection for Infrastructure Excellence. International Journal of Innovative Research in Science, Engineering and Technology, v. 12, n. 12, p. 14995–15007, 2023. Disponível em: <https://philarchive.org/rec/DIPRRI>. Acesso em: 04 ago. 2025.

TRESSIA, Gustavo; PRITZELWITZ, Philip; ALVES, Luiz Henrique Dias. Análise de trilho utilizado em regiões de JIC. Ouro Preto: Instituto Tecnológico Vale, 2022. (Relatório Técnico – PROD.TEC.ITV.MI – N003/2022). DOI: 10.29223/PROD.TEC.ITV.MI.2022.3.Tressia.

VOESTALPINE RAILWAY SYSTEMS. Junta isolada colada NRG. Disponível em: <https://www.voestalpine.com/railway-systems/pt/products/junta-isolada-colada-nrg>. Acesso em: 04 ago. 2025.

WANG, Qi; BU, Siqi; HE, Zhengyou. Achieving predictive and proactive maintenance for high-speed railway power equipment with LSTM-RNN. IEEE Transactions on Industrial Informatics, v. 16, n. 10, p. 6509–6517, out. 2020. DOI: 10.1109/TII.2020.2966033. Disponível em: https://ira.lib.polyu.edu.hk/bitstream/10397/100518/1/Wang_Achieving_Predictive_Proactive.pdf. Acesso em: 04 ago. 2025.