

**UNIVERSIDADE FEDERAL DA FRONTEIRA SUL
CAMPUS PASSO FUNDO – RS
CURSO DE MEDICINA**

MARIA CLARA DA SILVA MAIA

**FEBRE DO OROPOUCHE:
EFICÁCIA DE MODELOS MATEMÁTICOS E DE APRENDIZADO DE MÁQUINA NA
PREVISÃO DE CASOS**

**PASSO FUNDO – RS
2026**

MARIA CLARA DA SILVA MAIA

FEBRE DO OROPOUCHE:
EFICÁCIA DE MODELOS MATEMÁTICOS E DE APRENDIZADO DE MÁQUINA NA
PREVISÃO DE CASOS

Trabalho de Curso apresentado ao Curso de Medicina da Universidade Federal da Fronteira Sul (UFFS) – Campus Passo Fundo/RS, como requisito parcial para obtenção do título de médico.

Orientadora: Prof.^a Dr.^a. Renata dos Santos Rabello

Coorientador: Prof. Dr. Julio César Stobbe

Coorientador: Prof. Dr. Gustavo Olszanski Acrani

PASSO FUNDO – RS

2026

Bibliotecas da Universidade Federal da Fronteira Sul - UFFS

Maia, Maria Clara da Silva

FEBRE DO OROPOUCHE: EFICÁCIA DE MODELOS MATEMÁTICOS
E DE APRENDIZADO DE MÁQUINA NA PREVISÃO DE CASOS / Maria
Clara da Silva Maia, Julio César Stobbe, Gustavo
Olszanski Acrani, Renata dos Santos Rabello. -- 2026.
63 f.:il.

Orientadora: Doutorado Renata dos Santos Rabello

Coorientadores: Doutorado Julio César Stobbe,
Doutorado Gustavo Olszanski Acrani

Trabalho de Conclusão de Curso (Graduação) -
Universidade Federal da Fronteira Sul, Curso de
Bacharelado em Medicina, Passo Fundo, RS, 2026.

1. Medicina Tropical. 2. Infecções por Arbovírus. 3.
Algoritmos de Predição. 4. Modelos Preditivos de
Aprendizagem. 5. Planejamento em Saúde. I. Stobbe, Julio
César II. Acrani, Gustavo Olszanski III. Rabello,
Renata dos Santos IV. Rabello, Renata dos Santos,
orient. V. Stobbe, Julio César, co-orient. VI. Acrani,
Gustavo Olszanski, co-orient. VII. Universidade Federal

Elaborada pelo Sistema de Geração Automática de Ficha de Identificação da Obra pela UFFS
com os dados fornecidos pelo(a) autor(a).

MARIA CLARA DA SILVA MAIA

FEBRE DO OROPOUCHE:

**EFICÁCIA DE MODELOS MATEMÁTICOS E DE APRENDIZADO DE MÁQUINA NA
PREVISÃO DE CASOS**

Trabalho de Curso apresentado ao Curso de
Medicina da Universidade Federal da Fronteira Sul
(UFFS) – Campus Passo Fundo/RS, como
requisito parcial para obtenção do título de médico.

Este Trabalho de Curso foi defendido e aprovado pela banca em 17/06/2026

BANCA EXAMINADORA

Prof.^a Dr.^a Renata dos Santos Rabello
Orientadora

Prof. Esp. Hugo Vladimir Noal da Silva
Avaliador

Prof. Dr. Aníbal Lopes Guedes
Avaliador

AGRADECIMENTOS

À minha orientadora, Prof.^a Dr.^a. Renata dos Santos Rabello, pela fé e apoio a este projeto, que por vezes, ultrapassou o esperado por mim.

Ao meu coorientador, Prof. Dr. Júlio César Stobbe, pelas suas excelentes aulas, as quais me inspiraram a escolher o tema do projeto.

Ao meu coorientador, Prof. Dr. Gustavo Olszanski Acrani, pelo auxílio com seus conhecimentos específicos acerca do tema desse trabalho.

A todos aqueles que responderam com empolgação quando comentei sobre esse trabalho.

APRESENTAÇÃO

O presente estudo, intitulado “Febre do Oropouche: Eficácia de modelos matemáticos e de aprendizado de máquina na previsão de casos”, foi realizado pela acadêmica Maria Clara da Silva Maia, estudante do curso de medicina da Universidade Federal da Fronteira Sul (UFFS) – campus Passo Fundo/RS, sob a orientação da Prof.^a Dra. Renata dos Santos Rabello e coorientação do Prof. Dr. Julio César Stobbe e do Prof. Dr. Gustavo Olszanski Acrani. É requisito parcial para a obtenção de título de médico pela UFFS e está de acordo com as normas do Manual de Trabalhos Acadêmicos da UFFS, 3a edição, revisado e atualizado em 2024 e com o Regulamento de TC do Curso. Este volume é composto por três capítulos: Projeto de pesquisa, Relatório de pesquisa e Artigo científico. O primeiro trata-se do Projeto de Pesquisa, desenvolvido no Componente Curricular Trabalho de Curso I, no primeiro semestre letivo de 2025. O Relatório da pesquisa encontra-se no segundo capítulo, tendo sido desenvolvido durante o Componente Curricular Trabalho de Curso II, no segundo semestre letivo de 2025. O terceiro capítulo, formulado no primeiro semestre letivo de 2026, traz o Artigo Científico, elaborado a partir da análise dos dados obtidos. O propósito deste estudo foi investigar a eficácia dos modelos preditivos de média móvel integrada autorregressiva sazonal com variáveis exógenas (SARIMAX) e de memória de curto e longo prazo (LSTM) em estimar casos futuros de Febre do Oropouche no Brasil e deliberar sobre qual deles possui o melhor desempenho.

RESUMO

As mudanças climáticas amplificam a circulação de vírus como o Oropouche, da família *Orthobunyavirus*, causador da Febre do Oropouche (FO). No Brasil, ele restringia-se à Amazônia, até que se expandiu além de suas fronteiras, tornando-se uma doença emergente. Assim, o presente estudo objetivou estimar novos casos de FO no Brasil utilizando os modelos Autorregressivo Integrado de Médias Móveis Sazonal com Variáveis Exógenas (SARIMAX), de Memória de Curto-Longo Prazo (LSTM) multivariado e a média aritmética entre eles. Foram coletados dados do Painel Epidemiológico do Ministério da Saúde do *Prediction Of Worldwide Energy Resources* referentes aos casos de FO e às temperaturas médias de 2023 a 2025. Foi feita uma análise descritiva com valores absolutos e relativos de sexo, faixa etária e região de residência. Via testes de Dickey-Fuller aumentado, Mann-Kendall e Ljung-Box, foram analisadas estacionaridade, tendência e autocorrelação da série temporal, respectivamente. Ademais, a curva foi decomposta pelo modelo aditivo. Os modelos SARIMAX, LSTM e a média de ambos foram implementados em Python. Foram notificados 26.678 casos entre 2023 e 2025, predominando sexo masculino (52,36%) e faixa etária de 20 a 49 anos (56,90%). Observou-se concentração na região Norte (94,36%) em 2023, que transicionou para a região Sudeste (85,01%) em 2025. A série foi não estacionária, sem tendência detectável e com autocorrelação dos dados. Analisando-se a decomposição da curva, a tendência foi crescente de junho de 2023 até abril de 2025, quando tornou-se decrescente. A partir de julho, a série deixou de expressar tendência. Quanto à sazonalidade, as notificações aumentaram de setembro a fevereiro e diminuíram de março a outubro. Na validação dos modelos, o LSTM demonstrou o melhor desempenho. Seu Erro Absoluto Médio foi de três casos prevendo para 3 meses e, 515, prevendo para seis. Portanto, os resultados indicaram finalização da epidemia de 2023/2025. Porém, como a FO demonstrou potencial epidêmico e expansivo, necessitam-se ações governamentais no seu monitoramento e combate. Para tal, o LSTM pode auxiliar, mas necessita ainda calibragem.

Palavras-chave: Medicina Tropical; Infecções por Arbovírus; Algoritmos de Predição; Modelos Preditivos de Aprendizagem; Planejamento em Saúde.

ABSTRACT

Climate change amplifies the circulation of viruses such as Oropouche, from the *Orthobunyavirus* family, which causes Oropouche Fever (OF). In Brazil, it was previously confined to the Amazon region until it expanded beyond its borders, becoming an emerging disease. Thus, this study aimed to estimate new cases of OF in Brazil using the Autoregressive Integrated Moving Average with Exogenous Variables (SARIMAX) model, the multivariate Long-Term Memory (LSTM) model, and the arithmetic mean between them. Data from the Ministry of Health's Epidemiological Panel on Prediction of Worldwide Energy Resources regarding OF cases and average temperatures from 2023 to 2025 were collected. A descriptive analysis was performed with absolute and relative values of sex, age group and region of residence. Using the augmented Dickey-Fuller, Mann-Kendall, and Ljung-Box tests, stationarity, trend, and autocorrelation of the time series were analyzed, respectively. Furthermore, the curve was decomposed using an additive model. The SARIMAX and LSTM models, and the average of both, were implemented in Python. 26,678 cases were reported between 2023 and 2025, predominantly in males (52.36%) and the 20-49 age group (56.90%). A concentration was observed in the North region (94.36%) in 2023, which transitioned to the Southeast region (85.01%) in 2025. The series was non-stationary, with no detectable trend and with autocorrelation of the data. Analyzing the curve decomposition, the trend was increasing from June 2023 to April 2025, when it became decreasing. From July onwards, the series ceased to show a trend. Regarding seasonality, notifications increased from September to February and decreased from March to October. In the model validation, the LSTM demonstrated the best performance. Its Mean Absolute Error was three cases predicting for 3 months and 515 predicting for six. Therefore, the results indicated the end of the epidemic in 2023/2025. However, as the OF demonstrated epidemic and expansive potential, governmental actions are needed for its monitoring and control. For this, the LSTM can assist, but it still needs calibration.

Keywords: Tropical Medicine; Arbovirus Infections; Prediction Algorithms; Predictive Learning Models; Health Planning.

LISTA DE ABREVIATURAS E SIGLAS

ACF	Função De Autocorrelação
AIC	Critério De Informação De Akaike
AR	Modelo Autoregressivo
ARIMA	Modelo Autoregressivo Integrado de Média Móvel
BIC	Critério De Informação Bayesiano
Deep-ESN	Rede de Estado de Eco Profundo
ELM	Máquina de Aprendizado Extremo
ESM	Máquina de Estado de Eco
FO	Febre Do Oropouche
LR	Regressão Linear
LSTM	Modelo de Memória de Curto e Longo Prazo
MA	Modelo de Média Móvel
MAE	Erro Absoluto Médio
MAPE	Erro Percentual Médio Absoluto
MLP	Perceptrons Multicamadas
MSE	Erro Quadrático Médio
NASA	National Aeronautics and Space Administration
OROV	Vírus Oropouche
PACF	Função De Autocorrelação Parcial
POWER	Prediction Of Worldwide Energy Resources
RFR	Regressor de Floresta Aleatória
RMSE	Erro Quadrático Médio
RNN	Redes Neurais Recorrentes
SARIMA	Seasonal Autoregressive Integrated Moving Average
SARIMAX	Seasonal Autoregressive Integrated Moving Average with Exogenous Variables
SINAN	Sistema De Informação De Agravos De Notificação
SVM	Máquina de Vetores de Suporte
VAR	Autorregressão Vetorial
XGBoost	Reforço de Gradiente Extremo

SUMÁRIO

1 INTRODUÇÃO.....	10
2 DESENVOLVIMENTO.....	12
2.1 PROJETO DE PESQUISA.....	12
2.1.1 Tema.....	12
2.1.2 Problemas.....	12
2.1.3 Hipóteses.....	12
2.1.4 Objetivos.....	13
2.1.4.1 Objetivo geral.....	13
2.1.4.1 Objetivos específicos.....	13
2.1.5 Justificativa.....	13
2.1.6 Referencial teórico.....	14
2.1.6.1 Classificação e epidemiologia.....	14
2.1.6.2 Ciclo de transmissão, sintomas e diagnóstico.....	16
2.1.6.3 Modelos de previsão de séries temporais.....	18
2.1.6.4 Testes na predição de outras arboviroses.....	20
2.1.7 Metodologia.....	22
2.1.7.1 Tipo de estudo.....	22
2.1.7.2 Local e período de realização da coleta.....	22
2.1.7.3 População, amostragem e limitações do estudo.....	23
2.1.7.4 Variáveis e coleta de dados.....	23
2.1.7.5 Processamento, controle de qualidade e análise dos dados.....	24
2.1.7.6 Aspectos éticos.....	26
2.1.8 Recursos.....	27
2.1.9 Cronograma.....	27
2.1.10 Referências.....	27
Apêndice A – Código para horizonte preditivo de 3 meses.....	33
2.2 RELATÓRIO DE PESQUISA.....	41
3 ARTIGO CIENTÍFICO.....	43
4 CONSIDERAÇÕES FINAIS.....	62

1 INTRODUÇÃO

A Febre do Oropouche (FO) é causada pelo arbovírus *Orthobunyavirus oropoucheense* (OROV). Pertence à família viral *Peribunyaviridae*, ao gênero *Orthobunyavirus* e ao sorogrupo *Simbu* (CDC, 2025). Trata-se de um patógeno até então considerado pouco importante. Entretanto, a partir de 2024, o OROV começou a receber mais repercussão na mídia, principalmente após a confirmação de mortes associadas.

Tipicamente, o OROV está presente em um ciclo silvestre atingindo lagartas, primatas não humanos e pássaros. Porém, é possível circular em um ciclo urbano e infectar pessoas (Pastula; Beckham; Tyler, 2025). O principal vetor é o *Culicoides paraensis*, conhecido como maruim ou mosquito pólvora (BRASIL, 2024). Não há tratamento ou vacinas, portanto o combate ao mosquito é essencial (Pastula; Beckham; Tyler, 2025).

Após a contaminação, o período de incubação gira em torno de quatro a oito dias. Cerca de 60% das infecções são sintomáticas, causando febre abrupta, cefaleia, mialgia, artralgia, distúrbios gastro-intestinais, dor retro-ocular e/ou fotofobia durante 3 a 10 dias. É possível que a doença siga um curso bifásico, com o retorno dos sintomas uma semana depois de seu início (Okesanya et al., 2025; Pastula; Beckham; Tyler, 2025).

Alguns estudos indicam transmissão vertical com influência no desenvolvimento fetal, de forma similar ao Zika vírus (Chaves et al., 2025; Pastula; Beckham; Tyler, 2025). Também pode ocorrer meningite, encefalite e síndrome de Guillain-Barré (Pastula; Beckham; Tyler, 2025; González-Quevedo; Farrill; Becquer, 2025).

Há um aumento de casos de febre do OROV no cenário nacional, justificado pelo novo protocolo de busca ativa por FO e febre do Mayaro adotado após a descentralização do diagnóstico biomolecular. Os casos saltaram de 833 em 2023 para 13,795 em 2024. A FO foi adicionada às doenças de notificação compulsória, mais especificamente às doenças de notificação imediata, devido ao potencial epidêmico (BRASIL, 2023, 2024, 2025). Atualmente, o Ministério da Saúde recomenda o diagnóstico via RT-qPCR em casos negativos para outras arboviroses desde 2023 (Silva et al., 2025; Gomes et al., 2025).

O estudo epidemiológico sobre a FO permite classificá-la como uma doença negligenciada (Zhang, 2024). Uma outra possibilidade de classificação é como doença emergente (Mendoza-Landinez; Freyle-Roman; Rincón-Orozco, 2024; Silva; Coutinho, 2024). Ou seja, a FO não recebeu a devida atenção por estar restrita à bacia amazônica. Por isso, permitiu-se uma alta de casos para além do comum e sua expansão para novas regiões.

Nesse cenário, os modelos de previsão de séries temporais de doenças infecciosas se destacam como abordagem estratégica para apoiar a formulação de políticas públicas mais eficazes. Afinal, com a capacidade de identificar surtos de forma precoce, promove-se o uso otimizado de recursos e viabiliza-se a implementação de medidas preventivas antes que a situação alcance níveis críticos (Biggerstaff, 2018; Lutz, 2019).

Considerando a grande diversidade da aplicação de modelos preditivos em outras arboviroses, percebe-se um potencial de aplicabilidade dessa técnica para FO. Destarte, é importante que mais pesquisas epidemiológicas sejam feitas sobre a prevalência de OROV no Brasil, principalmente quanto à previsão da série temporal da notificação de casos. Nesse âmbito, deve-se comparar diferentes modelos para descobrir o mais adequado para estimar casos futuros.

2 DESENVOLVIMENTO

2.1 PROJETO DE PESQUISA

2.1.1 Tema

Avaliação da eficácia dos modelos de média móvel integrada autorregressiva sazonal com variáveis exógenas (SARIMAX) e rede de memória de curto e longo prazo (LSTM) na previsão de casos de Febre do Oropouche no Brasil.

2.1.2 Problemas

Existe presença de autocorrelação e estacionaridade nas notificações semanais de Febre do Oropouche no Brasil no período de janeiro de 2023 a dezembro de 2025 por gênero, por faixa etária, região de residência e total?

Qual a tendência das notificações semanais de Febre do Oropouche no Brasil no período de janeiro de 2023 a dezembro de 2025 por gênero, por faixa etária, região de residência e total?

O modelo matemático SARIMAX e o modelo de aprendizado de máquina LSTM possuem erro absoluto percentual médio (MAPE) menor ou igual a 30% na previsão das notificações semanais de novos casos de Febre do Oropouche no Brasil?

Entre o SARIMAX e o LSTM, qual possui o menor MAPE na previsão das notificações semanais de novos casos de Febre do Oropouche no Brasil?

2.1.3 Hipóteses

Existe presença de autocorrelação e estacionaridade nas notificações semanais de Febre do Oropouche no Brasil no período de janeiro de 2023 a dezembro de 2025 por gênero, por faixa etária, região de residência e total.

A tendência das notificações semanais de Febre do Oropouche no Brasil no período de janeiro de 2023 a dezembro de 2025 por gênero, faixa etária, região de residência e total é de estabilidade.

O modelo matemático SARIMAX e o modelo de aprendizado de máquina LSTM possuem erro absoluto percentual médio (MAPE) menor ou igual a 30% na previsão das notificações semanais de casos de Febre do Oropouche no Brasil.

O modelo matemático SARIMAX possui MAPE e demais indicadores menores na previsão das notificações semanais de novos casos de Febre do Oropouche no Brasil em comparação ao LSTM.

2.1.4 Objetivos

2.1.4.1 Objetivo geral

Avaliar comparativamente o desempenho dos modelos SARIMAX e LSTM na previsão de casos semanais futuros de Febre do Oropouche.

2.1.4.1 Objetivos específicos

Avaliar se existe presença de autocorrelação e estacionaridade nas notificações semanais de Febre do Oropouche no Brasil no período de janeiro de 2023 a dezembro de 2025 por gênero, por faixa etária, região de residência e total.

Observar se há tendência de aumento, diminuição ou estabilidade da curva de notificações semanais de FO no Brasil por gênero, por faixa etária, região de residência e total.

Calcular os indicadores erro percentual absoluto médio (MAPE), raiz do erro quadrático médio (RMSE) e erro absoluto médio (MAE) dos modelos SARIMAX e LSTM comparando seus valores antes e depois da inclusão dos dados climáticos.

Utilizar os modelos SARIMAX e LSTM para estimar pontos futuros dessa curva de notificações semanais.

2.1.5 Justificativa

As mudanças climáticas tendem a aumentar a incidência de Febre do Oropouche (FO) e sua área de abrangência na América do Sul (Okesanya et al., 2025; Zhang, 2024). Ademais, a sua similaridade com outras doenças e escassez de recursos diagnósticos, tornam a FO subnotificada. Por isso, é necessário um enfoque maior da saúde pública no que tange à vigilância, à educação pública e colaboração internacional no combate a essa doença (Okesanya et al., 2025). Tal necessidade foi reconhecida pelo Ministério da Saúde ao implementar o protocolo de diagnóstico e notificação já mencionados.

O Sistema de Informação de Agravos de Notificação (SINAN) é uma ferramenta eficaz para estimar dados epidemiológicos no Brasil. Isso posto, sua

manutenção e revisão são essenciais. No entanto, o pleno aproveitamento do sistema depende da organização desses dados brutos (Pereira et al., 2023). Nesse contexto, é válido e necessário o uso desses dados em domínio público para fins de pesquisa.

Entre tantas possibilidades de utilização para esse registro de dados governamentais, está a previsão de séries temporais (Lima et al., 2022). Essa atividade é importante para um melhor controle das campanhas preventivas em saúde pública (Zhu, 2022). Uma revisão sistemática sobre modelos de predição utilizados para arboviroses relatou que os modelos de Poisson, média móvel (ARIMA, SARIMA, ARIMAX, SARIMAX) e redes neurais artificiais, principalmente o LSTM, são os mais utilizados (Lima et al., 2022).

Portanto, fazer uma análise epidemiológica preditiva da notificação de casos de FO é relevante no contexto brasileiro, principalmente com a utilização dos modelos mais populares em outras arboviroses. Afinal, caso seja precisa, essa estimativa pode auxiliar a gestão pública a saber o quanto de enfoque deve ser dado ao combate dessa doença na agenda pública. Desse modo, é possível evitar epidemias sem dispensar energia em excesso para o *Orthobunyavirus oropoucheense* (OROV).

Considerando esse contexto, a proposta desse estudo possui relevância acadêmica por validar o uso desses modelos, já amplamente aplicados em outras doenças infecciosas, para FO. Afinal, não foram encontradas, em revisão de literatura, pesquisas com essa abordagem. Também possui relevância social, porque fornece informações necessárias ao planejamento em saúde pública.

2.1.6 Referencial teórico

2.1.6.1 Classificação e epidemiologia

Enquanto a dengue vem atingindo recordes de infecções, outra arbovirose pouco conhecida está espalhando e tornando os diagnósticos confusos: a febre do Oropouche (FO) (Schweitzer, 2025). Essa enfermidade é endêmica em partes da América do Sul, América Central e Caribe. Recentemente, surgiram casos nativos em Cuba e foram detectados casos importados na Europa e Estados Unidos (Pastula; Beckham; Tyler, 2025).

O estudo epidemiológico sobre a FO permite classificá-la como uma doença negligenciada (Zhang, 2024). O termo se refere a doenças causadas por agentes infecciosos que são endêmicas em populações de baixa renda. A nomenclatura “negligenciadas” faz alusão ao fato de que poderiam ser praticamente erradicadas, porém há pouco ou nenhum interesse da indústria farmacêutica e órgãos governamentais na causa (Souza, 2010, p. 1).

Uma outra possibilidade de classificação é a divisão em doenças emergentes e reemergentes. As primeiras são aquelas com incidência crescente no período recente e as segundas são aquelas com incidência crescente após um declínio no passado (Netto, 1997). A FO seria classificada como emergente. (Mendoza-Landinez; Freyle-Roman; Rincón-Orozco, 2024; Silva; Coutinho, 2024).

A FO é causada pelo arbovírus *Orthobunyavirus oropoucheense* (OROV). Pertence à família viral *Peribunyaviridae*, ao gênero *Orthobunyavirus* e ao sorogrupo *Simbu* (CDC, 2025). Trata-se de um vírus de RNA monocatenário de sentido negativo com 3 segmentos de diferentes comprimentos de nucleotídeos: S (pequeno), M (médio), and L (grande). O segmento L codifica a RNA polimerase; o segmento M codifica duas glicoproteínas externas (Gc e Gn) e uma proteína não-estrutural (PNE); o segmento S codifica o nucleocapsídeo e uma proteína não-estrutural (NS) (Okesanya et al., 2025; Puerto, 2024).

O OROV foi detectado pela primeira vez em 1955, em um trabalhador florestal de uma aldeia de Trinidad e Tobago chamada Vega de Oropouche. Ele é endêmico da bacia amazônica (CDC, 2025). Foi isolado no Brasil em 1960. Estava presente no sangue de um bicho-preguiça da espécie *Bradypus tridactylus*, encontrada durante as obras da construção da rodovia BR-153 (Brasil, 2024). O primeiro aconteceu em 1961 e contou com mais de 11 mil casos confirmados no Pará (Silva et al., 2025).

A FO teve surtos esporádicos em pequenas aldeias na Amazônia brasileira entre 1980 e 2005. Na década de 2000, a doença se espalhou para estados do norte e nordeste e foi identificada pela primeira vez em macacos no sudeste do Brasil. Casos surgiram em áreas não endêmicas, incluindo a região metropolitana de Salvador, Bahia, em amostras de pacientes com sintomas de infecção por arbovírus, sem histórico de viagem para regiões com circulação conhecida do vírus. Em 2016, o primeiro caso fora da Amazônia foi relatado, levantando preocupações sobre sua disseminação para novas áreas (Zhang et al., 2024).

Há um aumento de casos de febre do OROV no cenário nacional, justificado pelo novo protocolo de busca ativa por FO e febre do Mayaro adotado após a descentralização do diagnóstico biomolecular. Os casos saltaram de 833 em 2023 para 13,795 em 2024. A FO foi adicionada às doenças de notificação compulsória, mais especificamente às doenças de notificação imediata, devido ao potencial epidêmico (BRASIL, 2023, 2024, 2025).

No Brasil, não há variações importantes no número de casos entre os gêneros feminino e masculino até os 60 anos. A partir dos 60 anos, há uma leve predominância entre homens. As faixas etárias com mais casos notificados foram de 20 a 49 anos. Em 2023, foram notificados 833 casos. Os estados com mais notificações foram AM, AC e RR. Em 2024, foram notificados 13.791 casos, duas mortes confirmadas e uma em investigação. Os estados com mais casos notificados foram ES, AM, RO. Em 2025, foram notificados 7.756 casos e uma morte em investigação. Os estados com mais casos notificados foram ES, RJ, PB. Vários estados não notificaram casos (Brasil, 2025).

2.1.6.2 Ciclo de transmissão, sintomas e diagnóstico

As mudanças climáticas tendem a aumentar sua incidência e área de abrangência na América do Sul (Okesanya et al., 2025; Zhang, 2024). Afinal, desde o início, o OROV se relacionou com o desmatamento, urbanização desordenada e mudanças climáticas, pois tais fatores intensificam a interação entre humanos, vetores e hospedeiros silvestres (Silva et al., 2025; Zhang, 2024).

Tipicamente, está presente em um ciclo silvestre atingindo lagartas, primatas não humanos e pássaros. Porém, é possível circular em um ciclo urbano e afetar pessoas (Pastula; Beckham; Tyler, 2025). O principal vetor é o *Culicoides paraensis*, conhecido como maruim ou mosquito pólvora, porém o *Culex quinquefasciatus* pode transmitir também. Há registros de isolamento do OROV nos *Coquillettidia venezuelensis* e *Aedes serratus* (Brasil, 2024).

A capacidade do OROV de persistir em ambientes contribui para sua disseminação. Após a contaminação, o período de incubação gira em torno de quatro a oito dias. Cerca de 60% das infecções são sintomáticas, causando febre abrupta, cefaleia, mialgia, artralgia, distúrbios gastrointestinais, dor retro-ocular e/ou fotofobia durante 3 a 10 dias. É possível que a doença siga um curso bifásico, com o

retorno dos sintomas uma semana depois de seu início (Okesanya et al., 2025; Pastula; Beckham; Tyler, 2025).

Alguns estudos indicam transmissão vertical com influência no desenvolvimento fetal, de forma similar ao Zika vírus (Chaves et al., 2025; Pastula; Beckham; Tyler, 2025). Invasão neuronal do OROV foi observada em modelos animais. Em humanos, foi detectado ocorrência de meningite, encefalite e síndrome de Guillain-Barré (Pastula; Beckham; Tyler, 2025; González-Quevedo; Farrill; Becquer, 2025).

Devido a sua similaridade sintomática com outras doenças e a falta de recursos diagnósticos em áreas endêmicas, a FO é subnotificada. Isso torna difícil compreender seu real impacto. Por isso, é necessário um enfoque maior da saúde pública no que tange à vigilância, à educação pública e colaboração internacional no combate a essa doença (Okesanya et al., 2025).

O diagnóstico pode ser feito por detecção de ácidos nucleicos virais, imunoglobulinas M, neutralização de anticorpos no soro e/ou líquido cefalorraquidiano. Não há tratamento ou vacinas, portanto o combate ao mosquito é essencial (Pastula; Beckham; Tyler, 2025).

No Brasil, recomenda-se o diagnóstico via RT-qPCR em casos negativos para outras arboviroses desde 2023 (Silva et al., 2025; Gomes et al., 2025). A detecção de anticorpo IgM e IgG é possível, mas não possui sensibilidade e especificidade identificadas para OROV. Os métodos SHERLOCK (Specific High-Sensitivity Enzymatic Reporter UnLOCKing) e DETECTR (DNA Endonuclease-Targeted CRISPR Trans Reporter) da tecnologia CRISPR/Cas possuem potencial de serem novas ferramentas diagnósticas para FO. (Silva et al., 2025). Em uma revisão sistemática, foram analisadas sensibilidade e limites de detecção de métodos diagnósticos de OROV e “a ddPCR demonstrou maior sensibilidade que a RT-qPCR, sendo capaz de detectar até uma 1 cópia/ μL ” (Gomes et al., 2025).

Em setembro de 2024, foi detectado um caso de coinfeção de OROV e dengue sorotipo 3 em um paciente retornando de Cuba para a Itália. Para o diagnóstico, foi utilizado o método RT-PCR (Reação em Cadeia da Polimerase com Transcrição Reversa). Isso indica que coinfeções entre múltiplas arboviroses são possíveis e devem ser consideradas. Nesse contexto, o diagnóstico laboratorial é essencial. Assim, informações sobre em quais partes do corpo o OROV é

armazenado e expelido tornam-se chave para refinar métodos diagnósticos e orientar políticas públicas (Colavita et al., 2025).

2.1.6.3 Modelos de previsão de séries temporais

Nesse cenário, os modelos de previsão de séries temporais de doenças infecciosas se destacam como abordagem estratégica para apoiar a formulação de políticas públicas mais eficazes. Afinal, com a capacidade de identificar surtos de forma precoce, promove-se o uso otimizado de recursos e viabiliza-se a implementação de medidas preventivas antes que a situação alcance níveis críticos (Biggerstaff, 2018; Lutz, 2019). Uma revisão sistemática sobre modelos de predição utilizados para arboviroses relatou que os modelos de Poisson, média móvel (ARIMA, SARIMA, ARIMAX, SARIMAX) e redes neurais artificiais, principalmente o LSTM, são os mais utilizados (Lima et al., 2022).

Inicia-se a análise de uma série temporal pela decomposição em três componentes: tendência (T_t), sazonalidade (S_t) e irregularidade (I_t). Após isso, pode-se retornar à série original pela soma dos componentes em $Y_t = T_t + S_t + I_t$ (Wang, 2024). Para testar a estacionalidade, aplica-se o teste de Dickey–Fuller Aumentado (Hasan et al., 2024). Para analisar a tendência da curva, utiliza-se o teste de Mann–Kendall. A irregularidade representa a aleatoriedade, sendo impossível de analisar (Wang, 2024).

Modelo ARIMA é um acrônimo para modelo autorregressivo (AR) integrado (I) de média móvel (MA) (Hasan et al., 2024). As respectivas ordens de cada modelo são representadas por ARIMA (p, d, q). Ele pode ser adaptado para captar sazonalidade na série temporal, tornando-se o ARIMA sazonal (SARIMA). Nessa versão, é grafado como SARIMA (p, d, q) (P, D, Q) s . O componente “s” define o número de observações (Bhatnagar et al., 2012).

O modelo AR observa a influência entre os valores passados da série com os valores futuros, o modelo MA utiliza os erros anteriores para ajustar as previsões e o modelo I diferencia a série, tornando-a estacionária (Silva et al., 2023). As ordens dos modelos de autorregressão (AR) e de média móvel (MA) são identificadas via função de autocorrelação parcial (PACF) e função de autocorrelação (ACF) respectivamente (Bhatnagar et al., 2012).

Finalmente, compara-se os modelos pelo critério de informação de Akaike (AIC) e opta-se por aquele com o menor valor de AIC (Martinez; Silva, 2011a). O AIC

é uma medida de qualidade da adequação do modelo que penaliza a sua complexidade. Uma alternativa a ele é o Critério de Informação Bayesiano (BIC), que penaliza ainda mais a quantidade de parâmetros (Hasan et al., 2024). A seguir, verifica-se a qualidade do ajuste de cada modelo a partir da verificação de autocorrelação e autocorrelação parcial dos resíduos por meio do teste de Ljung-Box (Martinez; Silva, 2011b).

É possível também analisar a influência de fatores externos nos valores futuros da variável analisada. Isso é feito pela extensão do modelo SARIMA para SARIMAX (modelo autorregressivo integrado de média móvel sazonal com variáveis exógenas). Desse modo, tende-se a melhorar o desempenho preditivo por aumentar a análise das relações causais (Ferreira, 2024). Essa melhora de eficácia é percebida tanto em modelos estatísticos quanto de aprendizado de máquina na previsão de casos futuros de arboviroses. Afinal, temperatura, precipitação e humidade do ar estão relacionados com a proliferação dos mosquitos transmissores (Chen; Moraga, 2025).

As Redes Neurais Recorrentes (RNN) são uma categoria de rede neural adequada para analisar séries temporais. Elas possuem conexões que transmitem informações entre diferentes momentos no tempo, com cada neurônio recebendo dados da camada anterior e de sua própria ativação na passagem anterior. Esse mecanismo possibilita a utilização de autocorrelação dos dados na tomada de decisão da rede (Orssatto et al., 2022).

Os modelos baseados em RNN possuem como pontos negativos cálculos lentos, falta de consideração de entradas futuras e o problema do gradiente de fuga, que dificulta o ajuste dos pesos da rede. Para solucionar esse último ponto, criou-se um novo tipo de RNN: o modelo de Memória de Curto e Longo Prazo (LSTM). Sua arquitetura inclui três portas: esquecimento, entrada e saída (Campanhol, 2024).

A porta de esquecimento define quais informações excluir do estado da célula. Para tal, é utilizada uma função de ativação sigmoide, com saída binária. Se a saída for 1, os dados serão armazenados e, se for 0, os dados serão descartados. Já a porta de entrada, determina qual informação adicionar ao estado da célula. Este procedimento consiste em duas operações: Primeiro, computam-se os valores adicionáveis aos estados da célula. Em seguida, os valores de ativação da entrada e das portas são calculados. Os novos estados da célula são decididos com base nos resultados das etapas anteriores (ICCAI, 2019).

Após a célula de memória passar pela porta de entrada e esquecimento, a porta de saída gera a saída. Duas portas serão implementadas nas portas de saída: A primeira porta usa uma camada sigmoide para determinar quais partes do estado da célula serão geradas. Um valor será colocado na célula de memória a partir da função de ativação. Finalmente, as duas portas são multiplicadas para produzir um valor que será distribuído (Othman et al., 2022). As redes LSTM conseguem incorporar valores exógenos nas previsões, o que as torna uma boa opção para estimativas de casos futuros de arboviroses (Chen; Moraga, 2025).

Para avaliar o desempenho dos modelos preditivos, pode-se utilizar o erro percentual absoluto médio (MAPE), o erro quadrático médio (RMSE) e o erro absoluto médio (MAE) (Tang, 2022). O MAPE representa a porcentagem da diferença entre o valor previsto e o valor real em relação ao valor real. O RMSE trata-se da diferença média entre o valor previsto e o valor real. O MAE denota o desvio absoluto médio entre os valores previstos e os reais. Quanto menor forem tais índices, melhor o ajuste do modelo (Wang, 2024).

2.1.6.4 Testes na predição de outras arboviroses

O modelo SARIMA foi vastamente testado para previsão dos casos de dengue. Em um estudo na região de Paranavaí (PR), utilizou-se o teste de Ljung-Box para afirmar que o SARIMA previu casos para o ano subsequente com relativa precisão (Silva et al., 2023). Também a partir do teste de Ljung-Box e da função de autocorrelação (ACF), um estudo similar em Ribeirão Preto (SP) julgou o SARIMA adequado para estimativa futura de casos de dengue (Martinez; Silva, 2011a). A partir desses mesmos critérios de análise, um estudo em Campinas (SP) validou a eficácia do SARIMA (Martinez; Silva; Fabbro, 2011b).

Um estudo em Foz do Iguaçu comparou o ARIMA e o SARIMA na previsão de internações por dengue. O ARIMA se mostrou relativamente adequado, porém a necessidade de captar sazonalidade se refletiu em um melhor desempenho do SARIMA. Porém, mesmo o SARIMA se mostrou deficiente na modelagem de eventos extremos. Desse modo, o estudo recomenda a adoção de modelos mais completos (SLAEDR, 2024).

Assim como no Brasil, um estudo tailândes verificou, por meio do MAPE, que o SARIMA demonstrou confiabilidade de um ano para previsões de casos futuros de dengue (ICIT 2019: IoT and Smart City, 2020). Um estudo em Bangladesh também

confirmou a eficácia do SARIMA na predição dos casos de dengue via MAPE, teste de Ljung-Box e ACF (Hasan et al., 2024). Na Índia, o SARIMA foi validado via MAPE e teste de Ljung-Box (Bhatnagar, 2012). Outro estudo comparou as previsões do SARIMA para vários países (Brasil, México, Cingapura, Sri Lanka e Tailândia) com a incidência média histórica e descobriu maior precisão do SARIMA em todas as populações, principalmente em previsões para 1 a 2 meses (Riley et al., 2020).

Um estudo em Porto Alegre comparou o SARIMA com o SARIMAX, incluindo variáveis climáticas no modelo. Sua conclusão foi que essa adição contribuiu para aumentar a precisão das previsões. Afinal, a transmissão da dengue é influenciada por condições ambientais. O SARIMA obteve MAE = 53,14 e MAPE = 125,67% no primeiro período de análise e MAE = 56,35 e MAPE = 620,34% no segundo. Já o SARIMAX, obteve MAE = 53,02 e MAPE = 98,00% no primeiro período analisado e MAE = 56,16 e MAPE = 587,06% no segundo (Ferreira, 2024).

Em um estudo feito em Jambi, na Indonésia, foi comparado o SARIMA e o LSTM na predição de casos de Febre Hemorrágica da Dengue. O desempenho da previsão foi avaliado com base no MAE e no RMSE. Ambos apresentaram bom desempenho relativamente, porém o SARIMA se mostrou mais adequado. Seu RMSE foi de 30,07 e o MAE de 18,97, enquanto o LSTM obteve RMSE de 30,41 e MAE de 18,27 (ICCAI, 2019). Outro estudo da Indonésia, dessa vez em Surabaya, também indicou melhor desempenho do SARIMA em relação ao ARIMA e LSTM na predição de casos de dengue a partir das métricas RMSE, MAE e MAPE (Othman et al., 2022).

Um estudo de Bangladesh comparou modelos preditivos para internações hospitalares por dengue no país. Três modelos foram comparados: SARIMA, Reforço de Gradiente Extremo (XGBoost) e Regressor de Floresta Aleatória (RFR). O SARIMA mostrou melhor desempenho com o menor MSE (0,3011) e maior R^2 (0,9418) (COMNETSAT, 2024).

Um estudo em Cascavel (PR) comparou os modelos de Média Móvel, Suavização Exponencial, ARIMA, Máquina de Vetores de Suporte (SVM), Perceptrons Multicamadas (MLP), Florestas Aleatórias e Redes Neurais Recorrentes (RNN) para predição de casos de dengue. Utilizou-se o número de notificações no passado, a temperatura média, o índice pluviométrico e a umidade relativa do ar. O MLP teve o melhor desempenho, com RMSE de 3,163. Em segundo lugar, veio o

Floresta Aleatório (RMSE de 3,264) e, em terceiro, o de Média Móvel (RMSE de 4,123) (Orssatto et al., 2022).

Um estudo feito em Recife (PE), comparou os modelos SVM, Floresta Aleatória, MLP, Máquina de Aprendizado Extremo (ELM), Máquina de Estado de Eco (ESM), Rede de Estado de Eco Profundo (Deep-ESN) e Regressão Linear (LR) objetivando prever casos de dengue, chikungunya e zika. Foram utilizadas as séries históricas de casos e informações climáticas e ambientais (velocidades do vento, temperaturas e precipitação). O melhor desempenho foi da Floresta Aleatória de 10 árvores, com correlação de Pearson acima de 0,99 e RMSE (%) abaixo de 6% (Silva et al., 2022).

Um estudo feito no Rio de Janeiro (RJ) comparou os modelos Autorregressivo (AR), Média Móvel (MA), ARIMA, Suavização Exponencial, SARIMA, SARIMAX, Autorregressão Vetorial (VAR), Floresta Aleatória, XGBoost, SVM, LSTM e Prophet na predição de casos de dengue. Os modelos SARIMA e SARIMAX obtiveram os melhores desempenhos com MAE de 78,26 e 79,24 respectivamente e MAPE de 18,25% e 17,25% respectivamente (Chen; Moraga, 2025).

Um estudo em Porto Alegre (RS) utilizou para predição dos casos de dengue um modelo LSTM com discretização dos dados a partir de técnicas de SAX e qSAX para detectar surtos. Foi analisada também a influência das variáveis climáticas por meio de um modelo de Floresta Aleatória, concluindo que as maiores influências vinham do ponto de orvalho, da pressão e da chuva. Após a aplicação de diferentes configurações para as árvores de decisões, a maior acurácia atingida foi de 78,92% (Campanhol, 2024).

2.1.7 Metodologia

2.1.7.1 Tipo de estudo

Trata-se de um estudo observacional, ecológico, de caráter descritivo e analítico com abordagem quantitativa.

2.1.7.2 Local e período de realização da coleta

O estudo será realizado na cidade de Passo Fundo/RS, nas dependências da Universidade Federal da Fronteira Sul – *Campus* Passo Fundo/RS, junto ao curso de Medicina, no período de agosto de 2025 a julho de 2026.

2.1.7.3 População, amostragem e limitações do estudo

Esse estudo utilizará dados secundários disponíveis no Painel Epidemiológico do Ministério da Saúde referente às notificações de novos casos semanais de Febre do Oropouche (FO) no Brasil. A população do estudo será composta pelos casos de FO notificados entre janeiro de 2023 e dezembro de 2025.

Todos os registros presentes no Painel Epidemiológico serão incluídos na análise. Não haverá cálculo e definição de tamanho de amostra, pois, por incluir todos os registros, torna-se uma análise de um Censo. Espera-se incluir um total de 24.044 casos.

2.1.7.4 Variáveis e coleta de dados

A coleta de dados secundários será realizada pela acadêmica autora do projeto diretamente do Painel Epidemiológico do Ministério da Saúde seguindo o passo a passo: site do Ministério da Saúde <https://www.gov.br/saude/pt-br> > Assuntos > Saúde de A a Z > O > Oropouche > Painel Epidemiológico.

Esse estudo realizará uma análise das variáveis sociodemográficas disponíveis no Painel Epidemiológico referentes a sexo, faixa etária e região de residência. A organização será da seguinte forma: sexo, (feminino e masculino); faixa etária categorizada em anos (0-9; 10-19; 20-29; 30-39; 40-49; 50-59; 60 ou +), região de residência (norte, nordeste, sudeste, centro-oeste e sul). Serão descritas as proporções de casos de FO relatados em cada um desses subgrupos.

Também serão coletadas as variáveis climáticas de índices pluviométricos, umidade do ar e temperatura. Elas serão adicionadas aos modelos como variáveis exógenas, para que se leve em conta a sua influência no valor futuro das notificações junto com os valores anteriores da série temporal. Isso porque, no caso de arboviroses, os dados climáticos podem ter uma participação significativa entre os determinantes de prevalência. Assim, aumenta-se a precisão do modelo ao incluí-las no algoritmo.

Elas serão coletadas no site Prediction Of Worldwide Energy Resources (POWER) da National Aeronautics and Space Administration (NASA) seguindo o passo a passo: site do POWER <https://power.larc.nasa.gov/> > Regional. Afinal, a Febre do Oropouche, por ser uma arbovirose, tem sua disseminação influenciada de forma importante por esses elementos. Serão coletados dados do estado com o maior número de notificações da doença no período analisado.

2.1.7.5 Processamento, controle de qualidade e análise dos dados

Inicia-se a análise de uma série temporal pela decomposição em três componentes: tendência (T_t), sazonalidade (S_t) e irregularidade (I_t) (Wang, 2024). Para testar a estacionalidade, aplica-se o teste de Dickey–Fuller Aumentado (Hasan et al., 2024). Para analisar a tendência da curva, utiliza-se o teste de Mann–Kendall. A irregularidade não se analisa (Wang, 2024). Para implementação dessas etapas, serão utilizadas as bibliotecas StatsModels e pyMannKendall da linguagem de programação Python.

Modelo ARIMA é um acrônimo para modelo autorregressivo (AR) integrado (I) de média móvel (MA) (Hasan et al., 2024). As respectivas ordens de cada modelo são representadas por ARIMA (p, d, q). Ele pode ser adaptado para captar sazonalidade na série temporal, tornando-se o ARIMA sazonal (SARIMA). Nessa versão, é grafado como SARIMA (p, d, q) (P, D, Q) s . O componente “ s ” define o número de observações (Bhatnagar et al., 2012).

O modelo AR observa a influência entre os valores passados da série com os valores futuros, o modelo MA utiliza os erros anteriores para ajustar as previsões e o modelo I diferencia a série, tornando-a estacionária (Silva et al., 2023). As ordens dos modelos de autorregressão (AR) e de média móvel (MA) são identificadas via função de autocorrelação parcial (PACF) e função de autocorrelação (ACF) respectivamente (Bhatnagar et al., 2012).

Finalmente, compara-se os modelos pelo critério de informação de Akaike (AIC) e opta-se por aquele com o menor valor de AIC (Martinez; Silva, 2011a). O AIC é uma medida de qualidade da adequação do modelo que penaliza a sua complexidade. Uma alternativa a ele é o Critério de Informação Bayesiano (BIC), que penaliza ainda mais a quantidade de parâmetros (Hasan et al., 2024). A seguir, verifica-se a qualidade do ajuste de cada modelo a partir da verificação de autocorrelação e autocorrelação parcial dos resíduos por meio do teste de Ljung-Box (Martinez; Silva, 2011b).

É possível também analisar a influência de fatores externos nos valores futuros da variável analisada. Isso é feito pela extensão do modelo SARIMA para SARIMAX (modelo autorregressivo integrado de média móvel sazonal com variáveis exógenas). Desse modo, tende-se a melhorar o desempenho preditivo por aumentar a análise das relações causais (Ferreira, 2024). Afinal, temperatura, precipitação e

umidade do ar estão relacionados com a proliferação dos mosquitos transmissores (Chen; Moraga, 2025).

Para implementação do SARIMA e SARIMAX, será utilizada a biblioteca StatsForecast da linguagem de programação Python. Ela possui tanto ferramentas para fazer as previsões quanto para calcular automaticamente qual a ordem ideal do modelo.

As Redes Neurais Recorrentes (RNN) são uma categoria de rede neural adequada para analisar séries temporais. Elas possuem conexões que transmitem informações entre diferentes momentos no tempo, com cada neurônio recebendo dados da camada anterior e de sua própria ativação na passagem anterior. Esse mecanismo possibilita a utilização de autocorrelação dos dados na tomada de decisão da rede (Orssatto et al., 2022).

Os modelos baseados em RNN possuem como pontos negativos cálculos lentos, falta de consideração de entradas futuras e o problema do gradiente de fuga, que dificulta o ajuste dos pesos da rede. Para solucionar esse último ponto, criou-se um novo tipo de RNN: o modelo de Memória de Curto e Longo Prazo (LSTM). Sua arquitetura inclui três portas: esquecimento, entrada e saída (Campanhol, 2024).

A porta de esquecimento define quais informações excluir do estado da célula. Para tal, é utilizada uma função de ativação sigmoide, com saída binária. Se a saída for 1, os dados serão armazenados e, se for 0, os dados serão descartados. Já a porta de entrada, determina qual informação adicionar ao estado da célula. Este procedimento consiste em duas operações: Primeiro, computam-se os valores adicionáveis aos estados da célula. Em seguida, os valores de ativação da entrada e das portas são calculados. Os novos estados da célula são decididos com base nos resultados das etapas anteriores (ICCAI, 2019).

Após a célula de memória passar pela porta de entrada e esquecimento, a porta de saída gera a saída. Duas portas serão implementadas nas portas de saída: A primeira porta usa uma camada sigmoide para determinar quais partes do estado da célula serão geradas. Um valor será colocado na célula de memória a partir da função de ativação. Finalmente, as duas portas são multiplicadas para produzir um valor que será distribuído (Othman et al., 2022).

As redes LSTM conseguem incorporar valores exógenos nas previsões, o que as torna uma boa opção para estimativas de casos futuros de arboviroses (Chen;

Moraga, 2025). Para implementar o LSTM, será utilizada a biblioteca Keras da linguagem de programação Python.

Para avaliar o desempenho dos modelos preditivos, utiliza-se o erro percentual absoluto médio (MAPE), o erro quadrático médio (RMSE) e o erro absoluto médio (MAE) (Tang, 2022). O MAPE representa a porcentagem da diferença entre o valor previsto e o valor real em relação ao valor real. O RMSE trata-se da diferença média entre o valor previsto e o valor real. O MAE denota o desvio absoluto médio entre os valores previstos e os reais. Quanto menor forem tais índices, melhor o ajuste do modelo (Wang, 2024). Para avaliar a eficácia dos modelos, será utilizada a biblioteca SkLearn da linguagem de programação Python.

As bibliotecas escolhidas para utilização nessa pesquisa surgiram em um contexto de desenvolvimento de versões para Python de ferramentas estatísticas já presentes na linguagem R. O objetivo de suas criações é tornar o Python a principal linguagem para análise estatística (SCIPY, 2011). Nesse contexto, considera-se válida a experimentação de suas funções. Serão feitas previsões e validações dessas previsões para um ano e um semestre. O código será escrito e executado em nuvem por meio da ferramenta Google Colabs, que é gratuita.

2.1.7.6 Aspectos éticos

A coleta de dados será realizada pelo Painel Epidemiológico do Ministério da Saúde, utilizando dados secundários de domínio público. Por isso, será dispensada pela Comissão Nacional de Ética em Pesquisa (CONEP) a submissão ao Comitê de Ética em Pesquisa (CEP/CONEP) com Seres Humanos da UFFS, resolução nº510, de 07 de abril de 2016 do Conselho Nacional de Saúde (CNS).

Como os dados coletados são secundários e agregados por municípios, não existe risco de identificação pessoal dos participantes envolvidos. Os dados coletados serão armazenados no computador da acadêmica responsável, com acesso restrito por login e senha, durante cinco anos e, depois, serão excluídos.

Quanto aos benefícios do estudo, tem-se principalmente a possibilidade de otimização das campanhas públicas de saúde brasileiras com relação à prevenção da Febre do Oropouche. Afinal, não será necessário esperar que a situação aconteça de fato para mapeá-la e tentar modificá-la.

2.1.8 Recursos

Os materiais serão adquiridos com recursos próprios.

Quadro 1 – Orçamento

ITEM	QUANTIDADE	CUSTO
Notebook	1	R\$ 4.000,00
Energia elétrica	100 Kwh	R\$ 200,00
Internet	1 (240Mb)	R\$ 120,00
Total		R\$ 3.320,00

Fonte: Própria, 2025.

2.1.9 Cronograma

Quadro 2 – Cronograma de execução (agosto de 2025 a junho de 2026)

ATIVIDADE PREVISTA	PERÍODO
Revisão de literatura	Agosto de 2025 a julho de 2026
Processamento e análise de dados	Agosto de 2025 a dezembro de 2025
Redação do artigo	Janeiro de 2026 a junho de 2026
Divulgação dos resultados	Julho de 2026

Fonte: Própria, 2025.

2.1.10 Referências

ANNUAL SCIENTIFIC COMPUTING WITH PYTHON CONFERENCE (SCIPY), 2011, Austin. Tema: Time Series Analysis in Python with statsmodels. **Proceedings of SciPy** [...]. Proceedings of SciPy, 2011. DOI: <https://doi.org/10.25080/Majorebaa42b7-012>. Disponível em: https://www.researchgate.net/publication/340142416_Time_Series_Analysis_in_Python_with_statsmodels. Acesso em: 15 mai. 2025.

AUNG, Soe Htet et al. A SARIMA time series forecasting for dengue cases for reporting to Yangon Region, Myanmar. **Journal of Public Health and Development**, v. 22, n.1, 2024. DOI: <https://doi.org/10.55131/jphd/2024/220114>. Disponível em: https://www.researchgate.net/profile/Soe-Htet-Aung/publication/378302945_A_SARIMA_time_series_forecasting_for_dengue_cases_for_reporting_to_Yangon_Region_Myanmar/links/65d31175e51f606f9979e4de/A-SARIMA-time-series-forecasting-for-dengue-cases-for-reporting-to-Yangon-Region-Myanmar.pdf. Acesso em: 04 abr. 2024.

BHATNAGAR, Sunil et al. Forecasting Incidence of Dengue in Rajasthan, Using Time Series Analyses. **Indian Journal of Public Health**, v. 56, n. 4, p 281-285, 2012. DOI: <https://doi.org/10.4103/0019-557X.106415>. Disponível em: https://journals.lww.com/ijph/fulltext/2012/56040/forecasting_incidence_of_dengue_in_rajasthan,.5.aspx. Acesso em: 04 abr. 2025.

BIGGERSTAFF, Matthew et al. Results from the second year of a collaborative effort to forecast influenza seasons in the United States. *Epidemics*. **BMC Public Health**, v. 24, p. 26-33, 2018. DOI: <https://doi.org/10.1016/j.epidem.2018.02.003>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/29506911/>. Acesso em: 26 dez. 2024.

BRASIL, Ministério da Saúde. **Nota técnica nº 6/2024-CGAR/DEDT/SVSA/MS**. Orientações para a vigilância da Febre do Oropouche. Brasília: Ministério da Saúde, 2024.

BRASIL, Ministério da Saúde. **Plataforma GovBR**, 2025. Oropouche > Painel Epidemiológico. Disponível em: <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/o/oropouche/painel-epidemiologico>. Acesso em: 10 abr. 2025.

BRASIL, Ministério da Saúde. **Plataforma GovBR**, 2024. Oropouche. Disponível em: <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/o/oropouche>. Acesso em: 14 mar. 2025.

BRASIL. Ministério da Saúde. **Portaria GM/MS nº 217, de 1º de março de 2023**. Altera a Lista Nacional de Notificação Compulsória de Doenças, Agravos e Eventos de Saúde Pública. Diário Oficial da União: seção 1, Brasília, DF, 2 mar. 2023.4

CAMPANHOL, Andriza Maria da Cunha. **Explicabilidade de modelo Long Short-Term Memory (LSTM) para previsão de dengue em Porto Alegre**. 2024. 45 f. TCC (Graduação) – Curso de Ciência da Computação, Universidade Federal de Santa Maria, Santa Maria, RS. Disponível em: <https://repositorio.ufsm.br/handle/1/32817>. Acesso em: 3 mai. 2025.

CHAVES, Hellyângela Maria da Silva et al. Oropouche: Integrative review on the implications of vertical transmission in a human host. **Research, Society and Development**, [S. l.], v. 14, n. 2, p. e5414248219, 2025. DOI: <https://doi.org/10.33448/rsd-v14i2.48219>. Disponível em: <https://rsdjournal.org/index.php/rsd/article/view/48219>. Acesso em: 23 mar. 2025.

CHEN, Xiang; MORAGA, Paula. Assessing dengue forecasting methods: a comparative study of statistical models and machine learning techniques in Rio de Janeiro, Brazil. **Tropical Medicine and Health**, v. 53, n. 1, p. 52, 2025. DOI: <https://doi.org/10.1186/s41182-025-00723-7>. Disponível em: <https://tropmedhealth.biomedcentral.com/articles/10.1186/s41182-025-00723-7>. Acesso em: 3 mai. 2025.

COLAVITA, Francesca et al. Prolonged detection of Oropouche virus RNA in whole blood samples. **Lancet Infectious Diseases**, v. 25, n. 1, p. E-11-E12, 2025. DOI: [https://doi.org/10.1016/S1473-3099\(24\)00798-9](https://doi.org/10.1016/S1473-3099(24)00798-9). Disponível em: [https://www.thelancet.com/journals/laninf/article/PIIS1473-3099\(24\)00798-9/fulltext](https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(24)00798-9/fulltext). Acesso em: 23 mar. 2025.

ESTADOS UNIDOS, Departamento de Saúde e Serviços Humanos, Centros de Controle e Prevenção de Doenças. **Clinical Overview of Oropouche Virus**

Disease. Atlanta, 2025. Disponível em: <https://www.cdc.gov/oropouche/hcp/clinical-overview/index.html>. Acesso em: 14 mar. 2025.

FERREIRA, Raul Santos. **Predição semanal de números de casos de dengue em Porto Alegre por modelos de séries temporais.** 2024. 99 f. TCC (Graduação) – Curso de Engenharia de Controle e Automação, Universidade Federal de Santa Maria, Santa Maria, RS. Disponível em: <https://repositorio.ufsm.br/handle/1/32624>. Acesso em: 7 abr. 2025.

GOMES, Larissa Santos Carneiro. et al. Ensaios Diagnósticos para Detecção do Vírus Oropouche em Amostras Humanas: Revisão Sistemática. **Pleidade**, v. 19, n. 46, p. 115-118, 2025. DOI: <https://doi.org/10.32915/pleiade.v19i46.1141>. Disponível em: <https://pleiade.uniamerica.br/index.php/pleiade/article/view/1141/1343>. Acesso em: 31 mar. 2025.

GONZÁLEZ-QUEVEDO, Alina; FARRILL, Zurina Lestayo O; BECQUER, Reinaldo Mustelier. Oropouche virus – another antecedent event for Guillain–Barré syndrome?. **Pan American Journal of Public Health**, v. 49, n. 2, 2025. ISSN-e: 1680-5348. ISSN: 1020-4989. Disponível em: <https://dialnet.unirioja.es/servlet/articulo?codigo=10018764>. Acesso em: 31 mar. 2025.

HASAN, Pratyay et al. Predictive Insights From ARIMA and SARIMA Models in Bangladesh: A Time Series Analysis. **Health Science Reports**, v. 7, n. 12, p. e70276. DOI: <https://doi.org/10.1002/hsr2.70276>. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1002/hsr2.70276>. Acesso em: 04 abr. 2025.

INTERNATIONAL CONFERENCE ON COMPUTING AND APPLIED INFORMATICS (ICCAI 2019), 4., 2019, Medan. **Journal of Physics: Conference Series** [...]. Universidade Utara Malaysia, 2019. Tema: A comparison of SARIMA and LSTM in forecasting dengue hemorrhagic fever incidence in Jambi, Indonesia. DOI: <https://doi.org/10.1088/1742-6596/1566/1/012054>. Disponível em: <https://iopscience.iop.org/article/10.1088/1742-6596/1566/1/012054/meta>. Acesso em: 10 abr. 2025.

INTERNATIONAL CONFERENCE ON COMMUNICATION, NETWORKS AND SATELLITE (COMNETSAT), 2024, Cox's Bazar. **IEEE Xplore** [...]. IEEE, 2024. Tema: Enhancing Dengue Outbreak Prediction in Bangladesh: A Comparative Study of Advanced Predictive Models. DOI: <https://doi.org/10.1109/COMPAS60761.2024.10796965>. Disponível em: <https://ieeexplore.ieee.org/abstract/document/10796965>. Acesso em: 10 abr. 2025.

INTERNATIONAL CONFERENCE ON INFORMATION TECHNOLOGY: IOT AND SMART CITY, 7., 2019, Shanghai. **ICIT 2019: IoT and Smart City** [...]. New York: Association for Computing Machinery, 2020. Tema: Time Series Analysis of Dengue Fever Cases in Thailand Utilizing the SARIMA Model. DOI: <https://doi.org/10.1145/3377170.3377215>. Disponível em: <https://dl.acm.org/doi/abs/10.1145/3377170.3377215>. Acesso em: 06 abr. 2025.

LIMA, Clarisse Lins de et al. Temporal and Spatiotemporal Arboviruses Forecasting by Machine Learning: A Systematic Review. **Frontiers in Public Health**, v. 10, 2022. DOI: <https://doi.org/10.3389/fpubh.2022.900077>. Disponível em: <https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2022.900077/full#B45>. Acesso em: 31 mar. 2025.

LUTZ, Chelsea S. et al. Applying infectious disease forecasting to public health: a path forward using influenza forecasting examples. **BMC Public Health**, v. 19, n. 1659, 2019. DOI: <https://doi.org/10.1186/s12889-019-7966-8>. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC6902553/>. Acesso em: 18 dez. 2024.

MARTINEZ, Edson Zangiacomi; SILVA, Elisângela Aparecida Soares da. Previsão do número de casos de dengue em Ribeirão Preto, São Paulo, Brasil, por um modelo SARIMA. **Caderno de Saúde Pública**, v. 27, n. 9, p. 1809-1818, 2011a. DOI: <https://doi.org/10.1590/S0102-311X2011000900014>. Disponível em: <https://www.scielo.br/j/csp/a/KGP9rcchBH7DHHhRNftFzwM/>. Acesso em: 04 abr. 2025.

MARTINEZ, Edson Zangiacomi; SILVA, Elisângela Aparecida Soares da; FABBRO, Amaury Lelis Dal. Um modelo SARIMA para predição do número de casos de dengue em Campinas, Estado de São Paulo. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 44, n. 4, 2011b. DOI: <https://doi.org/10.1590/S0037-86822011000400007>. Disponível em: <https://www.scielo.br/j/rsbmt/a/wQt5dCYymC9nvYVzfRMjgpJ/?lang=en>. Acesso em: 06 abr. 2025.

MENDOZA-LANDINEZ, Brayan Fabian; FREYLE-ROMAN, Ivette Karina; RINCÓN-OROZCO, Bladimiro. Virus Oropouche, un arbovirus emergente en búsqueda de protagonismo en las Américas. **Revista de la Universidad Industrial de Santander Salud**, v. 56, p. e24030. DOI: <https://doi.org/10.18273/saluduis.56.e:24030>. Disponível em: http://www.scielo.org.co/scielo.php?pid=S0121-08072024000124030&script=sci_arttext. Acesso em: 31 mar. 2025.

NETTO, Antonio Ruffino. Brasil: doenças emergentes ou reemergentes?. **Medicina (Ribeirão Preto), Ribeirão Preto**, v. 30, n. 3, p. 405, 1997. DOI: <https://doi.org/10.11606/issn.2176-7262.v30i3p405>. Disponível em: <https://www.revistas.usp.br/rmrp/article/view/6792>. Acesso em: 31 mar. 2025.

OKESANYA, John Olalekan et al. Addressing the emerging threat of Oropouche virus: implications and public health responses for healthcare systems. **Trop Dis Travel Med Vaccines**, v. 11, n. 1, 2025. DOI: <https://doi.org/10.1186/s40794-024-00236-x>. Disponível em: <https://link.springer.com/article/10.1186/s40794-024-00236-x>. Acesso em: 14 mar. 2025.

ORSSATTO, Angelo José et al. Aplicação de séries temporais na estimação do número de casos de dengue em Cascavel-PR. **Revista Brasileira de Computação Aplicada**, [S. l.], v. 14, n. 3, p. 37-50, 2022. DOI: <https://doi.org/10.5335/rbca.v14i3.13483>. Disponível em: <https://seer.upf.br/index.php/rbca/article/view/13483>. Acesso em: 10 abr. 2025.

OTHMAN, Mahmud et al. Model Forecasting Development for Dengue Fever Incidence in Surabaya City Using Time Series Analysis. **Processes**, v. 10, n. 11, p. 2454, 2022. DOI: <https://doi.org/10.3390/pr10112454>. Disponível em: <https://www.mdpi.com/2227-9717/10/11/2454>. Acesso em: 10 abr. 2025.

PASTULA, Daniel M; BECKHAM, J. David; TYLER, Kenneth L. Oropouche Virus: An Emerging Neuroinvasive Arbovirus. **Annals of Neurology**, v. 97, n. 1, p. 28-33, 2024. DOI: <https://doi.org/10.1002/ana.27139>. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/ana.27139>. Acesso em: 14 mar. 2025.

PEREIRA, Vitória Rocha et al. Aspectos epidemiológicos da sífilis no Brasil. **Estudos Avançados sobre Saúde e Natureza**, v. 3, 2023. Disponível em: <https://www.periodicojs.com.br/index.php/easn/article/view/1111>. Acesso em: 18 dez. 2024.

PUERTO, Fernando I. La fiebre de Oropouche en América Latina: estrategias de prevención y control. **Revista Biomédica**, v. 35, n. 3, 2024. DOI: <https://doi.org/10.32776/revbiomed.v35i3.1252>. Disponível em: <https://revistabiomedica.uady.mx/index.php/revbiomed/article/view/1252>. Acesso em: 31 mar. 2025.

RILEY, Pete. SARIMA Forecasts of Dengue Incidence in Brazil, Mexico, Singapore, Sri Lanka, and Thailand: Model Performance and the Significance of Reporting Delays. **Med RX IV**, 2020. DOI: <https://doi.org/10.1101/2020.06.26.20141093>. Disponível em: <https://www.medrxiv.org/content/10.1101/2020.06.26.20141093v1.full-text>. Acesso em: 07 abr. 2025.

SCHWEITZER, Kate. Little-Known Oropouche Virus Is Spreading: What Clinicians Need to Know About the Insect-Borne Illness. **JAMA Medical News**, v. 333, n. 6, p. 447-448, 2025. DOI: [doi:10.1001/jama.2024.28216](https://doi.org/10.1001/jama.2024.28216). Disponível em: <https://jamanetwork.com/journals/jama/article-abstract/2829242>. Acesso em: 23 mar. 2025.

SILVA, Ana Paula da; COUTINHO, Diógenes José Gusmão. Além da amazônia: a dispersão e os impactos da Febre Oropouche no Brasil. **Revista Ibero-Americana de Humanidades**, v. 10, n. 12, p. 3627–3642, 2024. DOI: <https://doi.org/10.51891/rease.v10i12.17656>. Disponível em: <https://periodicorease.pro.br/rease/article/view/17656>. Acesso em: 31 mar. 2025.

SILVA, Cecilia Cordeiro da et al. Spatiotemporal forecasting for dengue, chikungunya fever and Zika using machine learning and artificial expert committees based on meta-heuristics. **Research on Biomedical Engineering**, v. 38, p. 499–537, 2022. DOI: <https://doi.org/10.1007/s42600-022-00202-6>. Disponível em: <https://link.springer.com/article/10.1007/s42600-022-00202-6>. Acesso em: 10 abr. 2025.

SILVA, Edilene da Costa et al. Febre do oropouche no Brasil: um framework integrado para abordagem epidemiológica, imunológica e terapêutica de uma arbovirose emergente e negligenciada. **Revista Contemporânea**, v. 5, n. 2, p. e7427, 2025. DOI: <https://doi.org/10.56083/RCV5N2-034>. Disponível em:

<https://ojs.revistacontemporanea.com/ojs/index.php/home/article/view/7427>. Acesso em: 14 mar. 2025.

SILVA, Gabriel Manoel da et al. Uma aplicação do modelo SARIMA para prever o número de notificações de dengue na região de Paranavaí – Paraná. **Tecnologias avançadas em Ciências Exatas: Desenvolvimentos recentes e perspectivas futuras**, v. 1, n. 4, 2023. DOI: <https://doi.org/10.47402/ed.ep.c231934272>. Disponível em: <https://editorapublicar.com.br/ojs/index.php/publicacoes/article/view/630>. Acesso em: 04 abr. 2025.

SIMPÓSIO LATINO-AMERICANO DE ESTUDOS DE DESENVOLVIMENTO REGIONAL (SLAEDR), 4., 2024, Ijuí. **Anais do IV Simpósio Latino-Americano de Estudos de Desenvolvimento Regional (SLAEDR)** [...]. Unijuí, 2025. Tema: Comparação de modelos ARIMA, SARIMA e PROPHET na previsão de internações por dengue em Foz do Iguaçu (2008-2024). Disponível em: <https://publicacoeseventos.unijui.edu.br/index.php/slaedr/article/view/26837>. Acesso em: 3 mai. 2025.

SOUZA, Wanderley de coord.. **Doenças negligenciadas**. Rio de Janeiro: Academia Brasileira de Ciências, 2010.

TANG, Xijia et al. The evaluation of preventive and control measures on congenital syphilis in Guangdong Province, China: a time series modeling study. **Infection**, v. 50, p. 1179–1190, 2022. DOI: <https://doi.org/10.1007/s15010-022-01791-1>. Disponível em: <https://link.springer.com/article/10.1007/s15010-022-01791-1>. Acesso em: 27 dez. 2024.

WANG, Zhen D. et al. Analysis and forecasting of syphilis trends in mainland China based on hybrid time series models. **Cambridge University Press of Epidemiology and Infection**, v. 152, n. 93, 2024. DOI: [10.1017/S0950268824000694](https://doi.org/10.1017/S0950268824000694). Disponível em: <https://www.cambridge.org/core/journals/epidemiology-and-infection/article/analysis-and-forecasting-of-syphilis-trends-in-mainland-china-based-on-hybrid-time-series-models/9953AFB9C8FB3A9EAD6BBBF30818FA2A>. Acesso em: 18 dez. 2024.

ZHANG, Yuli et al. Oropouche virus: A neglected global arboviral threat. **Virus Research**, v. 341, 2024. DOI: <https://doi.org/10.1016/j.virusres.2024.199318>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S016817022400011X?via%3Dihub>. Acesso em: 31 mar. 2025.

ZHU, Zhixin et al. Development and comparison of predictive models for sexually transmitted diseases—AIDS, gonorrhea, and syphilis in China, 2011–2021. **Frontiers in Public Health**, v. 10, 2022. DOI: <https://doi.org/10.3389/fpubh.2022.966813>. Disponível em: <https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2022.966813/full>. Acesso em: 27 dez. 2024.

Apêndice A – Código para horizonte preditivo de 3 meses

```

"""**Aqui vamos importar as bibliotecas necessárias para a análise estatística**"""

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import statsmodels.api as sm
from statsmodels.tsa.statespace.sarimax import SARIMAX
from statsforecast import StatsForecast
from statsforecast.models import AutoARIMA
import pymannkendall as mk
import datetime
import sklearn
from sklearn.impute import SimpleImputer
from sklearn.preprocessing import MinMaxScaler
from sklearn.decomposition import KernelPCA
import math
import keras
import tensorflow as tf
import warnings
warnings.filterwarnings('ignore')
tf.random.set_seed(99)

"""**Aqui vamos ler o arquivo com os dados e formatá-los**"""

dados = pd.read_csv('FOX.csv')
dados['Mes'] = pd.to_datetime(dados['Mes'])

#Cria tabela para o SARIMA
df_sarima = {
    'unique_id' : dados['unique_id'],
    'ds' : dados['Mes'],
    'y' : dados['Casos']
}
df_sarima = pd.DataFrame(df_sarima)

#Cria série para o LSTM
df_lstm = {
    'Casos' : dados['Casos'],
    'Temp2M' : dados['Temp2M'],
    'AmplitudeTemp' : dados['AmplitudeTemp']
}
df_lstm = pd.DataFrame(df_lstm)
df_lstm = df_lstm.reset_index(drop=True)
scaler = MinMaxScaler(feature_range=(0, 1))
df_scaled = scaler.fit_transform(df_lstm.to_numpy())
df_scaled = pd.DataFrame(df_scaled, columns=list(df_lstm.columns))

#Cria tabela para análise da série
df_serie = {
    'Data' : df_sarima['ds'],
    'Casos' : dados['Casos']
}
df_serie = pd.DataFrame(df_serie)
df_serie = df_serie.set_index('Data')

#Graficando série temporal
df_serie.plot()

```

```

plt.suptitle('S rie hist rica de Notifica  es', fontsize=10)
plt.xticks(rotation=45)
plt.xlabel('Data')
plt.ylabel('Casos')
plt.tight_layout()

"""*Aqui vamos definir a fun  o de vari veis ex genas*"""

def singleStepSampler(df, window):
    xRes = []
    yRes = []
    for i in range(0, len(df) - window):
        res = []
        for j in range(0, window):
            r = []
            for col in df.columns:
                r.append(df[col][i + j])
            res.append(r)
        xRes.append(res)
        yRes.append(df[['Casos', 'Temp2M', 'AmplitudeTemp']].iloc[i +
window].values)
    return np.array(xRes), np.array(yRes)

"""*Aqui vamos aplicar alguns testes para analisar a s rie temporal*"""

#Aplicando teste dickey-fuller aumentado
from statsmodels.tsa.stattools import adfuller
print("Teste Dickey-Fuller aumentado")
print(adfuller(df_serie), '\n')

#Aplicando o teste de Mann-Kendall
print("Teste de Mann-Kendall")
print(mk.original_test(df_serie), '\n')

#Aplicando o teste Q de Ljung-Box na s rie temporal
print("Teste Q de Ljung-Box na s rie temporal")
print(sm.stats.acorr_ljungbox(df_serie, lags=[5], return_df=True), '\n')

#Decompondo a s rie temporal
from statsmodels.tsa.seasonal import seasonal_decompose
decomposition = seasonal_decompose(df_serie, model='additive', period=10)
fig = decomposition.plot()
fig.suptitle('Decomposi  o da S rie Temporal', fontsize=10)
plt.tight_layout()
plt.subplots_adjust(top=0.95)
plt.xticks(rotation=45)
plt.show()

#Salvando figura
plt.savefig("decomposicaoFO_1200dpi.png", dpi=1200, bbox_inches='tight')

"""*Define o horizonte preditivo*"""

horizon = 3

"""*Ordem  tima SARIMA*"""

#Preparar o SARIMA para fazer as previs es
from statsforecast.arima import arima_string
season_length = 12
models = [AutoARIMA(season_length=season_length)]

```

```

sf = StatsForecast(
    models=models,
    freq=pd.offsets.MonthBegin(),
    n_jobs=1
)

#Removendo as linhas que queremos prever
df_sarima_completo = df_sarima
numLinhas = len(df_sarima) - 1
for x in range(0,horizon):
    cont_linha = numLinhas - x
    df_sarima = df_sarima.drop(cont_linha)

#Encontrando ordem 3tima
sf.fit(df_sarima)
forecast_df= sf.predict(h=horizon, level=[95])
print("Ordem 3tima do modelo SARIMA")
for model, uid in zip(sf.fitted_, forecast_df.index.unique()):
    print(uid, ' - ', arima_string(model[0].model_))

"""**Aqui vamos ajustar as vari3veis ex3genas**"""

#L3a vari3veis ex3genas
ex = pd.read_csv('FOX.csv')
ex['Mes'] = pd.to_datetime(ex['Mes'])
ex.drop('Casos', inplace=True, axis=1)

#Retira as linhas que vamos prever
ex_completo = ex
numLinhas = len(ex) - 1
for x in range(0,horizon):
    cont_linha = numLinhas - x
    ex = ex.drop(cont_linha)

#Cria nosso array para a fazer a previs3o
temp2M_df = {
    'unique_id' : ex['unique_id'],
    'ds' : ex['Mes'],
    'y' : ex['Temp2M']
}

#Converte para DataFrame
temp2M_df = pd.DataFrame(temp2M_df)

#Mandando o modelo fazer as previs3es
sf.fit(temp2M_df)
forecast_Temp= sf.predict(h=horizon, level=[95])

#Cria nosso array com as previs3es
prevEx = pd.DataFrame({
    'forecast_Temp' : forecast_Temp['AutoARIMA']
})
prevEx=prevEx.to_numpy()

ex.drop(columns=['unique_id'], inplace=True)
ex.drop(columns=['Mes'], inplace=True)
ex.drop(columns=['AmplitudeTemp'], inplace=True)
ex = ex.reset_index(drop=True)

"""**Testando SARIMAX**"""

```

```

#Montando SARIMAX
model = SARIMAX(df_sarima['y'].to_numpy(), exog=ex.to_numpy(),
order=(2,0,0), seasonal_order=(0,1,0, 12))
results = model.fit(dis=False)
print(results.summary())

#Previendo com SARIMAX
forecast = results.get_forecast(exog=prevEx, steps=horizon)
previsoes = forecast.predicted_mean
print("Previsões")
print(previsoes)

"""**LSTM**"""

#Separando teste e validação
SPLIT = 0.92
(xVal, yVal) = singleStepSampler(df_scaled, horizon + 1)
X_train = xVal[:int(SPLIT * len(xVal))]
y_train = yVal[:int(SPLIT * len(yVal))]
X_test = xVal[int(SPLIT * len(xVal)):]
y_test = yVal[int(SPLIT * len(yVal)):]

#Rodando LSTM
multivariate_lstm = keras.Sequential()
multivariate_lstm.add(keras.layers.LSTM(1000,
input_shape=(X_train.shape[1], X_train.shape[2])))
multivariate_lstm.add(keras.layers.Dropout(0.2))
multivariate_lstm.add(keras.layers.Dense(3, activation='linear'))
multivariate_lstm.compile(loss = 'MeanAbsolutePercentageError',
metrics=['MSE'], optimizer='Adam')
multivariate_lstm.summary()
history = multivariate_lstm.fit(X_train, y_train, epochs=400)
loss_per_epoch = multivariate_lstm.history.history['loss']
plt.plot(range(len(loss_per_epoch)), loss_per_epoch)
predicted_values = multivariate_lstm.predict(X_test)
true_predictions = scaler.inverse_transform(predicted_values)
true_test = scaler.inverse_transform(y_test)
print("Previsões")
print(true_predictions)

"""**Aqui vamos validar o modelo SARIMAX-LSTM**"""

#Criando tabela de desempenho
validacao = {
    'SARIMAX': previsoes,
    'LSTM': true_predictions[:, 0],
    'Modelo SARIMAX-LSTM': true_predictions[:, 0],
    'Casos Reais': true_test[:, 0],
    'Resíduos': true_test[:, 0]
}
validacao = pd.DataFrame(validacao)
validacao.index = df_lstm.index[-len(true_test):]

#Calculando erros do modelo híbrido
linhas = validacao[validacao.columns[0]].count()
for y in range (0,linhas):
    x = y + len(df_serie) - horizon
    validacao['Modelo SARIMAX-LSTM'][x] = (validacao['SARIMAX'][x] +
validacao['LSTM'][x])/2
    validacao['Resíduos'][x] = validacao['Modelo SARIMAX-LSTM'][x] -
validacao['Casos Reais'][x]

```

```

#Separando parte de treino e valida  o
dados = pd.read_csv('FOX.csv')
dados['Mes'] = pd.to_datetime(dados['Mes'])
df_serie = {
    'Data' : dados['Mes'],
    'Casos' : dados['Casos']
}
df_serie = pd.DataFrame(df_serie)
treino = df_serie
teste = df_serie
for x in range (0, len(df_serie)):
    if x < len(df_serie) - horizon:
        teste = teste.drop(index = x)
    else:
        treino = treino.drop(index = x)

#Aplicando o teste de Ljung-Box nos res  duos do modelo h  brido
print("Teste Q de Ljung-Box nos res  duos do modelo h  brido")
print(sm.stats.acorr_ljungbox(validacao['Res  duos'], lags=[2],
return_df=True), '\n')

#Calculando erros modelo SARIMAX
print("M  tricas de erros do modelo SARIMAX")
from sklearn.metrics import mean_squared_error,
mean_absolute_percentage_error, mean_absolute_error
MSE = mean_squared_error(teste['Casos'], validacao['SARIMAX'])
RMSE = np.sqrt(MSE)
print("RMSE = {:.2f}".format(RMSE))
MAPE = mean_absolute_percentage_error(teste['Casos'], validacao['SARIMAX'])
print("MAPE = {:.0f}%".format(MAPE*100))
MAE = mean_absolute_error(teste['Casos'], validacao['SARIMAX'])
print("MAE = {:.0f}\n".format(MAE))

#Calculando erros modelo LSTM
print("M  tricas de erros do modelo LSTM")
from sklearn.metrics import mean_squared_error,
mean_absolute_percentage_error, mean_absolute_error
MSE = mean_squared_error(teste['Casos'], validacao['LSTM'])
RMSE = np.sqrt(MSE)
print("RMSE = {:.2f}".format(RMSE))
MAPE = mean_absolute_percentage_error(teste['Casos'], validacao['LSTM'])
print("MAPE = {:.0f}%".format(MAPE*100))
MAE = mean_absolute_error(teste['Casos'], validacao['LSTM'])
print("MAE = {:.0f}\n".format(MAE))

#Calculando erros modelo h  brido
print("M  tricas de erros do modelo h  brido")
from sklearn.metrics import mean_squared_error,
mean_absolute_percentage_error, mean_absolute_error
MSE = mean_squared_error(teste['Casos'], validacao['Modelo SARIMAX-LSTM'])
RMSE = np.sqrt(MSE)
print("RMSE = {:.2f}".format(RMSE))
MAPE = mean_absolute_percentage_error(teste['Casos'], validacao['Modelo
SARIMAX-LSTM'])
print("MAPE = {:.0f}%".format(MAPE*100))
MAE = mean_absolute_error(teste['Casos'], validacao['Modelo SARIMAX-LSTM'])
print("MAE = {:.0f}\n".format(MAE))

#Graficando erros
plt.plot(treino.index, treino['Casos'], label='Treino')

```

```

plt.plot(teste.index, teste['Casos'], label='Teste')
plt.plot(teste.index, validacao['SARIMAX'], label='SARIMAX')
plt.plot(teste.index, validacao['LSTM'], label='LSTM')
plt.plot(teste.index, validacao['Modelo SARIMAX-LSTM'], label='SARIMAX-LSTM')
#plt.plot(teste.index, validacao['ResÃduos'], label='ResÃduos')
plt.xticks(rotation=45)
plt.legend()

"""**Aqui vamos ajustar SARIMAX para o futuro**"""

#Buscando ordem Ãtima SARIMA
sf.fit(df_sarima_completo)
forecast= sf.predict(h=horizon, level=[95])
print("Ordem Ãtima do modelo SARIMA")
for model, uid in zip(sf.fitted_, forecast_df.index.unique()):
    print(uid, ' - ', arima_string(model[0].model_))

#Cria
nosso array para a fazer a previsÃo
temp2M_df = {
    'unique_id' : ex_completo['unique_id'],
    'ds' : ex_completo['Mes'],
    'y' : ex_completo['Temp2M']
}

ex_completo.drop(columns=['unique_id'], inplace=True)
ex_completo.drop(columns=['Mes'], inplace=True)
ex_completo.drop(columns=['AmplitudeTemp'], inplace=True)
ex_completo = ex_completo.reset_index(drop=True)

#Converte para DataFrame
temp2M_df = pd.DataFrame(temp2M_df)

#Mandando o modelo fazer as previsÃes
sf.fit(temp2M_df)
forecast_Temp= sf.predict(h=horizon, level=[95])

#Cria nosso array com as previsÃes
prev = pd.DataFrame({
    'forecast_Temp' : forecast_Temp['AutoARIMA']
})
prev=prev.to_numpy()

"""**Aqui vamos prever com SARIMAX**"""

#Definindo SARIMAX
model = SARIMAX(df_sarima_completo['y'].to_numpy(),

```

```

exog=ex_completo.to_numpy(), order=(0,1,1), seasonal_order=(1,1,0, 12))
results = model.fit(dis=False)
forecast = results.get_forecast(exog=prev, steps=horizon)
previsoesSARIMAX = forecast.predicted_mean
print("Previsões SARIMAX")
print(previsoesSARIMAX)

```

*****Aqui vamos prever com o LSTM*****

```

(X_train, y_train) = singleStepSampler(df_scaled, horizon + 1)
multivariate_lstm = keras.Sequential()
multivariate_lstm.add(keras.layers.LSTM(1000,
input_shape=(X_train.shape[1], X_train.shape[2])))
multivariate_lstm.add(keras.layers.Dropout(0.2))
multivariate_lstm.add(keras.layers.Dense(3, activation='linear'))
multivariate_lstm.compile(loss = 'MeanAbsolutePercentageError',
metrics=['MSE'], optimizer='Adam')
multivariate_lstm.summary()
history = multivariate_lstm.fit(X_train, y_train, epochs=400)
loss_per_epoch = multivariate_lstm.history.history['loss']
plt.plot(range(len(loss_per_epoch)), loss_per_epoch)
predicted_values = multivariate_lstm.predict(X_test)
true_predictions = scaler.inverse_transform(predicted_values)
previsoesLSTM = true_predictions[:, 0]
true_test = scaler.inverse_transform(y_test)
print("Previsões LSTM")
print(previsoesLSTM)

```

*****Aqui vamos graficar as previsões do modelo híbrido*****

```

#Fazendo a média das previsões
hibrido = []
for x in range (0, len(previsoesLSTM)):
    hibrido.append((previsoesLSTM[x] + previsoesSARIMAX[x])/2)

#Gerar array com as datas para as quais fizemos previsões
datas = np.arange('2026-01-01', '2026-04-01', dtype='datetime64[M]')

#Imprime o gráfico com as previsões
plt.plot(df_sarima_completo['ds'], df_sarima_completo['y'], label='Treino')
plt.plot(datas, previsoesSARIMAX, label='SARIMAX')

```

```
plt.plot(datas,previsoesLSTM, label='LSTM')  
plt.plot(datas,hibrido, label='MÃ©dia')  
plt.xticks(rotation=45)  
plt.legend()
```

2.2 RELATÓRIO DE PESQUISA

Conforme proposto no componente curricular Trabalho de Curso I, II e III, o projeto de pesquisa “Febre do Oropouche: Eficácia de modelos matemáticos e de aprendizado de máquina na previsão de casos” foi desenvolvido durante o ano de 2025 e primeiro semestre de 2026. Ele foi orientado pela Prof^a. Dr^a. Renata dos Santos Rabello Bernardo e coorientado pelo Prof. Dr. Julio César Stobbe e pelo Prof. Dr. Gustavo Olszanski Acrani. Trata-se de um estudo observacional, ecológico, de caráter descritivo com abordagem quantitativa.

O objetivo principal da pesquisa foi implementar os modelos matemático autorregressivo integrado de médias móveis sazonal com variáveis exógenas (SARIMAX) e o de aprendizado de máquina de memória de longo e curto prazo (LSTM) para estimar as notificações futuras de casos de Febre do Oropouche (FO) e, por meio das métricas erro percentual absoluto médio (MAPE), raiz do erro quadrático médio (RMSE) e erro absoluto médio (MAE), deliberar sobre qual deles teve melhor desempenho na tarefa. Além disso, foi analisada a proporção de casos notificados em segmentos específicos da população, como sexo, faixa etária e região de residência.

A escolha do tema foi motivada pelo recente crescimento de casos de FO em nível global, inclusive em regiões para além das consideradas endêmicas para essa arbovirose. Esse acontecimento evidencia a necessidade de pesquisas na área de previsão de epidemias como ferramenta de gestão em saúde pública e do fortalecimento do monitoramento epidemiológico de doenças negligenciadas. Assim, o presente trabalho se propõe a levantar a discussão sobre a aplicabilidade dos modelos SARIMAX e LSTM em cumprir essa missão.

Foram utilizados dados secundários do Painel Epidemiológico do Ministério da Saúde referente às notificações de novos casos semanais de FO no Brasil entre janeiro de 2023 e dezembro de 2025. Foi seguido o passo a passo: site do Ministério da Saúde <https://www.gov.br/saude/pt-br> > Assuntos > Saúde de A a Z > O > Oropouche > Painel Epidemiológico.

Também foram coletadas as variáveis climáticas de índices pluviométricos, umidade do ar e temperatura site Prediction Of Worldwide Energy Resources (POWER) da National Aeronautics and Space Administration (NASA) seguindo o passo a passo: site do POWER <https://power.larc.nasa.gov/> > Regional. Elas foram

adicionadas aos modelos como variáveis exógenas, para que se levasse em conta a sua influência no valor futuro das notificações junto com os valores anteriores da série temporal. Isso porque, no caso de arboviroses, os dados climáticos podem ter uma participação significativa entre os determinantes de prevalência. Assim, aumenta-se a precisão do modelo ao incluí-las no algoritmo.

Todas as etapas da implementação dos modelos foram feitas por meio da linguagem de programação Python. Para decompor a série nos componentes tendência, sazonalidade e irregularidade foram utilizadas as bibliotecas StatsModels e pyMannKendall. Para execução do SARIMAX, a StatsForecast. Ela possui tanto ferramentas para fazer as previsões quanto para calcular automaticamente qual a ordem ideal do modelo. Para o LSTM, foi utilizada a biblioteca Keras. Para avaliar a eficácia dos modelos, a SkLearn. O código foi escrito e executado em nuvem por meio da ferramenta Google Colabs, que é gratuita.

Ao longo de sua execução, percebeu-se que a média aritmética das previsões dos modelos acompanhou melhor as oscilações da curva de notificações. Por isso, foi incluída essa outra análise para as métricas de erro, mesmo que não idealizada como parte do projeto no primeiro momento. Desse modo, a comparação de desempenho preditivo se deu entre o SARIMAX, o LSTM e a média aritmética de ambos.

Além disso, foi realizada uma análise das variáveis sociodemográficas disponíveis no Painel Epidemiológico referentes a sexo, faixa etária e região de residência. A organização foi da seguinte forma: sexo, (feminino e masculino); faixa etária categorizada em anos (0-9; 10-19; 20-29; 30-39; 40-49; 50-59; 60 ou +), região de residência (norte, nordeste, sudeste, centro-oeste e sul). Descreveram-se as proporções de casos de FO relatados em cada um desses subgrupos.

Ao final da pesquisa, foi redigido um artigo científico com base nas normas de submissão da revista Sustinere, disponíveis no endereço eletrônico: <https://www.e-publicacoes.uerj.br/index.php/sustinere/about/submissions#authorGuidelines>. Ela foi escolhida por seu foco em temas interdisciplinares e transdisciplinares em saúde e educação envolvendo ciência e tecnologia, conversando, portanto, com o presente trabalho.

3 ARTIGO CIENTÍFICO

Febre do Oropouche: eficácia de modelos matemáticos e de aprendizado de máquina na previsão de casos

Maria Clara da Silva Maia

Júlio César Stobbe

Gustavo Olszanski Acrani

Renata dos Santos Rabello

Resumo: Mudanças climáticas amplificam vírus como o *Orthobunyavirus oropoucheense*, causador da Febre do Oropouche (FO). No Brasil, ele restringia-se à Amazônia, até expandir-se além dela, tornando-se doença emergente. Assim, estimou-se novos casos de FO no Brasil utilizando os modelos Autorregressivo Integrado de Médias Móveis Sazonal com Variáveis Exógenas, de Memória de Curto-Longo Prazo (LSTM) e a média aritmética entre eles. Coletaram-se dados do Painel Epidemiológico do Ministério da Saúde e do *Prediction Of Worldwide Energy Resources* referentes aos casos de FO e às temperaturas médias de 2023 a 2025. Fez-se análise descritiva da amostra incluindo sexo, faixa etária e região. Via testes de Dickey-Fuller aumentado, Mann-Kendall e Ljung-Box, analisaram-se estacionaridade, tendência e autocorrelação da série temporal. Ademais, decompôs-se a curva no modelo aditivo. Os modelos preditivos foram implementados em Python. Notificaram-se 26.678 casos entre 2023 e 2025, predominando sexo masculino (52,36%) e faixa etária de 20 a 49 anos (56,90%). Houve concentração dos casos no Norte (94,36%) em 2023, que evoluiu para concentração no Sudeste em 2025 (85,01%). A série foi não estacionária, sem tendência e autocorrelacionada. A decomposição da curva evidenciou tendência crescente de junho/2023 a abril/2025, quando tornou-se decrescente. A partir de julho/2025, perdeu sua tendência. Quanto à sazonalidade, as notificações aumentaram de setembro a fevereiro e diminuíram de março a outubro. Na validação dos modelos, o LSTM demonstrou melhor desempenho. Seu Erro Absoluto Médio foi três casos prevendo 3 meses e, 515, prevendo seis. Portanto, LSTM pode auxiliar no combate à FO. Porém, necessita ainda calibragem.

Palavras-chave: Medicina Tropical; Infecções por Arbovírus; Algoritmos de Predição; Modelos Preditivos de Aprendizagem; Planejamento em Saúde.

Oropouche fever: effectiveness of mathematical models and machine learning in predicting cases

Abstract: Climate change amplifies viruses such as *Orthobunyavirus oropoucheense*, which causes Oropouche Fever (OF). In Brazil, it was confined to the Amazon until it expanded beyond its borders, becoming emerging disease. Therefore, new cases of OF in Brazil were estimated using the Autoregressive Integrated Moving Average model with exogenous variables, the Long-Short Term Memory (LSTM) model, and the arithmetic mean between them. Data from the Epidemiological Panel of the Ministry of Health and the Prediction of Worldwide Energy Resources regarding OF cases and average temperatures from 2023 to 2025 were collected. A descriptive analysis of the sample was performed, including sex, age group and region. Stationarity, trend, and autocorrelation of the time series were analyzed using the augmented Dickey-Fuller, Mann-Kendall, and Ljung-Box tests. Furthermore, the curve was decomposed using the additive model. Predictive models were implemented in Python. 26,678 cases were reported between 2023 and 2025, predominantly male (52.36%)

and aged 20 to 49 years (56.90%). Cases were concentrated in the North (94.36%) in 2023, shifting to the Southeast in 2025 (85.01%). The series was non-stationary, trendless, and autocorrelated. Curve decomposition showed an increasing trend from June/2023 to April/2025, after which it became decreasing. From July/2025 onwards, the trend disappeared. Regarding seasonality, notifications increased from September to February and decreased from March to October. In model validation, LSTM demonstrated the best performance. Its Mean Absolute Error was three cases predicting 3 months and 515 predicting six. Therefore, LSTM can assist in combating FO. However, it still needs calibration.

Keywords: Tropical Medicine; Arbovirus Infections; Prediction Algorithms; Predictive Learning Models; Health Planning.

Fiebre de Oropouche: efectividad de los modelos matemáticos y el aprendizaje automático en la predicción de casos

Resumen: El cambio climático amplifica virus como el *Orthobunyavirus oropoucheense*, causante de la Fiebre Oropouche (FO). En Brasil, se limitaba a la Amazonía hasta que se extendió más allá, convirtiéndose en enfermedad emergente. Así, se estimaron nuevos casos de FO en Brasil utilizando el modelo de media móvil integrada autorregresiva con variables exógenas, el modelo de memoria a largo y corto plazo (LSTM) y la media aritmética entre ellos. Se recopilaron datos del Panel Epidemiológico del Ministerio de Salud y del *Prediction Of Worldwide Energy Resources* sobre casos de FO y temperaturas promedio de 2023 a 2025. Se realizó un análisis descriptivo de la muestra, incluyendo sexo, edad y región. La estacionariedad, tendencia y autocorrelación se analizaron mediante pruebas de Dickey-Fuller aumentada, Mann-Kendall y Ljung-Box. Además, la curva se descompuso utilizando el modelo aditivo. Los modelos predictivos se implementaron en Python. Se notificaron 26.678 casos, predominantemente hombres (52,36%) y edad de 20-49 años (56,90%). Los casos se concentraron en el Norte (94,36%) en 2023, pasando al Sureste (85,01%) en 2025. La serie fue no estacionaria, sin tendencia y autocorrelacionada. La descomposición mostró tendencia creciente de junio/2023 a abril/2025, luego decreciente hasta desaparecer desde julio/2025. En cuanto a la estacionalidad, las notificaciones aumentaron de septiembre a febrero y disminuyeron de marzo a octubre. En la validación del modelo, LSTM demostró el mejor rendimiento. Su error absoluto medio fue de tres casos prediciendo 3 meses y 515 prediciendo seis. Por lo tanto, LSTM puede ayudar a combatir la FO, pero aún necesita calibración.

Palabras clave: Medicina Tropical; Infecciones por Arbovirus; Algoritmos de Predicción; Modelos de Aprendizaje Predictivo; Planificación en Salud.

INTRODUÇÃO

A Febre do Oropouche (FO) é causada pelo arbovírus *Orthobunyavirus oropoucheense* (OROV). Esse pertence à família viral *Peribunyaviridae*, ao gênero *Orthobunyavirus* e ao sorogrupo *Simbu* (CDC, 2025). Sendo negligenciado, causa surtos na América Central e do Sul desde a década de 1950 (Hua et al., 2026). Desde então, pelo menos 30 epidemias e mais de 500 mil casos de infecção foram relatados, especificamente na região amazônica (Bai et al., 2025). Entretanto, a partir de 2024, começou a receber mais atenção dos gestores de saúde pública devido à expansão para áreas não endêmicas (Penha et al., 2026).

Tipicamente, o OROV está presente em um ciclo silvestre atingindo mamíferos não humanos (Bradypus, Primatas, Marsupiais), lagartas e pássaros (Porwal et al., 2025). Porém, é possível circular em um ciclo urbano e infectar pessoas (Pastula; Beckham; Tyler, 2025). O principal vetor é o *Culicoides paraensis*, conhecido como maruim ou mosquito pólvora (BRASIL, 2024). Apesar disso, o *Aedes serratus* e *Culex quinquefasciatus* também podem ser vetores (Porwal et al., 2025). Não há tratamento e a prevenção se dá pelo combate ao mosquito (Torres et al., 2026).

Após a contaminação, o período de incubação gira em torno de quatro a oito dias. Cerca de 60% das infecções são sintomáticas, causando febre abrupta, cefaleia, mialgia, artralgia, distúrbios gastrointestinais, dor retro-ocular e/ou fotofobia durante 3 a 10 dias. É possível que a doença siga um curso bifásico, com o retorno dos sintomas uma semana depois de seu início (Okesanya et al., 2025).

Também pode ocorrer meningite, encefalite e síndrome de Guillain-Barré (González-Quevedo; Farrill; Becquer, 2025). Esse acometimento neurológico associado a relatos de transmissão vertical indicam que a doença pode ser mais severa do que parece à primeira vista (Wang et al., 2026).

As mudanças climáticas contribuem para o aumento de surtos de doenças como a FO, com o Brasil sendo um território mais vulnerável para tal (Lorenz et al., 2026). Apesar do impacto considerável, os casos de OROV permanecem negligenciados e subnotificados (Oliveira et al, 2026). Por isso, é necessário um enfoque maior da saúde pública no que tange à vigilância, à educação pública e colaboração internacional no combate a essa doença (Okesanya et al., 2025). Afinal, por ela não receber a devida atenção quando estava restrita à bacia amazônica, permitiu-se uma epidemia e sua expansão para novas regiões.

Tal necessidade foi reconhecida pelo Ministério da Saúde ao implementar o novo protocolo de busca ativa por FO e febre do Mayaro adotado após a descentralização do diagnóstico biomolecular. A FO foi adicionada às doenças de notificação compulsória imediata (BRASIL, 2024). Atualmente, recomenda-se o diagnóstico via RT-qPCR em casos negativos para outras arboviroses desde 2023 (Gomes et al., 2025).

Nesse cenário, os modelos de previsão de doenças infecciosas são abordagem estratégica para formular políticas públicas. Afinal, identificando-se surtos precocemente, promove-se o uso otimizado de recursos e medidas preventivas (Lutz et al., 2019). Além de

análise da curva de incidência, pode-se observar a influência de fatores externos, melhorando o desempenho preditivo (Ferreira, 2024). Isso é fundamental no caso de arboviroses, pois temperatura, precipitação e umidade do ar estão relacionados com a proliferação dos mosquitos transmissores (Chen; Moraga, 2025). Em 2024, dados de temperatura foram os fatores mais fortemente associados à presença de OROV (Lorenz et al., 2025).

Desse modo, o estudo em questão objetivou utilizar os modelos matemático autorregressivo integrado de médias móveis sazonal com variáveis exógenas (SARIMAX) e o de aprendizado de máquina de memória de longo e curto prazo (LSTM) para estimar as notificações futuras FO e deliberar sobre qual deles teve melhor desempenho na tarefa. Ademais, visou-se analisar a proporção de casos notificados em segmentos específicos da população.

METODOLOGIA

Trata-se de um estudo ecológico, realizado a partir de dados secundários do Painel Epidemiológico do Ministério da Saúde referente às notificações de novos casos semanais de Febre do Oropouche (FO) no Brasil entre janeiro de 2023 e dezembro de 2025. As semanas epidemiológicas de um mesmo mês foram aglutinadas e suas notificações correspondentes foram somadas para melhor aproveitamento da análise, que transformou-se em mensal.

A coleta de dados secundários foi realizada diretamente do Painel Epidemiológico do Ministério da Saúde seguindo o passo a passo: site do Ministério da Saúde <https://www.gov.br/saude/pt-br> > Assuntos > Saúde de A a Z > O > Oropouche > Painel Epidemiológico. Realizou-se uma análise das variáveis sociodemográficas referentes a sexo, faixa etária e região de residência. A organização foi da seguinte forma: sexo, (feminino e masculino); faixa etária categorizada em anos (0-9; 10-19; 20-29; 30-39; 40-49; 50-59; 60 ou +), região de residência (norte, nordeste, sudeste, centro-oeste e sul). Foram descritas as proporções de casos de FO relatados em cada um desses subgrupos.

Também foram coletadas as variáveis climáticas de temperatura do solo e amplitude térmica. Elas foram adicionadas aos modelos como variáveis exógenas, para que se leve em conta a sua influência no valor futuro das notificações junto com os valores anteriores da série temporal. Essas variáveis foram coletadas no site Prediction Of Worldwide Energy Resources (POWER) da National Aeronautics and Space Administration (NASA) seguindo o passo a

passo: site do POWER <https://power.larc.nasa.gov/> > Regional. Foram coletados dados do estado com o maior número de notificações da doença no período analisado.

Iniciou-se a análise da série temporal pela decomposição em três componentes: tendência, sazonalidade e irregularidade. Para testar a estacionalidade, aplicou-se o teste de Dickey–Fuller Aumentado e, para analisar a tendência da curva, o teste de Mann–Kendall. Também verificou-se a presença de autocorrelação dos dados via teste de Ljung-Box. A implementação dessas etapas foi feita a partir das bibliotecas StatsModels e pyMannKendall da linguagem de programação Python.

O modelo ARIMA é um acrônimo para modelo autorregressivo (AR) integrado (I) de média móvel (MA) (Hasan et al., 2024). As respectivas ordens de cada modelo são representadas por ARIMA (p, d, q). Ele pode ser adaptado para captar sazonalidade na série temporal, tornando-se o ARIMA sazonal (SARIMA). Nessa versão, é grafado como SARIMA (p, d, q) (P, D, Q)s. O componente “s” define o número de observações (Bhatnagar et al., 2012). O modelo AR observa a influência entre os valores passados da série com os valores futuros, o modelo MA utiliza os erros anteriores para ajustar as previsões e o modelo I diferencia a série, tornando-a estacionária (Silva et al., 2023). É possível também analisar a influência de fatores externos nos valores futuros da variável analisada. Isso é feito pela extensão do modelo SARIMA para SARIMAX (SARIMA com variáveis exógenas) (Ferreira, 2024).

Para implementação do SARIMAX, foi utilizada a biblioteca StatsForecast da linguagem de programação Python. Ela possui tanto ferramentas para fazer as previsões quanto para calcular automaticamente qual a ordem ideal do modelo.

O modelo de Memória de Curto e Longo Prazo (LSTM) é um tipo de Rede Neural Recorrente (Campanhol, 2024). Essa é uma categoria de rede neural capaz de detectar autocorrelação entre dados de uma série temporal, pois possui conexões transmissoras entre diferentes momentos no tempo. Seus neurônios recebem dados da camada anterior e da sua própria ativação na passagem pregressa (Orssatto et al., 2022). Sua arquitetura inclui três portas: esquecimento, entrada e saída (Campanhol, 2024). A porta de esquecimento define quais informações excluir do estado da célula. Já a de entrada, determina qual informação adicionar ao estado da célula (ICCAI, 2019). Após a célula de memória passar pela porta de entrada e esquecimento, a porta de saída gera a previsão do próximo valor a partir do produto de uma camada sigmoide e uma função de ativação (Othman et al., 2022).

Para implementar o LSTM, foi utilizada a biblioteca Keras da linguagem de programação Python.

A fim de avaliar o desempenho dos modelos preditivos, utilizou-se o erro percentual absoluto médio (MAPE), a raiz do erro quadrático médio (RMSE) e o erro absoluto médio (MAE). Também verificou-se a qualidade do ajuste via teste de Ljung-Box. Esses foram calculados com a biblioteca SkLearn da linguagem de programação Python. Foram feitas previsões e validações dessas previsões para um semestre. O código foi escrito e executado em nuvem por meio da ferramenta Google Colabs, que é gratuita.

Como foram utilizados dados secundários de domínio público, dispensa-se pela Comissão Nacional de Ética em Pesquisa (CONEP) a submissão ao Comitê de Ética em Pesquisa (CEP/CONEP) com Seres Humanos da UFFS, resolução nº510, de 07 de abril de 2016 do Conselho Nacional de Saúde (CNS). Considerando tratar-se de dados secundários e agregados por municípios, não existe risco de identificação pessoal dos participantes envolvidos.

RESULTADOS

Em 2023, foram notificados 834 casos de Febre do Oropouche (FO). A distribuição das notificações ao longo do ano se manteve relativamente constante, com média de 32 casos por mês de janeiro a novembro. Porém, houve uma rápida escalada no mês de dezembro, que contou com 480 casos. O estado com o maior número de ocorrências foi o Amazonas (54,79%), seguido do Acre (21,35%) e Roraima (18,22%). Como esses três estados juntos representaram 94,36% do total, percebe-se uma concentração da incidência na região Norte do país. O sexo masculino representou 52,28% dos casos notificados. Em relação à faixa etária, a população até 10 anos registrou 3,85% dos casos; de 10 a 19 anos, 16,31%; de 20 a 29 anos, 17,86%; de 30 a 39 anos, 26,74%; de 40 a 49 anos, 17,86%; 50 a 59 anos, 9,23% e, na população de 60 anos ou mais, 8,15%. Até então, não havia registros de óbitos.

Em 2024, constatarem-se 13.856 casos, um aumento de 16,61 vezes. Os casos continuaram em um patamar alto até abril, quando lentamente começaram a cair. Em setembro, já estavam no mesmo patamar de 2023. Porém, já em outubro, o padrão de escalada exponencial se repetiu de forma ainda mais intensa. A doença não se restringiu mais na região Norte, tendo o Espírito Santo como estado com o maior número de notificações: 42,35% dos

casos. O Amazonas ficou em segundo lugar, com 23,32% dos casos e, em terceiro, ficou Rondônia, com 12,35% dos casos. Esses estados somados concentraram 78,01% do total de casos. Uma menor porcentagem em comparação ao trio de 2023, devido ao maior número de estados com ocorrências de FO. Estados esses pertencentes às cinco regiões do país, mostrando uma expansão do vírus. O sexo masculino novamente foi predominante (52,51% dos casos notificados). Em relação à faixa etária, a população até 10 anos registrou 2,74% dos casos; de 10 a 19 anos, 12,41%; de 20 a 29 anos, 18,88%; de 30 a 39 anos, 19,87%; de 40 a 49 anos, 18,60%; 50 a 59 anos, 13,86% e, na população de 60 anos ou mais, 13,64%. Além disso, houve 4 óbitos confirmados e 3 suspeitos, indicando que a FO pode sim ser letal se negligenciada.

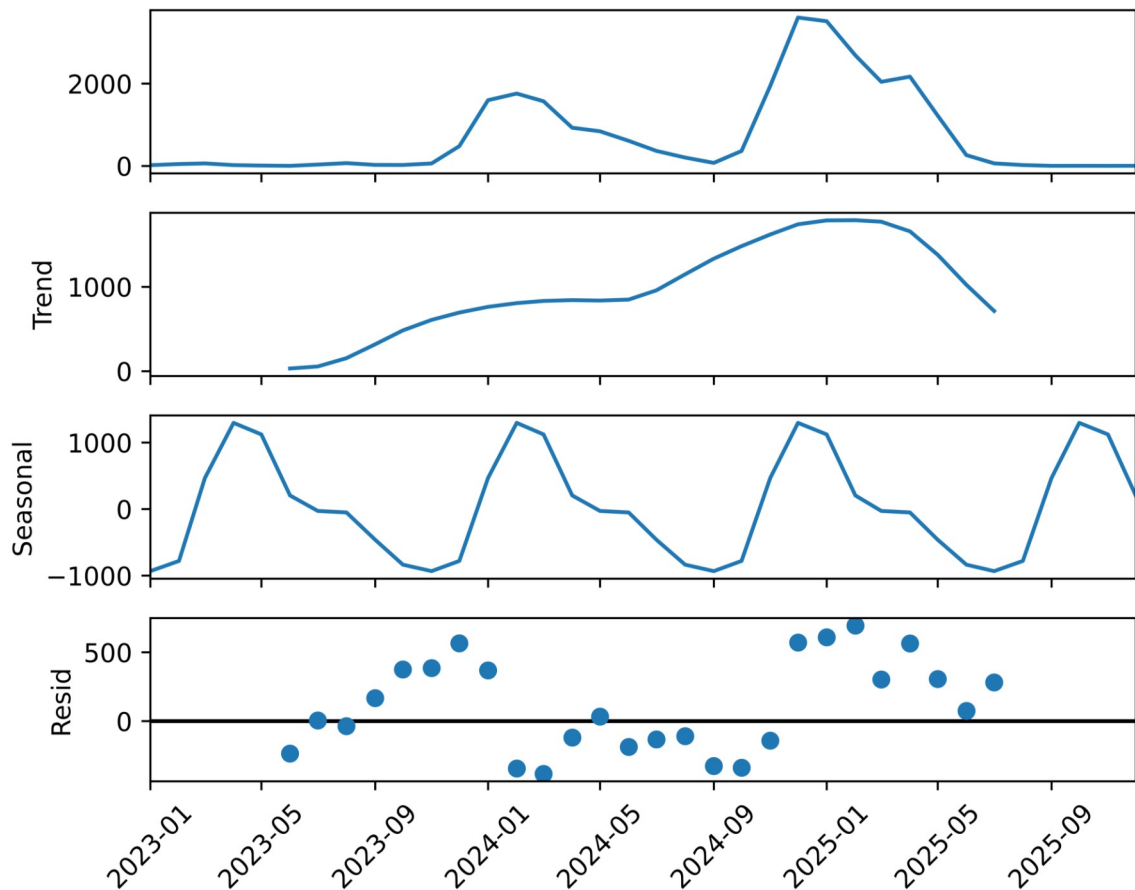
Em 2025, foram notificados 11.988 casos, um patamar similar ao do ano anterior. Porém, existiu uma mudança notável na distribuição desses casos ao longo do ano: a incidência se concentrou no primeiro semestre (99,27% dos casos até junho). O Espírito Santo continuou como o líder de casos, com 52,77% das notificações, seguido do Rio de Janeiro (20,88%) e Minas Gerais (11,40%). Assim, observa-se uma concentração de 85,01% dos casos na região sudeste. Isso confirma que a FO tem potencial para ser endêmica em outras partes do Brasil e não é mais uma doença apenas Amazônica. O sexo masculino seguiu tendo mais casos notificados (52,41%). Em relação à faixa etária, a população até 10 anos registrou 2,22% dos casos; de 10 a 19 anos, 9,97%; de 20 a 29 anos, 18,21%; de 30 a 39 anos, 19,04%; de 40 a 49 anos, 18,74%; 50 a 59 anos, 14,69% e, na população de 60 anos ou mais, 17,13%. Confirmaram-se 5 mortes e 2 ficaram com *status* de suspeita.

Na Figura 1, consta a decomposição da série temporal de casos de FO no Brasil entre 2023 e 2025. No topo, observa-se a curva original, seguida do componente tendência, sazonalidade e aleatoriedade, respectivamente. A curva completa permite perceber um patamar endêmico em 2023, seguido de duas ondas epidêmicas crescentes (2024 e 2025) e um retorno a patamares semelhantes aos do início da série na segunda metade de 2025. A curva de 2025 cresceu aproximadamente na mesma inclinação que 2024, porém sua queda foi mais vertiginosa e não houve novo crescimento a partir da primavera.

Analisando-se o componente tendência, percebe-se crescimento a partir de junho de 2023, o qual só parou em abril de 2025, quando a tendência se tornou decrescente. A partir de julho, a série deixou de expressar uma tendência clara, o que pode impactar na eficácia dos modelos preditivos. Quanto à sazonalidade, percebe-se uma crescente de casos a partir da

primavera, com pico aproximadamente em fevereiro, seguida de uma redução até setembro, quando as notificações voltam a crescer. Já os resíduos, oscilaram mais de 500 casos, refletindo variações não explicadas pela tendência ou sazonalidade, mas sim por variáveis exógenas.

Figura 1 – Série temporal dos casos de Febre do Oropouche no Brasil entre 2023 e 2025 e sua decomposição em tendência, sazonalidade e resíduos respectivamente



Legenda: o eixo Y representa a quantidade de casos de FO e o eixo X representa a data desses casos.

Fonte: Própria., 2026.

O teste de Dickey-Fuller aumentado indicou que a série não era estacionária ($p = 0,08$); o teste de Mann-Kendall confirmou a ausência de tendência da série temporal. O teste de Ljung-Box ($p = 1,79 \times 10^{-8}$) apontou a autocorrelação dos dados, reforçando a possibilidade de aplicar métodos de previsão.

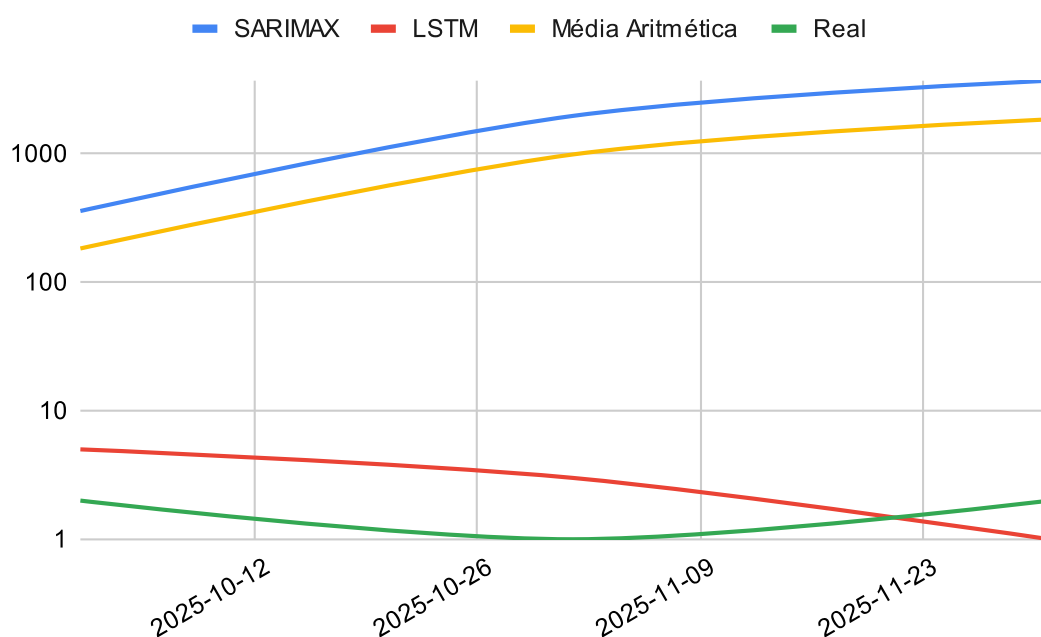
A Figura 2 apresenta a comparação entre as previsões dos modelos SARIMAX (que obteve melhor desempenho levando em conta a temperatura média do estado com maior número de notificações), LSTM (que obteve melhor desempenho levando em conta a temperatura média e a amplitude térmica do estado com maior número de notificações) e a

média aritmética dos dois com os valores reais observados no período de outubro de 2025 a dezembro de 2025 (horizonte preditivo de 3 meses).

Nota-se que o modelo SARIMAX previu uma terceira onda da doença, iniciando no mesmo período do ano que as duas anteriores. Já o LSTM, previu que os casos não voltariam a patamares epidêmicos. Na validação das previsões, percebe-se que quem melhor acompanhou as oscilações da curva foi o LSTM, enquanto o SARIMAX se mostrou inadequado para fazer estimativas.

As métricas de erros do modelo SARIMAX foram $RMSE = 2.428$ casos; $MAPE = 133.852\%$ e $MAE = 2007$ casos. As métricas de erros do modelo LSTM foram $RMSE = 3$ casos; $MAPE = 182\%$ e $MAE = 3$ casos. As métricas de erros calculando a média aritmética ficaram em $RMSE = 1214$ casos; $MAPE = 66.630\%$ e $MAE = 1.004$ casos. A análise dos resíduos pelo teste Q de Ljung-Box indicou independência na distribuição dos ruídos para ambos os modelos, com $p = 0,15$ para o SARIMAX e $p = 0,25$ para o LSTM. A média aritmética manteve esse resultado, com $p = 0,15$. Isso indicaria boa adequação dos modelos à curva, o que, no caso do LSTM, é reforçado pelas métricas MAE e RMSE. Já no caso do SARIMAX, demonstra que observar a aleatoriedade do erro isoladamente pode ser uma forma inadequada de avaliar modelos preditivos.

Figura 2 – Validação do SARIMAX, LSTM e média aritmética para os casos de Febre do Oropouche no Brasil partindo de janeiro de 2026 com horizonte preditivo de 3 meses



Legenda: o eixo Y representa a quantidade de casos de FO e o eixo X representa a data desses casos.

Fonte: Própria., 2026.

Fazendo a mesma comparação anterior entre os modelos, porém com horizonte preditivo de 4 meses (setembro de 2025 a dezembro de 2025), notou-se que ambos tiveram pior desempenho preditivo, mas o LSTM manteve resultados confiáveis. Por esse motivo, o cálculo da média aritmética novamente não agregou. As métricas de erros do modelo SARIMAX foram $RMSE = 2.166$ casos; $MAPE = 106.165\%$ e $MAE = 1.591$ casos. As métricas de erros do modelo LSTM foram $RMSE = 10$ casos; $MAPE = 648\%$ e $MAE = 9$ casos. As métricas de erros calculando a média aritmética ficaram em $RMSE = 1.088$ casos; $MAPE = 53.406\%$ e $MAE = 800$ casos. A análise dos resíduos pelo teste Q de Ljung-Box indicou independência na distribuição dos ruídos para ambos os modelos e para a média aritmética, com $p = 0,16$ para o LSTM; $p = 0,12$ para o SARIMAX e $p = 0,33$ para a média aritmética.

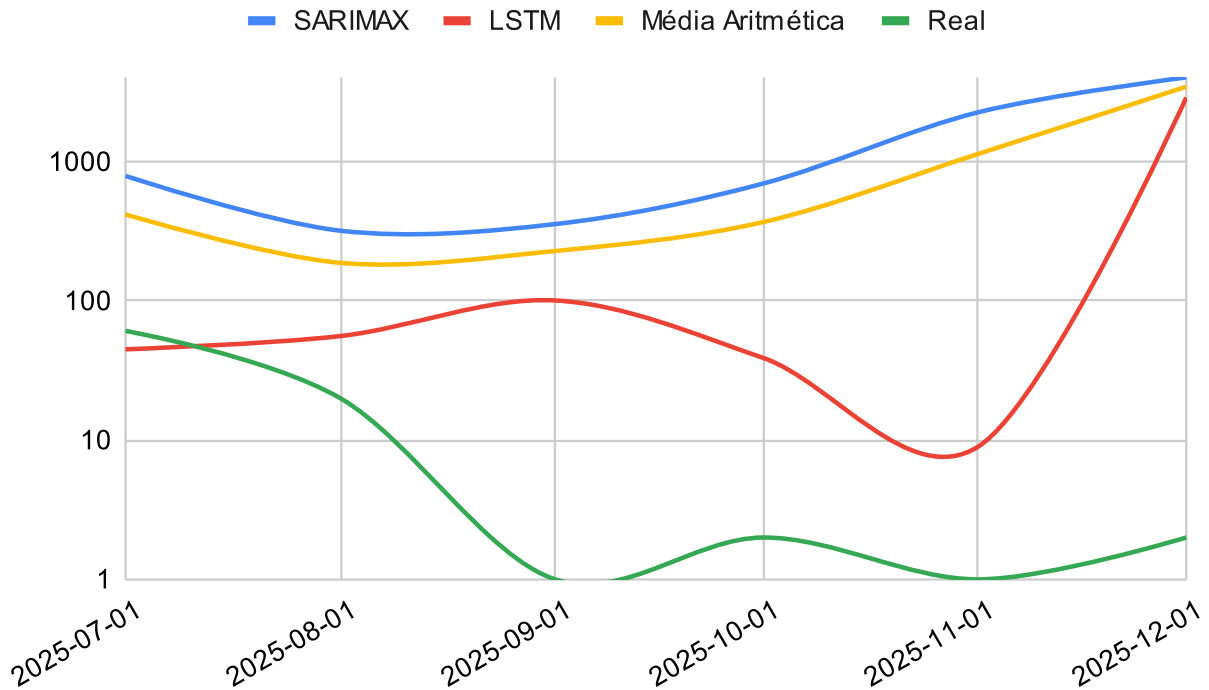
Realizando novamente aquela comparação entre os modelos, porém com horizonte preditivo de 5 meses (agosto de 2025 a dezembro de 2025), constatou-se que, dessa vez, o LSTM diminuiu de forma importante sua confiabilidade. Como anteriormente, a média aritmética não possuiu informações complementares. As métricas de erros do modelo SARIMAX foram $RMSE = 2.225$ casos, $MAPE = 113.982\%$ e $MAE = 1.747$ casos. As métricas de erros do modelo LSTM foram $RMSE = 55$ casos; $MAPE = 3.177\%$ e $MAE = 46$ casos. As métricas de erros calculando a média aritmética ficaram em $RMSE = 1.139$ casos; $MAPE = 58.580\%$ e $MAE = 896$ casos. A análise dos resíduos pelo teste Q de Ljung-Box indicou independência na distribuição dos ruídos para o SARIMAX ($p = 0,17$) e para a média aritmética ($p = 0,51$). O LSTM obteve $p = 0,047$, indicando autocorrelação nos ruídos. Essa métrica corrobora com a piora da confiabilidade preditiva no horizonte de 5 meses.

Analizando pela última vez a mesma comparação no horizonte preditivo de 6 meses (julho de 2025 a dezembro de 2025) observou-se que, enfim, o LSTM passou também a ficar inadequado, pois previu a terceira onda que o SARIMAX já mostrava. Consequentemente, a média aritmética tampouco obteve desempenho satisfatório. A Figura 3 ilustra esse cenário. Nesse contexto, o horizonte preditivo de 7 meses não foi testado.

As métricas de erros do modelo SARIMAX foram $RMSE = 1.943$ casos; $MAPE = 83.621\%$ e $MAE = 1.398$ casos. As métricas de erros do modelo LSTM foram $RMSE = 1.282$ casos; $MAPE = 26.272\%$ e $MAE = 515$ casos. As métricas de erros calculando a média aritmética ficaram em $RMSE = 1.508$ casos; $MAPE = 54.942\%$ e $MAE = 954$ casos. A análise dos resíduos pelo teste Q de Ljung-Box indicou independência na distribuição dos ruídos para

ambos os modelos e para a média aritmética, com $p = 0,83$ para o LSTM, $p = 0,24$ para o SARIMAX e $p = 0,75$ para a média aritmética.

Figura 3 – Validação do SARIMAX, LSTM e média aritmética para os casos de Febre do Oropouche no Brasil partindo de janeiro de 2026 com horizonte preditivo de 6 meses



Legenda: o eixo Y representa a quantidade de casos de FO e o eixo X representa a data desses casos.

Fonte: Própria., 2026.

A Figura 4 e a Tabela 1 ilustram as projeções das notificações mensais de Febre do Oropouche no Brasil entre janeiro de 2026 e junho de 2026. O LSTM prevê um aumento dos casos em relação aos 4 meses anteriores, mas ainda dentro dos níveis endêmicos, sem novas ondas epidêmicas. O SARIMAX prevê um aumento importante, apesar de menor que os vistos em 2024 e 2025, com rápida queda. A média aritmética suaviza os picos de ambos os modelos, mantendo o desenho estimado pelo SARIMAX, mas com números mais baixos.

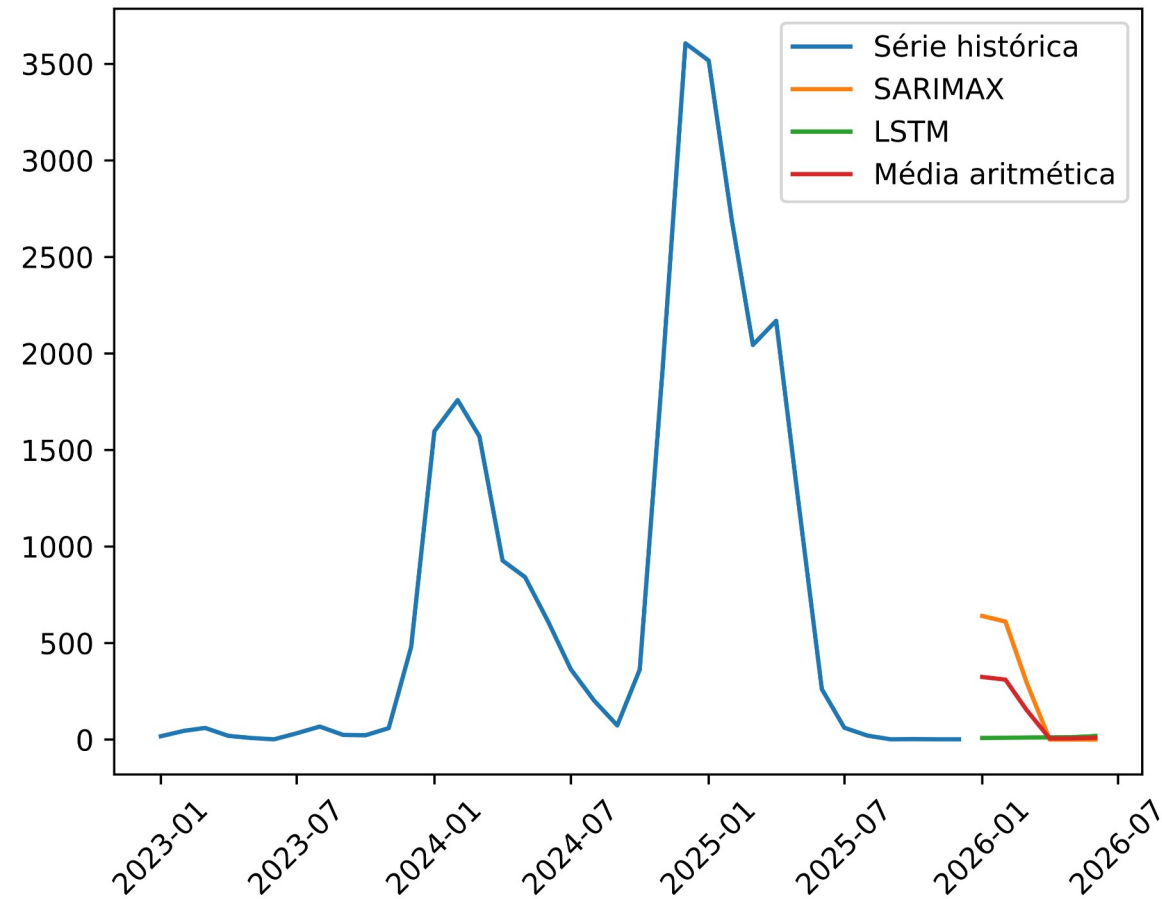
Tabela 1 – Projeções de casos mensais de Febre do Oropouche no Brasil de janeiro de 2026 e junho de 2026 pelos modelos SARIMAX, LSTM e a média aritmética de suas previsões

Data	LSTM	SARIMAX	Média Aritmética
Jan/2026	5	642	324
Fev/2026	2	612	307
Mar/2026	2	298	150
Abr/2026	7	0	4

Mai/2026	4	0	2
Jun/2026	10	0	5

Fonte: Própria., 2026.

Figura 6 – Projeções de casos de Febre do Oropouche no Brasil de janeiro de 2026 e junho de 2026 pelos modelos SARIMAX, LSTM e a média aritmética de suas previsões



Fonte: Própria., 2026.

DISCUSSÃO

Uma pesquisa realizou testes diagnósticos para OROV em doadores de sangue em Manaus (AM) entre 2023 e 2024. A proporção de soropositividade permitia estimar 390.000 infecções na cidade, um número substancialmente maior do que o oficial (Milani et al., 2026). Outro estudo parecido estimou a soroprevalência de IgG para o OROV em 11,4% em doadores de sangue em novembro de 2023, 17,8% em junho de 2024 e 25,7% em novembro de 2024. Em contraste, a taxa de soropositividade para IgM foi de 1,8% em novembro de 2023, 1,6% em junho de 2024 e 1,8% em novembro de 2024 (Manuli et al., 2026). Isso indica altos níveis de subnotificação mesmo quando a doença estava nos holofotes. Portanto, é

necessário analisar criticamente o comportamento da curva encontrado no presente estudo. Pois, é possível que a queda de casos no segundo semestre de 2025 seja resultado mais da falta de testagem do que da ausência da doença. Desse modo, tem-se como limitação para análise atual a ausência de testagem e notificação dessa doença negligenciada.

Um estudo coletou 6.735 casos de FO em 2020, 2021 e 2024 oriundos de 170 municípios e identificou quatro agrupamentos de alto risco para incidência dessa doença: a região Amazônica, os estados do Espírito Santo, da Bahia e do Rio de Janeiro (Lorenz et al., 2025). Outra pesquisa analisou a distribuição geoespacial dos casos de OROV com dados secundários do Gerenciador do Ambiente Laboratorial e do Instituto Brasileiro de Geografia e Estatística e revelou aglomerados de alta incidência na Região Norte e no estado do Espírito Santo (Oliveira et al., 2026). O presente estudo encontrou maior número de casos entre esses locais citados (Espírito Santo, região Norte e Rio de Janeiro), cada um com protagonismo em um momento diferente da epidemia. Porém, a Bahia não esteve entre os destaques.

Uma revisão sistemática de estudos observacionais na América Latina trouxe que a distribuição de FO por sexo foi comparável (51% homens) e foi mais comuns entre 20 e 39 anos (38%) (Wang et al., 2026). A partir os dados de vigilância do Departamento de Saúde de Minas Gerais, um estudo analisou os casos confirmados de OROV na Fundação Ezequiel Dias. As mulheres representaram 55,0% dos casos, com maior incidência entre adultos de 20 a 59 anos (Flores et al., 2026). No Espírito Santo, outra pesquisa indicou que a enfermidade foi mais prevalente entre homens (53,5%) em 2024 (Cola et al., 2026). Uma pesquisa com dados de Santa Catarina mostrou que a maioria dos afetados era do sexo masculino (58%), e a faixa etária predominante foi de 30 a 49 anos (41%) (Carl et al., 2026). Um estudo feito em Manaus (AM) encontrou que 53,1% dos casos eram do sexo masculino a mediana de idade foi de 35 anos (Manuli et al., 2026). No presente estudo, observou-se maior acometimento no sexo masculino, concordando com os resultados da América Latina, Santa Catarina, Manaus e Espírito Santo, mas divergindo dos dados de Minas Gerais. Quanto à faixa etária, o intervalo de 20 a 49 anos, que cobre as faixas etárias de maior incidência encontradas na América Latina, Santa Catarina, Manaus e Minas Gerais representou 56,90% dos casos.

Em uma pesquisa de Porto Alegre (RS), analisou a curva de notificações de dengue de janeiro de 2015 até março do ano de 2024. Os resultados mostram que a série temporal era estacionária, sazonal e autocorrelacionada (Ferreira, 2024). Outro estudo em Paranavaí (PR) observou a curva de notificações de dengue de janeiro de 2010 a dezembro de 2020. Também

encontram uma série sazonal e autocorrelacionada, porém, não estacionária (Silva, 2023). Essas mesmas características foram encontradas em um estudo da curva de notificações de dengue em Ribeirão Preto (SP) de 2000 a 2009 (Martinez; Silva, 2011). O presente estudo, também encontrou essas mesmas características de Paranavaí e Ribeirão Preto na curva de FO. Isso pode indicar um padrão comum em séries temporais de notificações de arboviroses. Já a presença de estacionaridade da curva de Porto Alegre, demonstra a não estagnação dessas características ao longo do tempo e espaço.

Utilizando dados ambientais como variáveis exógenas, uma pesquisa elaborou um modelo preditivo para casos de infecções do trato respiratório superior. O horizonte testado foi de oito semanas e os dados coletados se referiam ao período de 2012 a 2022 (Lim et al., 2026). O presente estudo conseguiu um baixo erro para até 3 meses, o que também representa apenas um curto prazo em termos de políticas públicas. Isso reforça a dificuldade de realizar previsões precisas no longo prazo, em especial se há demanda de variáveis exógenas. Desse modo, uma limitação desses e de outros estudos de modelos preditivos trata-se da queda da confiabilidade quando o horizonte preditivo é muito longo.

O modelo SARIMA foi vastamente testado para previsão dos casos de dengue. Em um estudo em Paranavaí (PR), utilizou-se o teste de Ljung-Box para validar a adequação do SARIMA para tal fim (Silva et al., 2023). Também a partir do teste de Ljung-Box e da função de autocorrelação, outro estudo aprovou o SARIMA em Ribeirão Preto (SP) (Martinez; Silva, 2011). Como foi encontrado no presente estudo um teste de Ljung-Box indicando boa adequação do SARIMAX e LSTM à curva da FO, poder-se-ia afirmar que ambos os modelos são apropriados. Porém, ao observar outras métricas de erro da análise atual, vê-se que o teste de Ljung-Box isoladamente é insuficiente para afirmá-lo. Outros estudos fazem uma análise mais ampla de métricas. Em Bangladesh, uma pesquisa confirmou a eficácia do SARIMA na predição dos casos de dengue via MAPE, teste de Ljung-Box e ACF (Hasan et al., 2024). Em um estudo indiano, o SARIMA foi validado para estimativas de incidência de dengue via MAPE e teste de Ljung-Box (Bhatnagar et al., 2012).

Em uma pesquisa de Porto Alegre (RS), comparou-se o SARIMA com o SARIMAX para prever os casos de dengue. O SARIMA obteve MAE = 53,14 e MAPE = 125,67% no primeiro período de análise e MAE = 56,35 e MAPE = 620,34% no segundo. Já o SARIMAX, obteve MAE = 53,02 e MAPE = 98,00% no primeiro período analisado e MAE = 56,16 e MAPE = 587,06% no segundo (Ferreira, 2024). Esses resultados reforçam a

importância do uso das variáveis exógenas na predição de arboviroses. Quanto às métricas de erro, expõe que não só para FO que há uma dificuldade em encontrar bons ajustes para os modelos preditivos. No presente estudo, o MAPE para três meses foi de 182% para o LSTM; 133.852% para o SARIMAX e 66.630% para a média aritmética. Assim, talvez pela maior incidência, percebe-se que os modelos captam melhor as variações da dengue.

Em Jambi, na Indonésia, um estudo comparou o SARIMA e o LSTM na predição de casos de Febre Hemorrágica da Dengue. Ambos apresentaram bom desempenho, porém o SARIMA se mostrou mais adequado (ICCAI, 2019). Em Surabaya, também na Indonésia, uma pesquisa reiterou o melhor desempenho do SARIMA em relação ao ARIMA e LSTM na predição de casos de dengue (Othman et al., 2022). Tais resultados diferem dos encontrados no presente estudo, em que o LSTM demonstrou desempenho superior ao SARIMAX. Essa situação pode decorrer de diferenças entre dengue e FO, entre Brasil e Indonésia, entre os períodos analisados, o horizonte preditivo, entre a proporção de treino e validação e etc. Desse modo, existe uma grande variabilidade na confiabilidade dos modelos, que requerem validação constante.

No Rio de Janeiro (RJ), um estudo comparou os modelos Autorregressivo, Média Móvel, ARIMA, Suavização Exponencial, SARIMA, SARIMAX, Autorregressão Vetorial, Floresta Aleatória, XGBoost, SVM, LSTM e Prophet na predição de casos de dengue. Os modelos SARIMA e SARIMAX obtiveram os melhores desempenhos com MAE de 78,26 e 79,24 respectivamente e MAPE de 18,25% e 17,25% respectivamente (Chen; Moraga, 2025). Novamente, percebe-se o SARIMAX com melhor desempenho que o LSTM, divergindo do presente estudo. Ademais, cabe destacar as métricas de erro, as quais ficaram bem menores. Assim, nota-se que é possível encontrar bons ajustes a partir da testagem massiva de diversos modelos para arboviroses. Desse modo, percebe-se a importância de continuar a calibragem do modelo apresentado no presente estudo para que, no longo prazo, possa-se ter estimativas mais precisas para períodos maiores de FO.

Um estudo propôs um modelo híbrido de equações diferenciais fracionárias e inteligência artificial para prever FO. Isso pois a simplicidade dos modelos compartimentais dificulta a captura das complexidades da dinâmica dessa arbovirose. A combinação resultou em melhorias consideráveis na previsão (Raghavendran; Gunasekar, 2026). Portanto, é reconhecida a necessidade e viabilidade da construção de modelos preditivos para FO, assim como as limitações inerentes à tarefa. O presente estudo corrobora com essa tese e reforça a

necessidade de contínua testagem de técnicas preditivas buscando menores erros e maiores horizontes para FO.

CONCLUSÃO

As mudanças climáticas proporcionam a emergência e a reemergência de doenças infecciosas, como a Febre do Oropouche. O Brasil já presenciou o impacto dessa situação na epidemia de dezembro de 2023 até junho de 2025. Portanto, é necessária a criação de políticas públicas que visem o controle dessa enfermidade. Para tal, o modelo LSTM associado a variáveis exógenas se mostrou com potencial para otimizar a definição dessas ações governamentais. Porém, novos estudos são necessários para melhorar a calibragem do modelo visando alcançar maior precisão e horizonte preditivo.

REFERÊNCIAS BIBLIOGRÁFICAS

- BAI, Fengwei et al. A Comprehensive Review of the Neglected and Emerging Oropouche Virus. **Viruses**, v. 17, n. 3, p. 439, 2025. DOI: <https://doi.org/10.3390/v17030439>. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11945866/#sec1-viruses-17-00439>. Acesso em 30 mar. de 2026.
- BHATNAGAR, Sunil et al. Forecasting Incidence of Dengue in Rajasthan, Using Time Series Analyses. **Indian Journal of Public Health**, v. 56, n. 4, p 281-285, 2012. DOI: <https://doi.org/10.4103/0019-557X.106415>. Disponível em: https://journals.lww.com/ijph/fulltext/2012/56040/forecasting_incidence_of_dengue_in_rajasthan,.5.aspx. Acesso em: 04 abr. 2025.
- BRASIL, Ministério da Saúde. **Nota técnica nº 6/2024-CGAR/DEDT/SVSA/MS**. Orientações para a vigilância da Febre do Oropouche. Brasília: Ministério da Saúde, 2024.
- BRASIL, Ministério da Saúde. **Plataforma GovBR**, 2025. Oropouche > Painel Epidemiológico. Disponível em: <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/o/oropouche/painel-epidemiologico>. Acesso em: 10 abr. 2025.
- CAMPANHOL, Andriza Maria da Cunha. **Explicabilidade de modelo Long Short-Term Memory (LSTM) para previsão de dengue em Porto Alegre**. 2024. 45 f. TCC (Graduação) – Curso de Ciência da Computação, Universidade Federal de Santa Maria, Santa Maria, RS. Disponível em: <https://repositorio.ufsm.br/handle/1/32817>. Acesso em: 3 mai. 2025.
- CARL, Isadora Amorim et al. **Clinical-epidemiological profile of oropouche fever cases in Santa Catarina between 2024 and 2025**, v. 30, n. 1, 2026. DOI: <https://doi.org/10.1016/j.bjid.2026.105606>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1413867026009967>. Acesso em: 31 mar. 2026.
- CHEN, Xiang; MORAGA, Paula. Assessing dengue forecasting methods: a comparative study of statistical models and machine learning techniques in Rio de Janeiro, Brazil. **Tropical Medicine and Health**, v. 53, n. 1, p. 52, 2025. DOI: <https://doi.org/10.1186/s41182-025-00723-7>. Disponível em: <https://tropmedhealth.biomedcentral.com/articles/10.1186/s41182-025-00723-7>. Acesso em: 3 mai. 2025.
- COLA, João Paulo et al. Sociodemographic characteristics, clinical manifestations, and concurrent outcomes of Oropouche fever and dengue in Espírito Santo, Brazil. **Revista Panamericana de Salud Pública**, v. 49, 2026.

DOI: <https://doi.org/10.26633/RPSP.2025.127>. Disponível em:
<https://www.scielo.org/article/rpsp/2025.v49/e127/en/>. Acesso em: 12 mar. 2026.

ESTADOS UNIDOS, Departamento de Saúde e Serviços Humanos, Centros de Controle e Prevenção de Doenças. **Clinical Overview of Oropouche Virus Disease**. Atlanta, 2025. Disponível em:
<https://www.cdc.gov/oropouche/hcp/clinical-overview/index.html>. Acesso em: 14 mar. 2025.

FERREIRA, Raul Santos. **Predição semanal de números de casos de dengue em Porto Alegre por modelos de séries temporais**. 2024. 99 f. TCC (Graduação) – Curso de Engenharia de Controle e Automação, Universidade Federal de Santa Maria, Santa Maria, RS. Disponível em:
<https://repositorio.ufsm.br/handle/1/32624>. Acesso em: 7 abr. 2025.

FLORES, Edmundo Rinolino Magalhães et al. Oropouche fever in Minas Gerais, Brazil, 2024: Epidemiological characterization, spatial distribution, and clinical implications. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 59, 2026. DOI: <https://doi.org/10.1590/0037-8682-0109-2025>. Disponível em:
<https://pmc.ncbi.nlm.nih.gov/articles/PMC12990782/>. Acesso em: 12 mar. 2026.

GOMES, Larissa Santos Carneiro. et al. Ensaios Diagnósticos para Detecção do Vírus Oropouche em Amostras Humanas: Revisão Sistemática. **Pleiade**, v. 19, n. 46, p. 115-118, 2025. DOI:
<https://doi.org/10.32915/pleiade.v19i46.1141>. Disponível em:
<https://pleiade.uniamerica.br/index.php/pleiade/article/view/1141/1343>. Acesso em: 31 mar. 2025.

GONZÁLEZ-QUEVEDO, Alina; FARRILL, Zurina Lestay O; BECQUER, Reinaldo Mustelier. Oropouche virus – another antecedent event for Guillain–Barré syndrome?. **Pan American Journal of Public Health**, v. 49, n. 2, 2025. ISSN-e: 1680-5348. ISSN: 1020-4989. Disponível em: <https://dialnet.unirioja.es/servlet/articulo?codigo=10018764>. Acesso em: 31 mar. 2025.

HASAN, Pratyay et al. Predictive Insights From ARIMA and SARIMA Models in Bangladesh: A Time Series Analysis. **Health Science Reports**, v. 7, n. 12, p. e70276. DOI: <https://doi.org/10.1002/hsr2.70276>. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1002/hsr2.70276>. Acesso em: 04 abr. 2025.

HUA, Xinyi et al. Ecological and demographic drivers of Oropouche virus transmission. **Nature Health**, 2026. DOI: <https://doi.org/10.1038/s44360-026-00065-6>. Disponível em: <https://www.nature.com/articles/s44360-026-00065-6#citeas>. Acesso em: 30 mar. de 2026.

INTERNATIONAL CONFERENCE ON COMPUTING AND APPLIED INFORMATICS (ICCAI 2019), 4., 2019, Medan. **Journal of Physics: Conference Series** [...]. Universidade Utara Malaysia, 2019. Tema: A comparison of SARIMA and LSTM in forecasting dengue hemorrhagic fever incidence in Jambi, Indonesia. DOI: <https://doi.org/10.1088/1742-6596/1566/1/012054>. Disponível em:
<https://iopscience.iop.org/article/10.1088/1742-6596/1566/1/012054/meta>. Acesso em: 10 abr. 2025.

LIM, Jue Tao et al. Forecasting upper respiratory tract infection burden using high-dimensional time series data and forecast combinations. **PloS Computational Biology**, v. 19, n. 2, p. E1010892. DOI:
<https://doi.org/10.1371/journal.pcbi.1010892>. Disponível em: <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1010892>. Acesso em: 12 mar. de 2026.

LORENZ, Camila et al. Oropouche fever outbreak in Brazil: Key factors behind the largest epidemic in history. **PLOS One**, v. 20, n. 7, p. e0327845, 2025. DOI: <https://doi.org/10.1371/journal.pone.0327845>. Disponível em:
<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0327845>. Acesso em: 23 ago. 2025.

LORENZ, Camila et al. Brazil is already experiencing the brutal impacts of climate change. **Revista de Saúde Pública**, v. 60, n. 13, 2026. DOI: <https://doi.org/10.11606/s1518-8787.2026060007009>. Disponível em:
<https://www.scielo.org/article/rsp/2026.v60/e6/en/>. Acesso em: 12 mar. 2026.

LUTZ, Chelsea S. et al. Applying infectious disease forecasting to public health: a path forward using influenza forecasting examples. **BMC Public Health**, v. 19, n. 1659, 2019. DOI: <https://doi.org/10.1186/s12889-019-7966-8>. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC6902553/>. Acesso em: 18 dez. 2024.

MANULI, Erika R. et al. Transmission dynamics of Oropouche virus in Latin America and the Caribbean. **Nature Medicine**, 2026. DOI: <https://doi.org/10.1038/s41591-026-04221-z>. Disponível em: <https://www.nature.com/articles/s41591-026-04221-z>. Acesso em 31 mar. 2026.

MARTINEZ, Edson Zangiacomi; SILVA, Elisângela Aparecida Soares da. Previsão do número de casos de dengue em Ribeirão Preto, São Paulo, Brasil, por um modelo SARIMA. **Caderno de Saúde Pública**, v. 27, n. 9, p. 1809-1818, 2011a. DOI: <https://doi.org/10.1590/S0102-311X2011000900014>. Disponível em: <https://www.scielo.br/j/csp/a/KGP9rcHBH7DHHhRNftFzwM/>. Acesso em: 04 abr. 2025.

MILANI, Pamela et al. Molecular and Serological Evidence of Oropouche Virus Circulation in Asymptomatic Blood Donors During the 2023–2024 Outbreak in Manaus, Brazil. **The Journal of Infectious Diseases**, v. 26, 2026. DOI: <https://doi.org/10.1093/infdis/jiag088>. Disponível em: <https://academic.oup.com/jid/advance-article/doi/10.1093/infdis/jiag088/8487534>. Acesso em: 12 mar. 2026.

OKESANYA, John Olalekan et al. Addressing the emerging threat of Oropouche virus: implications and public health responses for healthcare systems. **Trop Dis Travel Med Vaccines**, v. 11, n. 1, 2025. DOI: <https://doi.org/10.1186/s40794-024-00236-x>. Disponível em: <https://link.springer.com/article/10.1186/s40794-024-00236-x>. Acesso em: 14 mar. 2025.

OLIVEIRA, Gabriele Nascimento de et al. Entre o desconhecido e o emergente: mapeando os vírus Oropouche e Mayaro no Brasil. **Cadernos de saúde pública**, v. 42, p. e00067525, 2026. DOI: <https://doi.org/10.1590/0102-311XPT067525>. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC12956284/>. Acesso em: 12 de mar. de 2026.

ORSSATTO, Angelo José et al. Aplicação de séries temporais na estimação do número de casos de dengue em Cascavel-PR. **Revista Brasileira de Computação Aplicada**, [S. l.], v. 14, n. 3, p. 37-50, 2022. DOI: <https://doi.org/10.5335/rbca.v14i3.13483>. Disponível em: <https://seer.upf.br/index.php/rbca/article/view/13483>. Acesso em: 10 abr. 2025.

OTHMAN, Mahmod et al. Model Forecasting Development for Dengue Fever Incidence in Surabaya City Using Time Series Analysis. **Processes**, v. 10, n. 11, p. 2454, 2022. DOI: <https://doi.org/10.3390/pr10112454>. Disponível em: <https://www.mdpi.com/2227-9717/10/11/2454>. Acesso em: 10 abr. 2025.

PASTULA, Daniel M; BECKHAM, J. David; TYLER, Kenneth L. Oropouche Virus: An Emerging Neuroinvasive Arbovirus. **Annals of Neurology**, v. 97, n. 1, p. 28-33, 2024. DOI: <https://doi.org/10.1002/ana.27139>. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/ana.27139>. Acesso em: 14 mar. 2025.

PENHA, Gabriele Barbosa et al. Culicoides (Diptera: Ceratopogonidae) in Extra-Amazonian Oropouche Outbreak Areas of Minas Gerais, Brazil: Ecological Insights into Virus Transmission. **Viruses**, v. 18, n. 3, p. 361, 2026. DOI: <https://doi.org/10.3390/v18030361>. Disponível em: <https://www.mdpi.com/1999-4915/18/3/361>. Acesso em: 30 mar. de 2026.

PORWAL, Sejal et al. Mysterious Oropouche virus: Transmission, symptoms, and control. **Infectious Medicine**, v. 4, n. 2, p. 100177, 2025. DOI: <https://doi.org/10.1016/j.imj.2025.100177>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2772431X25000164>. Acesso em: 30 mar. de 2026.

RAGHAVENDRAN, Prabakaran; GUNASEKAR, Tharmalingam. Chapter 8 – AI-driven fractional order models for emerging viral epidemics: Application to Oropouche virus outbreak forecasting. **Mathematical Modeling and AI-Driven Computational Techniques for Epidemiology and Disease Dynamics**, p. 211-242, 2026. DOI: <https://doi.org/10.1016/B978-0-44-333234-0.00018-2>. Disponível em: <https://www.sciencedirect.com/science/chapter/edited-volume/abs/pii/B9780443332340000182>. Acesso em: 12 mar. 2026.

SILVA, Gabriel Manoel da et al. Uma aplicação do modelo SARIMA para prever o número de notificações de dengue na região de Paranavaí – Paraná. **Tecnologias avançadas em Ciências Exatas: Desenvolvimentos recentes e perspectivas futuras**, v. 1, n. 4, 2023. DOI: <https://doi.org/10.47402/ed.ep.c231934272>. Disponível em: <https://editorapublicar.com.br/ojs/index.php/publicacoes/article/view/630>. Acesso em: 04 abr. 2025.

TORRES, Jaime R. et al. Epidemiología, diagnóstico y tratamiento de las infecciones por virus Mayaro y Oropouche: implicaciones para la práctica clínica y la salud pública. **Enfermedades infecciosas y microbiología clínica**, n. 503055, 2026. DOI: <https://doi.org/10.1016/j.eimce.2025.503055>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S2529993X25002357>. Acesso em: 12 mar. de 2026.

WANG, Xin et al. Comprehensive analysis of epidemiological and clinical features of Oropouche virus infection (1975 to 2025): a systematic review. **New Microbes and New Infections**, v. 70, n. 101705, 2026. DOI: <https://doi.org/10.1016/j.nmni.2026.101725>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2052297526000296>. Acesso em: 12 mar. de 2026.

4 CONSIDERAÇÕES FINAIS

O presente Trabalho de Curso representa o resultado de três semestres de aprendizado sobre como conduzir uma pesquisa científica na área da saúde. Os conteúdos teóricos foram aplicados na prática na forma de um estudo sobre modelos preditivos na estimativa de casos de Febre do Oropouche (FO) no Brasil. Foram utilizados dados do Painel Epidemiológico do Ministério da Saúde, que recém incluiu FO entre as doenças de notificação compulsória.

Os resultados mostraram expansão de uma doença anteriormente apenas amazônica para outras partes do país, em especial a Região Sudeste. As explicações por trás desse fenômeno, segundo a literatura científica, giram em torno das mudanças climáticas e desmatamento, mostrando como as ações humanas no meio ambiente podem impactar em novas epidemias. Ademais, percebeu-se o desafio de implementar modelos preditivos em doenças emergentes, mostrando a complexidade da técnica abordada e a necessidade de pesquisas continuadas no longo prazo.

Portanto, por meio desse componente curricular, foi possível compreender os desafios em produzir pesquisa acadêmica e, em especial, o de transformá-la em ferramentas úteis para auxiliar na prática de gestão em saúde pública. Afinal, por mais que os casos de FO tenham diminuído e as previsões não refiram provável aumento, é só questão de tempo para uma nova epidemia, a qual pode e deve ser antecipada com modelos preditivos.