

Análise de Séries Temporais de Acidentes nas Rodovias Federais de Pernambuco

Time Series Analysis of Accidents on Federal Highways in Pernambuco

Welton L. Santos¹
orcid.org/0009-0009-0650-3614

Paulo Victor¹
orcid.org/0009-0007-3595-3839

Andina Lerma¹
orcid.org/0000-0001-7904-7773

Durval Neto¹
orcid.org/0009-0003-4273-6104

¹Escola Escola Politécnica de Pernambuco, Universidade de Pernambuco, Recife, Brasil. E-mail: wlgs@ecomp.poli.br

DOI: [10.25286/repa.v11i1.3532](https://doi.org/10.25286/repa.v11i1.3532)

Esta obra apresenta Licença Creative Commons Atribuição-Não Comercial 4.0 Internacional.

Como citar este artigo pela NBR 6023/2018: Welton L. Santos; Paulo Victor; Andina Lerma; Durval Neto. Análise de Séries Temporais de Acidentes nas Rodovias Federais de Pernambuco Revista de Engenharia e Pesquisa Aplicada, v.11, n. 1, p. 31-40, 2026.

RESUMO

A pesquisa sobre acidentes de trânsito no Brasil, com ênfase nas rodovias federais de Pernambuco, tem como objetivo analisar e investigar acidentes por meio de análise de dados. O propósito central é prever eventos futuros em diferentes tipos de pistas — simples, duplas e múltiplas — com base em dados históricos de 2017 até fevereiro de 2024. O estudo busca apoiar o desenvolvimento de estratégias e fundamentar a tomada de decisões baseadas em métricas, utilizando um protocolo experimental com o uso de técnicas e modelos de análise de dados, como Isolated Forest, Prophet e SARIMAX. Cada etapa envolve a suavização dos dados com técnicas de média e mediana móvel para gerar novas métricas, identificando a melhor configuração para cada tipo de pista.

PALAVRAS-CHAVE: temporais; Acidentes em rodovias; Mineração de dados.

ABSTRACT

The research on traffic accidents in Brazil, with an emphasis on federal highways in Pernambuco, aims to analyze and investigate accidents through data analysis. The main purpose is to predict future events on different types of lanes — single, double and multiple — based on historical data from 2017 to February 2024. The study seeks to support the development of strategies and support decision-making based on metrics, using an experimental protocol using data analysis techniques and models, such as Isolated Forest, Prophet and SARIMAX. Each step involves smoothing the data with moving average and median techniques to generate new metrics, identifying the best configuration for each type of lane.

KEY-WORDS: Time series; Highway accidents; Data mining.

1 INTRODUÇÃO

1.1 CONTEXTUALIZAÇÃO

O Brasil possui uma extensa malha viária federal, composta por mais de 70 mil quilômetros de rodovias sob fiscalização da Polícia Rodoviária Federal (PRF) [1]. Devido à predominância do transporte rodoviário, acidentes de trânsito são uma realidade frequente. Dados do Datasus (Departamento de Informática do Sistema Único de Saúde) indicam que cerca de 33 mil pessoas morreram (2022) em acidentes de transporte terrestre no Brasil [2], enquanto, em média, 642 pessoas foram internadas por dia na rede pública de saúde com lesões causadas por sinistros de trânsito (2021) [3]. Este cenário reforça a urgência de abordar a segurança no trânsito. A ONU, em 2021, lançou a Segunda Década de Ação pela Segurança no Trânsito (2021-2030) [3], com a meta de reduzir pela metade as mortes e feridos no trânsito. O Brasil, signatário desta iniciativa, enfrenta o desafio de cumprir essa meta.

1.2 DESCRIÇÃO DO PROBLEMA

O Estado de Pernambuco enfrenta desafios significativos na segurança de suas rodovias federais. Com uma malha de aproximadamente 2 mil quilômetros [4] e pouco mais de 500 Policiais Rodoviários Federais em atividade [5], a capacidade de fiscalização é insuficiente. A Polícia Rodoviária Federal (PRF) detém uma extensa base de dados de acidentes de trânsito [6]. Esta vasta quantidade de informações, que acompanha detalhadamente cada acidente, representa um recurso valioso para a análise e prevenção de acidentes graves.

1.3 OBJETIVOS

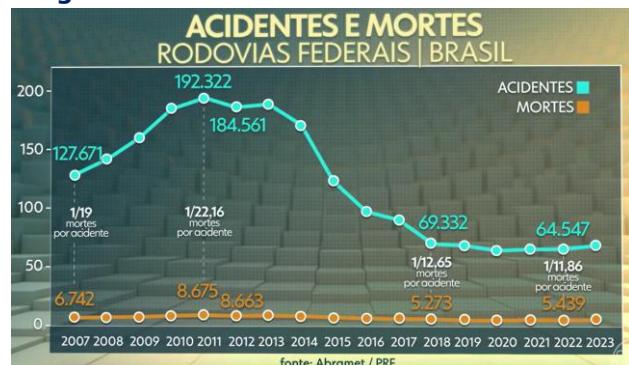
Este artigo visa o estudo dos acidentes graves nas rodovias federais de Pernambuco, buscando entender e perceber padrões com base nos dados. Utilizando técnicas de mineração de dados, busca-se identificar fatores de risco e padrões que contribuem para os acidentes. A exploração da extensa base de dados da PRF permitirá prever locais e principais causas de acidentes, a exemplo das pistas do tipo **simples** (BR 423 - Garanhuns), **dupla** (BR 232 - entre Recife e Caruaru) e **múltipla** (BR 101 - Abreu e Lima). Ademais, é importante

identificar áreas críticas e otimizar os recursos de fiscalização. Com base nos *insights* obtidos, poderão ser promovidas campanhas de conscientização a públicos direcionados.

1.4 JUSTIFICATIVA

Nos últimos 15 anos, a proporção de mortes em relação ao número de acidentes aumentou 50%, embora a quantidade de acidentes tenha caído [7]. Acidentes de trânsito sobrecarregam o sistema de saúde, ocupando mais de 60% dos leitos do SUS [8].

Figura 1 - Os acidentes nas rodovias do Brasil.



Fonte: [7]

Em Pernambuco, o custo com vítimas de trânsito chega a R\$6 bilhões anuais, 4,8% do gasto nacional. Desse total, R\$1,2 bilhão é gasto com pedestres, R\$3 bilhões com motociclistas e R\$1 bilhão com ocupantes de automóveis, conforme o Relatório do Observatório Nacional de Segurança Viária de 2022 [9].

No Estado, temos um cenário de rodovias federais do tipo simples, dupla e múltipla, onde 57% dos acidentes graves (lesões graves ou óbitos) ocorrem em pistas do tipo simples e 38% em pistas do tipo dupla.

1.5 ESCOPO NEGATIVO

O estudo focou na análise dos dados de acidentes de trânsito das rodovias federais a partir de 2017, excluindo dados de crimes e infrações de trânsito. A área geográfica será restrita ao Estado de Pernambuco, excluindo vias estaduais e municipais.

2 FUNDAMENTAÇÃO TEÓRICA

Acidente de trânsito é todo fato não premeditado que produz lesões em pessoas ou danos patrimoniais, ou ambientais como consequência da movimentação de veículos. São passíveis de registro pela PRF todos os acidentes de trânsito ocorridos em sua circunscrição.

Todo acidente de trânsito que tenha como resultado ao menos uma pessoa com lesão corporal grave ou morta é considerado acidente grave. Acresce algumas importantes definições:

Lesionado: Pessoa que apresenta sinais de lesões, inconsciência, estado de desorientação ou foi removida para atendimento médico. As lesões são classificadas em:

Lesões Leves: Lesões aparentes como:

- Pequenos cortes, contusões e escoriações;
- Fratura dos dentes;
- Pequenas hemorragias externas;
- Pequenas entorses;
- Luxações e/ou fraturas fechadas ou abertas dos dedos.

Lesões Graves: Lesões não classificadas como leves e que não resultam em óbito.

Morto: Pessoa que veio a óbito no local do acidente, constatado por um profissional habilitado ou com sinais evidentes de morte, incluindo:

- Estado de decomposição;
- Decapitação ou segmentação do tronco;
- Esmagamento do corpo ou crânio com perda de massa encefálica;
- Carbonização do corpo;
- Rigidez cadavérica;
- Entre outros.

Nas próximas seções trataremos de acidentes de trânsito classificados como graves pela PRF, mais **especificamente que resultem em morte.**

2.1 ANÁLISE DE SÉRIES TEMPORAIS

A predição de Séries Temporais (ST) permite estimar valores futuros de uma variável quantitativa. Baseia-se exclusivamente em seus dados históricos e nas inter-relações entre esses dados. Este método utiliza modelos matemáticos para analisar padrões ao longo do tempo. Assim, projeta tendências e futuros comportamentos da variável [10].

2.2 TRABALHOS RELACIONADOS

Pesquisas em séries temporais têm se destacado por suas aplicações em diversos campos, como engenharia, economia, climatologia e saúde pública. Em particular, esta pesquisa utiliza modelos ARIMA, SARIMA, SARIMAX, LSTM e Prophet para modelagem e previsão (Tabela 1). Esses avanços permitem uma melhor compreensão e previsão de fenômenos temporais, contribuindo significativamente para a tomada de decisões.

Tabela 1 - Benchmark das publicações relevantes nos últimos 5 anos envolvendo acidentes em rodovias.

| TÍTULO | ANO | MODELO UTILIZADO |
|---|------|-----------------------------|
| [11] SARIMA Modelling Approach for Forecasting of Traffic Accidents | 2022 | SARIMA,SARIM AX |
| [12] The comparative analysis of SARIMA, Facebook Prophet, and LSTM for road traffic injury prediction in Northeast China | 2022 | SARIMA, Prophet, LSTM |
| [13] A New Time Series Forecasting Using Decomposition Method with SARIMAX Model | 2022 | SARIMAX |
| [14] Time Series Modeling of Road Traffic Accidents in Amhara Region | 2021 | ARIMA, análise de tendência |
| [15] Time Series Forecasting Using a Hybrid Prophet and Long Short-Term Memory Model | 2022 | Prophet |

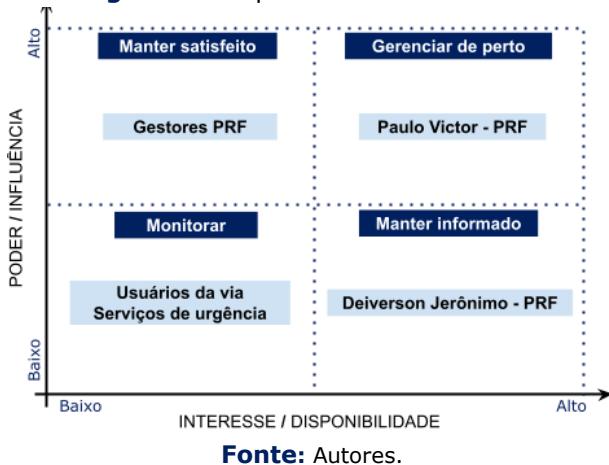
Fonte: Autores.

Os artigos orientarão nossa proposta de mineração de dados, fornecendo *insights* valiosos.

3 MATERIAIS E MÉTODOS

Os Stakeholders envolvidos no projeto são Policiais Rodoviários Federais que atuam nos setores de segurança viária e de tecnologia da informação e comunicação, respectivamente (Figura 2). A importância desses stakeholders para a sociedade reside no fortalecimento da segurança viária e na modernização dos processos de fiscalização e gestão do trânsito. A atuação dos Policiais Rodoviários Federais contribui diretamente para a redução de acidentes, preservação de vidas e garantia da ordem nas rodovias federais.

Figura 2 - Mapa de Stakeholders finais.



Fonte: Autores.

3.1 DESCRIÇÃO DA BASE DE DADOS

A base de dados utilizada nesta análise foi extraída do portal de dados abertos da Polícia Rodoviária Federal (PRF) [6]. Para compreender a estrutura e o significado das variáveis presentes na base de dados, utilizamos o dicionário de dados fornecido pela PRF [16]. Esta base de dados contém informações detalhadas sobre os acidentes registrados pela PRF, permitindo uma análise aprofundada sobre seus diversos aspectos. O dataset original possui 37 atributos no total, incluindo informações de localização dos acidentes, como município, coordenadas geográficas e quilômetro do registro. Também abrange condições climáticas, como condição meteorológica e fase do dia, além de informações que ajudam a entender o cenário dos acidentes, como o tipo de acidente e sua causa presumível, o tipo de veículo envolvido e o estado físico das pessoas envolvidas.

Devido à natureza da atividade de fiscalização e acompanhamento das ocorrências de acidentes nas rodovias, todos os atributos devem ser preenchidos, com pelo menos uma opção informada para cada atributo no momento da ocorrência. Por exemplo, idade e sexo podem ser registrados como “não informado” quando essa informação não está disponível.

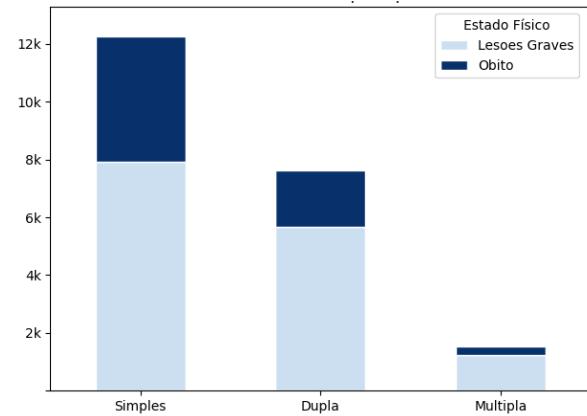
3.2 ANÁLISE DESCRIPTIVA DOS DADOS

A análise descritiva dos dados de acidentes de trânsito entre 2017 e 2024 focou em Acidentes Graves, que resultaram em pelo menos um ferido grave ou em óbito. Os cenários foram avaliados para três tipos de pista: simples, dupla e múltipla.

Entre os tipos de pista, a pista simples é a mais letal, respondendo por 65,9% dos óbitos, enquanto a pista dupla por 29,7% dos óbitos. Ademais, as pistas simples registram 57% de todos os Acidentes no Estado de Pernambuco.

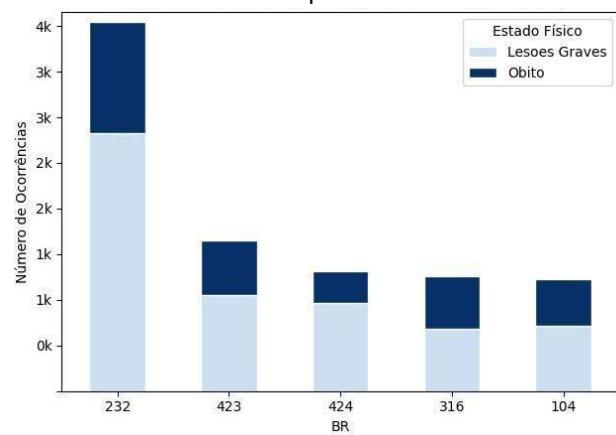
A fim de evitar citações demasiadas, pontuamos que as letras **S**, **D** e **M** representam o tipo de pista **Simples**, **Dupla** e **Múltipla**.

Figura 3 – Acidentes graves por tipo de pista (Simples, Dupla e Múltipla).



Fonte: Autores.

Figura 4 – Acidentes graves por BR do tipo pista simples.



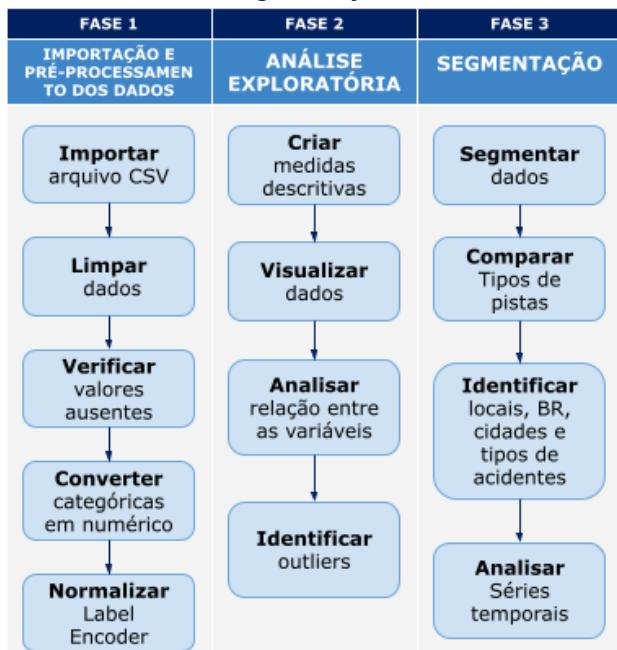
Fonte: Autores.

Diante da alta letalidade das pistas simples (Figura 2), foram identificadas as rodovias federais de Pernambuco (BRs) com maior incidência de acidentes graves (Figura 3). Nota-se que a BR 232 se destaca tanto em óbitos quanto em lesões graves, fato corroborado pelo grande fluxo de veículos que a utilizam. Destaca-se também a BR 423, rodovia que liga as cidades de Caruaru e Garanhuns até Paulo Afonso.

3.3 PRÉ-PROCESSAMENTO DOS DADOS

Para tratar adequadamente as informações, o pré-processamento foi segmentado em três fases (Quadro 1).

Quadro 1 - Fases de importação, análise e segmentação.



Fonte: Autores.

Após o pré-processamento dos dados, obteve-se o dicionário final (Tabela 2).

Tabela 2 - Dicionário final de dados.

| ATRIBUTOS | DESCRIÇÃO |
|---------------|--|
| dia semana | Dia da semana da ocorrência. Ex.: Segunda, Terça, etc STRING; NOT NULL |
| br | Variável com valores numéricos, representando o identificador da BR do acidente. INTEGER; NOT NULL |
| km | Identificação do quilômetro onde ocorreu o acidente, com valor mínimo de 0,1 km e com a casa decimal separada por ponto DECIMAL; NOT NULL |
| município | Nome do município de ocorrência do acidente STRING; NOT NULL |
| tipo acidente | Identificação do tipo de acidente. Ex.: Colisão frontal, Saída de pista, etc. Neste conjunto de dados são excluídos os tipos de acidentes com ordem maior ou igual a dois. STRING; NOT NULL |
| fase dia | Fase do dia no momento do acidente. Ex. Amanhecer, pleno dia, etc STRING; NOT NULL |

| | |
|------------------------|---|
| condicao meteorologica | Condição meteorológica no momento do acidente: Céu claro, chuva, vento, etc. STRING; NOT NULL |
| tipo pista | Tipo da pista considerando a quantidade de faixas: Dupla, simples ou múltipla STRING; NOT NULL |
| mortos | Número de pessoas mortas INTEGER; NOT NULL |
| feridos graves | Número de pessoas com ferimentos graves INTEGER; NOT NULL |
| latitude | Latitude do local do acidente em formato geodésico decimal DECIMAL; NOT NULL |
| longitude | Longitude do local do acidente em formato geodésico decimal DECIMAL; NOT NULL |

Fonte: [6].

O dicionário (Tabela 2) representa a seleção das mais importantes *features* de acordo com sua utilização no desenvolvimento das análises.

3.4 METODOLOGIA EXPERIMENTAL

O protocolo experimental detalha os procedimentos e métodos e modelo, garantindo replicabilidade e validade. Ele descreve as etapas para a coleta e análise das métricas (Quadro 2).

Para cada etapa de zero a onze, conforme ilustrado na (Quadro 2), foram utilizadas as seguintes métricas: MAE (*Mean Absolute Error*), que indica a média das diferenças absolutas entre valores previstos e observados; RMSE (*Root Mean Square Error*), que calcula a raiz quadrada da média dos quadrados das diferenças, penalizando erros maiores; e MASE (*Mean Absolute Scaled Error*), que compara o erro do modelo com o erro de uma previsão, ajustando para a escala dos dados. [18].

Para suavizar as flutuações mensais e detectar tendências sazonais de curto prazo, foi utilizada uma janela de três meses para calcular a média e mediana móvel. A média suaviza os dados, enquanto a mediana reduz o impacto de valores atípicos [17].

Isolation Forest é uma técnica de aprendizado de máquina usada para detecção de anomalias que se baseia no conceito de isolamento de instâncias em um conjunto de dados, construindo várias árvores de decisão de forma aleatória e as instâncias que exigem menos divisões para serem isoladas são consideradas anomalias por ser uma abordagem eficiente e escalável [18]. O IsolationForest é eficaz em identificar padrões de dados com valores discrepantes.

Prophet é uma ferramenta de modelagem de séries temporais desenvolvida pelo Facebook,

projetada para lidar com dados sazonais e cílicos. Utilizando uma abordagem aditiva, combina componentes de tendência, sazonalidade e feriados, sendo robusta a dados ausentes e *outliers*. A simplicidade de uso e a capacidade de gerar previsões confiáveis com pouca necessidade de pré-processamento tornam o Prophet popular entre analistas e cientistas de dados [19].

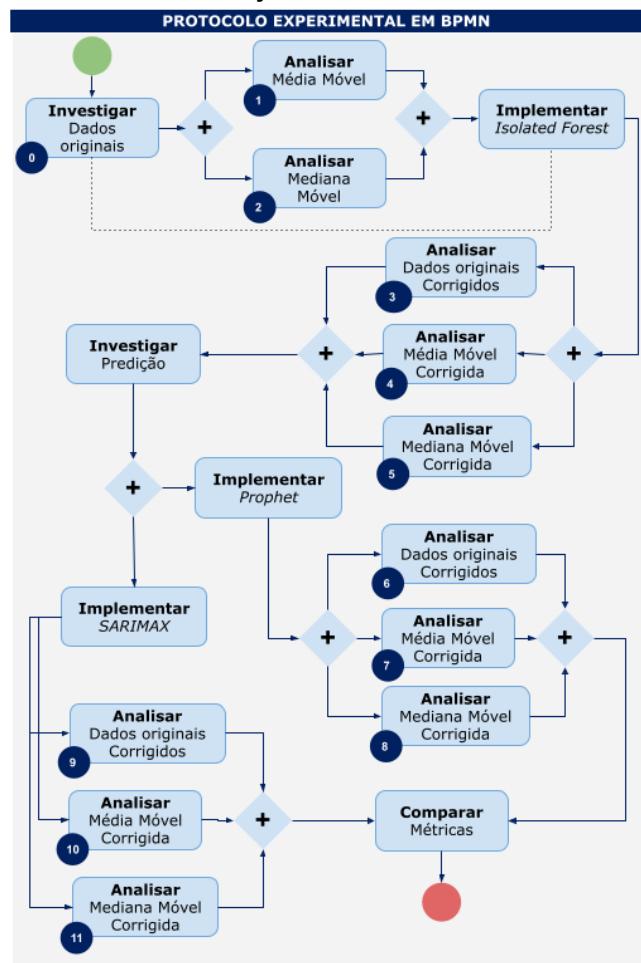
SARIMAX (*Seasonal AutoRegressive Integrated Moving Average with eXogenous factors*) é uma extensão do ARIMA que incorpora sazonalidade e variáveis exógenas. Ele é usado para análise e previsão de séries temporais que apresentam padrões sazonais e podem ser influenciadas por fatores externos. Ao combinar componentes autorregressivos, de médias móveis, integrais e sazonais, o SARIMAX fornece uma abordagem robusta para modelar dados temporais complexos, permitindo previsões mais precisas e úteis em diversos contextos aplicacionais [20].

O **protocolo experimental** (Quadro 2) possibilita a execução de técnicas de suavização para detectar e corrigir anomalias, além de aplicar modelos preditivos para investigar o melhor método para cada tipo de pista: Simples (**S**), Dupla (**D**) e múltipla (**M**).

Conclui-se que a combinação de técnicas de detecção de anomalias e modelos preditivos fortalece a análise de séries temporais em contextos complexos. O protocolo experimental proposto viabiliza a comparação sistemática dessas abordagens. Tal estrutura permite avaliar o desempenho conforme o tipo de pista analisado. Dessa forma, amplia-se a compreensão sobre a adequação de cada método. Os resultados contribuem para decisões mais precisas em aplicações práticas.

O protocolo experimental apresenta um fluxo detalhado de investigação, análise e implementação utilizando diferentes técnicas de análise de dados, incluindo modelos Isolated Forest, Prophet e SARIMAX. Cada etapa envolve a correção e análise dos dados com métodos de média e mediana móvel para garantir a precisão dos resultados. O processo culmina na comparação das métricas obtidas, avaliando o melhor resultado conforme o tipo de pista identificando a melhor configuração.

Quadro 2 - Processo experimental para obtenção de métricas.



Fonte: Autores.

4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Os resultados em cada etapa destacam-se entre os modelos Prophet e SARIMAX. Após o protocolo experimental (Quadro 2), foram retiradas, no início, as primeiras métricas dos dados originais para a pista Simples (**S**), onde foram obtidos **MAE**: 16.863716, **RMSE**: 23.203896 e **MASE**: 0.811214, valores bastante elevados para previsão (Tabela 4). Novas métricas foram realizadas para cada etapa, com média e mediana móvel, a fim de obter melhores resultados. Em seguida, foi utilizado o algoritmo Isolated Forest para detectar valores discrepantes (anomalias) e suavizá-los para depois predizer com os algoritmos Prophet e SARIMAX e investigar melhores métricas que indiquem com maior certeza a quantidade de mortos ao longo do tempo até o final de 2024 para cada tipo de pista, baseada nas melhores métricas.

Tabela 4 - Resultado das métricas.

| ETAPA | TIPO DE PISTA | MÉTRICAS | | |
|---|---------------|-----------|-----------|----------|
| | | MAE | RMSE | MASE |
| Dados Originais | | | | |
| ETAPA 0 | S | 16.863716 | 23.203896 | 0.811214 |
| | D | 8.089393 | 10.524502 | 0.754773 |
| | M | 2.731534 | 3.698673 | 0.667185 |
| Técnica Média Móvel | | | | |
| ETAPA 1 | S | 13.067 | 19.447 | 0.616 |
| | D | 5.762 | 7.847 | 0.537 |
| | M | 2.381 | 3.160 | 0.581 |
| Técnica Mediana Móvel | | | | |
| ETAPA 2 | S | 12.333 | 22.534 | 0.582 |
| | D | 5.095 | 9.429 | 0.475 |
| | M | 1.929 | 3.086 | 0.471 |
| Uso do Modelo Isolated Forest | | | | |
| Dados originais + Isolate Forest | | | | |
| ETAPA 3 | S | 3.372093 | 15.127366 | 0.162212 |
| | D | 0.825581 | 3.774147 | 0.077030 |
| | M | 0.662791 | 2.842780 | 0.161889 |
| Médias Móveis + Isolate Forest | | | | |
| ETAPA 4 | S | 0.848837 | 3.655285 | 0.116812 |
| | D | 0.236434 | 0.992218 | 0.059694 |
| | M | 0.052326 | 0.290349 | 0.040311 |
| Medianas Móveis + Isolate Forest | | | | |
| ETAPA 5 | S | 1.348837 | 7.039341 | 0.167619 |
| | D | 0.430233 | 2.117205 | 0.089852 |
| | M | 0.116279 | 0.555103 | 0.081684 |
| Uso do Modelo Prophet | | | | |
| PROPHET = Dados originais + Isolate Forest | | | | |
| | S | 12.222326 | 15.423249 | 0.781714 |

| | | | | |
|---|---|------------|------------|-----------|
| ETAPA 6 | D | 5.932457 | 7.675985 | 0.608274 |
| | M | 1.703929 | 2.220494 | 0.570212 |
| PROPHET = Médias Móveis + Isolate Forest | | | | |
| ETAPA 7 | S | 8.596514 | 10.875804 | 1.381292 |
| | D | 4.518649 | 5.846542 | 1.219318 |
| | M | 1.297564 | 1.571013 | 1.137040 |
| PROPHET = Medianas Móveis + Isolate Forest | | | | |
| ETAPA 8 | S | 9.480127 | 11.891361 | 1.406302 |
| | D | 4.718307 | 6.305520 | 1.081014 |
| | M | 1.329397 | 1.584433 | 1.018007 |
| Uso do Modelo Sarimax | | | | |
| SARIMAX = Dados originais + Isolate Forest | | | | |
| ETAPA 9 | S | 27.3620839 | 30.3892954 | 0.9634537 |
| | D | 11.2448029 | 14.8351456 | 1.0040003 |
| | M | 2.9964016 | 3.3667137 | 0.713429 |
| SARIMAX = Médias Móveis + Isolate Forest | | | | |
| ETAPA 10 | S | 6.4295028 | 6.9645228 | 1.0540169 |
| | D | 10.1823653 | 12.410757 | 2.3679919 |
| | M | 1.9411909 | 2.5326997 | 1.3865649 |
| SARIMAX = Medianas Móveis + Isolate Forest | | | | |
| ETAPA 11 | S | 32.4148826 | 37.3387381 | 2.7012402 |
| | D | 7.5135134 | 8.6258572 | 1.5333701 |
| | M | 1.7844353 | 2.2575038 | 0.7435147 |

Fonte: Autores.

Após a análise das métricas **MAE**, **RMSE** e **MASE** para cada tipo de pista Simples, Dupla e Múltipla nos modelos Prophet e SARIMAX, destacamos o resultado final com as melhores métricas do melhor modelo (Tabela 5).

Tabela 5 - Melhores métricas para cada tipo de pista.

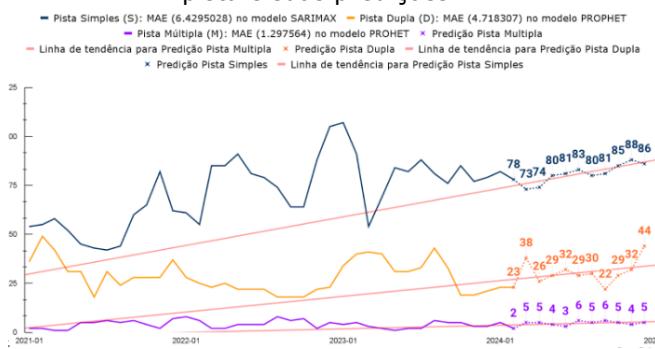
| ETAPA | MOD | TIPO DE PISTA | MÉTRICA |
|-------|---------|---------------|-----------------|
| 7 | SARIMAX | S | 6.4295028 (MAE) |
| 11 | PROPHET | D | 4.718307 (MAE) |
| 6 | PROPHET | M | 1.297564 (MAE) |

Fonte: Autores.

Para a pista simples, o modelo SARIMAX apresentou a melhor métrica, com **MAE** de 6.4295028 (Médias Móveis + Isolate Forest), indicando menor erro absoluto médio em comparação ao PROPHET, que teve **MAE** de 4.518649. Na pista dupla, o modelo PROPHET se destacou com **MAE** de 4.718307 (Medianas Móveis + Isolate Forest), mostrando menor erro absoluto escalado que o SARIMAX com **MASE** de 1.5333701. Para a pista múltipla, o modelo PROPHET obteve a menor **MAE**, com 1.297564

(Médias Móveis + Isolate Forest), demonstrando maior precisão em relação ao SARIMAX que teve **MAE** de 1.9411909. Observa-se ao longo do tempo as três séries temporais com suas respectivas métricas e dados preditivos até o final de 2024 (Figura 4).

Figura 4 – Séries Temporais para cada tipo de pista e suas previsões.



Fonte: Autores.

5 CONCLUSÕES

Como demonstrado nas análises, as pistas simples têm um grau de letalidade superior às demais, com um alto índice de acidentes graves que resultam em óbitos. Para a Polícia Rodoviária Federal (PRF), esses dados são essenciais para identificar áreas críticas e compreender as causas desses acidentes, permitindo fiscalizações e campanhas educativas mais eficazes.

Como demonstrado, essas informações podem fundamentar estudos que indiquem a necessidade de duplicação de trechos de rodovias simples e a implementação de melhorias na infraestrutura viária, como sinalização, iluminação e condições do asfalto.

Em última análise, a preservação da vida é o objetivo maior. A aplicação de dados precisos e medidas informadas é fundamental para transformar a segurança viária e garantir um ambiente rodoviário mais seguro e sustentável.

5.1 TRABALHOS FUTUROS

Para trabalhos futuros na análise de dados, é recomendável realizar experimentos para aprimorar modelos existentes. No contexto de segurança viária, isso é crucial, pois dados de acidentes podem revelar padrões importantes para prevenção. Experimentos podem incluir aplicações de técnicas de normalização para uniformizar os dados e a adição de novos modelos preditivos, como o modelo Kan.

Outro algoritmo recomendado é o Apriori que pode auxiliar na análise de acidentes de trânsito, identificando padrões e associações entre fatores como condições climáticas, características das vias, comportamentos dos motoristas e tipos de veículos. Ele permite entender interações complexas e fornece *insights* valiosos para desenvolver estratégias de prevenção mais eficazes.

A conclusão é o encerramento do seu trabalho, esse é o espaço que o(s) autor(es) do artigo mais aparecem. Essa parte do texto permite rever os pontos principais do manuscrito, explicando os resultados mais importantes, sugerindo aplicações e extensões. Um excelente manual de estilo e fonte de informação para autores está proposto em Young [12].

REFERÊNCIAS

- [1] Portal do Ministério da Justiça e Segurança Pública. Polícia Rodoviária Federal. Institucional. Disponível em: <https://www.gov.br/prf/pt-br/acesso-a-informacao/institucional>. Acesso em: 17 jul. 2024.
- [2] ONSV - Observatório Nacional de Segurança Viária. Dados Consolidados de Óbitos no Trânsito Brasileiro - 2022. 2024. Disponível em: <https://www.onsv.org.br/pdi/analise-datasus-2022>. Acesso em: 21 jul. 2024.
- [3] ONSV - Observatório Nacional de Segurança Viária. Relatório ONSV 2022. 2022. Disponível em: https://www.onsv.org.br/source/files/originals/Relatorio_ONSV_2022_25.7_internet-089959.pdf. Acesso em: 17 jul. 2024.
- [4] Portal do Ministério da Justiça e Segurança Pública. Polícia Rodoviária Federal. PRF/PE promove ações em comemoração aos seus 95 anos. 2023. Disponível em: <https://www.gov.br/prf/pt-br/noticias/estaduais/pernambuco/2023/julho/prf-pe-promove-acoes-em-comemoracao-aos-seus-95-anos>. Acesso em: 17 jul. 2024.
- [5] Ministério dos Transportes. DNIT - Departamento Nacional de Infraestrutura de Transportes. DNIT moderniza a malha rodoviária do Nordeste com ampliação dos investimentos. 2023. Disponível em: <https://www.gov.br/dnit/pt-br/assuntos/noticias/dnit-moderniza-a-malha-rodoviaria-do-nordeste-com-ampliacao-dos-investimentos>. Acesso em: 17 jul. 2024.
- [6] Portal do Ministério da Justiça e Segurança Pública. Polícia Rodoviária Federal. Dados abertos da PRF. 2022. Disponível em: <https://www.gov.br/prf/pt-br/acesso-a-informacao/dados-abertos/dados-abertos-da-prf>. Acesso em: 17 jul. 2024.
- [7] Jornal Hoje. Quantidade de acidentes nas rodovias federais cai, mas número de mortes aumenta no país. Globo.com. 2024. Disponível em: <https://g1.globo.com/df/distrito-federal/noticia/2024/01/16/quantidade-de-acidentes-nas-rodovias-federais-cai-mas-> numero-de-mortes-aumenta-no-pais.ghtml. Acesso em: 17 jul. 2024.
- [8] RAFAEL, Pedro. Brasil reduz mortes no trânsito, mas está longe da meta para 2020. Agência Brasil. 2018. Disponível em: <https://agenciabrasil.ebc.com.br/geral/noticia/2018-09/brasil-reduz-mortes-no-transito-mas-esta-longe-da-meta-para-2020>. Acesso em: 17 jul. 2024.
- [9] ONSV - Observatório Nacional de Segurança Viária. Relatório ONSV 2022. 2022. Disponível em: https://www.onsv.org.br/source/files/originals/Relatorio_ONSV_2022_25.7_internet-089959.pdf. Acesso em: 17 jul. 2024.
- [10] SORJAMAA, A.; HAO, J.; REYHANI, N.; JI, Y.; LENDASSE, A. Methodology for long-term prediction of time series. Neurocomputing, v. 70, p. 2861-2869, 2007.
- [11] DERETIĆ, N.; STANIMIROVIĆ, D.; AWADH, M. A.; VUJANOVIĆ, N.; DJUKIĆ, A. SARIMA Modelling Approach for Forecasting of Traffic Accidents.
- [12] LI, J.; WANG, Y.; ZHANG, H. The comparative analysis of SARIMA, Facebook Prophet, and LSTM for road traffic injury prediction in Northeast China. *Frontiers in Public Health*, 2022.
- [13] KUMAR, S.; TOSHNIWAL, D. A novel framework to analyze road accident time series data. Journal of Big Data, v. 3, 05 2016.
- [14] ALEMTESEG, K. Time Series Modeling of Road Traffic Accidents in Amhara Region. 2021.
- [15] Kong, Yih Hern & Lim, Khai & Chin, Wan Yoke. (2021). Time Series Forecasting Using a Hybrid Prophet and Long Short-Term Memory Model. 10.1007/978-981-16-7334-4_14.
- [16] Portal do Ministério da Justiça e Segurança Pública. Polícia Rodoviária Federal. Dicionário de dados - Infrações. 2020. Disponível em: <https://www.gov.br/prf/pt-br/acesso-a-informacao/dados-abertos/dicionario-infracoes>. Acesso em: 28 mai. 2024.
- [17] SCIKIT-LEARN. Regression metrics. Disponível em: https://scikit-learn.org/stable/modules/model_evaluation.html#regression-metrics. Acesso em: 9 jul. 2024.

[18] SCIKIT-LEARN. sklearn ensemble

IsolationForest. Disponível em: <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.IsolationForest.html>. Acesso em: 9 jul. 2024.

[19] TAYLOR, Sean J.; LETHAM, Benjamin.

Prophet: Forecasting at scale. Disponível em: <https://facebook.github.io/prophet/>. Acesso em: 28 jul. 2024.

[20] STATS MODELS. SARIMAX: Seasonal

AutoRegressive Integrated Moving Average
with eXogenous factors. Disponível em:
<https://www.statsmodels.org/dev/generated/statsmodels.tsa.statespace.sarimax.SARIMAX.html>. Acesso em: 24 jul. 2024.