



UNIVERSIDADE DO ESTADO DA BAHIA
DEPARTAMENTO DE CIÊNCIAS EXATAS E DA TERRA
CURSO DE GRADUAÇÃO EM SISTEMAS DE INFORMAÇÃO

VITOR MANOEL PEREIRA DOS SANTOS

**MODELANDO A DINÂMICA DE TRANSMISSÃO DO MPOX ENTRE HUMANOS:
UM MODELO COMPARTIMENTAL COM REDES DE POPULAÇÕES CONECTADAS**

SALVADOR, BAHIA, BRASIL

2023

VITOR MANOEL PEREIRA DOS SANTOS

MODELANDO A DINÂMICA DE TRANSMISSÃO DO MPOX ENTRE HUMANOS: UM
MODELO COMPARTIMENTAL COM REDES DE POPULAÇÕES CONECTADAS

Monografia apresentada ao curso de Sistemas de Informação do Departamento de Ciências Exatas e da Terra da Universidade do Estado da Bahia - UNEB, como requisito parcial à obtenção do grau de bacharel em Sistemas de Informação. Área de Concentração: Ciência da Computação

Orientador: Prof. Dr. Diego Gervasio Frías Suárez

SALVADOR, BAHIA, BRASIL

2023

Termo de Anuência do Orientador

Declaro para os devidos fins que li e revisei este trabalho e atesto sua qualidade como resultado final desta monografia. Confirmando que o referencial teórico apresentado é completo e suficiente para fundamentar os objetivos propostos e que a metodologia científica utilizada e os resultados finais são consistentes e com qualidade suficiente para submissão à banca examinadora final do Trabalho de Conclusão de Curso II do curso de Bacharelado em Sistemas de Informação.



Prof. Dr. Diego Gervasio Frías Suárez

VITOR MANOEL PEREIRA DOS SANTOS

MODELANDO A DINÂMICA DE TRANSMISSÃO DO MPOX ENTRE HUMANOS: UM
MODELO COMPARTIMENTAL COM REDES DE POPULAÇÕES CONECTADAS

Monografia apresentada ao curso de Sistemas de Informação do Departamento de Ciências Exatas e da Terra da Universidade do Estado da Bahia - UNEB, como requisito parcial à obtenção do grau de bacharel em Sistemas de Informação. Área de Concentração: Ciência da Computação

Aprovada em: 06 de Dezembro de 2023

BANCA EXAMINADORA

Prof. Dr. Diego Gervasio Frías Suárez (Orientador)
Universidade do Estado da Bahia – UNEB

Prof. Dra. Maria Inês Valderrama Restovic
Universidade do Estado da Bahia - UNEB

Prof. Dr. Vagner de Souza Fonseca
Universidade do Estado da Bahia - UNEB

Prof. Dr. Wildo Navegantes de Araújo
Universidade de Brasília - UNB



Documento assinado digitalmente

WILDO NAVEGANTES DE ARAUJO

Data: 19/12/2023 16:23:15-0300

Verifique em <https://validar.iti.gov.br>

Aos meus pais, que dedicaram seu amor e esperança em mim, acreditando que eu chegaria até aqui. E também à minha família e amigos pelo apoio imensurável que me proporcionaram ao longo desta jornada.

AGRADECIMENTOS

Aos meus pais, Manoel e Jurilda, que são a fonte de todas as minhas conquistas e a razão pela qual tenho a determinação para prosseguir. A eles, devo não apenas minha vida, mas também minha vontade incessante de crescer. Seu apoio inabalável e amor constante me fornecem força e propósito a cada dia. À minha querida "mãedrinha" Eliete, agradeço por todo suporte e esperança que sempre depositou em mim, desde o meu nascimento. Seu carinho e cuidado moldaram quem sou hoje. Expresso minha gratidão às minhas tias Eliene e Neuma, que desempenham o papel de mães em minha vida, nunca hesitando em me apoiar e torcer por mim. Aos meus tios, primos e primas, minha família, por serem essenciais na minha jornada e por contribuírem para a pessoa que me tornei.

Gostaria de estender meu agradecimento a Diego Frias, cuja contribuição foi crucial para o desenvolvimento deste trabalho. Sua orientação, suporte técnico e emocional foram inestimáveis, e eu sou profundamente grato por sua compaixão diante dos meus erros e falhas. Agradeço ao professor Alexandre Lenz pelos conselhos sábios e pela paciência ao me corrigir. Ao professor Marco Simões, agradeço pelo conhecimento transmitido. À professora Inês, minha gratidão por me introduzir ao mundo da computação e pelos fundamentos essenciais que compartilhou. A todos os outros professores do curso, reconheço a importância de suas contribuições, que foram basilares para que eu alcançasse este momento.

Aos meus amigos, que se tornaram minha família, Aron, Evelyn, Rafael, Rickson, Jorge, Ludmila, Vanessa, Tiago, Lázaro, Jadson e Gabriel, agradeço por estarem ao meu lado durante toda essa jornada.

Por fim, dedico um agradecimento especial aos meus irmãos Júnior, João, Nando, Camile, Reinilson, Filipe e Jéssica, por quem guardo um carinho imensurável e com quem espero compartilhar muitas outras conquistas no futuro.

Meu carinho por todos vocês é infinito.

"A verdadeira sabedoria está em reconhecer a própria ignorância."

(Sócrates)

RESUMO

Neste estudo, concentramos nossa atenção no mpox, um vírus que desencadeou preocupações significativas nas autoridades de saúde global, especialmente após o surto de 2022 que afetou regiões não endêmicas, especialmente entre a população de HSH na América Latina. Nosso objetivo foi desenvolver um modelo matemático diferencial compartimental baseado em redes de contato para descrever a dinâmica do mpox. Utilizando dados dos casos ocorridos durante o surto de 2022 em países latino-americanos, nossa pesquisa buscou compreender e prever o comportamento da doença. Os resultados obtidos indicam que nossa abordagem inovadora foi eficaz na descrição da dinâmica epidemiológica do mpox. Conseguimos reproduzir com precisão os casos reais no Brasil, Colômbia, México, Peru e Estados Unidos, países analisados neste estudo. Além disso, pudemos estimar parâmetros epidemiológicos por meio da análise e mineração de dados reais fornecidos por órgãos de saúde. No entanto, enfatizamos a necessidade de dados mais confiáveis para uma melhoria contínua dos resultados. Além disso, ressaltamos a importância de investigações futuras para aprimorar a estimação do número de redes populacionais conectadas, a fim de fortalecer nossa compreensão e gestão eficaz de epidemias como o mpox.

Palavras-chave: Mpox. Modelagem epidemiológica. Redes conectadas. Equações diferenciais. Previsão de epidemias. Modelo compartimental

ABSTRACT

In this study, our focus was on mpox, a virus that raised significant concerns among global health authorities, especially after the 2022 outbreak that affected non-endemic regions, particularly among the MSM population in Latin America. Our aim was to develop a compartmental differential mathematical model based on contact networks to describe mpox dynamics. Utilizing data from cases during the 2022 outbreak in Latin American countries, our research sought to understand and predict the disease's behavior. The obtained results indicate that our innovative approach was effective in describing mpox epidemiological dynamics. We were able to accurately reproduce real cases in Brazil, Colombia, Mexico, Peru, and the United States, countries analyzed in this study. Furthermore, we could estimate epidemiological parameters through the analysis and mining of real data provided by health organizations. However, we emphasize the need for more reliable data for ongoing improvement of results. Additionally, we highlight the importance of future investigations to enhance the estimation of the number of connected population networks, in order to strengthen our understanding and effective management of epidemics like mpox.

Keywords: Mpox. Epidemiological modeling. Connected networks. Differential equations. Epidemic forecasting. Compartmental model

LISTA DE FIGURAS

| | |
|--|-----|
| Figura 1 – Principais eventos históricos e geográficos relacionados à infecção pelo vírus mpox | 27 |
| Figura 2 – Diagrama do modelo compartimental SIR | 32 |
| Figura 3 – Diagrama do modelo SEIQR apresentado por Bragazzi et al. (2023) | 39 |
| Figura 4 – Modelo compartimental desenvolvido por Al-Shomrani et al. (2023) | 41 |
| Figura 5 – Identificação dos pontos de transição de ondas no surto de Covid-19 na Índia | 42 |
| Figura 6 – Etapas do DSRM | 43 |
| Figura 7 – Algoritmo de suavização de dados por média móvel ponderada | 50 |
| Figura 8 – Processo de seleção da combinação inicial de variáveis | 57 |
| Figura 9 – Processo de otimização dos parâmetros epidemiológicos | 58 |
| Figura 10 – Modelo SPLICE para mpox | 60 |
| Figura 11 – Distribuição de casos de Varíola dos macacos (mpox) nas Américas por data de início dos casos. | 66 |
| Figura 12 – Series temporais suavizadas por país | 69 |
| Figura 13 – Distribuição da quantidade de óbitos por tempo até a morte | 71 |
| Figura 14 – Distribuição da quantidade de casos por tempo até a detecção | 72 |
| Figura 15 – Identificação do número de dias até a detecção para 90% dos casos | 73 |
| Figura 16 – Distribuição da quantidade de casos por tempo até a detecção dos 90% dos casos detectados | 73 |
| Figura 17 – Identificação dos picos nas ondas epidemiológicas dos países Americanos | 74 |
| Figura 18 – Resultados da simulação do mpox no Brasil | 80 |
| Figura 19 – Resultados da simulação do mpox na Colômbia | 80 |
| Figura 20 – Resultados da simulação do mpox no México | 80 |
| Figura 21 – Resultados da simulação do mpox no Peru | 81 |
| Figura 22 – Resultados da simulação do mpox nos Estados Unidos | 81 |
| Figura 23 – Resultado da simulação considerando conjunto variado de parâmetros por país | 82 |
| Figura 24 – Número de casos por idade dos países das Américas. | 100 |
| Figura 25 – (Continuação) Número de casos por idade dos países das Américas. | 101 |
| Figura 26 – (Continuação) Número de casos por idade dos países das Américas. | 102 |

| | |
|---|-----|
| Figura 27 – Resultados da simulação do mpox no Brasil | 106 |
| Figura 28 – Resultados da simulação do mpox no Peru | 106 |
| Figura 29 – Resultado da simulação considerando conjunto variado de parâmetros por país | 107 |
| Figura 30 – Resultado das simulações dos diferentes cenários de taxa de vacinação . . . | 108 |
| Figura 31 – Distribuição da população entre os compartimentos para cada cenário de vacinação estudado. | 109 |

LISTA DE TABELAS

| | |
|---|-----|
| Tabela 1 – Compartimentos utilizados nos modelos compartimentais do mpox encontrados na literatura | 37 |
| Tabela 2 – Correlação dos compartimentos clássicos com supercompartimentos do splice | 59 |
| Tabela 3 – Principais parâmetros do modelo <i>Susceptible-Protected-Latent-Infectious-Controlled-Ended</i> (SPLICE) para mpox | 64 |
| Tabela 4 – Informações gerais da epidemia por país | 68 |
| Tabela 5 – Informações relacionadas ao óbito por mpox inferidas a partir do <i>dataset</i> da Organização Pan Americana de Saúde (OPAS) | 71 |
| Tabela 6 – cenários das redes de populações conectadas nas Américas | 75 |
| Tabela 7 – Configuração inicial dos parâmetros | 77 |
| Tabela 8 – Hiperparâmetros para processo de otimização dos parâmetros epidemiológicos | 78 |
| Tabela 9 – Parâmetros epidemiológicos resultantes da otimização | 78 |
| Tabela 10 – Parâmetros vinculados as redes conectadas resultantes | 80 |
| Tabela 11 – Taxa de erro da simulação para os países selecionados | 81 |
| Tabela 12 – Taxa de variação das previsões do modelo para os países selecionados . . . | 83 |
| Tabela 13 – cenários das redes de populações conectadas nas Américas | 105 |
| Tabela 14 – Parâmetros vinculados as redes conectadas resultantes | 106 |
| Tabela 15 – Taxa de erro da simulação para os países selecionados | 107 |
| Tabela 16 – Cobertura da epidemia para cada cenário de vacinação estudado | 109 |
| Tabela 17 – Total de casos em cada cenário de vacinação estudado | 110 |

LISTA DE ABREVIATURAS E SIGLAS

| | |
|--------|---|
| DSR | <i>Design Science Research</i> |
| DSRM | <i>Design Science Research Methodology</i> |
| EDO | Equação Diferencial Ordinária |
| HRG | Grupo de Alto Risco |
| HSH | homens que fazem sexo com homens |
| LRG | Grupo de Baixo Risco |
| mpox | Varíola dos macacos |
| NCDC | Centros de Controle e Prevenção de Doenças da Nigéria |
| nRMSE | Erro Quadrático Médio Normalizado |
| OMS | Organização Mundial de Saúde |
| OPAS | Organização Pan Americana de Saúde |
| R0 | Número de Reprodução Básico |
| SEIQR | Suscetível-Exposto-Infetado-Quarentenado-Recuperado |
| SEIR | Suscetível-Exposto-Infetado-Recuperado |
| SIR | Suscetíveis-Infetados-Recuperados |
| SPLICE | <i>Susceptible-Protected-Latent-Infectious-Controlled-Ended</i> |
| SVEIR | Suscetível-Vacinado-Exposto-Infetado-Recuperado |

LISTA DE SÍMBOLOS

| | |
|----------|--|
| <i>S</i> | Compartimento dos indivíduos suscetíveis |
| <i>I</i> | Compartimento dos indivíduos infectados |
| <i>R</i> | Compartimento dos indivíduos recuperados |
| <i>Q</i> | Compartimento dos indivíduos em quarentena |
| <i>V</i> | Compartimento dos indivíduos vacinados |
| <i>E</i> | Compartimento dos indivíduos expostos a doença |
| <i>P</i> | Compartimento dos indivíduos protegidos |
| <i>C</i> | Compartimento dos indivíduos com a doença controlada |
| <i>E</i> | Compartimento dos indivíduos em estágio finalizador |

SUMÁRIO

| | | |
|------------|---|-----------|
| 1 | INTRODUÇÃO | 17 |
| 1.1 | Objetivos | 21 |
| 1.2 | Justificativa | 21 |
| 2 | FUNDAMENTAÇÃO TEÓRICA | 24 |
| 2.1 | mpox | 24 |
| 2.1.1 | Transmissão e sintomas da doença | 25 |
| 2.1.2 | Epidemiologia | 26 |
| 2.1.2.1 | Surto de 2022 | 26 |
| 2.2 | Modelagem Epidemiológica | 27 |
| 2.2.1 | Equações Diferenciais Ordinárias - EDOs | 29 |
| 2.2.2 | Modelos Compartimentais | 31 |
| 2.2.3 | Redes de Contato | 33 |
| 2.3 | Análise de Dados | 34 |
| 2.4 | Trabalhos Correlatos | 36 |
| 2.4.1 | Bragazzi - Modelo com estratificação por comportamento sexual | 39 |
| 2.4.2 | Al-Shomrani - Modelagem dos estágios do mpox incluindo população animal | 40 |
| 2.4.3 | Gosh - Um modelo SIR para epidemias de múltiplas ondas | 41 |
| 3 | METODOLOGIA | 43 |
| 3.1 | Materiais e Métodos | 46 |
| 3.1.1 | Minação de dados | 47 |
| 3.1.2 | Suavização dos dados | 48 |
| 3.1.3 | Busca em grade | 50 |
| 3.1.4 | Avaliação do modelo | 52 |
| 3.1.4.1 | Função de Erro Utilizada | 52 |
| 3.1.4.2 | Justificativa da Função de Erro | 52 |
| 3.1.4.3 | Normalização dos Dados | 53 |
| 3.1.4.4 | Métrica de Perda de Forma ('ShLoss') | 54 |
| 3.1.4.5 | Métrica de Perda de Distância ('DtLoss') | 55 |

| | | |
|------------|--|------------|
| 3.1.4.6 | Combinação Ponderada de Métricas | 55 |
| 3.1.5 | Inicialização e Otimização dos parâmetros epidemiológicos | 55 |
| 3.1.5.1 | Escolhendo a combinação inicial de variáveis | 55 |
| 3.1.5.2 | Otimização dos parâmetros | 57 |
| 4 | MODELO SPLICE | 59 |
| 4.1 | Modelando mpox com SPLICE | 60 |
| 4.1.1 | Redes de Populações Conectadas | 62 |
| 4.1.2 | Sistemas de EDOs do modelo | 63 |
| 5 | DESENVOLVIMENTO | 65 |
| 5.1 | Caracterização dos dados | 65 |
| 5.2 | Series Suavizadas | 67 |
| 5.3 | Inferência de parâmetros e Mineração de dados | 69 |
| 5.3.1 | Informações e parâmetros relacionados a morte do individuo | 70 |
| 5.3.2 | Informações e parâmetros relacionados ao tempo de detecção dos casos . . . | 71 |
| 5.4 | Estimativa do número de redes dos países americanos | 73 |
| 5.5 | Simulação da epidemia e Estimativa dos parâmetros epidemiológicos para os países americanos | 75 |
| 5.5.1 | Estimativa dos parâmetros epidemiológicos | 76 |
| 5.5.2 | Simulação do mpox com SPLICE | 79 |
| 5.6 | Análise de Sensibilidade | 81 |
| 6 | ANÁLISE E DISCUSSÃO DOS RESULTADOS | 84 |
| 6.1 | Análise dos resultados da simulação | 88 |
| 7 | CONSIDERAÇÕES FINAIS | 92 |
| | REFERÊNCIAS | 93 |
| | APÊNDICES | 99 |
| | APÊNDICE A – Distribuição de casos por faixa etária nos países das Américas | 100 |
| | APÊNDICE B – Distribuição da população por compartimentos do SPLICE | 103 |
| | APÊNDICE C – Explorando Novos Cenários: | 105 |
| C.1 | Avaliando diferentes cenários de redes conectadas | 105 |
| C.2 | Avaliando Impactos de Diferentes Estratégias de Vacinação | 108 |
| | APÊNDICE D – Script python de otimização dos parâmetros, simulação e avaliação do modelo | 111 |

| | | |
|------------------|---|------------|
| <i>D.0.0.0.1</i> | <i>SPLICE MPX SIMULATOR - PAHO ACADEMIC STUDY</i> | <i>111</i> |
| <i>D.0.0.0.2</i> | <i>IMPORTING PYTHON LIBRARIES</i> | <i>111</i> |
| <i>D.0.0.0.3</i> | <i>SMOOTHED TIME SERIES PREPROCESSING FUNCTIONS</i> | <i>111</i> |
| <i>D.0.0.0.4</i> | <i>READING, VISUALIZING AND PLOTTING EPIDEMIC RAW DATA</i> | <i>116</i> |
| <i>D.0.0.0.5</i> | <i>READING, VISUALIZING AND PLOTTING EPIDEMIC SMOOTHED TIME SERIES</i> | <i>118</i> |
| <i>D.0.0.0.6</i> | <i>OPTIMIZER SECTION</i> | <i>127</i> |
| <i>D.0.0.0.7</i> | <i>GRADIENT UPDATER</i> | <i>132</i> |
| <i>D.0.0.0.8</i> | <i>POINT EVALUATOR: SIMULATION OF MODEL AT A GIVEN POINT IN THE PARAMETRIC GRID</i> | <i>136</i> |
| <i>D.0.0.1</i> | <i>COUNTRY SELECTION & DATA COLLECTION</i> | <i>152</i> |
| <i>D.0.0.2</i> | <i>DATA VISUALIZATION AND PREPROCESSING</i> | <i>152</i> |
| <i>D.0.0.2.1</i> | <i>HYPER-PARAMETER CONFIGURATION</i> | <i>153</i> |
| <i>D.0.1</i> | <i>BSO-CIV:</i> | <i>166</i> |
| <i>D.0.2</i> | <i>REPETITIONS OF SEARCHING FOR OPTIMAL SOLUTIONS</i> | <i>168</i> |
| <i>D.0.3</i> | <i>UNCERTAINTY ANALYSIS SECTION</i> | <i>170</i> |
| <i>D.0.4</i> | <i>SENSITIVITY ANALYSIS SECTION</i> | <i>172</i> |
| <i>D.0.4.1</i> | <i>Show results of simulation with best params</i> | <i>177</i> |

1 INTRODUÇÃO

A varíola dos macacos, conhecida atualmente como mpox (World Health Organization, 2022b), é uma doença infecciosa que tem suscitado grande preocupação entre as autoridades de saúde pública em todo o mundo. Especialmente após o surto ocorrido em 2022, que despertou a atenção por apresentar casos da doença em pessoas que não haviam viajado para países endêmicos anteriormente (KLUGE; AMMON, 2022). Trata-se de uma zoonose causada pelo vírus monkeypox, pertencente à família Poxviridae, sendo transmitida aos seres humanos por meio do contato com lesões ou fluidos biológicos de animais infectados, tais como roedores e símios não humanos. A transmissão entre humanos ocorre principalmente por exposição a secreções corporais, lesões cutâneas infectadas dos pacientes e inalação de gotículas respiratórias, embora materiais contaminados também possam ser fonte de infecção (GESSAIN et al., 2022).

Historicamente, os surtos de mpox estavam restritos aos países africanos, onde a doença é endêmica. No entanto, a partir do surto de 2022, novos casos da doença foram detectados em outras regiões. Até abril de 2023, foram confirmados 87.078 casos, dos quais mais de 98,3% originaram-se em países não endêmicos (Centers for Disease Control and Prevention, 2022). Os países mais afetados foram os Estados Unidos, o Brasil e a Espanha. Além disso, estudos realizados pelo Centro Europeu de Prevenção e Controle de Doenças (ECDC) e pela Organização Mundial de Saúde (OMS) em outubro de 2022 revelaram que os novos casos da doença estavam concentrados em populações de homens que fazem sexo com homens (HSH) (KUPFERSCHMIDT, 2022).

Embora a vacina contra a varíola seja eficaz no combate à mpox, sua administração foi interrompida após a erradicação da varíola por volta de 1980 (FINE et al., 1988). Além disso, a doença costumava ser negligenciada devido à sua ocorrência concentrada em regiões da África (NAKOUNE; OLLIARO, 2022).

Em 11 de maio de 2023, a doença deixou de ser categorizada como uma emergência de saúde global. Todavia, nas palavras do diretor-geral da OMS, Tedros Adhanom, "No entanto, como na Covid-19, isso não significa que o trabalho acabou. mpox continua a representar desafios significativos de saúde pública que precisam de uma resposta robusta, proativa e

sustentável"(CNN (Cable News Network), 2023).

Neste contexto, a modelagem matemática de epidemias desempenha um papel crucial como uma ferramenta importante para entender a dinâmica de transmissão de doenças virais e prever possíveis cenários. A construção de modelos matemáticos permite simular o comportamento da doença em diferentes populações, avaliar o impacto de medidas de controle e prevenção, bem como orientar a tomada de decisões em saúde pública (RHODES et al., 2020). Essa abordagem é fundamental para compreender e combater epidemias, contribuindo para a promoção da saúde e o bem-estar da população.

Dentre os diversos tipos de modelos epidêmicos existentes, os modelos compartimentais são amplamente utilizados para modelar epidemias e prever o comportamento da propagação de doenças infecciosas. Isso ocorre devido à capacidade desses modelos de descrever a dinâmica de como as pessoas são infectadas, adoecem e se recuperam de uma doença. Esses modelos dividem a população em diferentes compartimentos, permitindo a simulação de diversos cenários e estratégias de intervenção, como vacinação, isolamento social e tratamentos, a fim de avaliar sua eficácia na prevenção e controle da epidemia. Além disso, eles possibilitam a quantificação de parâmetros epidemiológicos, como o número básico de reprodução, a taxa de mortalidade e a duração da infecção, que são essenciais para compreender a evolução da epidemia e o impacto das intervenções de saúde pública.

Os modelos compartimentais podem variar em sua complexidade, geralmente derivando do modelo clássico que contém três compartimentos principais: Suscetíveis-Infectados-Recuperados (SIR) (KERMACK; MCKENDRICK, 1927 apud MOLLA et al., 2023). Esses modelos, conhecidos como "SIR-like", podem ser modificados ou expandidos ao adicionar compartimentos adicionais ou modificar as características dos compartimentos existentes para representar a heterogeneidade da população, como idade, sexo, propensão à vacinação e localização geográfica dos indivíduos. Essas modificações permitem uma representação mais realista da dinâmica da epidemia, considerando as características específicas de diferentes grupos populacionais e suas interações.

A utilização de modelos mais complexos, como modelos Suscetível-Exposto-Infectado-Recuperado (SEIR) (GAO et al., 2023; YANG et al., 2023), Suscetível-Exposto-Infectado-Quarentenado-Recuperado (SEIQR) (BRAGAZZI et al., 2023) e outros, permite incorporar estágios adicionais da doença, como o período de exposição antes do desenvolvimento da in-

fecção ativa. Esses modelos são particularmente úteis para descrever doenças com um período de incubação significativo, onde os indivíduos expostos não são imediatamente infecciosos. A inclusão desses compartimentos adicionais e a consideração de fatores como a transmissão assintomática e a presença de portadores assintomáticos podem fornecer informações importantes sobre a propagação da doença e o impacto das medidas de controle.

Além disso, os modelos compartimentais podem ser adaptados para incorporar outros fatores relevantes, como a interação entre diferentes regiões geográficas, a mobilidade da população, a distribuição espacial da doença e a eficácia variável de diferentes intervenções. Essas adaptações permitem simular e avaliar estratégias de intervenção específicas para diferentes contextos, contribuindo para a tomada de decisões informadas e eficazes em saúde pública.

A vigilância contínua, o desenvolvimento de estratégias de prevenção e controle, bem como o fortalecimento dos sistemas de saúde são fundamentais para enfrentar essa doença e garantir a segurança global.

A modelagem matemática da dinâmica de transmissão da mpox (mpox) tem sido explorada em diversos estudos, nos quais foram utilizados modelos compartimentais existentes ou desenvolvidos novos modelos para caracterizar essa doença. Em um levantamento realizado por Molla et al. (2023), são descritas as características gerais dos modelos abordados por alguns desses estudos.

Os trabalhos mais recentes têm se concentrado na descrição da dinâmica de transmissão do novo surto da doença, ocorrido em 2022. Eles enfocaram as novas características da doença, como a transmissão majoritariamente entre humanos, a presença em países não endêmicos e a segmentação da população em grupos de risco com base na orientação sexual ou na quantidade de contatos sexuais. Por exemplo, em estudos como os de Knight et al. (2022), Bragazzi et al. (2023) e Yang et al. (2023), foram construídos modelos compartimentais juntamente com redes de transmissão parcialmente conectadas, utilizando dados de casos de mpox em Ontário, Canadá. O modelo apresentado por Knight et al. (2022), em particular, inclui compartimentos para representar a população vacinada e isolada. Além das redes de contato, outras estruturas podem ser incorporadas aos modelos compartimentais para auxiliar na análise da dinâmica da doença. Por exemplo, no estudo realizado por Augsburger et al. (2022), uma estrutura baseada na teoria dos jogos é adicionada a um modelo compartimental Suscetível-Vacinado-Exposto-Infectado-Recuperado (SVEIR). A teoria dos jogos tem como

objetivo prever o comportamento individual em um cenário onde as estratégias escolhidas pelos indivíduos dependem das estratégias adotadas por outros na população. Nesse estudo, é integrada ao modelo uma estrutura de jogo que descreve o comportamento da população em relação às políticas de vacinação voluntária, com o intuito de analisar a eficácia da vacinação voluntária na prevenção do surto de mpox. A incorporação da teoria dos jogos é interessante, uma vez que o comportamento individual da população tem um impacto direto na dinâmica de transmissão da doença, resultando na redução do número de infecções. Em (AUGSBURGER et al., 2022), os indivíduos do compartimento suscetível participam do jogo de vacinação, decidindo se vão se vacinar ou não, com base em parâmetros que descrevem as consequências da vacinação e da infecção por mpox, bem como a taxa de infecção para indivíduos vacinados e não vacinados.

Embora já existam estudos na literatura que utilizam modelos compartimentais para descrever a dinâmica de transmissão da mpox, alguns problemas foram identificados nessas abordagens, especialmente nos estudos voltados para a modelagem do novo surto da doença. Um desses problemas, destacado por Korosec et al. (2023), refere-se à dificuldade de obtenção de dados específicos da doença para aprimorar a precisão dos modelos. A maioria dos autores acaba assumindo valores arbitrários ou baseando-se em dados disponíveis na literatura para estimar os parâmetros dos modelos. A escassez de dados sobre os casos do novo surto também dificulta a capacidade dos modelos de realizar previsões, como observado no estudo de Bragazzi et al. (2023), que se baseou em dados coletados em um período de apenas 7 dias.

Por fim, há uma escassez de modelos compartimentais direcionados ao estudo da dinâmica da mpox em países não endêmicos da América Latina. Além disso, não parece haver estudos que considerem que diferentes segmentos da população possam estar em diferentes momentos epidemiológicos, o que pode ajudar a explicar o comportamento observado nas curvas epidemiológicas da mpox durante o surto de 2022 em alguns países (MATHIEU et al., 2022). Essas lacunas destacam a importância de realizar pesquisas adicionais e desenvolver modelos mais abrangentes e contextualizados para a compreensão e previsão da dinâmica da mpox em diferentes cenários epidemiológicos.

Dessa forma, esse estudo visa contribuir para o avanço da compreensão da dinâmica de transmissão de doenças infecciosas, especialmente no caso do mpox, fornecendo um modelo robusto e computacionalmente viável para a previsão e controle de epidemias. A aplicação desse modelo em países da Américas, principalmente da América Latina, onde o surto de

2022 teve impacto significativo, será de extrema relevância para a saúde pública e auxiliará na implementação de estratégias eficientes de intervenção e prevenção.

1.1 OBJETIVOS

Este estudo visa formular, desenvolver e validar um modelo compartimental com redes de populações conectadas, descrita por um sistema de equações diferenciais, para a modelagem de epidemias, utilizando a mpox como caso de estudo.

- Realizar um estudo abrangente sobre modelos baseados em compartimentos para entender a dinâmica de transmissão do mpox entre humanos.
- Coletar e tratar dados das organizações de saúde para criar uma base informativa sobre a dinâmica da epidemia de mpox nas Américas durante o surto de 2022.
- Desenvolver um modelo matemático com equações diferenciais para descrever com precisão a dinâmica de transmissão da doença.
- Construir um modelo computacional baseado no modelo matemático e nos dados processados para prever os episódios epidêmicos nos países das Américas.
- Validar o modelo proposto comparando os dados previstos com os casos reais, analisando a evolução diária de novos casos da doença.

1.2 JUSTIFICATIVA

A execução deste trabalho, que envolve a formulação, desenvolvimento e validação de um modelo de equações diferenciais baseado em redes de contato para a modelagem da epidemia da mpox, apresenta uma série de motivos e relevâncias sociais, econômicas e acadêmicas. Embora a doença não seja mais considerada uma emergência sanitária global, ainda existem razões cruciais para a continuidade dos estudos relacionados a ela. A seguir, são destacados os principais pontos que justificam a importância deste projeto:

- **Impacto na saúde pública:** A mpox, agora denominada mpox, continua sendo uma doença infecciosa que representa desafios significativos para a saúde pública. Embora tenha deixado de ser categorizada como uma emergência global, ainda é essencial compreender sua dinâmica de transmissão, prever possíveis cenários e desenvolver estratégias eficazes de controle e prevenção. Ao formular e validar um modelo matemático para a epidemia,

este projeto contribuirá para a elaboração de medidas de saúde pública mais eficientes e direcionadas, visando mitigar os impactos da doença e proteger a população.

- **Contribuição científica e acadêmica:** O desenvolvimento de um modelo de equações diferenciais baseado em redes de contato para a mpox representa uma contribuição significativa para o campo da epidemiologia e modelagem de doenças infecciosas. Essa abordagem permite uma compreensão mais aprofundada da dinâmica de transmissão da doença em diferentes populações e contextos, levando em consideração as interações entre os indivíduos. Essa contribuição cumulativa ao conhecimento científico é essencial para melhorar as estratégias de controle e prevenção não apenas para a mpox, mas também para outras doenças infecciosas que compartilham mecanismos semelhantes de transmissão.
- **Preenchimento de lacunas no conhecimento:** Embora existam estudos relatados na literatura que utilizam modelos compartimentais para descrever a dinâmica de transmissão da mpox, há lacunas e desafios específicos que ainda precisam ser abordados. Uma das principais lacunas é a escassez de modelos que se concentrem na dinâmica da doença em países não endêmicos, como os países da América Latina. Além disso, a segmentação da população em grupos de risco com base na orientação sexual ou na quantidade de contatos sexuais é um aspecto que requer uma abordagem mais aprofundada. Ao realizar estudos de validação do modelo em países selecionados das Américas e explorar essas lacunas, este projeto busca preencher essas lacunas de conhecimento e fornecer informações valiosas para a compreensão e controle da doença.

Em resumo, a realização deste projeto contribuirá para o conhecimento científico sobre a mpox, fornecendo informações valiosas para a saúde pública e auxiliando na formulação de estratégias eficazes de controle e prevenção da doença. Além disso, o desenvolvimento de um modelo baseado em redes de contato preenche lacunas existentes na compreensão da dinâmica da doença em países não endêmicos e considerando segmentos específicos da população. Dessa forma, o projeto apresenta uma relevância social, econômica e acadêmica significativa ao contribuir para a superação de desafios e lacunas no conhecimento sobre a mpox e outras doenças infecciosas.

Na seções a seguir, este trabalho se aprofunda nos objetivos e justificativas que fundamentam esta monografia. No capítulo 2, estabelecemos o referencial teórico, explorando conceitos e informações que guiam nossa pesquisa sobre o mpox, modelagem epidemiológica

e análise de dados. Dentro deste capítulo, na seção 2.4, delineamos o estado atual da arte relacionado à modelagem epidemiológica do mpox por meio de modelos compartimentais. Além disso, nas subseções subsequentes, apresentamos trabalhos correlatos que serviram como base, inspiração ou estão de alguma forma relacionados ao nosso estudo.

O capítulo 3 descreve a metodologia científica empregada, detalhando os materiais, métodos e metodologias utilizados durante o desenvolvimento para alcançar nossos objetivos. Em seguida, no capítulo 4, introduzimos o cerne desta monografia: o modelo epidemiológico proposto e a abordagem de modelagem desenvolvida.

No capítulo 5, apresentamos o processo de desenvolvimento e os resultados obtidos em nosso trabalho. O capítulo 6 é dedicado à discussão desses resultados, levantando limitações e ameaças à validação do trabalho e propondo pontos de melhoria e direções para futuras pesquisas com base nas descobertas até o momento.

Por fim, no capítulo 7, elaboramos as considerações finais, consolidando as contribuições do nosso estudo e delineando possíveis caminhos para pesquisas subsequentes nesta área.

2 FUNDAMENTAÇÃO TEÓRICA

Nesta seção, são apresentados os principais conceitos que embasam este trabalho, fornecendo uma fundamentação teórica sólida para o desenvolvimento da pesquisa. São abordados os conceitos-chave relacionados à dinâmica de transmissão de doenças, modelos compartimentais e redes de contato. Além disso, são descritos trabalhos relacionados ao tema proposto nesta pesquisa, destacando suas contribuições, limitações e lacunas a serem preenchidas. A revisão da literatura tem como objetivo contextualizar o trabalho dentro do cenário científico atual e destacar a originalidade e relevância da abordagem adotada.

2.1 MPOX

O monkeypox foi identificada pela primeira vez em 1958, sendo isolada a partir de amostras de macacos de Singapura que estavam infectados. A primeira ocorrência em humanos foi registrada em 1970, quando o vírus foi detectado em uma criança da República Democrática do Congo.

Quanto à etimologia, o vírus foi nomeado como monkeypox, justamente por ter sido descoberto em espécies primatas. No entanto, em 2022, durante o surto global de varíola dos macacos, causado pelo vírus, após a observação do uso de linguagem racista e estigmatizante associado à doença e diversas solicitações de alteração do nome da doença, a OMS passou a adotar e recomendar a utilização de um novo sinônimo em inglês para a varíola dos macacos (World Health Organization, 2022b). Assim, o termo "mpox" se tornou o termo preferencial, substituindo o termo estigmatizante "monkeypox". As considerações para as recomendações incluíram justificativa, adequação científica, extensão do uso atual, pronúncia, usabilidade em diferentes idiomas, ausência de referências geográficas ou zoológicas e facilidade de recuperação de informações científicas históricas.

Quanto à etiologia, o mpox pertence à família Poxviridae e ao gênero Orthopoxvirus. Além do mpox, várias outras espécies de poxvírus podem causar infecções em humanos, tais como Varíola bovina, Varíola do camelo, Alaskapox, Yaba monkey tumor virus, Tanapox virus, Orf virus, Pseudocowpox virus, Bovine papular stomatitis virus, Buffalopox e Molluscum

contagiosum (TITANJI et al., 2022). O mpox é uma zoonose e, embora tenha sido identificada pela primeira vez em macacos, seu reservatório natural ainda é desconhecido. No entanto, diversos pequenos mamíferos são suscetíveis ao vírus, como esquilos, ratos, cães da pradaria e camundongos encontrados nas regiões africanas.

2.1.1 Transmissão e sintomas da doença

O mpox pode ser contraído por humanos através de duas rotas de transmissão: animal-humano ou humano-humano.

As transmissões de animal para humano ocorrem a partir de exposições a animais infectados, contato com animais infectados, seja através de mordidas e arranhões ou da manipulação e consumo de produtos dos animais.

A transmissão de humano para humano do vírus pode ocorrer por secreções respiratórias, contato direto, transmissão vertical, transmissão percutânea ou contato indireto através de fômites. A transmissão respiratória ocorre quando grandes gotículas respiratórias do hospedeiro transmissor se depositam nas membranas mucosas da boca e do nariz do hospedeiro receptor. O contato cara a cara prolongado, como o contato doméstico, pode ser necessário para que a transmissão ocorra por essa via. As atividades que resultam na ressuspensão do material seco das lesões também podem representar um risco e devem ser evitadas. O vírus mpox ainda pode se espalhar durante atividades que incluem contato próximo e íntimo com um indivíduo infectado, e a transmissão pode ser facilitada por uma brecha na pele ou mucosa do receptor, como abrasões microscópicas que ocorrem durante atividade sexual.

Os indivíduos infectados com mpox passam por três estágios da infecção: o período de incubação, o prodrômico e o de erupção cutânea (DOU et al., 2023). O período de incubação costuma durar entre uma e duas semanas, é quando o vírus se replica no local de inoculação e depois se espalha para os linfonodos locais. Em seguida, uma viremia inicial leva à disseminação viral e à disseminação para outros órgãos. Na fase prodrômica, os sintomas incluem febre, dores de cabeça, calafrios, dores de garganta, gânglios linfáticos inchados, mal-estar e linfadenopatia. Por fim, no último estágio da doença, o indivíduo começa a apresentar erupções cutâneas, que costumam aparecer em regiões do corpo como palmas das mãos e plantas dos pés, rosto, boca, garganta, virilha, áreas genitais e ânus.

Por fim, é importante esclarecer que pessoas com mpox são infecciosas e podem transmitir a doença para outras pessoas até que todas as feridas tenham cicatrizado e uma nova camada de pele tenha se formado.

2.1.2 Epidemiologia

Historicamente, os casos de mpox estiveram relacionados aos países africanos, nos quais a doença é endêmica. Até 2003, não havia relatos registrados de infecção humana fora desses países endêmicos, e a transmissão do vírus ocorre principalmente de animais para humanos.

No entanto, em 2003, o vírus foi identificado pela primeira vez nos EUA, infectando humanos e cães de pradaria. Nesse momento, a origem do vírus havia sido associada a roedores que haviam sido importados de Gana e que posteriormente infectaram os cães de pradaria.

Em 2017, ocorreu o maior surto de mpox já registrado na Nigéria. Esse surto atingiu 228 pessoas e foi caracterizado pela predominante transmissão entre humanos. Após o surto na Nigéria, a doença continuou a ocorrer em diversos países africanos. Segundo um relatório da OMS de 2022, os seguintes países são classificados como endêmicos: Benin, Camarões, República Centro-Africana, República Democrática do Congo, Gabão, Gana (com casos identificados apenas em animais), Costa do Marfim, Libéria, Nigéria, Serra Leoa e Sudão do Sul. Além disso, surtos esporádicos ocorreram em países não endêmicos, sempre relacionados a viagens e importação de animais provenientes de países africanos (World Health Organization, 2022a).

Essa disseminação da mpox além das fronteiras dos países endêmicos tem sido um desafio para os esforços de controle e prevenção da doença.

2.1.2.1 Surto de 2022

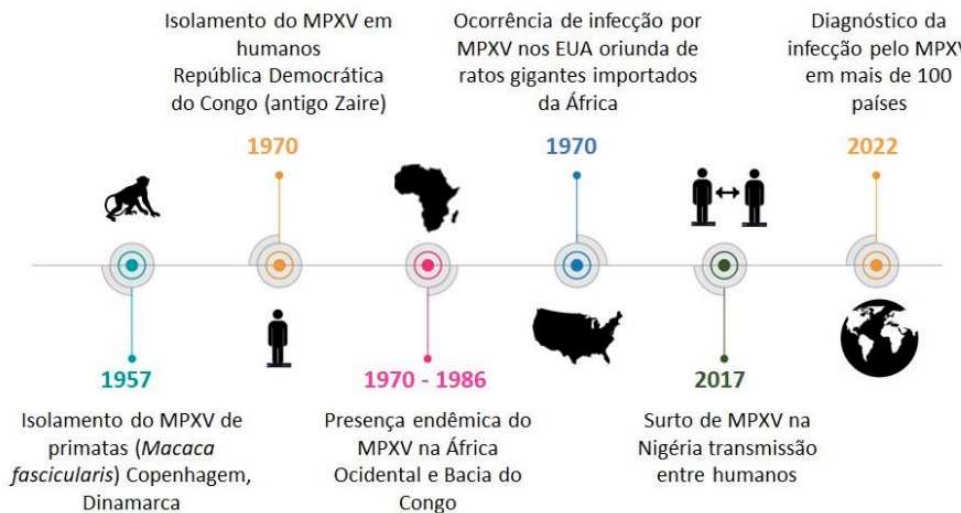
Em maio de 2022, foram registrados os primeiros casos de um novo surto, que posteriormente se tornou uma epidemia global. Desde então, o número de casos tem continuado a aumentar em todo o mundo. Em julho de 2022, a Organização Mundial da Saúde (OMS) declarou o surto de mpox como uma emergência de saúde pública de interesse internacional. Diferentemente dos casos anteriores observados, esse novo surto afetou principalmente os países

da Europa e das Américas, regiões onde a doença não é endêmica. Não há evidências que relacionem os casos relatados às áreas da África Ocidental e Central, o que sugere a possibilidade de transmissão não detectada por algum tempo.

Os 10 países mais afetados globalmente são: Estados Unidos da América (30.267 casos), Brasil (10.950 casos), Espanha (7.559 casos), França (4.146 casos), Colômbia (4.090 casos), México (4.026 casos), Peru (3.800 casos), Reino Unido (3.753 casos), Alemanha (3.691 casos) e Canadá (1.496 casos). Esses países juntos representam 83,9% dos casos notificados globalmente.

Além de afetar principalmente países não endêmicos, o surto de 2022 também se diferencia dos surtos anteriores pela caracterização da população afetada. Neste cenário, a grande maioria dos casos ocorreu em indivíduos masculinos adultos que mantinham relações sexuais com outros homens. De acordo com dados do relatório do CDC (Center for Disease Control and Prevention, 2023), nos Estados Unidos, o país mais afetado pela doença, mais de 95% dos casos registrados envolvem indivíduos do sexo masculino. A idade média dos casos é de 34 anos. Além disso, 75,1% dos casos com informações sobre o contato sexual são resultado de relações entre homens.

Figura 1 – Principais eventos históricos e geográficos relacionados à infecção pelo vírus mpox



Fonte: Retirada de Kreutz et al. (2022)

2.2 MODELAGEM EPIDEMIOLÓGICA

Os modelos têm sido úteis para adquirir conhecimento sobre os sistemas-alvo, explicar os fenômenos observados, organizar o conhecimento e desenvolver conceitos, entre

outros propósitos.

A modelagem epidemiológica refere-se ao conjunto de abordagens que utilizam ferramentas matemáticas, estatísticas e computacionais para estudar a propagação de patógenos transmissíveis em populações hospedeiras. Os modelos são construídos com base em dados e hipóteses capazes de descrever as características da epidemia, como processos demográficos, dinâmica de transmissão e consequências para a saúde (GARNER et al., 2011).

Os modelos matemáticos têm desempenhado um papel fundamental na compreensão e previsão de diversos fenômenos complexos em diferentes áreas do conhecimento. No campo da saúde pública, os modelos epidemiológicos são especialmente importantes para compreender a propagação de doenças infecciosas e planejar medidas de controle e prevenção. Esses modelos, ao combinar equações matemáticas, estatísticas e dados, oferecem uma representação simplificada do sistema em estudo, permitindo a investigação de diferentes cenários e a avaliação de intervenções.

Em particular, os modelos epidemiológicos são desenvolvidos para capturar a dinâmica da transmissão de doenças em populações. Eles são construídos com base em princípios matemáticos e estatísticos, incorporando dados epidemiológicos para estimar parâmetros importantes, como taxa de transmissão, taxa de recuperação e suscetibilidade da população. Esses modelos fornecem uma estrutura formal para analisar a propagação de doenças, permitindo simulações computacionais que podem ajudar a prever a disseminação de uma doença em diferentes cenários populacionais.

Um dos principais motivos para utilizar modelos epidemiológicos é a capacidade de obter *insights* sobre o curso potencial de uma epidemia e a eficácia de diferentes intervenções de controle. Esses modelos podem ser utilizados para prever a magnitude e a duração de uma epidemia, identificar grupos populacionais de alto risco e avaliar estratégias de mitigação. Além disso, os modelos epidemiológicos são essenciais para a tomada de decisões em saúde pública, fornecendo informações valiosas para a alocação de recursos, planejamento de vacinação em larga escala e implementação de medidas de distanciamento social.

Ao longo dos anos, diversos modelos epidemiológicos têm sido desenvolvidos e aprimorados por pesquisadores ao redor do mundo. Anderson e May (1991) e Wormser e Pourbahloul (2008) contribuíram significativamente para o campo dos modelos epidêmicos,

fornecendo fundamentos teóricos sólidos e insights importantes sobre as dinâmicas de propagação de doenças. A grande maioria dos modelos adota uma abordagem compartimental, que descreve a doença como uma sequência de diferentes estágios, desde a infecção até a cura ou morte (ANDERSON; MAY, 1991).

Um estudo de revisão sistemática realizado por Molla et al. (2023) examinou a modelagem matemática do vírus mpox, fornecendo uma visão abrangente dos trabalhos desenvolvidos nessa área. Os resultados revelaram que a maioria dos estudos utilizou modelos compartimentais, como o modelo SIR, para descrever a dinâmica de transmissão da doença. Esses modelos permitiram a análise quantitativa da propagação do vírus e a avaliação de estratégias de controle. No entanto, o estudo também destacou a utilização de outros modelos menos frequentes, como modelos de Monte Carlo (estocásticos), baseados em agentes e em redes. Essa diversidade de abordagens modelísticas evidencia a busca contínua por modelos mais precisos e adequados para descrever a dinâmica do vírus mpox, contribuindo para o avanço do conhecimento sobre a doença.

Entretanto, é importante destacar que os modelos epidemiológicos têm suas limitações. Eles são simplificações da complexa realidade, assumindo homogeneidade da população, transmissão aleatória e outros pressupostos simplificadores. Além disso, a precisão dos resultados depende da qualidade e disponibilidade dos dados utilizados, bem como da precisão dos parâmetros estimados. Portanto, é essencial ter cautela na interpretação dos resultados e considerar as incertezas associadas aos modelos.

2.2.1 Equações Diferenciais Ordinárias - EDOs

A Equação Diferencial Ordinária (EDO) desempenha um papel fundamental na modelagem matemática de fenômenos dinâmicos, incluindo a propagação de doenças em populações. Uma EDO é uma equação que relaciona uma função desconhecida com suas derivadas em relação a uma variável independente. No contexto dos modelos epidemiológicos, as EDOs são utilizadas para descrever as taxas de variação dos compartimentos em um sistema compartimental, como o modelo SIR. (KERMACK; MCKENDRICK, 1932)

$$\frac{dS}{dt} = -\beta S I \quad (2.1)$$

$$\frac{dI}{dt} = \beta S I - \gamma I \quad (2.2)$$

$$\frac{dR}{dt} = \gamma I \quad (2.3)$$

Onde:

dS/dt : Taxa de variação de suscetível em relação ao tempo.

dI/dt : Taxa de variação de infectados em relação ao tempo.

dR/dt : Taxa de variação de Recuperados em relação ao tempo.

β é a taxa de transmissão e γ a taxa de recuperação.

Os modelos epidemiológicos baseados em EDOs são amplamente utilizados devido à sua capacidade de capturar a dinâmica temporal das doenças infecciosas. Esses modelos são formulados como um sistema de equações diferenciais que descreve as mudanças nas populações de suscetíveis, infectados e recuperadas ao longo do tempo. Estas equações relacionam as taxas de variação dos compartimentos com base em parâmetros epidemiológicos, como a taxa de transmissão e a taxa de recuperação.

A solução das EDOs permite prever a evolução temporal da epidemia e entender como os diferentes compartimentos da população são afetados. Através da análise das soluções, é possível obter informações sobre a propagação da doença, como o pico da epidemia, o tempo necessário para a contenção da doença e a eficácia de intervenções de controle.

Vários métodos numéricos e analíticos podem ser aplicados para resolver as EDOs que governam os modelos epidemiológicos. Métodos clássicos, como o método de Euler e o método de Runge-Kutta, são comumente utilizados para aproximar numericamente as soluções das EDOs. Além disso, técnicas analíticas, como a linearização e a análise de estabilidade, permitem obter resultados teóricos sobre o comportamento do sistema ao longo do tempo.

A relação entre as EDOs e os modelos epidemiológicos é essencial para entender a propagação de doenças infecciosas e auxiliar na tomada de decisões em saúde pública. Ao formular as EDOs apropriadas com base nos princípios do modelo epidemiológico escolhido, é possível simular e prever cenários epidêmicos, avaliar a eficácia de estratégias de controle e

apoiar a tomada de decisões embasadas em evidências científicas.

Alguns trabalhos, por exemplo, Brauer et al. (2019) e Murray (2002), discutem detalhadamente a relação entre as EDOs e os modelos epidemiológicos, apresentando métodos e técnicas para a resolução e análise dessas equações no contexto das doenças infecciosas.

2.2.2 Modelos Compartmentais

Os modelos compartimentais são amplamente utilizados na epidemiologia para estudar a propagação de doenças infecciosas em populações, são representados por estruturas matemáticas que se baseiam na compreensão dos processos biológicos subjacentes à propagação de doenças em uma população. Esses modelos fornecem uma estrutura simplificada para representar a interação entre diferentes grupos de indivíduos, permitindo uma análise sistemática da dinâmica da doença ao longo do tempo (ANDERSON; MAY, 1991). Eles dividem a população em diferentes compartimentos, representando os vários estados possíveis em relação à doença, como suscetíveis, infectados, recuperados ou expostos. Essa estrutura permite a simulação da dinâmica da transmissão da doença ao longo do tempo, levando em consideração fatores como taxas de contágio, período de incubação e recuperação, entre outros.

Um dos modelos compartimentais mais comuns é o modelo SIR, que divide a população em três compartimentos: suscetíveis, infectados e recuperados. Os indivíduos suscetíveis são aqueles que ainda não possuem imunidade ao agente infeccioso. Os infectados são os indivíduos que estão atualmente infectados e têm a capacidade de transmitir a doença aos suscetíveis. Já os recuperados são os indivíduos que se recuperaram da doença e se tornaram imunes.

Em uma população suscetível, qualquer novo caso de infecção introduzido pode levar à transmissão da doença aos indivíduos suscetíveis durante o período infeccioso. Durante esse período, uma pessoa infectada é capaz de transmitir o patógeno a um hospedeiro suscetível, resultando no aumento da população de infectados e, conseqüentemente, na diminuição da população suscetível. Após o período infeccioso, os indivíduos infectados se recuperam e adquirem imunidade, o que leva ao aumento da população de recuperados.

Nesse modelo, a população se mantém constante, ou seja, a soma dos indivíduos suscetíveis, infectados e recuperados é igual ao tamanho total da população ($N = S + I + R$). A transição entre os compartimentos é governada por taxas de transmissão e recuperação. Na figura

2, o símbolo β denota a taxa de infecção por contato de indivíduos suscetíveis com pessoas infectadas, enquanto γ representa a taxa de recuperação dos indivíduos que estavam no estado I, passando para o estado R.

Figura 2 – Diagrama do modelo compartimental SIR



Fonte: Retirada de Carrión (2017)

Outros modelos compartimentais, como o modelo SEIR, adicionam um compartimento adicional para representar indivíduos expostos, que foram infectados, mas ainda não apresentam sintomas (WORMSER; POURBOHLOUL, 2008). Essa extensão permite a modelagem de doenças com um período de incubação, sendo útil para entender a disseminação de doenças como o COVID-19.

Esses modelos são fundamentados na interação entre os indivíduos da população e na propagação do agente infeccioso, possibilitando prever o número de indivíduos suscetíveis, infectados e recuperados ao longo do tempo. Eles são valiosos para a modelagem de epidemias de doenças virais transmissíveis, uma vez que permitem explorar diferentes cenários e estratégias de controle, como a implementação de vacinação em massa, medidas de distanciamento social ou intervenções farmacológicas.

A abordagem dos modelos compartimentais tem sido amplamente utilizada devido à sua capacidade de representar a dinâmica da doença em uma população de maneira quantitativa. Esses modelos podem ser analisados matematicamente, fornecendo uma compreensão mais profunda da propagação da doença e permitindo a previsão de resultados em diferentes cenários (HETHCOTE, 2000).

Kermack e McKendrick (1927) foram pioneiros na modelagem epidemiológica, estabelecendo as bases para os modelos compartimentais. Seu trabalho fundamental no modelo SIR permitiu uma melhor compreensão da propagação de doenças infecciosas e influenciou gerações subsequentes de pesquisadores na área.

Embora os modelos compartimentais sejam uma ferramenta valiosa, é importante

reconhecer suas limitações. Esses modelos são baseados em simplificações da realidade complexa e a precisão dos resultados depende da qualidade dos dados utilizados e da precisão dos parâmetros estimados (DIEKMANN et al., 1990).

Outra limitação dos modelos compartimentais clássicos é a incapacidade de representação de ondas epidemiológicas sobrepostas, em surtos de doenças de múltiplas ondas, essa limitação é inerente a natureza das EDOs que descrevem os modelos pois assumem que uma vez que a população infectada começou a diminuir, ela não aumentará novamente, o que é inevitável uma vez que a população suscetível vai diminuindo em número com o tempo, o que é denotado pelo sinal negativo na equação 2.2, do modelo. Isso sugere que o modelo SIR tradicional sempre prevê que, uma vez que uma única onda ocorra, a doença desaparecerá. Infelizmente, não houve uma única onda na maioria das regiões dos Estados Unidos e em muitos outros países. (PERAKIS et al., 2023) Existem diversos modos de contornar esta limitação, alguns autores incluem fatores de reinfeção no modelo, permitindo que indivíduos já infectados tenham a probabilidade de retornar para o estágio de suscetibilidade (GHOSH; GHOSH, 2022). Outros autores, como Perakis et al. (2023) utilizaram os dados para que seus modelos aprendessem momentos de mudança de comportamento dos indivíduos, sem se preocupar com os motivos que impulsionam a mudança, que culminam no ajuste dos parâmetros do modelo aumentando ou diminuindo a taxa de infecção. Por fim, outra abordagem é a utilização de redes de contato, que visam descrever as relações entre diferentes segmentos populacionais com características epidemiológicas diferentes.

2.2.3 Redes de Contato

As redes representam as interações entre os indivíduos, estabelecendo conexões que podem influenciar a disseminação de uma doença. Em um contexto epidemiológico, as redes de contato descrevem os padrões de interação social e contatos físicos entre os indivíduos, permitindo uma representação mais realista da dinâmica das epidemias (LÓPEZ-FLORES et al., 2021, ch. 1.9).

Uma rede de contato pode ser definida como um conjunto de nós (representando os indivíduos) conectados por arestas (representando os contatos entre os indivíduos). Essa estrutura de rede permite estudar como a propagação de uma doença ocorre através das conexões entre os indivíduos. Em uma rede de contato, um indivíduo infectado pode transmitir a doença para

seus contatos diretos, que por sua vez podem infectar outros indivíduos com os quais eles estão conectados. Esse processo de transmissão é modelado matematicamente por meio de modelos epidemiológicos baseados em redes de contato. (PASTOR-SATORRAS et al., 2015)

A representação de epidemias multiondas é uma das vantagens das redes de contato. Em uma epidemia multionda, a doença pode apresentar picos de atividade repetidos ao longo do tempo. As redes de contato permitem estudar a disseminação da doença em diferentes ondas, considerando as mudanças na conectividade e na interação social ao longo do tempo. Isso é especialmente relevante para doenças com características sazonais ou com surtos recorrentes.

Uma abordagem comum para modelar epidemias em redes de contato é por meio de grafos (MAGAL et al., 2018). Em um modelo baseado em grafos, cada nó representa um indivíduo, e as arestas representam os contatos entre os indivíduos. As características do grafo, como a distribuição de grau (número de conexões de um nó) e a estrutura da rede, têm um impacto significativo na propagação da doença. A análise de grafos permite investigar como a topologia da rede influencia a dinâmica epidêmica e pode fornecer *insights* sobre estratégias de controle e intervenções eficazes.

Além disso, as redes de contato também podem ser modeladas como um conjunto de n sistemas de equações compartimentais, em que os indivíduos de diferentes redes (sendo cada rede representada por um dos EDOs) infectam-se mutuamente. Esses modelos, conhecidos como modelos de interação ou modelos bipartidos, dividem a população em dois ou mais grupos distintos que interagem entre si. Por exemplo, em uma epidemia de doença animal que afeta os humanos, um sistema compartimental representa a população animal e o outro representa a população humana. A interação entre os dois sistemas permite estudar a transmissão cruzada da doença e a influência mútua entre as populações.

2.3 ANÁLISE DE DADOS

A compreensão e modelagem de epidemias dependem em grande parte da análise de dados (SCHOENBACH et al., 1999). Os modelos epidemiológicos são alimentados por informações coletadas em diversos estágios da disseminação da doença, desempenhando um papel fundamental na estimativa de parâmetros, validação do modelo e embasamento das decisões com base em evidências. A análise de dados desempenha permite a identificação de padrões, tendências e características relevantes associadas à propagação da doença. É por meio dessa

análise minuciosa que se torna possível compreender os padrões de transmissão, identificar fatores de risco, monitorar o impacto de intervenções e prever a evolução futura da epidemia Sofonea et al. (2022).

Os dados utilizados em modelos epidemiológicos podem incluir informações sobre casos confirmados, número de hospitalizações, óbitos, contatos de indivíduos infectados, entre outros. Esses dados são analisados estatisticamente para identificar padrões de transmissão e estimar valores para parâmetros, como a taxa de reprodução da doença. A análise de dados também pode ser usada para detectar surtos precoces, monitorar a eficácia das intervenções e prever a evolução futura da epidemia.

Além disso, a análise de dados pode ser combinada com técnicas de modelagem estatística para desenvolver modelos epidemiológicos mais sofisticados. Modelos estatísticos, como regressão, séries temporais e modelos de sobrevivência, permitem explorar as relações entre variáveis epidemiológicas e fornecer estimativas mais precisas dos parâmetros do modelo (MOLLISON et al., 1994).

A análise de dados também desempenha um papel importante na validação dos modelos epidemiológicos. Ao comparar as previsões do modelo com os dados observados, é possível avaliar a precisão do modelo e identificar possíveis limitações. A análise de dados permite aperfeiçoar e calibrar os modelos, ajustando seus parâmetros para melhor se adequarem aos dados reais. Essa retroalimentação entre modelos e dados é essencial para melhorar a compreensão e a previsão das epidemias.

Estudos recentes, como o trabalho de Lauer et al. (2020), têm demonstrado a importância da análise de dados na modelagem epidemiológica durante a pandemia de COVID-19. Esse estudo investigou o período de incubação da doença com base em casos confirmados publicamente e ressaltou a relevância da análise de dados para compreender a propagação do vírus. Da mesma forma, Salje et al. (2020) enfatizaram a necessidade de estimar a carga de SARS-CoV-2 na população, utilizando técnicas de análise de dados para avaliar a disseminação da doença. Essas abordagens destacam como a análise de dados desempenha um importante papel na tomada de decisões em saúde pública e no desenvolvimento de modelos epidemiológicos mais precisos.

2.4 TRABALHOS CORRELATOS

Diversos trabalhos recentes têm se dedicado à modelagem epidemiológica do vírus mpox por meio de modelos compartimentais, evidenciando o interesse em compreender e prever a propagação dessa doença infecciosa. Um levantamento abrangente desses estudos foi realizado nas bases de literatura acadêmica e científica, PubMed e Google Scholar. Estas plataformas foram escolhidas devido à sua extensa cobertura de publicações científicas nas áreas da saúde e biologia. Limitamos nossos resultados aos trabalhos publicados entre janeiro de 2020 e junho de 2023, uma vez que estávamos interessados em trabalhos que analisassem, predominantemente a dinâmica de transmissão entre humanos e preferencialmente estudos sobre o novo surto que ocorreu em 2022. Consideramos apenas os estudos que contivessem informações relevantes sobre a transmissão do vírus mpox entre seres humanos ou que utilizassem modelos compartimentais para analisar a dinâmica dessa transmissão na população humana.

Para proporcionar uma visão abrangente dessas pesquisas, apresentamos a Tabela 1, que resume os diversos estados da doença representados por compartimentos nos modelos utilizados por cada autor. Essa tabela oferece uma maneira eficiente de comparar sistematicamente os estudos e identificar os fatores relevantes na dinâmica de transmissão da doença, conforme discutido pelos respectivos autores. Esta abordagem torna mais acessível a compreensão dos trabalhos e a análise de suas contribuições para o campo.

No geral, apesar da maioria dos casos de mpox ocorrerem em países latino-americanos, como o Brasil, que atualmente é o segundo país em número de casos do surto de 2022 (Centers for Disease Control and Prevention, 2022), nenhum dos estudos encontrados nesta revisão sistemática concentrou-se em analisar especificamente a dinâmica da doença nessa região. A maioria dos estudos concentrou-se em analisar casos nos Estados Unidos, Reino Unido e outros países europeus. Além disso, apenas cinco dos estudos, que utilizaram modelos compartimentais, incluíram o comportamento da comunidade homens que fazem sexo com homens (HSH) nos modelos propostos (KNIGHT et al., 2022), (BRAGAZZI et al., 2023), (YANG et al., 2023), (BETTI et al., 2022), (XIRIDOU et al., 2023), o que é particularmente preocupante, já que a população HSH tem uma maior probabilidade de contrair a doença. Essa falta de análise de casos em regiões afetadas e de consideração da população HSH em muitos dos estudos limita a aplicabilidade dos resultados encontrados na revisão sistemática para outras regiões e populações.

Tabela 1 – Compartimentos utilizados nos modelos compartimentais do mpox encontrados na literatura

| Compartimentos | Definição |
|--|--|
| $S_i, E_i, P_i, I_i, R, Q, Q_v, Q_{si}, \bar{V}_2, \bar{V}_2, \underline{V}_1$ | Suscetível(S_i), Exposto(E_i), Infectado prodromal(P_i), Infectado agudo(I_i), Recuperado(R), Quarentenado(Q), Quarentenado e vacinado(Q_v), Quarentenado e Suscetível(Q_{si}), Grupo de Alto Risco (HRG) parcialmente vacinado(\bar{V}_2), HRG completamente vacinado(\bar{V}_2) e Grupo de Baixo Risco (LRG) completamente vacinado(\underline{V}_1). Com i sendo o número de populações, que neste estudo é representado por dois grupos, o LRG e HRG.(YUAN et al., 2022) |
| S, V, E, I, R | Suscetível(S), Vacinado(V), Exposto(E), Infectado(I), Recuperado(R). (AUGSBURGER et al., 2022) |
| S, I, R | Suscetível(S), Infectado(I), Recuperado(R). (LIAO et al., 2023) e (BETTI et al., 2022) |
| S, V, E, I, H, R | Suscetível(S), Vacinado(V), Exposto(E), Infectado(I), Isolado(H), Recuperado(R). (KNIGHT et al., 2022) e (XIRIDOU et al., 2023) |
| $S_h, E_h, I_h, Q_h, R_h, S_r, E_r, I_r$ | Humano, Suscetível(S_h), Exposto(E_h), Infectado(I_h), Quarentenado(H_h), Recuperado(R_h) e roedor, Suscetível(S_r), Exposto(E_r), Infectado(I_r). (NGUNGU et al., 2023) |
| S, I, R, D | Suscetível(S), Infectado(I), Recuperado(R), Morto(D). (AL-RAEEL, 2023) |
| S, E, I, R | Suscetível(S), Exposto(E), Infectado(I), Recuperado(R). (GAO et al., 2023) e (YANG et al., 2023) |
| S_i, E_i, I_i, Q_i, R_i | Suscetível(S_i), Exposto(E_i), Infectado(I_i), Quarentenado(Q_i), Recuperado(R_i). Com i sendo o número de populações, que neste estudo é representado por dois grupos, o de baixo risco (LRG) e alto risco (HRG). (BRAGAZZI et al., 2023) |
| $S_l, S_h, E_h, P, I_1, I_2, H, R_h, S_r, E_r, I_r, R_r$ | Humano, Suscetível de baixo risco(S_l), Suscetível de alto risco(S_l), Exposto(E_i), Infectado prodromal(P), Infectado leve(I_1), Infectado agudo(I_2), Hospitalizado(H), Recuperado(R). E animal, Suscetível(S_r), Exposto(E_r), Infectado(I_r), Recuperado(R_r)(AL-SHOMRANI et al., 2023) |
| $S_h, E_h, I_h, Q_h, H_h, R_h, V_h, S_r, I_r$ | Humano, Suscetível(S_h), Exposto(E_h), Infectado(I_h), Quarentenado(H_h), Hospitalizado(H_h), Recuperado(R_h) e Vacinado(V_h). E animal, Suscetível(S_r), Infectado(S_r). (MAJEE et al., 2023) |

Fonte: O autor

Com o levantamento, observamos que alguns dos estudos selecionados utilizaram seus modelos para prever possíveis novos surtos de mpox. Entre estes, destacam-se os estudos em (LIAO et al., 2023), (YANG et al., 2023), (XIRIDOU et al., 2023) e (AL-SHOMRANI et al., 2023). O estudo (YANG et al., 2023), em particular, realizou previsões para cinco países diferentes. As previsões são importantes ferramentas para o planejamento de estratégias de prevenção e controle de doenças infecciosas, permitindo a adoção de medidas preventivas antes que um surto ocorra.

Após a leitura completa dos artigos selecionados, observou-se uma limitação comum nos modelos compartimentais utilizados para modelar a dinâmica do mpox. Esta limitação refere-se às fontes de dados utilizadas para estimar os parâmetros do modelo. A maioria dos autores assumiu valores arbitrários ou baseou-se em dados disponíveis na literatura para estimar esses parâmetros. Devido à falta de dados específicos sobre o mpox, essa abordagem pode levar à imprecisão nos resultados obtidos. Essa limitação foi observada por autores em (KOROSEC et al., 2023), que destacaram a importância da obtenção de dados específicos sobre a doença para melhorar a precisão dos modelos.

Outra limitação observada nos estudos selecionados é a disponibilidade limitada de dados sobre a doença. Em (BRAGAZZI et al., 2023), por exemplo, os dados coletados estavam restritos a uma janela de apenas 7 dias. Essa falta de dados específicos sobre o mpox torna difícil a previsão da dinâmica da doença e a avaliação da eficácia de medidas preventivas, como a vacinação. A obtenção de dados mais específicos sobre o mpox é fundamental para melhorar a precisão e a confiabilidade dos modelos epidemiológicos utilizados para prever a propagação da doença e avaliar medidas preventivas.

Por fim nenhum estudo considerou a ocorrência ondas epidêmicas em diferentes regiões em redes de contato distintas que ocorrem em momentos diferentes em um determinado país. Embora muitos dos estudos apresentados nesta revisão tenham considerado a transmissão da doença entre indivíduos dentro da mesma rede de contato, poucos consideraram a possibilidade de diferentes regiões dentro de um mesmo país estarem em diferentes estágios de epidemia. Essa limitação pode ter um impacto significativo nas previsões e análises dos modelos, uma vez que a dinâmica da transmissão da doença pode variar de acordo com a região e o momento da epidemia. O destaque fundamental do trabalho apresentado nesta monografia reside no contexto específico ao qual é aplicado. Aqui, propomos um modelo compartimental inovador, elaborado para representar a transmissão do mpox em países latino-americanos, utilizando dados de casos reais como base. O diferencial deste modelo é a incorporação de uma estrutura de rede de contatos, que é capaz de capturar os impactos de epidemias que ocorrem em momentos distintos em um ou mais grupos populacionais de uma região específica. Além disso, o modelo segmenta a população em diferentes grupos de risco, distinguindo entre a população de baixo risco e a população de alto risco (HSH). Essa abordagem estratificada permite uma análise mais aprofundada e precisa da dinâmica da transmissão do mpox, considerando não apenas a

variabilidade nos grupos populacionais, mas também os padrões de contato entre eles.

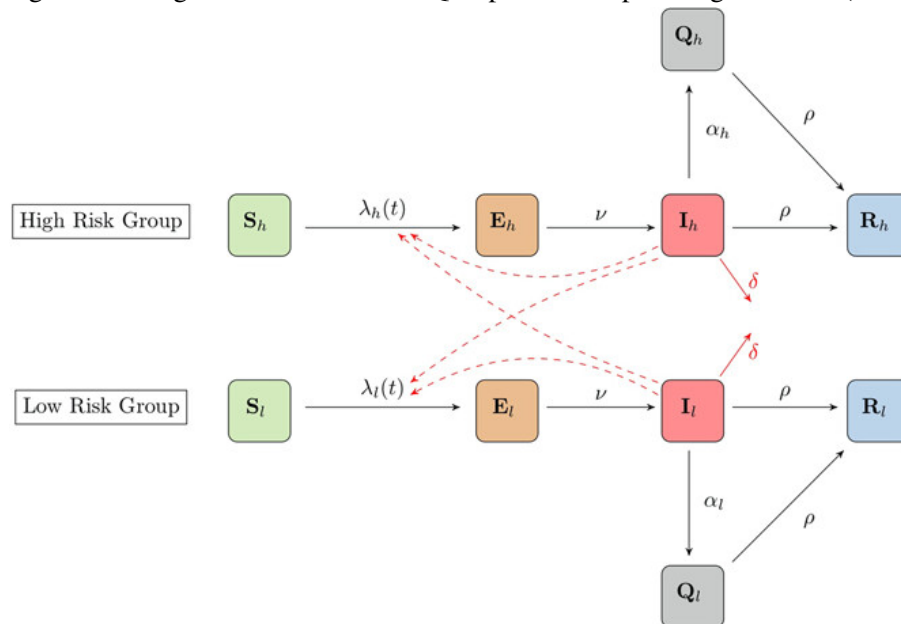
A seguir, nesta seção, serão discutidos com maior ênfase alguns dos trabalhos que se aproximam dos objetivos almejados pelo autor desta pesquisa.

2.4.1 Bragazzi - Modelo com estratificação por comportamento sexual

Em (BRAGAZZI et al., 2023) foi desenvolvido um modelo SEIQR para estudar a dinâmica de transmissão do mpox do surto de 2022, se concentrando em analisar os registros de casos, de um intervalo de 7 dias, de episódios da doença no Canadá.

O principal diferencial deste trabalho é a definição de diferentes grupos populacionais, estratificados pelo comportamento sexual, como definido no diagrama da figura 3, o sistema SEIQR é duplicado, para a população de alto risco e de baixo risco que interagem entre si e podem causar a transmissão de doenças entre os grupos.

Figura 3 – Diagrama do modelo SEIQR apresentado por Bragazzi et al. (2023)



Fonte: Retirada de Bragazzi et al. (2023)

A infecção dos indivíduos é regida por taxas de contato sexual e não sexual. Os contatos sexuais são considerados apenas no grupo de alto risco, pois acredita-se que a maioria dos casos de varíola símia nesse grupo ocorre por contato sexual, e a transmissão por contato sexual no grupo de baixo risco é insignificante.

Os dados coletados foram utilizados para definir algumas características do modelo,

como o tamanho da população do grupo de baixo risco e alto risco. Entretanto, os autores não utilizaram técnicas como regressão linear, para se obter valores mais precisos para parâmetros, como a taxa de infecção nos grupos.

A solução de modelagem epidemiológica apresentada nesta monografia incorpora o conceito de grupos populacionais distintos, notadamente de alto e baixo risco, conforme proposto pelo autor Bragazzi et al. (2023). Consideramos essa abordagem como um fator relevante para a compreensão do processo de transmissibilidade do vírus mpox. Entretanto, nosso trabalho vai além, estendendo as fronteiras das limitações identificadas no estudo do autor. Passamos a considerar não apenas surtos de uma única onda, mas também a possibilidade de ocorrência de múltiplas ondas em diferentes populações ou redes. Adicionalmente, introduzimos outros compartimentos no modelo, tais como o de vacinados e indivíduos em abstinência, como minuciosamente abordado no capítulo 4. Essas expansões e refinamentos contribuem para um entendimento mais completo e abrangente da dinâmica do vírus mpox.

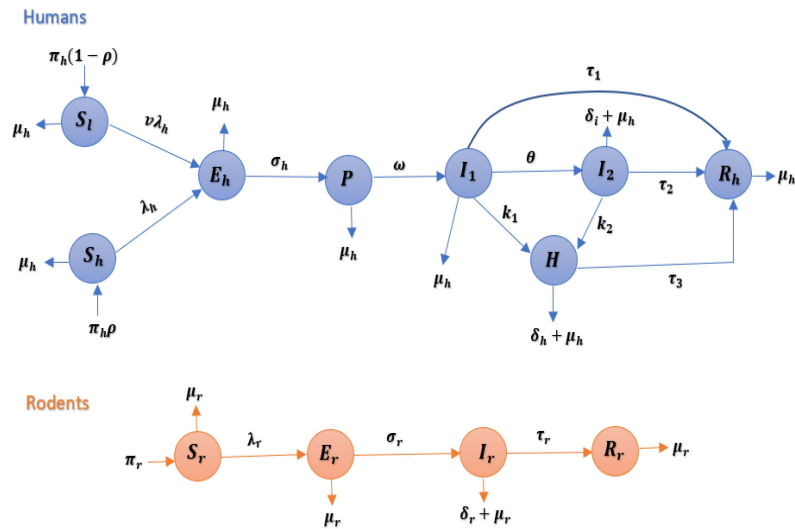
2.4.2 Al-Shomrani - Modelagem dos estágios do mpox incluindo população animal

O trabalho desenvolvido por Al-Shomrani et al. (2023) define diversos compartimentos que representam os diferentes estágios do mpox, o modelo abrange indivíduos na fase prodromal e com grau de infecção leve e agudo. Assim como o trabalho citado anteriormente ((BRAGAZZI et al., 2023)), também são considerados os grupos de baixo e alto risco, mas neste caso eles são agrupados em compartimentos de indivíduos suscetíveis, sendo S_{ho} grupo de indivíduos suscetíveis de alto risco e S_l a população suscetível de baixo risco, que se tornam expostos à uma taxa diferente como demonstrado no diagrama 4 abaixo.

O modelo proposto por autores também inclui o reservatório animal da doença, construindo um outro sistema de compartimentos para representar a transmissão da doença entre a população de roedores e a infecção cruzada entre roedores e humanos.

Por fim os autores também realizaram uma análise preditiva para validação do modelo, utilizando dados do mpox obtidos pelo Centros de Controle e Prevenção de Doenças da Nigéria (NCDC), obtendo bons resultados para os surtos da Nigéria.

Figura 4 – Modelo compartimental desenvolvido por Al-Shomrani et al. (2023)



Fonte: Retirada de Al-Shomrani et al. (2023)

2.4.3 Gosh - Um modelo SIR para epidemias de múltiplas ondas

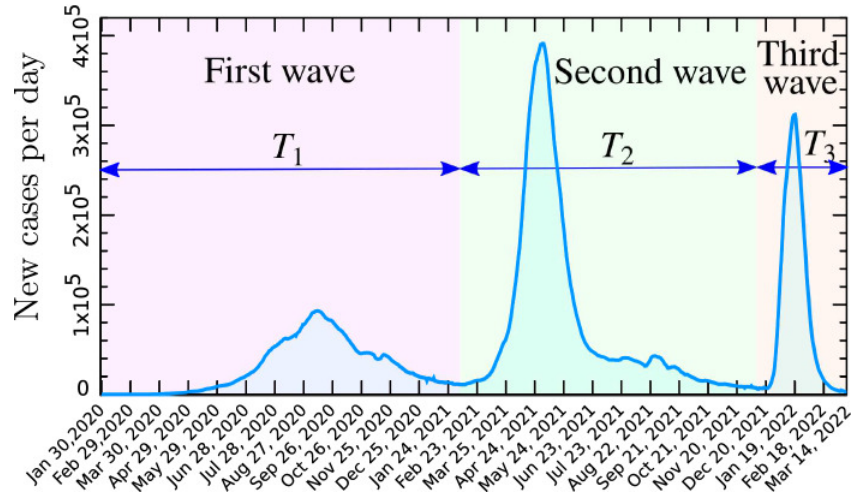
No estudo realizado por Ghosh e Ghosh (2022), um modelo compartimental foi desenvolvido com o objetivo de descrever o surto de COVID-19 que ocorreu na Índia em 2019. Diferentemente de outros estudos que se concentraram na extensão do modelo compartimental SIR padrão, os autores procuraram adicionar ao modelo a capacidade de representar múltiplas ondas de contaminação, incluindo a possibilidade de reinfeccção dos indivíduos.

Embora o vírus da varíola mpox não compartilhe a mesma característica epidemiológica do COVID-19, que permite a reinfeccção dos indivíduos, a estratégia adotada por Ghosh e Ghosh (2022) para definir os períodos de transição entre as ondas, através da análise minuciosa dos dados do surto, se mostra relevante para o escopo desta monografia. Estamos particularmente interessados em explorar a possibilidade de ocorrência de múltiplas ondas durante o surto de mpox de 2022, nos países americanos. Uma distinção clara entre o trabalho apresentado nesta monografia e o estudo conduzido por Ghosh e Ghosh (2022) reside no modelo utilizado para representar as epidemias de múltiplas ondas. Enquanto o modelo proposto por (GHOSH; GHOSH, 2022) se baseia na análise de reinfeccções, nossa abordagem propõe uma solução fundamentada em redes de populações interconectadas, como detalhado na subseção 4.1.1.

A figura 5 apresentada abaixo ilustra a curva epidemiológica da doença, utilizando os dados de casos diários. Essa análise permitiu identificar os pontos de transição da doença, correspondentes aos períodos de reinfeccção. Com base nessas informações, os autores ajustaram

os parâmetros do modelo durante a simulação, alterando a taxa de reinfeção dos indivíduos. Esse ajuste resultou em um aumento periódico da população infectada, algo que não seria possível considerando o modelo SIR padrão descrito na seção 2.2.2.

Figura 5 – Identificação dos pontos de transição de ondas no surto de Covid-19 na Índia



Fonte: Retirada de Ghosh e Ghosh (2022)

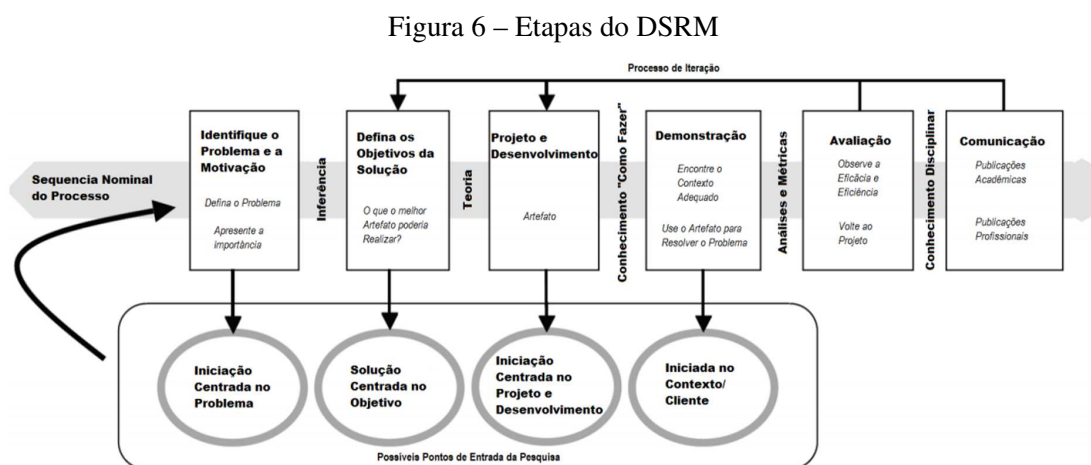
Essa abordagem, baseada na análise dos dados e na adaptação do modelo, proporciona uma compreensão mais precisa da dinâmica da epidemia e pode ser aplicada para estudar diferentes cenários epidemiológicos, inclusive aqueles que não envolvem a reinfeção de indivíduos.

3 METODOLOGIA

A execução deste trabalho segue a abordagem da *Design Science Research* (DSR), utilizando as diretrizes da *Design Science Research Methodology* (DSRM) propostas por (PEFFERS et al., 2007). Essa metodologia baseia-se em abordagens práticas para solução de problemas e criação de artefatos como resultado da pesquisa. A DSR é amplamente aplicada em estudos de ciência da computação e sistemas de informação, buscando criar soluções inovadoras por meio do desenvolvimento de artefatos que resolvam problemas reais.

No contexto deste trabalho, um artefato é definido como um produto ou construção concreta que visa solucionar um problema prático. Isso pode incluir softwares, modelos teóricos, frameworks ou qualquer elemento tangível desenvolvido durante o processo de pesquisa (HEVNER et al., 2010). A DSRM, por sua vez, oferece uma estrutura metodológica para a produção e apresentação de pesquisas em Design Science em Sistemas de Informação (PEFFERS et al., 2007).

A aplicação da DSRM neste projeto possibilitará a criação de artefatos voltados para a compreensão e enfrentamento da epidemia de mpox na América Latina durante o surto de 2022. Esses artefatos podem englobar modelos matemáticos, algoritmos, sistemas de informação ou outras soluções práticas que contribuam para a prevenção, controle e gestão da doença.



Fonte: traduzida de Peffers et al. (2007)

Conforme apresentado na figura 6, o DSRM é composta por um conjunto de seis

processos iterativos e interligados, que orientam o pesquisador em cada etapa do projeto:

O primeiro processo da DSRM é a identificação do problema e oportunidade de pesquisa. Nessa etapa, o pesquisador analisa a situação atual, identifica a lacuna de conhecimento e define o problema de pesquisa a ser abordado. Neste sentido, a primeira etapa da metodologia será dedicada à formulação das bases teórico-conceituais do método e do modelo matemático/algóritmico. Será realizado um estudo aprofundado e sistemático sobre a modelagem de epidemias, incluindo a revisão de literatura de trabalhos relacionados e o levantamento das características da transmissão do mpox. A partir dessa revisão, serão estabelecidos os fundamentos teóricos necessários para a construção do modelo e a definição dos compartimentos e parâmetros relevantes.

Após a identificação do problema, define-se os objetivos do projeto. Nessa etapa, o pesquisador estabelece os objetivos gerais e específicos que o trabalho pretende alcançar. No caso deste projeto, os objetivos estão relacionados à formulação e desenvolvimento de um modelo matemático descrito por equações diferenciais, bem como à criação de um modelo computacional capaz de reproduzir os episódios epidêmicos da doença na América Latina.

O terceiro processo é o design e desenvolvimento do artefato. Essa etapa envolve a concepção, criação e implementação do artefato de design, que no caso deste projeto é o modelo matemático e o modelo computacional. O pesquisador utiliza seus conhecimentos teóricos, dados coletados e ferramentas de desenvolvimento para criar soluções práticas e inovadoras que atendam aos objetivos estabelecidos.

O processo de construção do artefato envolve a definição do modelo matemático e sua representação computacional através de algoritmos capazes de representar os diferentes compartimentos definidos pelas equações ordinárias do modelo, bem como a dinâmica da relação entre os indivíduos por meio de uma estrutura de rede de contatos.

Em paralelo, serão processados dos dados da epidemia da mpox para a validação e treinamento do modelo. Os dados epidemiológicos, como casos confirmados, localização geográfica e dados demográficos, serão fornecidos pela OPAS. Serão aplicadas técnicas de análise exploratória de dados para identificar padrões, tendências e características da epidemia. Além disso, serão realizadas análises estatísticas para verificar a consistência e confiabilidade dos dados.

Ainda na etapa de desenvolvimento serão realizadas simulações utilizando o modelo desenvolvido e os dados processados da mpox. Será empregada a técnica de *Grid Search* com gradiente para ajustar os parâmetros do modelo aos dados reais de casos confirmados. O *Grid Search* refere-se a uma técnica usada para encontrar os melhores conjuntos de parâmetros para um modelo, ajustando diferentes combinações de valores de parâmetros em uma grade (ou grade, em português). Esta técnica é frequentemente utilizada em otimização de hiperparâmetros, onde o objetivo é encontrar a melhor configuração de parâmetros para um modelo de aprendizado de máquina ou, no caso de modelos epidemiológicos, os melhores parâmetros para descrever a propagação de uma doença, conforme detalhado na subsecção 3.1.3.

O quarto processo da DSRM é a demonstração do artefato. Nessa etapa, o pesquisador apresenta e valida o artefato desenvolvido, buscando evidências de sua eficácia e relevância. No contexto deste trabalho, a demonstração do artefato ocorrerá por meio da reprodução de episódios epidêmicos em países selecionados da América Latina como casos de estudo para avaliar a capacidade do modelo em reproduzir episódios epidêmicos nessas regiões. Além da análise de indicadores epidemiológicos, como taxa de Número de Reprodução Básico (R_0), taxa de mortalidade e duração da epidemia, para comparar as previsões do modelo com os dados reais observados. E análise estatística para avaliar a precisão e confiabilidade do modelo em diferentes contextos. Além disso, a capacidade preditiva do modelo em relação ao pico e ao fim de uma onda epidêmica será determinada. Serão comparadas as previsões do modelo com os dados reais de pico e fim da epidemia, utilizando técnicas estatísticas para avaliar a precisão das previsões.

O quinto processo é a avaliação do artefato. Nessa etapa, o pesquisador analisa os resultados obtidos, avalia a eficácia e a usabilidade do artefato e identifica possíveis melhorias. A avaliação do projeto será feita por meio da comparação dos resultados obtidos pelo modelo com dados reais da epidemia da mpox. A consistência e confiabilidade do modelo serão avaliadas por meio de análises estatísticas, que permitirão verificar a aderência do modelo aos dados reais e a capacidade de reprodução dos episódios epidêmicos. A validação também será feita por meio da comparação das previsões do modelo com os dados reais de pico, fim e detecção de novas ondas epidêmicas.

Por fim deve ser realizada a comunicação dos resultados. Nessa etapa, o pesquisador documenta e apresenta os resultados obtidos, por meio de relatórios, artigos científicos, apresentações em conferências e outras formas de divulgação. A comunicação dos resultados é

fundamental para compartilhar conhecimentos, validar o trabalho realizado e contribuir para o avanço da área de pesquisa.

Finalmente, a capacidade de detecção de uma nova onda epidêmica superposta será avaliada. Serão utilizados dados simulados ou reais para testar a capacidade do modelo em identificar e representar uma nova onda epidêmica que ocorra durante ou após a epidemia atual.

No desenvolvimento do trabalho, serão utilizadas tecnologias como linguagens de programação para implementação do modelo, ferramentas de visualização de dados para análise exploratória e software estatístico para a realização das análises estatísticas.

Dessa forma, a metodologia proposta permitirá formular, desenvolver e validar um modelo de equações diferenciais baseado em redes de contato para a modelagem da mpox. A abordagem DSR será aplicada ao longo de todo o processo, integrando a criação do modelo com a sua validação em cenários reais, por meio de técnicas estatísticas e utilização de dados fornecidos pela OPAS.

3.1 MATERIAIS E MÉTODOS

Neste trabalho, a modelagem epidêmica é apoiada pelo uso da linguagem de programação Python e suas bibliotecas especializadas, como NumPy, Pandas e Matplotlib.

Python é uma linguagem de programação de alto nível e de código aberto amplamente utilizada na comunidade científica devido à sua facilidade de uso, versatilidade e vasta gama de bibliotecas disponíveis (ROSSUM; DRAKE, 2009). No âmbito da modelagem epidêmica, Python oferece uma sintaxe clara e intuitiva, facilitando o desenvolvimento e implementação de algoritmos para simulação de modelos epidemiológicos. Sua flexibilidade permite a criação de estruturas de dados personalizadas e a implementação de métodos específicos para atender às necessidades de cada modelo em particular. O Python é usado neste projeto como a principal linguagem de programação para desenvolver e executar os algoritmos de simulação do modelo epidêmico (ANDERSON; MAY, 1991).

Além do Python, uma biblioteca fundamental para a modelagem epidêmica é o NumPy. O NumPy é uma biblioteca em Python que fornece suporte para cálculos numéricos eficientes e manipulação de arrays multidimensionais (HARRIS et al., 2020). Ele é amplamente utilizado na área da epidemiologia para realizar operações matemáticas e estatísticas

essenciais, como cálculo de médias, desvios padrão e geração de amostras aleatórias. Com suas funções otimizadas e estruturas de dados eficientes, o NumPy permite uma manipulação e processamento rápidos dos dados utilizados na simulação do modelo epidêmico (WORMSER; POURBOHLOUL, 2008).

Para a manipulação e análise de dados tabulares, a biblioteca Pandas desempenha um papel fundamental. O Pandas (TEAM, 2020) oferece estruturas de dados flexíveis, como o DataFrame, que permitem a organização eficiente de dados epidemiológicos em formato tabular. Com suas funções avançadas de manipulação e limpeza de dados, o Pandas facilita a extração de informações relevantes para a modelagem epidêmica. No presente estudo, o Pandas é utilizado para coletar e tratar os dados epidemiológicos disponibilizados pela OPAS (Organização Pan-Americana da Saúde), permitindo uma análise mais detalhada e precisa dos padrões de propagação da doença.

Além disso, a biblioteca Matplotlib (HUNTER, 2007) é amplamente utilizada para a visualização de dados em Python. Com suas diversas funcionalidades gráficas, o Matplotlib permite a criação de gráficos, plots e outras informações visuais que facilitam a análise e comunicação dos resultados obtidos na modelagem epidemiológica. O Matplotlib é empregado neste projeto para gerar gráficos informativos que fornecem uma representação visual clara e compreensível dos padrões de disseminação da doença e outros aspectos relevantes. Essa visualização é essencial para auxiliar na interpretação dos resultados e na tomada de decisões embasadas em evidências.

Em resumo, o Python e suas bibliotecas, como NumPy, Pandas e Matplotlib, desempenham um papel fundamental na modelagem epidêmica. O Python fornece uma linguagem de programação flexível e poderosa para a implementação dos algoritmos de simulação, enquanto o NumPy permite cálculos numéricos eficientes. O Pandas facilita a manipulação e análise de dados epidemiológicos, enquanto o Matplotlib é essencial para a visualização dos resultados. A utilização dessas ferramentas contribui para a obtenção de resultados precisos e uma compreensão aprofundada dos padrões de disseminação de doenças.

3.1.1 Mineração de dados

A mineração de dados é uma técnica essencial na análise de grandes conjuntos de dados, visando extrair informações relevantes, identificar padrões ocultos e construir modelos

que possam fornecer informações significativas. No contexto deste trabalho, a mineração de dados é aplicada na definição dos parâmetros epidemiológicos a serem incorporados no modelo compartimental da dinâmica de transmissão do mpox. Utilizando dados fornecidos pela OPAS, essa abordagem permite identificar padrões complexos nos dados epidemiológicos, desvendando relações e correlações que podem passar despercebidas em análises convencionais.

Um dos objetivos dos modelos epidemiológicos é medir a propagação de doenças infecciosas, definindo os fatores que a fazem transitar da endemia para a epidemia e, finalmente, para a pandemia, lembrando que a tarefa de caracterizar a dinâmica de um vírus geralmente se torna complexo ao decifrar o conhecimento com fins preditivos. Neste sentido, a mineração de dados desempenha um papel importante na descoberta e extração de padrões desconhecidos, sendo este um aspecto fundamental na criação de modelos que definam a mecânica da propagação de uma doença, o que implica fazer uso de algoritmos que têm raízes no aprendizado de máquina. (CORTÉS-MARTÍNEZ et al., 2022)

Através da mineração de dados, nosso objetivo é realizar uma análise profunda desses dados epidemiológicos, identificando tendências, correlações e comportamentos que possam ser cruciais para uma representação precisa da dinâmica de transmissão do mpox. Utilizamos métodos estatísticos avançados para explorar os dados, compreender as variáveis relevantes e quantificar os fatores que influenciam a propagação da doença.

Além disso, a mineração de dados também desempenha um papel fundamental na seleção e definição dos parâmetros a serem incorporados em nosso modelo compartimental. Ao identificar os padrões emergentes nos dados epidemiológicos, conseguimos ajustar os parâmetros do modelo de forma mais precisa, o que, por sua vez, aumenta a confiabilidade das projeções da dinâmica de transmissão do mpox.

3.1.2 Suavização dos dados

A suavização de dados é uma técnica amplamente utilizada na análise de séries temporais, desempenhando um papel fundamental na extração de tendências e na redução de flutuações de curto prazo em um conjunto de dados. Quando se lida com dados temporais, muitas vezes, a presença de variações irregulares de curto prazo pode dificultar a identificação de padrões de longo prazo, tornando a suavização uma ferramenta valiosa. Uma técnica comum para alcançar esse objetivo é a suavização por média móvel, que calcula médias ponderadas ao longo do tempo (NORTH, 2003).

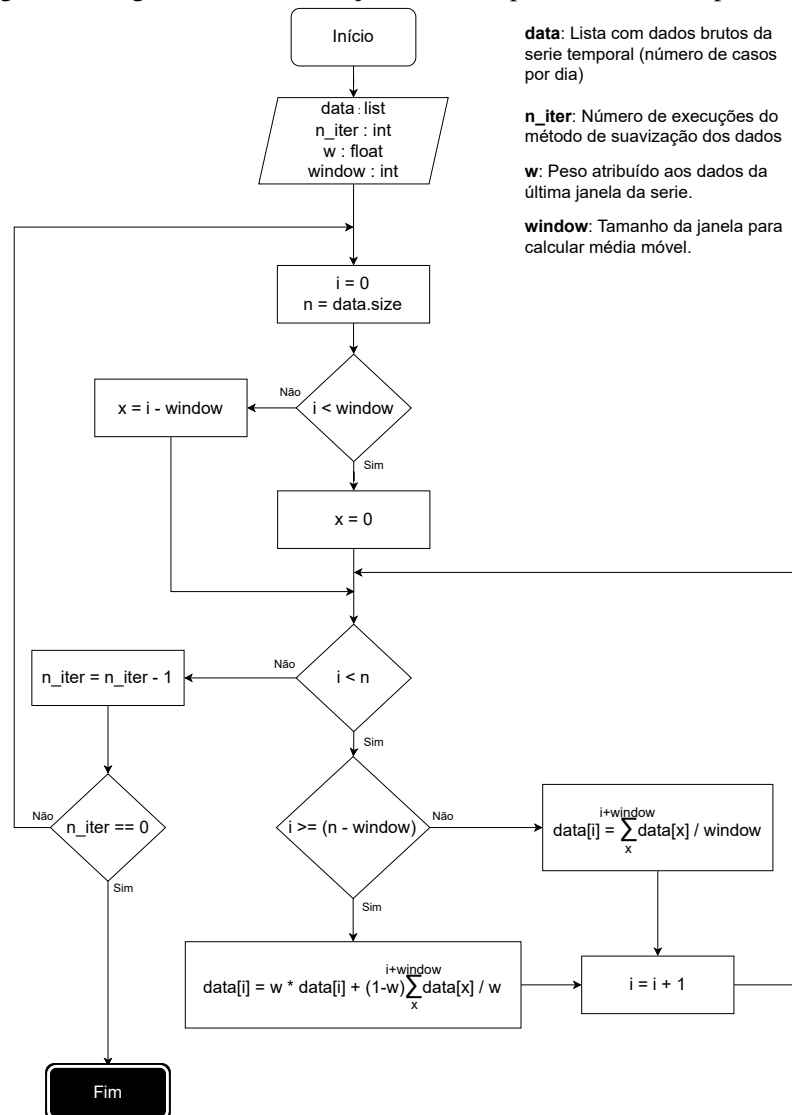
A suavização de dados, sob a perspectiva da modelagem epidemiológica, auxilia na mitigação de erros e na obtenção de informações mais claras a partir dos dados disponíveis. Muitas vezes, os registros epidemiológicos estão sujeitos a irregularidades, como atualizações tardias de casos, relatórios em datas variáveis, e sub-notificação de casos. Essas variações podem dificultar a interpretação dos dados e a detecção de tendências reais. Portanto, é essencial aplicar técnicas de suavização para tornar os dados mais úteis e confiáveis na modelagem epidemiológica (DEMONGEOT et al., 2020).

Para a suavização dos dados neste estudo, foi empregada uma técnica de suavização por média móvel, a partir de um novo método de suavização, desenvolvido pelo grupo de pesquisa G2BC da Universidade do Estado da Bahia (UNEB), demonstrado no fluxograma da figura 7, com uma abordagem que incorpora uma janela simétrica de tamanho $2w + 1$. A utilização de uma janela simétrica é uma escolha estratégica, pois traz vários benefícios para a análise de séries temporais. Primeiramente, a janela simétrica ajuda a preservar a simetria nos cálculos, considerando uma quantidade igual de pontos de dados à esquerda e à direita de cada ponto. Essa simetria é valiosa para manter padrões e tendências no conjunto de dados, minimizando o viés em qualquer direção específica.

A utilização de uma janela simétrica também permite a preservação da informação nos extremos da série temporal. Ao final da série, o método utiliza uma média ponderada que considera tanto o ponto de dados mais recente quanto a média dos pontos anteriores. Isso é particularmente útil para destacar os dados mais recentes, que podem ser mais relevantes para previsões futuras e auxiliar na identificação de mudanças recentes nas tendências. A média ponderada no final da série oferece uma abordagem equilibrada, levando em consideração tanto os dados mais recentes quanto a tendência histórica, proporcionando uma visão mais completa da evolução dos dados epidemiológicos.

Essa abordagem de suavização por média móvel com múltiplas iterações e média ponderada na última janela foi escolhida devido à sua capacidade de reduzir flutuações indesejadas, fornecendo uma representação mais estável dos dados. A técnica permite destacar tendências de longo prazo, reduzindo o impacto de dados ruidosos e flutuações de curto prazo.

Figura 7 – Algoritmo de suavização de dados por média móvel ponderada



Fonte: O autor

3.1.3 Busca em grade

A Busca em Grade (*Grid Search*, em inglês) representa uma ferramenta fundamental na ciência de dados e na modelagem estatística, sendo empregada para determinar os valores ótimos de parâmetros de um modelo matemático ou estatístico. Estes parâmetros, muitas vezes, não podem ser diretamente derivados dos dados disponíveis. O procedimento consiste em criar uma grade de valores possíveis para cada parâmetro e, subsequentemente, avaliar o desempenho do modelo para todas as combinações destes valores. O objetivo primordial é identificar a combinação que resulta no melhor ajuste do modelo aos dados observados ou na mais alta performance, medida por uma métrica específica. (BERGSTRA; BENGIO, 2012)

Além deste estudo, a aplicação da técnica de *Grid Search* com gradiente para

otimização de parâmetros epidemiológicos tem sido adotada em diversas pesquisas. Um exemplo notável é o trabalho realizado por Stochițoiu et al. (2020), que conduziu uma comparação entre o *Grid Search* com gradiente e uma abordagem baseada em redes neurais para otimizar os parâmetros epidemiológicos de um modelo compartimental desenvolvido para estimar os casos de Covid-19 na Romênia.

No âmbito deste estudo, a aplicação do *Grid Search* se mostra essencial para determinar os valores ideais dos parâmetros do modelo, dados não diretamente acessíveis através dos dados fornecidos pela OPAS (Organização Pan-Americana da Saúde). Um exemplo prático disso é apresentado no trabalho de Canto e Avila-Vales (2020), onde o *Grid Search* foi utilizado para encontrar a taxa de exposição dos indivíduos ao vírus da COVID-19, essencial para o modelo SEIR proposto.

Uma implementação amplamente utilizada desta técnica está disponível na biblioteca de aprendizado de máquina para Python, *Scikit-Learn* (PEDREGOSA et al., 2011). O método *GridSearchCV* encapsula um algoritmo de busca em grade com validação cruzada, facilitando a busca pelos melhores valores dos hiperparâmetros de um modelo (Scikit-Learn Contributors, 2023). Vale ressaltar que a versão adotada neste estudo é baseada em gradiente e foi fornecida pelo grupo de pesquisa G2BC da Universidade do Estado da Bahia (UNEB). O *Grid Search* baseado em gradiente não testa todas as combinações de hiperparâmetros; em vez disso, utiliza o gradiente para determinar a direção na qual os hiperparâmetros devem ser ajustados, economizando tempo computacional, especialmente quando o espaço de busca é vasto. Ao combinar *Grid Search* com validação cruzada, o *Grid Search CV* assegura uma busca robusta e otimizada em termos de tempo computacional, orientada pelo gradiente para explorar eficazmente o espaço dos hiperparâmetros.

Essa abordagem sistemática não apenas aprimora a qualidade e validade dos resultados obtidos neste estudo, mas também confere robustez e aplicabilidade aos contextos específicos relacionados à transmissão do vírus mpox em países latino-americanos. Ao utilizar técnicas sofisticadas como o *Grid Search* baseado em gradiente, este estudo está na vanguarda da pesquisa, garantindo não apenas resultados precisos, mas também uma metodologia avançada e eficiente.

3.1.4 Avaliação do modelo

Nesta seção, será discutida a função de erro utilizada para comparar os casos reais de *monkey pox* (mpox) com os casos previstos pelo modelo compartimental desenvolvido neste trabalho. A função de erro desempenha um papel fundamental na avaliação da precisão do modelo, fornecendo informações sobre quão bem o modelo está se ajustando aos dados reais.

3.1.4.1 Função de Erro Utilizada

Para comparar os dados reais de mpox com as previsões do modelo, desenvolvemos neste estudo uma função de **erro composta**, que leva em consideração tanto a forma das curvas quanto a distância entre elas. A função de erro é composta por duas métricas principais: a métrica de perda de forma ('ShLoss') e a métrica de perda de distância ('DtLoss').

3.1.4.2 Justificativa da Função de Erro

A função de erro proposta fornece uma avaliação abrangente da precisão do modelo, considerando tanto a forma quanto a distância entre as curvas do modelo e os dados reais de monkey pox (mpox). A métrica de perda de forma ('ShLoss') assegura que o modelo seja avaliado em relação à sua habilidade de reproduzir a forma geral da epidemia. Ao mesmo tempo, a métrica de perda de distância ('DtLoss') avalia a precisão das previsões diárias, oferecendo uma compreensão detalhada das discrepâncias entre as previsões do modelo e os dados reais.

Buscamos valores próximos de zero para as métricas de perda ('ShLoss' e 'DtLoss'). Um valor próximo de zero para 'ShLoss' indica que o modelo está reproduzindo fielmente a forma da curva epidêmica, enquanto um valor próximo de zero para 'DtLoss' indica que as previsões diárias do modelo estão se aproximando dos números reais de casos. Em outras palavras, quanto mais próximos de zero esses valores, mais preciso é o modelo em capturar tanto a forma geral quanto a quantidade exata de casos diários de mpox.

Além disso, a análise separada de 'ShLoss' e 'DtLoss' torna perceptível diferentes aspectos do desempenho do modelo. 'ShLoss' ajuda a identificar se o modelo está capturando corretamente a tendência geral da epidemia ao longo do tempo. Por outro lado, 'DtLoss' indica se o modelo está prevendo com precisão o número exato de casos diários. Ao analisar essas métricas individualmente, é possível identificar áreas específicas em que o modelo pode ser

aprimorado.

Por exemplo, se ‘ShLoss’ for baixo, mas ‘DtLoss’ for alto, pode indicar que o modelo está reproduzindo bem a forma geral da curva, mas há discrepâncias significativas nas previsões diárias. Nesse caso, é importante investigar as razões por trás dessas discrepâncias diárias, como variações sazonais ou eventos específicos que o modelo pode não estar capturando adequadamente.

Por outro lado, se ‘DtLoss’ for baixo, mas ‘ShLoss’ for alto, pode indicar que o modelo está prevendo bem o número de casos diários, mas não está reproduzindo com precisão a tendência geral da epidemia. Nesse cenário, é crucial revisar a estrutura do modelo ou os parâmetros utilizados para garantir uma representação mais precisa da forma da curva epidêmica ao longo do tempo.

Assim, a análise separada de ‘ShLoss’ e ‘DtLoss’ permite ajustes específicos para melhorar tanto a forma quanto a precisão das previsões. A combinação dessas análises oferece uma avaliação abrangente e fundamentada sobre a precisão do modelo de compartimentação em relação aos dados reais de monkey pox, contribuindo para aprimoramentos contínuos e conclusões confiáveis deste estudo.

3.1.4.3 Normalização dos Dados

Antes de detalhar as métricas de perda, é importante compreender o processo de normalização dos dados utilizados neste estudo. A normalização é uma etapa fundamental que coloca os dados em uma escala padronizada, permitindo uma comparação justa e significativa entre diferentes variáveis. No contexto deste estudo sobre a propagação da doença de mpox, a normalização é aplicada para garantir que os valores de casos diários estejam em uma faixa consistente e adimensional.

Os dados de casos diários de mpox são normalizados dividindo cada valor pelo valor máximo de casos diários observado durante o período de estudo. A normalização é feita da seguinte forma para cada valor diário X_i :

$$X_{i,\text{normalizado}} = \frac{X_i}{\text{Máximo de Casos Diários Observados}}$$

Essa abordagem de normalização pelo máximo garante que todos os valores estejam na faixa de 0 a 1. Normalizar os dados garante que a magnitude dos valores não influencie indevidamente a comparação entre os dados reais e as previsões do modelo. A normalização dos dados é um procedimento padrão para análises comparativas em epidemiologia e fornece uma base sólida para as análises subsequentes das métricas de perda apresentadas neste estudo.

3.1.4.4 Métrica de Perda de Forma ('ShLoss')

A métrica de perda de forma avalia a semelhança de forma entre as curvas suavizadas normalizadas do modelo e dos dados reais de mpox. Esta métrica é calculada como a soma das diferenças absolutas ponderadas entre as curvas normalizadas. A fórmula para 'ShLoss' é:

$$ShLoss = \frac{\sum_{i=1}^n |w_i \times (nmd_i - ngd_i)|}{\sum_{i=1}^n ngd_i}$$

Onde:

- nmd_i é o valor normalizado do modelo no dia i ,
- ngd_i é o valor normalizado dos dados reais suavizados no dia i ,
- w_i é o peso associado ao dia i ,
- n é o número total de dias com dados.

O peso (w_i) associado a cada dia (i) é calculado com base nas diferenças entre os dias com dados (gtD). Esses pesos representam a importância relativa de cada ponto de dados na avaliação da forma da curva. O objetivo do uso de pesos é garantir que os dias com menos dados disponíveis tenham uma contribuição proporcionalmente menor para a métrica de perda de forma. A inclusão dos pesos permite que o modelo seja mais sensível às áreas da curva onde há mais dados disponíveis, aumentando a precisão da análise.

$$w_i = \frac{gtD_i}{2.0} \text{ para } i = 1$$

$$w_i = \frac{(gtD_{i-1} + gtD_i)}{2.0} \text{ para } i > 1$$

3.1.4.5 Métrica de Perda de Distância ('DtLoss')

A métrica de perda de distância mede a distância absoluta entre as curvas do modelo e dos dados reais, sem normalização. Ela é calculada da mesma forma que 'ShLoss', mas usando as curvas não normalizadas. A fórmula para 'DtLoss' é a mesma que a de 'ShLoss', mas substituindo nmd_i e ngt_i pelos valores não normalizados.

$$DtLoss = \frac{\sum_{i=1}^n |w_i \times (mY_i - gtY_i)|}{\sum_{i=1}^n gtY_i}$$

Onde:

- mY_i é o valor do modelo no dia i ,
- gtY_i é o valor dos dados reais no dia i ,
- w_i é o peso associado ao dia i ,
- n é o número total de dias com dados.

Cada diferença é ponderada pelo peso w_i associado a cada dia, o que permite avaliar a precisão das previsões diárias de maneira proporcional à disponibilidade de dados em cada ponto da curva.

3.1.4.6 Combinação Ponderada de Métricas

Para criar uma função de erro composta, é calculado um fator de ponderação (W) como a diferença da razão entre 'ShLoss' e 'DtLoss', e 0,5, ajustado para manter W entre 0 e 1. Este fator de ponderação representa a importância relativa das métricas de forma e distância na função de erro final. A função de erro final ($Loss$) é então calculada como:

$$Loss = 50.0 \times (W \times ShLoss + (1 - W) \times DtLoss)$$

3.1.5 Inicialização e Otimização dos parâmetros epidemiológicos

3.1.5.1 Escolhendo a combinação inicial de variáveis

Na fase inicial do processo de modelagem, é determinada uma combinação inicial de variáveis que sirva como ponto de partida para a busca sistemática por soluções eficazes

no espaço de parâmetros definido. Para atingir esse objetivo, adotamos uma abordagem de inicialização aleatória, que envolve a seleção aleatória dos valores de cada variável presente no *grid*, de acordo com os intervalos de valores aceitos para cada uma delas.

Esta escolha estratégica de uma combinação inicial aleatória oferece várias vantagens significativas. Em primeiro lugar, promove a diversificação nas regiões exploradas dentro do espaço de parâmetros, evitando a estagnação da busca em mínimos locais. Ao evitar padrões fixos, nosso processo de modelagem é mais dinâmico, explorando diversas configurações que podem levar a resultados ótimos.

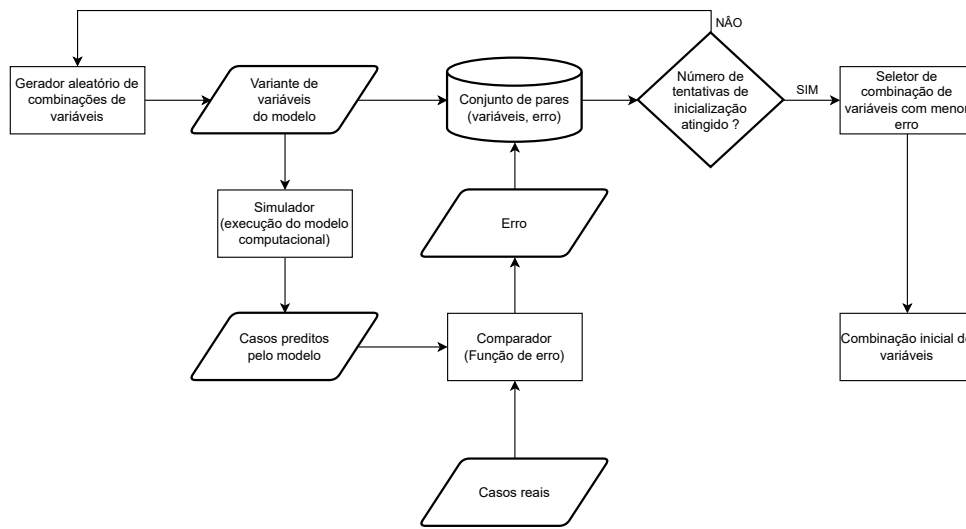
Além disso, essa abordagem também reduz qualquer viés potencial na busca, já que o modelo não parte sempre do mesmo ponto inicial. Isso é garante uma exploração ampla e imparcial do espaço de parâmetros, permitindo que o modelo avalie uma variedade de configurações e escolha aquela que melhor se ajusta aos dados reais dos casos.

O processo completo de seleção da combinação inicial de variáveis é meticulosamente delineado. Inicialmente, os valores de cada variável são escolhidos aleatoriamente, considerando as restrições definidas pelos intervalos aceitáveis. Em seguida, realizamos simulações da epidemia usando o modelo computacional derivado do modelo matemático proposto. Os resultados dessas simulações são então comparados com os dados reais dos casos, utilizando a função de erro mencionada anteriormente.

Esse processo é repetido várias vezes, de acordo com um número pré-definido de iterações. A combinação de variáveis que resulta no menor erro associado à simulação é então escolhida como a combinação inicial de variáveis para o modelo. Todo esse procedimento é visualizado de maneira clara e concisa no fluxograma apresentado na Figura 8.

Dessa forma, ao adotar uma estratégia de inicialização aleatória cuidadosamente planejada, garantimos uma abordagem imparcial e dinâmica na busca pela configuração inicial mais promissora. Essa escolha estratégica é fundamental para lançar as bases sólidas de nosso modelo, permitindo que ele explore eficientemente o espaço de parâmetros e, ao final, ofereça soluções precisas e bem ajustadas à dinâmica da epidemia de mpox nas Américas

Figura 8 – Processo de seleção da combinação inicial de variáveis



Fonte: O autor

3.1.5.2 Otimização dos parâmetros

Na busca pela melhor configuração de variáveis epidemiológicas que descrevem a dinâmica de transmissão do mpox nas Américas, implementamos um processo de otimização detalhado e estruturado. Esse processo, fundamental para a predição precisa dos casos de mpox, envolveu uma série de etapas cuidadosamente planejadas, com o objetivo de encontrar a combinação ideal de parâmetros dispersos na *grid*. A melhor configuração resultante deste processo servirá como base para a representação da epidemia por meio das previsões realizadas pelo modelo matemático/computacional proposto, alinhadas com os dados reais fornecidos pela OPAS.

Inicialmente, as entradas para o processo de otimização incluem o modelo matemático, os dados reais de casos e um conjunto inicial de variáveis adquirido através do processo de inicialização previamente estabelecido, conforme detalhado na seção anterior.

O processo prossegue com a etapa de simulação, em que o modelo matemático, representado computacionalmente, é utilizado juntamente com o conjunto inicial de variáveis para gerar uma nova série temporal. Essa série temporal representa os casos diários previstos pelo modelo para o país das Américas em questão.

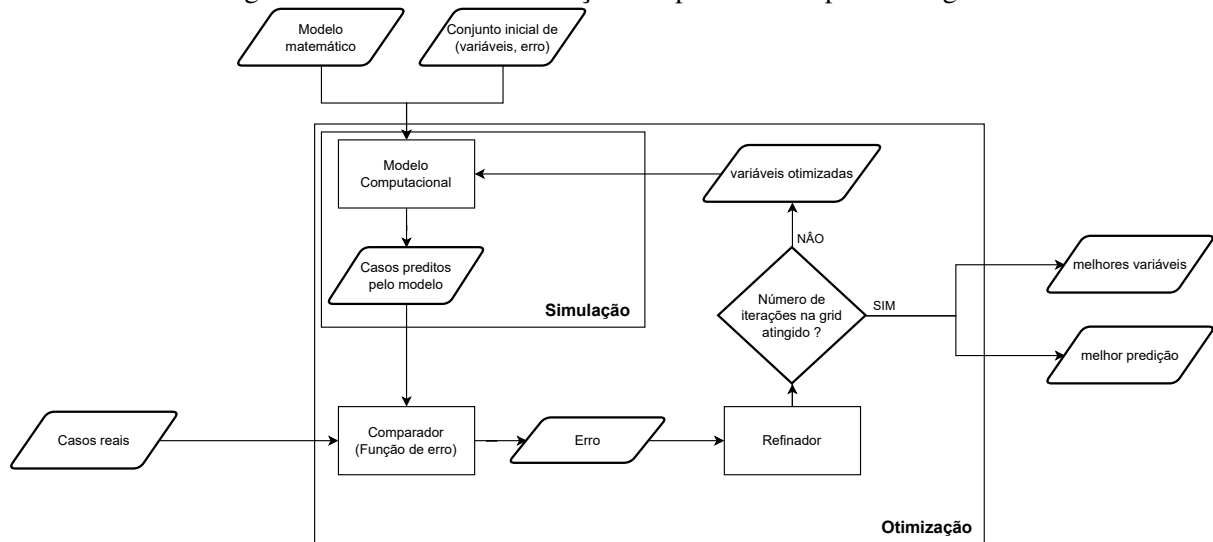
Em seguida, esses dados previstos são comparados com os dados reais por meio do módulo comparador, que aplica uma função de erro para quantificar a proximidade das previsões com os dados reais. A partir dessa comparação, as variáveis são refinadas com base no erro

obtido na etapa anterior.

O processo de refinamento, conhecido como a caminhada pela *grid*, emprega o método do gradiente descendente para determinar em qual direção da *grid* o próximo passo deve ser dado, visando minimizar o erro obtido. Ao percorrer a *grid*, obtemos um novo conjunto de variáveis, que é então retroalimentado no processo, servindo como entrada para o modelo computacional durante a etapa de simulação subsequente.

O fluxo de execução deste processo é claramente delineado no diagrama apresentado na Figura 9 e é descrito em detalhes a seguir:

Figura 9 – Processo de otimização dos parâmetros epidemiológicos



Fonte: O autor

Todo esse ciclo de simulação, comparação, determinação de erro e refinamento, através da busca por parâmetros com menor erro na *grid*, é repetido um número predefinido de vezes ou até que o mínimo global seja alcançado. Esse processo iterativo e meticuloso nos permite encontrar a combinação de variáveis que melhor se ajusta aos dados reais, fornecendo assim uma base sólida para a modelagem eficaz da epidemia de *mpox* nas Américas.

4 MODELO SPLICE

Neste trabalho foi desenvolvido um modelo compartimental generalista denominado SPLICE. O modelo recebe esse nome devido à sua principal característica, que é a capacidade de combinar diversos compartimentos de outros modelos, como seus subcompartimentos, permitindo representar facilmente qualquer outro modelo compartimental. Essa propriedade do modelo SPLICE é demonstrada na tabela 2, na qual cada compartimento que constitui os modelos compartimentais construídos em trabalhos correlatos, para a modelagem da dinâmica de transmissão, é alocado no compartimento correspondente do SPLICE.

Tabela 2 – Correlação dos compartimentos clássicos com supercompartimentos do splice

| | | | |
|-------------------------------|-------------------|-------------------|-----------|
| S (<i>Susceptible</i>) | S (Suscetível) | | |
| P (<i>Protected</i>) | I(Isolado) | | |
| L (<i>Latent</i>) | E(Exposto) | | |
| I (<i>Infectious</i>) | I(Infetado) | P (Prodromal) | |
| C (<i>Controlled</i>) | H (Hospitalizado) | Q (Em quarentena) | |
| E (<i>Endend</i>) | V (Vacinado) | R (Recuperado) | D (Morto) |

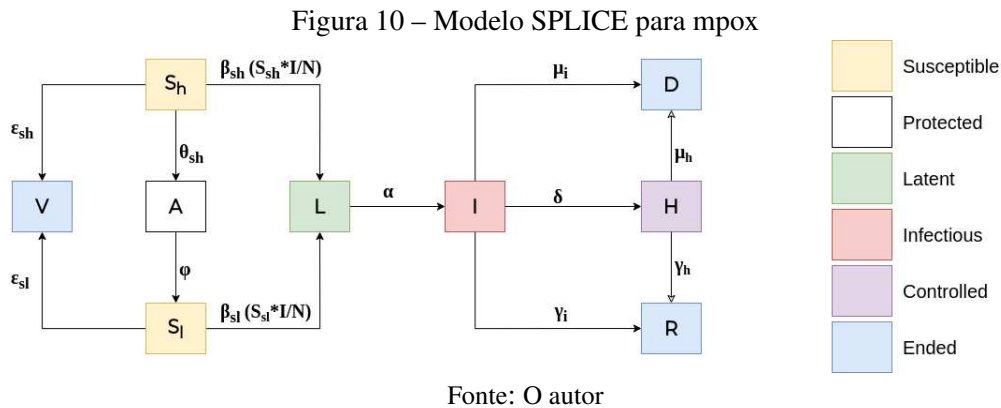
Na tabela são exibidos, por linha, exemplos de compartimentos e seus respectivos supercompartimentos -isto é, o compartimento de maior grau, no qual está inserido

Fonte: O autor

No modelo SPLICE, existem 5 supercompartimentos que podem agrupar subcompartimentos. O primeiro é o compartimento dos suscetíveis, que abrange qualquer compartimento destinado a representar indivíduos que podem contrair a doença. O segundo é o segmento dos protegidos, que representa os indivíduos suscetíveis que adotam medidas para evitar contrair a doença, como o isolamento e a abstinência. O terceiro é o compartimento latente, que engloba os indivíduos que contraíram a doença, mas não apresentam sintomas e não são contagiosos, estando na fase de incubação (expostos). O quarto é o compartimento infeccioso, que contém os indivíduos contagiosos que transmitem a doença para os suscetíveis. O quinto é o grupo dos controlados, que inclui os indivíduos protegidos e que não contraem mais a doença, como os em quarentena e os hospitalizados. Por fim, há o agrupamento dos compartimentos finalizadores, chamado de "*ended*"(final), que representa os indivíduos que estão protegidos e não contraem mais a doença, como os vacinados, os recuperados e os falecidos.

4.1 MODELANDO MPOX COM SPLICE

Para a modelagem da dinâmica de transmissão do mpox, foram considerados os diversos modelos compartimentais utilizados na literatura, bem como informações sobre a doença obtidas dos canais de comunicação da OMS, a fim de determinar os subcompartimentos do modelo. A figura 10 apresenta a versão parcial do modelo SPLICE para mpox.



O modelo realiza abstrações dos compartimentos, em comparação com outros modelos. Por exemplo, considera-se apenas um compartimento para o grupo de indivíduos infectados, diferente do que foi proposto por Al-Shomrani et al. (2023). No entanto, outras características importantes do surto de mpox em 2022 são levadas em consideração para representar o maior grau de contágio da doença ao qual certas populações, como a população de HSH, estão expostas. O modelo é composto por dois grupos de indivíduos suscetíveis: de baixo risco (S_l) e de alto risco (S_h). Cada grupo apresenta uma propensão diferente de exposição à doença (estágio latente) e um nível diferente de adesão à vacinação. Esses aspectos são baseados em dados obtidos da OMS, e também são discutidos por autores como Kupferschmidt (2022).

Para representar o comportamento de alguns indivíduos da população HSH que praticam abstinência sexual ou reduzem consideravelmente o contato sexual durante um surto, o modelo possui um compartimento chamado "*abstinence*" (abstinência), cuja transição se ocorre a partir dos indivíduos suscetíveis do grupo de alto risco. Esse estado reflete o período em que tais indivíduos optam por diminuir ou interromper temporariamente suas atividades sexuais com indivíduos terceiros a fim de reduzir a possibilidade de infecção pela doença. Após o período de abstinência, devido à mudança de comportamento do indivíduo, ele pode retornar ao grupo suscetível de baixo risco. Essa dinâmica permite capturar a variação no comportamento sexual de determinados indivíduos durante a propagação da epidemia.

Por fim, outra simplificação diz respeito ao processo de transição para o estado finalizador "death"(morte). Nessa modelagem, não consideramos a morte dos indivíduos por outras causas além da doença, fazendo com que a mudança para o compartimento "death" seja possível apenas após a infecção (passagem pelo compartimento infeccioso). Além do estágio finalizador de morte, um indivíduo infectado pode ser hospitalizado, que é um estágio intermediário no qual ele não representa mais risco de infecção para outros indivíduos, pois sua capacidade de transmitir a doença está controlada e seu contato com outras pessoas é reduzido. Após a conclusão desse estágio, o indivíduo transita para um dos estágios finalizadores: recuperado ou morto.

No modelo SPLICE, a transição do estado Suscetível (S) para o estado Latente (L) ocorre devido ao contato direto entre um indivíduo infeccioso e um suscetível, com a taxa de infecção sendo denotada como β . Observou-se, durante o surto de mpox de 2022, que essa taxa é mais alta para o subgrupo de HSH da população (KUPFERSCHMIDT, 2022). No modelo, um indivíduo tem a probabilidade P_{hg} de pertencer ao grupo de alto risco da população. Portanto, assumimos a existência de duas taxas de infecção: β_{sh} para os indivíduos suscetíveis do grupo de alto risco (S_{hs}) e β_{sl} para o grupo de baixo risco (S_l).

Considerando campanhas públicas de vacinação e conscientização da população, incorporamos transições para o estado de vacinação (V) com duas taxas diferentes: ϵ_{sh} e ϵ_{sl} , para o grupo de alto risco e baixo risco, respectivamente. Além disso, dado que a população de alto risco está ciente do maior risco de contrair a doença, indivíduos desse grupo podem se isolar e adotar medidas de abstinência para se protegerem. A transição para este estado (A) ocorre a uma taxa diária de θ_{sh} , que depende da probabilidade do indivíduo adotar medidas de auto-isolamento e abstinência (P_a) e a duração média, em dias, do isolamento (T_a). Quando essas medidas são relaxadas, os indivíduos voltam a ser suscetíveis ao mpox, embora com uma taxa de infecção menor, resultando em sua transição do grupo de alto risco para o grupo de baixo risco. Esse retorno ao estado suscetível após a abstinência ocorre a uma taxa ϕ , que é inversamente proporcional ao tempo de isolamento (T_a).

Após a exposição à doença no estado Latente, os indivíduos passam para o estado infeccioso (I) com uma taxa de transformação infecciosa α , que é inversamente proporcional ao tempo de incubação da doença (T_{inc}). Do estado I, os indivíduos podem transitar para os estados Hospitalizado (H), Recuperado (R) ou Morto (D). A transição para o estado H ocorre alguns dias

após o aparecimento dos sintomas (tempo de detecção - T_{det}), dependendo da disponibilidade dos serviços de saúde para diagnóstico adequado e encaminhamento para quarentena (probabilidade de detecção P_{det}). A combinação desses fatores é descrita pela taxa de transferência de I para H, denotada por δ .

A transição do estado I para R é inversamente proporcional ao tempo de recuperação (T_{rec}) e à sobrevivência (1 menos a probabilidade de morte pela doença, definida como P_d). A combinação desses fatores é representada pela taxa γ_i . A transição do estado I para D depende da probabilidade de morte devido à doença e do tempo até a morte (T_d) desde o início dos sintomas. A combinação desses fatores é representada pela taxa de mortalidade μ_i .

Os indivíduos no estado H podem transitar para os estados R ou D. A transição do estado H para R é inversamente proporcional ao tempo de recuperação, descontado o tempo de detecção, e à probabilidade de sobrevivência. A combinação desses fatores é representada pela taxa de recuperação da quarentena (γ_h). A transição do estado H para D depende da probabilidade de morte devido à doença e do tempo até a morte desde o início dos sintomas, descontado o tempo de detecção. A combinação desses fatores é representada pela taxa de mortalidade durante a hospitalização (μ_h).

4.1.1 Redes de Populações Conectadas

Com base nos dados fornecidos pela OPAS, observa-se que o comportamento de múltiplas ondas sobrepostas durante o surto de mpox ocorre em países latino-americanos. Como é sabido que os indivíduos que contraem mpox adquirem imunidade permanente, não é possível adotar a abordagem de modelagem de múltiplas ondas com reinfecção. Portanto, a estratégia adotada neste trabalho para explicar o fenômeno das múltiplas ondas durante o surto envolve a consideração de populações segmentadas, denominadas redes, que apresentam diferentes momentos epidemiológicos.

Considera-se que a doença começa em uma rede e, em seguida, se espalha para a segunda rede, da segunda para a terceira e assim por diante, até envolver um número de redes. Supõe-se que um determinado número de dias tenha transcorrido desde o início da transmissão em uma rede até a propagação para a próxima rede. Denotando por t o tempo decorrido desde o início da epidemia ($t = 0$) até o último dia considerado no estudo ($t = T$), o dia de início da pandemia em uma rede $m \in [1, M]$ é representado por $t_m^0 \in [0, T]$, desde que $t_{m+1}^0 \geq t_m^0$ para

$m \in [1, M]$. A m -ésima rede é composta por N_m pessoas que participam da cadeia de transmissão da doença, iniciada pela entrada de $E_m^0 > 0$ pessoas infectadas assintomáticas no dia t_m^0 . Na prática, as variáveis descritas não são conhecidas a priori, exceto, por definição, que $t_1^0 = 0$. Portanto, nossa lista inicial de variáveis desconhecidas no problema começa com $3M$ membros:

- M - número de redes envolvidas na epidemia,
- $(N_m, \quad m = 1, 2, \dots, M)$ - tamanho da população em M redes,
- $(t_m^0, \quad m = 2, 3, \dots, M)$ - Dia do início da epidemia nas $M-1$ redes que sucederam a primeira,
- $(E_m^0, \quad m = 1, 2, \dots, M)$ - Número de infectados iniciais nas M redes.

A partir destas premissas escrevemos o sistema de equações diferenciais do modelo SPLICE para cada rede.

4.1.2 Sistemas de EDOs do modelo

Para a m -ésima rede, o seguinte sistema pode ser resolvido para $t \geq t_m^0$ usando as condições iniciais definidas abaixo:

$$\frac{dSh_m(t)}{dt} = -\beta_h \frac{Sh_m(t)}{N_m} I_m(t) - Sh_m(t) \theta_h - Sh_m(t) \varepsilon_h \quad (4.1)$$

$$\frac{dSl_m(t)}{dt} = -\beta_l \frac{Sl_m(t)}{N_m} I_m(t) - Sl_m(t) \varepsilon_l + A_m(t) \phi \quad (4.2)$$

$$\frac{dL_m(t)}{dt} = \left(\beta_h \frac{Sh_m(t)}{N_m} + \beta_l \frac{Sl_m(t)}{N_m} \right) I_m(t) - \alpha L_m(t) \quad (4.3)$$

$$\frac{dA_m(t)}{dt} = Sh_m(t) \theta_h - A_m(t) \phi \quad (4.4)$$

$$\frac{dV_m(t)}{dt} = Sh_m(t) \varepsilon_h + Sl_m(t) \varepsilon_l \quad (4.5)$$

$$\frac{dI_m(t)}{dt} = \alpha L_m(t) - I_m(t) (\mu_i + \delta + \gamma_i) \quad (4.6)$$

$$\frac{dH_m(t)}{dt} = \delta I_m(t) - H_m(t) (\mu_h + \gamma_h) \quad (4.7)$$

$$\frac{dR_m(t)}{dt} = \gamma_i I_m(t) + \gamma_h H_m(t) \quad (4.8)$$

$$\frac{dD_m(t)}{dt} = \mu_i I_m(t) + \mu_h H_m(t) \quad (4.9)$$

onde:

$$\alpha = \frac{1}{T_{inc}}, \quad \gamma_i = \frac{(1 - P_d)(1 - P_{det})}{T_{rec}}, \quad \delta = \frac{(1 - P_d)P_{det}}{T_{det}}, \quad \theta_{sh} = \frac{P_a}{T_a}$$

$$\mu_i = \frac{(1 - P_{det})P_d}{T_d}, \quad \gamma_h = \frac{(1 - P_d)P_{det}}{T_{rec} - T_{det}}, \quad \mu_h = \frac{P_d}{T_d - T_{det}}, \quad \phi = \frac{1}{T_a}$$

As condições iniciais são:

$$H_m(t) = I_m(t) = R_m(t) = D_m(t) = V_m(t) = A_m(t) = 0 \text{ para } t \leq t_m^0,$$

$$L_m(t) = 0 \text{ para } t < t_m^0, \text{ e } L_m(t_m^0) = L_m^0.$$

$$Sh_m(t) = N_m * P_{hg} \text{ para } t < t_m^0, \text{ e } Sh_m(t_m^0) = (N_m - L_m^0) * P_{hg}.$$

$$Sl_m(t) = N_m * (1 - P_{hg}) \text{ para } t < t_m^0, \text{ e } Sl_m(t_m^0) = (N_m - L_m^0) * (1 - P_{hg}).$$

A descrição dos parâmetros e variáveis do modelo, são listados na tabela 3

Tabela 3 – Principais parâmetros do modelo SPLICE para mpox

| Parâmetro | Descrição |
|-----------------|---|
| P_{hg} | Probabilidade do indivíduo pertencer ao grupo de alto risco. |
| ϵ_{sh} | Taxa de vacinação dos indivíduos suscetíveis do grupo de alto risco. |
| ϵ_{sl} | Taxa de vacinação dos indivíduos suscetíveis do grupo de baixo risco. |
| θ_{sh} | Taxa de indivíduos do grupo de alto risco que adotam medidas de abstinência e isolamento. |
| ϕ | Grau com que os indivíduos em abstinência ou isolamento, relaxam as medidas e voltam a se tornar suscetíveis. |
| β_{sh} | Taxa de infecção dos indivíduos do grupo de alto risco. |
| β_{sl} | Taxa de infecção dos indivíduos do grupo de baixo risco. |
| α | taxa de indivíduos em estágio latente sendo infectados. |
| δ | taxa de indivíduos infectados detectados e hospitalizados. |
| γ_i | taxa com que os indivíduos infectados se recuperam da doença. |
| γ_h | taxa com que os indivíduos hospitalizados se recuperam da doença. |
| μ_i | taxa de mortalidade pela doença dos indivíduos não hospitalizados. |
| μ_h | taxa de mortalidade pela doença após hospitalização. |

Fonte: O autor

5 DESENVOLVIMENTO

Nesta seção, são apresentados os principais resultados obtidos durante o desenvolvimento do trabalho. Os resultados são analisados e discutidos em relação aos objetivos da pesquisa, fornecendo *insights* sobre a dinâmica de transmissão do mpox e a eficácia do modelo SPLICE na representação desses surtos epidêmicos. A análise dos resultados visa validar o modelo proposto, identificar padrões epidemiológicos e fornecer subsídios para a tomada de decisões e o planejamento de estratégias de controle da doença.

5.1 CARACTERIZAÇÃO DOS DADOS

A caracterização dos dados desempenha um papel essencial na compreensão do contexto subjacente a um conjunto de informações valiosas. No caso desta pesquisa, os dados coletados dizem respeito a casos de mpox, uma doença de significativa importância epidemiológica, registrados nos países das Américas durante o surto global de 2022. Esses dados foram disponibilizados pela Organização Pan-Americana da Saúde (OPAS), contendo informações essenciais que incluem a idade dos indivíduos afetados, a data da exportação do caso, a data de início dos sintomas, o país de relato e uma indicação sobre a ocorrência de morte.

O conjunto de dados original, em sua forma bruta, conforme fornecido pela OPAS, abrange um total de 57.297 registros de casos de mpox. Estes registros abrangem casos com datas de início variando de 5 de agosto de 1922 a 29 de janeiro de 2023. Esse extenso conjunto de dados apresentou uma oportunidade valiosa, mas também a necessidade de processamento e limpeza, uma vez que é fundamental que as informações sejam confiáveis e completas.

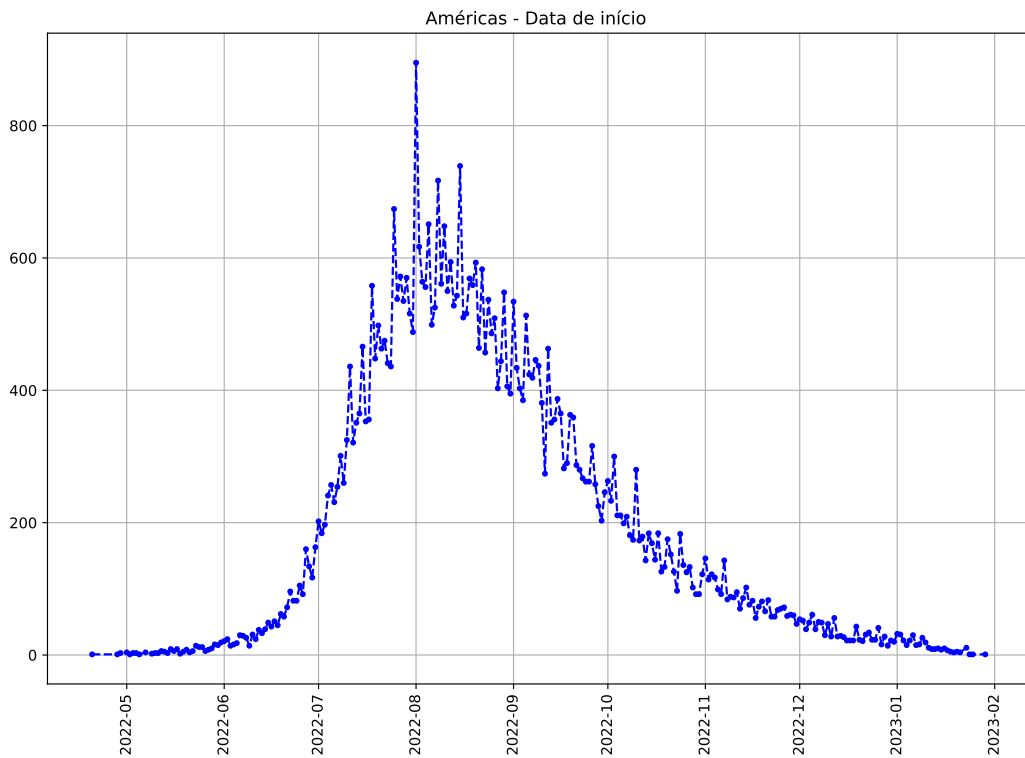
O processo de caracterização dos dados começou com a mineração de dados, que consistiu em uma filtragem inicial do conjunto de dados. Durante essa etapa, foram removidos os registros que não continham informações vitais, tais como o país de reportação, a data do paciente e a data de início dos sintomas. Essas informações são cruciais para a obtenção de métricas epidemiológicas significativas, como o número diário de novos casos de cada país.

Após a aplicação dessa discriminação de dados, o conjunto resultante continha um total de 51.054 registros, o que representa uma significativa redução em relação ao conjunto

de dados original. A menor data de início de sintomas identificada foi em 1º de abril de 2022. Essa etapa de filtragem garantiu que o conjunto de dados fosse composto apenas por registros que continham informações completas e relevantes, tornando-o mais apropriado para análises subsequentes.

A Figura 11 oferece uma representação visual da distribuição de casos de mpox nas Américas por data de início dos sintomas. Esse gráfico ajuda a compreender como os casos se distribuíram ao longo do período de tempo analisado, destacando qualquer tendência temporal notável.

Figura 11 – Distribuição de casos de mpox nas Américas por data de início dos casos.



Fonte: O autor

Além disso, as figuras 24, 25, 26, presentes no apêndice A, fornecem uma análise detalhada da distribuição de casos por faixa etária, segmentados por país. Para serem incluídos nesta análise, os países tiveram que reportar no mínimo nove casos. Estes gráficos proporcionam uma visão detalhada das diferentes faixas etárias afetadas em cada país, permitindo uma compreensão mais aprofundada das variações regionais nos padrões de infecção.

Após a etapa inicial de filtragem e discriminação dos dados, a aplicação da regra de Pareto revelou informações valiosas sobre os segmentos da população mais afetados pela

doença. Este método permitiu a identificação do grupo etário predominante nos registros, o que tem implicações significativas para a formulação de estratégias eficazes no combate à doença. Focando nos indivíduos que representam pelo menos 90% dos casos do *dataset*, agrupamos a população em intervalos regulares de 3 anos, como mostrado na 11. O resultado demonstraram que aproximadamente 91,98% dos casos estavam concentrados na faixa etária de 21 a 53 anos. Estes casos, que correspondem a esse grupo específico da população, tornaram-se o foco central desta pesquisa.

Selecionamos os casos que envolviam indivíduos entre 21 e 53 anos para continuar nossa investigação, uma vez que, conforme descrito acima, representam 91,98% dos totais de caso, além de compreender a população em idade sexualmente ativa, e que portanto tem maiores taxas de contato e propensão a contrair o mpox. Registros provenientes de países com menos de 9 casos da doença foram prontamente removidos do *dataset*, assegurando que a análise se concentrasse em locais com uma incidência relativamente substancial. Além disso, através de uma minuciosa análise dos dados, estabelecemos um critério para a data inicial do surto em cada país. Essa data foi definida como o momento em que os casos ultrapassaram 2 ou 0,002% do número total de casos reportados, indicando o início significativo da epidemia nesse país específico.

Os resultados obtidos, provenientes dos países que atendiam a essas condições rigorosas, foram consolidados na Tabela 4. Nessa tabela, a coluna "Duração" representa a diferença em dias entre o início e a última data relatada dos casos, fornecendo uma visão do período em que a doença esteve ativa em cada país. A coluna "Dias com Dados" indica quantos dias durante a epidemia foram realmente registrados como novos casos, destacando a continuidade da coleta de dados ao longo do tempo. A coluna "Cobertura" representa o percentual de dias com novos dados registrados. A coluna "Nº de Casos" apresenta o total de casos ocorridos no país durante a epidemia, revelando a escala do impacto da doença. Por fim, a coluna "Data Inicial" exibe a data em que a epidemia começou em cada país, proporcionando uma linha do tempo clara para análises longitudinais.

5.2 SERIES SUAVIZADAS

Ao aplicar o método de suavização de dados por média móvel ponderada, conforme detalhado na subseção 3.1.2, realizamos uma transformação valiosa nos dados que nos permitiu

Tabela 4 – Informações gerais da epidemia por país

| País | Duração (dias) | Dias com dados | cobertura | Nºde casos | Início |
|----------------|-----------------------|-----------------------|------------------|-------------------|---------------|
| ARGENTINA | 187 | 172 | 0,9198 | 915 | 23/07/2022 |
| BRASIL | 219 | 216 | 0,9863 | 9418 | 11/06/2022 |
| CANADÁ | 211 | 159 | 0,7536 | 1226 | 07/05/2022 |
| CHILE | 202 | 182 | 0,9010 | 1331 | 30/06/2022 |
| COLÔMBIA | 194 | 171 | 0,8814 | 3820 | 14/07/2022 |
| GUATEMALA | 131 | 122 | 0,9313 | 317 | 12/09/2022 |
| MÉXICO | 223 | 212 | 0,9507 | 3565 | 05/06/2022 |
| PERU | 174 | 170 | 0,9770 | 3413 | 21/06/2022 |
| ESTADOS UNIDOS | 253 | 247 | 0,9763 | 22020 | 16/05/2022 |

Fonte: O autor

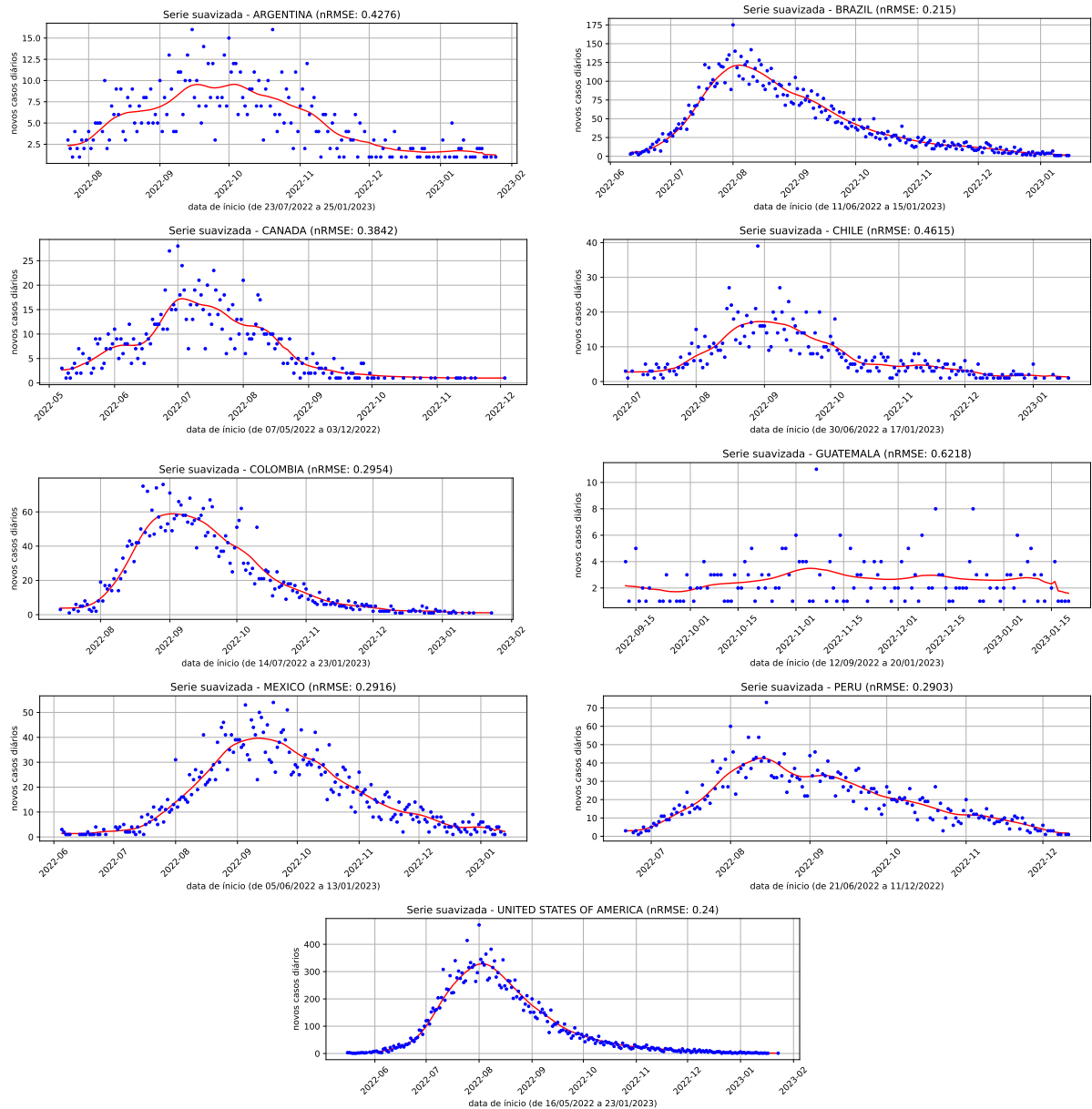
obter uma série temporal suavizada que descreve o número de casos diários por país, como apresentado na Tabela 4.

A aplicação bem-sucedida do método de suavização exigiu a configuração de parâmetros específicos, sendo eles $window = 5$, $n_iter = 3$ e $w = 0,3$. Essas configurações possibilitaram a criação de uma série suavizada com o Erro Quadrático Médio Normalizado (nRMSE) médio de 0,3586. Esse indicador é fundamental para avaliar a eficácia do processo de suavização, pois nos fornece uma medida da discrepância entre os valores observados e os valores suavizados.

Os resultados da suavização por país e os valores nRMSE associados a cada um podem ser visualizados de forma clara nos gráficos apresentados na Figura 12. Ao analisar esses gráficos, observamos que a curva suavizada se ajustou de forma satisfatória à maioria das séries temporais. No entanto, é importante destacar que, em alguns casos, a suavização não se mostrou tão eficaz, resultando em valores de nRMSE elevados, geralmente superiores a 3.

Especificamente, os países Argentina, Canadá, Chile e Guatemala se destacam como aqueles em que a curva suavizada não se ajustou adequadamente aos dados originais, conforme evidenciado pelos valores de nRMSE significativamente superiores a 3. Essas discrepâncias podem ser indicativas de particularidades na evolução dos casos nesses países ou problemas com a reportagem dos casos, exigindo uma análise mais aprofundada para compreender os motivos subjacentes a essas variações.

Figura 12 – Series temporais suavizadas por país



Fonte: O autor

5.3 INFERÊNCIA DE PARÂMETROS E MINERAÇÃO DE DADOS

Nesta seção, nosso objetivo foi realizar uma análise aprofundada dos dados disponíveis, focando na inferência de parâmetros essenciais para compreender a dinâmica do mpox nas Américas. O propósito foi ir além das análises tradicionais, explorando elementos pouco discutidos na literatura existente, especialmente em relação à detecção de indivíduos que contraíram a doença e à mortalidade associada aos casos de mpox na população da região.

A relevância deste estudo reside na necessidade de obter informações mais precisas e detalhadas sobre o mpox nas Américas, uma vez que aspectos específicos como detecção e

mortalidade podem variar significativamente de uma região para outra e ao longo do tempo. Nosso foco foi identificar e inferir valores para parâmetros que não foram claramente discutidos na literatura existente, proporcionando uma compreensão mais robusta da progressão da doença na população estudada.

Além disso, os resultados obtidos nesta seção são fundamentais para fornecer subsídios ao modelo compartimental SPLICE, uma ferramenta matemática valiosa utilizada para simular e prever a propagação de doenças. Ao identificar e compreender os parâmetros relacionados à detecção de casos e à mortalidade, conseguimos aprimorar a precisão do modelo, tornando-o mais eficaz para representar a realidade da situação epidemiológica do mpox nas Américas.

Dessa forma, esta análise não apenas contribui para o avanço da compreensão científica sobre o mpox, mas também oferece uma base sólida para futuras pesquisas e intervenções de saúde pública. Ao obter novas informações relevantes e detalhadas sobre a dinâmica da doença na população das Américas, este estudo é um passo significativo em direção ao desenvolvimento de estratégias mais eficazes de prevenção, controle e resposta a surtos de mpox na região.

5.3.1 Informações e parâmetros relacionados a morte do indivíduo

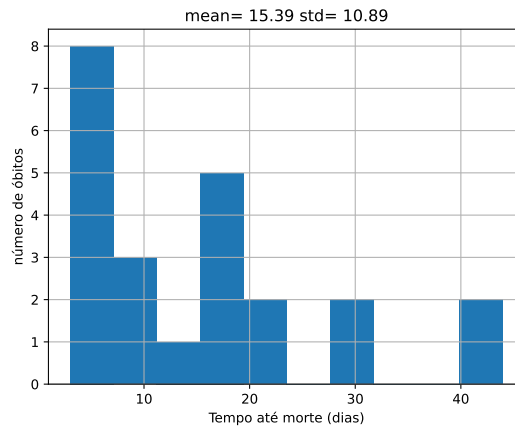
No processo de mineração de dados realizado, focamos em estimar informações cruciais sobre os casos fatais de indivíduos infectados com mpox nas Américas, utilizando o *dataset* detalhadamente caracterizado na seção 5.1.

Inicialmente, identificamos 25 registros nos quais o valor da coluna *OUTCOME* indicava óbito ('D'). Estes casos estavam distribuídos em diferentes países da região: 5 no Peru, 2 no Equador, 1 em Cuba, 2 na Argentina e 15 no Brasil. Para compreender o tempo até a morte, calculamos a diferença entre os valores das colunas *ONSET_DATE* (data de início dos sintomas) e *REPORT_DATE* (data do registro).

Após a eliminação de valores extremos (*outliers*) que poderiam distorcer a análise, que incluíam tempos até a morte de 2 e 92 dias, obtivemos resultados mais precisos. A média dos dias até a morte foi de aproximadamente 15,39 dias, com um desvio padrão de 10,89 dias e uma variância de 0,7078. Esses valores oferecem uma visão mais detalhada e confiável sobre a distribuição dos óbitos em relação ao tempo decorrido desde o início dos sintomas, como

ilustrado na Figura 13.

Figura 13 – Distribuição da quantidade de óbitos por tempo até a morte



Fonte: O autor

Para uma análise mais específica e precisa, restringimos nossa investigação aos países que possuíam um número mínimo de casos em determinados grupos etários. Nesse contexto, nos concentramos em Peru, Argentina e Brasil. A soma total de casos nesses três países foi de 13.746, enquanto o número de óbitos foi de 22. Isso nos proporcionou uma taxa de mortalidade de 1,6005 mortes por 1000 infectados, uma métrica essencial para entender a gravidade da doença em diferentes regiões e grupos populacionais.

As informações resultantes sobre os óbitos, obtidos através da mineração dos dados, são exibidas de forma sumarizada no quadro 5 abaixo:

Tabela 5 – Informações relacionadas ao óbito por mpox inferidas a partir do *dataset* da OPAS

| Probabilidade de Morte (indivíduo) | Dias até Morte | Desvio Padrão | Fonte |
|------------------------------------|----------------|---------------|-----------------------|
| 0,0016 | 15,3913 | 10,8937 | Peru,Argentina,Brasil |

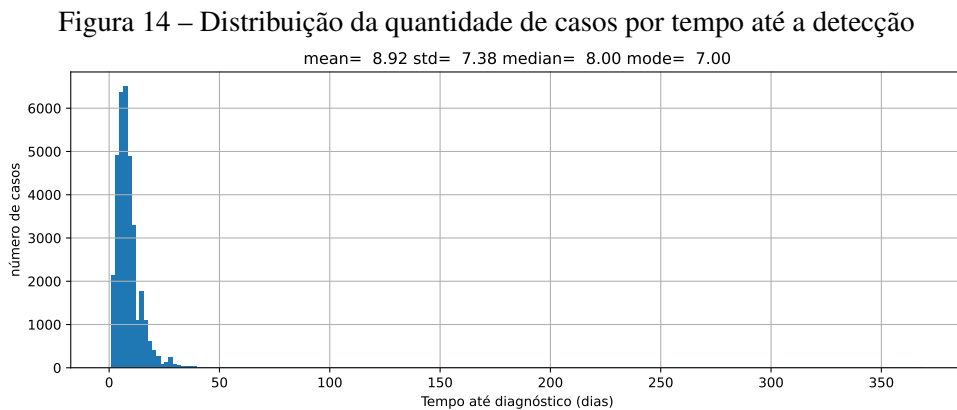
Fonte: O autor

5.3.2 Informações e parâmetros relacionados ao tempo de detecção dos casos

No âmbito da mineração de dados, continuamos nossa investigação com o objetivo de compreender e extrair informações relevantes do *dataset* relacionadas à detecção de novos casos de mpox pelas autoridades de saúde. Nossa abordagem buscou verificar se a coluna de data de diagnóstico (DATE_DIAGNOSIS) continha a informação referente à data de detecção da doença. Para validar essa hipótese, realizamos uma análise minuciosa, comparando as datas de diagnóstico com as datas de início dos sintomas (ONSET_DATE). Os resultados demonstraram

que, para a grande maioria dos casos registrados, a data de diagnóstico estava de fato posterior à data de início dos sintomas, com uma incidência de 95,8446% (ou 34.352 registros) dos casos em que a data de diagnóstico estava disponível, de um total de 35.737 registros.

Com base nessa observação, pudemos inferir que o tempo médio decorrido até a detecção da doença é de aproximadamente 8,92 dias, com um desvio padrão de 7,38 e uma moda de 7 dias. No entanto, vale destacar que esses valores sugerem uma possível distorção nos registros, como evidenciado de forma mais clara na Figura 14.

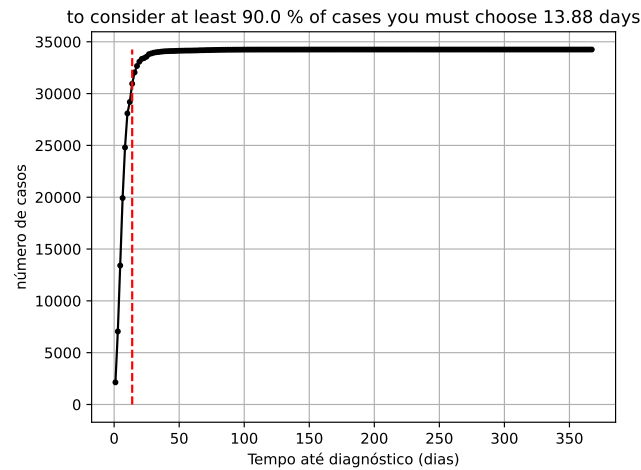


Fonte: O autor

A análise dos dados revelou a presença de *outliers*, ou seja, valores discrepantes que estavam afetando negativamente nosso estudo sobre o tempo até o diagnóstico. Alguns casos indicavam um período de mais de 300 dias entre o início dos sintomas e o diagnóstico, o que levantava dúvidas sobre a precisão dessas informações. Para mitigar esse efeito, aplicamos novamente a regra de Pareto, visando determinar a duração média até a detecção para pelo menos 90% dos casos.

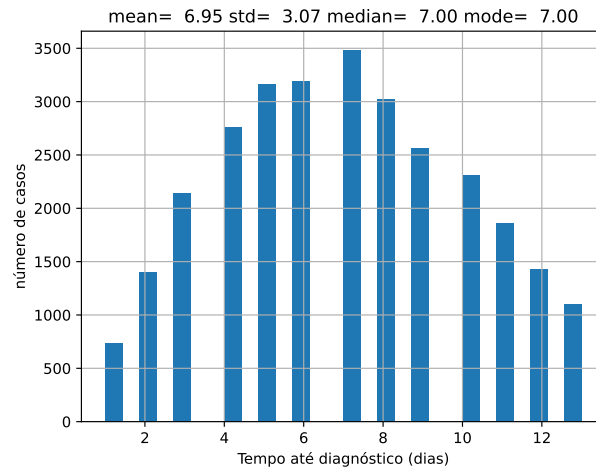
Os resultados obtidos após a aplicação da regra de Pareto revelaram que 90% dos casos de mpox foram detectados em um prazo de até 13,88 dias, como ilustrado na Figura 15. Utilizando métricas estatísticas, obtivemos que o tempo médio até a detecção de novos casos de mpox é de 6,95 dias, com um desvio padrão de 3,07 dias e uma moda de 7 dias. Esses valores oferecem uma visão mais precisa da distribuição dos casos em relação ao tempo de detecção, como representado na Figura 16.

Figura 15 – Identificação do número de dias até a detecção para 90% dos casos



Fonte: O autor

Figura 16 – Distribuição da quantidade de casos por tempo até a detecção dos 90% dos casos detectados



Fonte: O autor

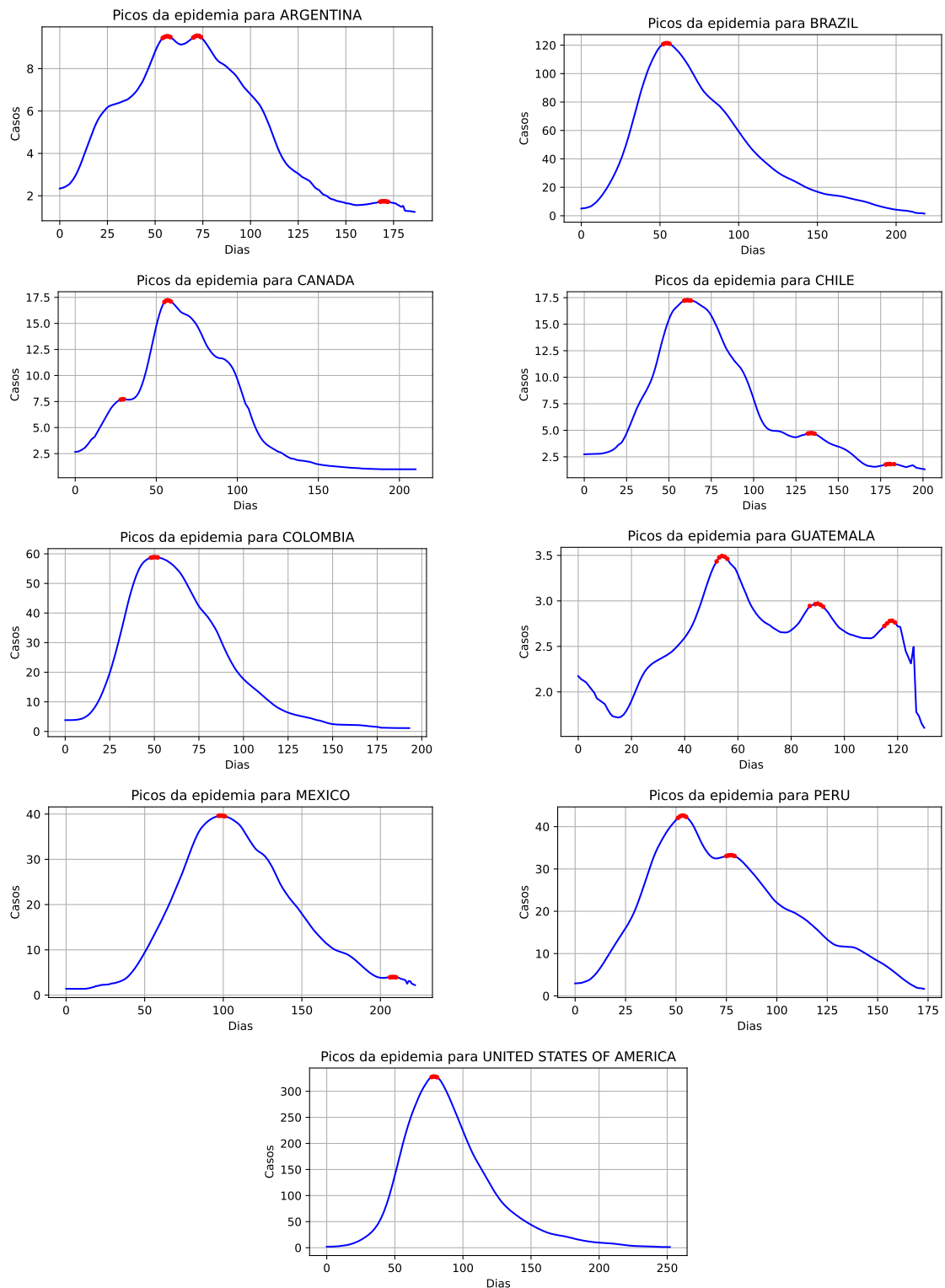
5.4 ESTIMATIVA DO NÚMERO DE REDES DOS PAÍSES AMERICANOS

Para estimar o número de redes, realizamos uma análise das formas das séries temporais dos casos relatados e disponibilizados pela Organização Pan-Americana da Saúde (OPAS). Nosso foco foi identificar possíveis ondas epidêmicas sobrepostas, considerando que cada onda pode ter sido causada por uma rede de transmissão diferente no mesmo país.

Para essa estimativa, utilizamos os dados coletados e aplicamos técnicas de análise de séries temporais para identificar os picos e declínios nas curvas suavizadas de casos (figura 17). Cada pico identificado foi considerado como uma possível indicação de uma onda epidêmica, e supomos que cada uma dessas ondas foi causada por uma rede de transmissão específica.

Com base nessa abordagem, na tabela 6 apresentamos o número estimado de redes

Figura 17 – Identificação dos picos nas ondas epidemiológicas dos países Americanos



Fonte: O autor

simuladas em cada país estudado. Cada valor na tabela representa o número de redes distintas que supomos terem contribuído para as diferentes ondas epidêmicas observadas em cada país. Essa estimativa nos permite compreender melhor a complexidade das interações entre as redes

de transmissão e a dinâmica da propagação da doença em cada contexto específico.

| País | Número de Redes |
|----------------|-----------------|
| Argentina | 3 |
| Brasil | 1 |
| Canadá | 2 |
| Chile | 3 |
| Colômbia | 1 |
| Guatemala | 3 |
| México | 2 |
| Peru | 2 |
| Estados Unidos | 1 |

Tabela 6 – cenários das redes de populações conectadas nas Américas

5.5 SIMULAÇÃO DA EPIDEMIA E ESTIMATIVA DOS PARÂMETROS EPIDEMIO-LÓGICOS PARA OS PAÍSES AMERICANOS

Nesta seção, apresentamos os resultados obtidos por meio da simulação do modelo SPLICE e a otimização dos parâmetros epidemiológicos que descrevem a dinâmica de transmissão do mpox nos países americanos, utilizando os dados fornecidos pela OPAS como base.

Vale ressaltar que a simulação e a estimativa dos parâmetros foram realizadas exclusivamente para os países cuja série temporal suavizada apresentou nRMSE inferior a 3. Este limite foi estabelecido para evitar a propagação do erro na simulação, uma vez que a precisão do modelo em simular a dinâmica de transmissão do mpox é calculada a partir da comparação com os dados suavizados (subseção 3.1.4). Dessa forma, os resultados apresentados a seguir referem-se à simulação e aos parâmetros epidemiológicos específicos para o Brasil, Colômbia, México, Peru e Estados Unidos da América.

Ao considerar apenas os países que atendem ao critério do nRMSE inferior a 3, pudemos realizar uma análise detalhada da dinâmica da epidemia de mpox nesses países. Os resultados das simulações oferecem informações valiosas sobre a propagação da doença, permitindo uma compreensão mais profunda das tendências de transmissão ao longo do tempo. Além disso, a otimização dos parâmetros epidemiológicos proporciona uma visão precisa das características específicas de cada país, incluindo taxas de transmissão, períodos de incubação e taxas de detecção.

A seguir, apresentaremos os resultados obtidos para cada um dos países mencionados, destacando as principais descobertas derivadas da simulação do modelo SPLICE e da otimização dos parâmetros. Cada análise é fundamentada nos dados reais suavizados e nas previsões geradas pelo modelo, fornecendo uma visão abrangente e detalhada da dinâmica da epidemia de mpox nos países estudados.

5.5.1 Estimativa dos parâmetros epidemiológicos

A identificação precisa dos parâmetros epidemiológicos é de fundamental importância para a compreensão detalhada da dinâmica de qualquer epidemia, incluindo o mpox. A definição cuidadosa desses parâmetros é impactam a precisão das previsões e simulações, fornecendo informações valiosas que orientam estratégias de controle e intervenção. Nesta seção, descrevemos o processo detalhado de estimativa dos parâmetros epidemiológicos para os países estudados, que se deu por meio de dois processos consecutivos e complementares.

Inicialmente, os valores dos parâmetros foram definidos com base em uma variedade de fontes confiáveis. Utilizamos informações extraídas do *dataset* da OPAS, como detalhado na seção de Inferência de Parâmetros e Mineração de dados (5.3), bem como valores obtidos por outros pesquisadores em trabalhos similares. Além disso, incorporamos dados oficiais disponibilizados por órgãos de saúde respeitáveis, como a OPAS e a OMS, em seus canais de comunicação. Para parâmetros que não foram claramente definidos na literatura, recorreremos a valores estimados, estabelecidos com base no conhecimento dos especialistas envolvidos no estudo. Todos esses valores de referência iniciais estão detalhados na Tabela 7.

Para iniciar o processo de estimação dos parâmetros, implementamos a otimização por meio da simulação da epidemia usando o modelo SPLICE, conforme detalhado na Subseção 3.1.5.2. Utilizamos a técnica de grid search, conforme explicado na Subseção 3.1.3, para explorar sistematicamente os intervalos definidos na Tabela 7. A definição dos intervalos foi feita a partir da seguinte abordagem: Para os parâmetros com valores de referência constantes, aplicamos uma variação mínima de $\pm 0,1\%$ em relação ao valor de referência, com o passo (step) definido como a média dos valores máximo e mínimo do intervalo. Essa abordagem foi adotada para manter a coerência dos valores resultantes da busca com os valores de referência, garantindo consistência e relevância aos resultados obtidos.

Para os parâmetros cujos valores de referência foram definidos como intervalos,

Tabela 7 – Configuração inicial dos parâmetros

| Parâmetro | Referência | Intervalo de Busca | Step | Fonte |
|---|-----------------------------|---------------------------------|----------|------------------------------|
| Probabilidade do indivíduo pertencer ao grupo de alto risco | 0,4 | $0,2 \leq x \leq 0,8$ | 0,01 | (BRAGAZZI et al., 2023) |
| Taxa de infecção no grupo de alto risco | - | $0,1 \leq x \leq 4$ | 0,1 | Assumido |
| Taxa de infecção no grupo de baixo risco | - | $0,05 \leq x \leq 2$ | 0,01 | Assumido |
| Taxa de vacinação no grupo de alto risco | - | $0,0005 \leq x \leq 0,0008$ | 0,0001 | Assumido |
| Taxa de vacinação no grupo de baixo risco | $0,0003 \leq x \leq 0,0005$ | $0,0003 \leq x \leq 0,0005$ | 0,0001 | (XIRIDOU et al., 2023) |
| Probabilidade de abstinência e isolamento | $0,3 \leq x \leq 0,5$ | $0,3 \leq x \leq 0,5$ | 0,1 | (YUAN et al., 2023) |
| Duração da abstinência e isolamento | 21 | $20,979 \leq x \leq 21,021$ | 0,021 | (YUAN et al., 2023) |
| Tempo até a recuperação (dias) | 21 | $20,979 \leq x \leq 21,021$ | 0,021 | (CONTROL; PREVENTION, 2023b) |
| Tempo de incubação (dias) | 8 | $7,992 \leq x \leq 8,008$ | 0,008 | (CONTROL; PREVENTION, 2023a) |
| Tempo até a detecção (dias) | 6,95 | $6,94305 \leq x \leq 6,95695$ | 0,000331 | Mineração de dados |
| Probabilidade de detecção | - | $0,4995 \leq x \leq 0,5005$ | 0,0005 | Assumido |
| Tempo até a morte (dias) | 68 | $67,932 \leq x \leq 68,068$ | 0,0017 | (RISER et al., 2023) |
| Probabilidade de morte | 0,0013 | $0,001299 \leq x \leq 0,001301$ | 0,00001 | Mineração de dados |

Fonte: O autor

utilizamos o próprio intervalo estabelecido, considerando os limites como os valores possíveis para o parâmetro na grid. O valor de incremento, tanto para os parâmetros assumidos como constantes quanto para aqueles definidos como intervalos, foi determinado conforme a referência da literatura utilizada. Especificamente, o valor de incremento foi calculado usando a expressão $x/(x \times 10^n)$, em que n representa o número de casas decimais de x , e x é o valor mínimo do

intervalo. Esse cuidado na definição dos intervalos e incrementos é fundamental para assegurar a precisão e confiabilidade do processo de busca pelos parâmetros ótimos.

Durante a execução da simulação e busca pelos parâmetros ideais, também configuramos os hiperparâmetros necessários para a simulação, conforme detalhado na Tabela 8.

Tabela 8 – Hiperparâmetros para processo de otimização dos parâmetros epidemiológicos

| Hiperparâmetro | Valor |
|---|--|
| Número de tentativas de inicialização aleatória das variáveis | 30 |
| Erro limiar da inicialização das variáveis | 0,25 |
| Número de iterações na <i>grid</i> | $50 * \sqrt{(\text{Número de redes})}$ |

Fonte: O autor

Tabela 9 – Parâmetros epidemiológicos resultantes da otimização

| Parâmetro | Brasil | Colômbia | México | Peru | Estados Unidos |
|---|--------|----------|--------|---------|----------------|
| Probabilidade do indivíduo pertencer ao grupo de alto risco | 0,38 | 0,41 | 0,61 | 0,27 | 0,8 |
| Taxa de infecção no grupo de alto risco | 1,0 | 1,2 | 0,5 | 4,0 | 0,5 |
| Taxa de infecção no grupo de baixo risco | 0,25 | 0,05 | 0,15 | 0,2 | 0,05 |
| Taxa de vacinação no grupo de alto risco | 0,0008 | 0,0005 | 0,0005 | 0,0008 | 0,0005 |
| Taxa de vacinação no grupo de baixo risco | 0,0001 | 0,0001 | 0,0005 | 0,0005 | 0,0001 |
| Probabilidade de abstinência e isolamento | 0,5 | 0,42 | 0,3 | 0,3 | 0,5 |
| Duração da abstinência e isolamento | 20,979 | 20,979 | 21,02 | 21,02 | 21,02 |
| Tempo até a recuperação (dias) | 21,020 | 21,020 | 21,02 | 21,02 | 21,02 |
| Tempo de incubação (dias) | 8,008 | 8,008 | 8,008 | 8,008 | 8,008 |
| Tempo até a detecção (dias) | 6,9499 | 6,9639 | 6,9639 | 6,9499 | 6,9639 |
| Probabilidade de detecção | 0,5005 | 0,4995 | 0,4995 | 0,5005 | 0,4995 |
| Tempo até a morte (dias) | 68,068 | 68,068 | 68,068 | 67,932 | 68,068 |
| Probabilidade de morte | 0,0013 | 0,0013 | 0,0013 | 0,00013 | 0,00013 |

Fonte: O autor

Os resultados ótimos obtidos para o modelo SPLICE em cada um dos países estudados são apresentados na tabela 9 acima, oferecendo uma visão detalhada e precisa dos parâmetros epidemiológicos estimados para o mpox nos países da América. Esses valores refinados são essenciais para uma representação acurada da dinâmica da epidemia e constituem a base sólida

para análises futuras e tomadas de decisão relacionadas ao controle da doença.

5.5.2 Simulação do mpox com SPLICE

No processo de simulação, o modelo SPLICE utiliza os parâmetros epidemiológicos definidos para reproduzir os episódios da doença, com base no modelo computacional derivado de equações diferenciais ordinárias. Para avaliar a capacidade do modelo em representar o surto, realizamos uma comparação entre os resultados obtidos - uma nova série temporal com os casos previstos - e os dados reais dos casos, após suavização, conforme explicado na Seção 3.1.4.

Além dos parâmetros epidemiológicos detalhados na Tabela 7, outras variáveis também precisam ser definidas para a execução do modelo SPLICE com redes de populações conectadas. Embora essas variáveis não estejam diretamente ligadas ao mpox, elas impactam o processo de simulação. Para cada país, esses parâmetros incluem o número de redes conectadas (ou ondas sobrepostas), como definido na Seção 17, a duração da simulação em dias - obtida por meio da análise de dados e apresentada na Tabela 4 -, o tamanho do grupo populacional (número de indivíduos) em cada uma das redes, o intervalo de tempo até o início do surto em cada rede e o número de indivíduos em estágio latente, que representam os disseminadores iniciais da doença em cada rede. Os intervalos de valores possíveis para esses três últimos parâmetros foram estimados para posterior otimização através da busca na *grid*. O número de indivíduos em cada rede varia entre 10 e 500 vezes o maior número de ocorrência de casos reais durante o período da epidemia. A data de início do surto por sua vez varia entre 1 e 0,85 vezes a duração da epidemia no país. Por fim, o número de disseminadores iniciais estão no intervalo de 1 à 0,05 o tamanho máximo que a população pode atingir.

Os valores finais para cada uma dessas variáveis, após a otimização para os países simulados, estão apresentados na Tabela 10.

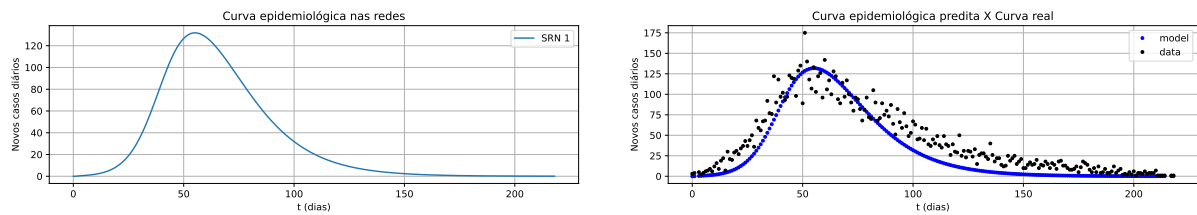
Tabela 10 – Parâmetros vinculados as redes conectadas resultantes

| Parâmetro de cada rede N | Brasil | Colômbia | México | Peru | Estados Unidos |
|-----------------------------|-----------------|-----------------|------------------------------|------------------------------|------------------|
| Tamanho da população | N1:10500 | N1:15960 | N1: 3780 N2: 3240 | N1: 2190 N2: 2920 | N1: 80070 |
| Início do surto | N1:0 | N1:0 | N1: 0 N2: 0 | N1: 0 N2: 12 | N1: 0 |
| Disseminadores iniciais | N1:25 | N1:79 | N1: 5 N2: 5 | N1: 39 N2: 1 | N1: 155 |

Fonte: O autor

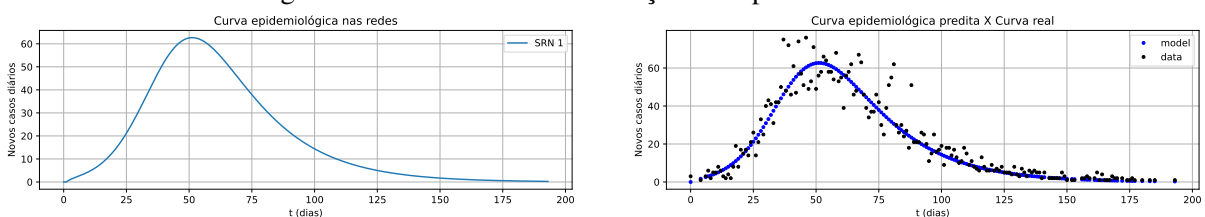
Os resultados das simulações são apresentados em dois tipos de gráficos diferentes, conforme exibido nas figuras 18, 19, 20, 21, 22. O primeiro tipo de gráfico mostra a curva epidemiológica predita para cada uma das redes conectadas. Por fim, o terceiro tipo de gráfico apresenta a curva epidemiológica dos casos preditos, comparada com os dados reais suavizados, proporcionando uma visualização clara da precisão do modelo.

Figura 18 – Resultados da simulação do mpox no Brasil



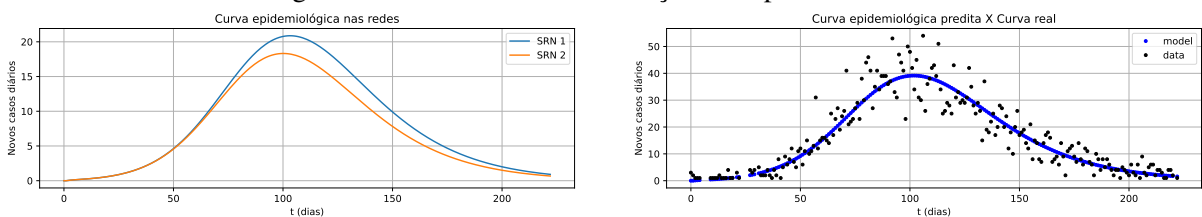
Fonte: O autor

Figura 19 – Resultados da simulação do mpox na Colômbia



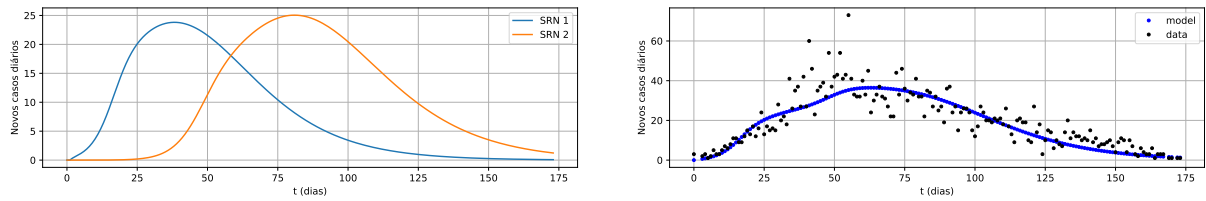
Fonte: O autor

Figura 20 – Resultados da simulação do mpox no México



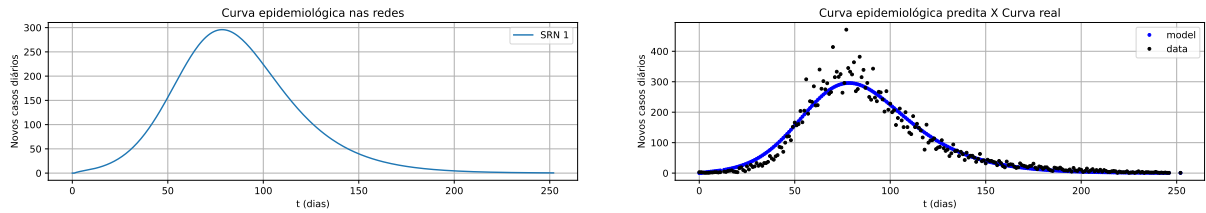
Fonte: O autor

Figura 21 – Resultados da simulação do mpox no Peru



Fonte: O autor

Figura 22 – Resultados da simulação do mpox nos Estados Unidos



Fonte: O autor

O quadro 11 resume os erros obtidos para a simulação do mpox com SPLICE em cada um dos países estudados, oferecendo uma visão abrangente da acurácia do modelo para cada cenário simulado.

Tabela 11 – Taxa de erro da simulação para os países selecionados

| Métrica de Erro | Brasil | Colômbia | México | Peru | Estados Unidos |
|------------------------------|--------|----------|--------|--------|----------------|
| Erro de Forma (`ShLoss`) | 0,0015 | 0,0007 | 0,0002 | 0,0010 | 0,0005 |
| Erro de Distância (`DtLoss`) | 0,0015 | 0,0010 | 0,0009 | 0,0014 | 0,0007 |
| Erro Total | 0,0748 | 0,0475 | 0,0481 | 0,0665 | 0,0314 |

Fonte: O autor

5.6 ANÁLISE DE SENSIBILIDADE

A análise de sensibilidade é uma técnica empregada para compreender como variações nos parâmetros de um modelo afetam suas saídas (WU et al., 2013). No contexto do nosso estudo sobre a propagação da mpox em países das Américas, identificamos as melhores combinações de parâmetros para representar adequadamente a dinâmica de transmissão da doença, como previamente detalhado em outra seção do trabalho.

Para compreender melhor a influência desses parâmetros nas previsões do modelo, aplicamos a análise de sensibilidade. Realizamos 20 simulações adicionais para cada país, variando os melhores parâmetros encontrados conforme definido na tabela 9. Em cada uma das 20 combinações diferentes de parâmetros variados, optamos por variar 50% dos parâmetros em até 15% do seu valor original. Isso nos permitiu avaliar como a variação simultânea de todos os

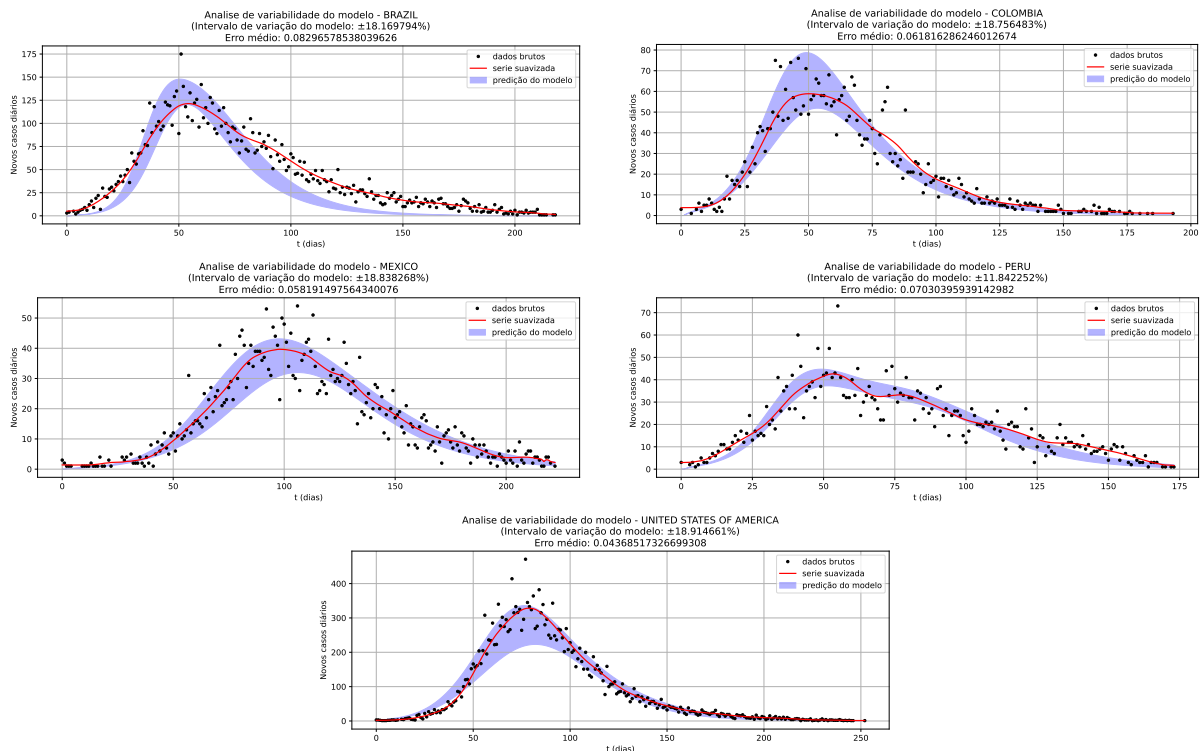
parâmetros afeta as previsões do modelo para cada país.

O objetivo principal ao utilizar essa abordagem foi avaliar a sensibilidade do modelo a diferentes cenários epidemiológicos, explorando como a variação conjunta dos parâmetros impacta as previsões da propagação da mpox. Essa metodologia foi escolhida para considerar a incerteza associada aos valores dos parâmetros e entender melhor a interação entre eles.

Os benefícios dessa análise de sensibilidade incluem a avaliação da estabilidade do modelo diante de incertezas nos parâmetros e a identificação de possíveis cenários extremos que podem impactar significativamente as previsões. A estratégia de variar os parâmetros dos modelos permitiu uma compreensão mais ampla da dinâmica da doença e das variações nas previsões do modelo em resposta a mudanças nos parâmetros.

A escolha de variar os parâmetros em conjunto dentro de um intervalo de 15% foi estratégica para explorar a variabilidade do modelo de maneira abrangente, contribuindo para uma compreensão mais robusta da propagação da mpox em diferentes países das Américas.

Figura 23 – Resultado da simulação considerando conjunto variado de parâmetros por país



Análise de sensibilidade realizada como a combinação da onda epidemiológica resultante da simulação, considerando 20 combinações de parâmetros ótimos, variando 50% dos parâmetros a uma taxa de $\pm 15\%$

Fonte: O autor

Os resultados da análise de sensibilidade são exibidos na figura apêndice 23, em

forma de gráficos. Cada gráfico representa a curva epidemiológica resultante da variação das predições do modelo em função de cada um dos parâmetros variados no processo de análise de sensibilidade. Além disto, na tabela 12, abaixo, apresentamos a taxa de variação do modelo para cada um dos países selecionados para o estudo.

Tabela 12 – Taxa de variação das predições do modelo para os países selecionados

| Brasil | Colômbia | México | Peru | Estados Unidos |
|--------|----------|--------|--------|----------------|
| 18,17% | 18,76% | 18,83% | 11,84% | 18,91% |

Fonte: O autor

6 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Neste capítulo, apresentamos uma discussão abrangente sobre os resultados obtidos em nosso estudo, abordando as descobertas relevantes, a identificação de limitações e ameaças à validação do trabalho realizado. Além disso, fornecemos sugestões para possíveis pontos de melhoria e delineamos áreas para futuras pesquisas, construindo uma visão abrangente do valor deste trabalho e das direções para o desenvolvimento futuro.

Por meio da análise e mineração dos dados fornecidos pela OPAS, nosso objetivo foi obter informações essenciais para a simulação do modelo epidemiológico proposto (4) e criar uma base sólida de dados capaz de caracterizar a dinâmica de transmissão do mpox nas Américas durante o surto de 2022. A análise revelou que os adultos entre 21 e 53 anos foram o grupo mais afetado, representando aproximadamente 92% dos casos registrados. Com base nessas descobertas e nos gráficos de distribuição de casos por faixa etária em cada país (figuras 24, 25, 26), as instituições de saúde podem adotar medidas específicas voltadas para esse grupo, visando mitigar os efeitos de possíveis surtos futuros de mpox.

No entanto, durante o período do surto coberto pelos registros do *dataset* da OPAS, observamos vários dias sem novos casos ou dias consecutivos com diferenças significativas no número de casos registrados. Isso sugere a presença de casos subnotificados de mpox, indicando a possibilidade de indivíduos que foram expostos à doença, mas não foram diagnosticados ou notificados pelas autoridades competentes. Essa subnotificação, juntamente com os intervalos de tempo sem dados e a falta de informações detalhadas sobre o comportamento dos pacientes, impacta negativamente a capacidade preditiva dos modelos epidemiológicos, conforme já discutido e enfrentado em (KOROSEC et al., 2023). Esses fatores são evidenciados pelos resultados do processo de suavização dos dados apresentados na Figura 12, onde observamos uma onda suavizada com maior erro quadrático médio normalizado para países com maior número de dias sem dados (4) ou com dados esparsos.

Além disso, a insuficiência de informações relevantes sobre os casos no *dataset* limitou nossa capacidade de extrair parâmetros epidemiológicos. Por exemplo, ao analisar os parâmetros relacionados às mortes de indivíduos com mpox na Subseção 5.3.1, identificamos

que apenas cinco países tinham informações sobre mortes, excluindo os Estados Unidos, que possui o maior número de casos registrados no *dataset*. Isso sugere que a OPAS pode não ter tido acesso aos dados de mortalidade dos pacientes neste país ou que as informações ainda não estavam disponíveis durante a construção do *dataset*. Portanto, optamos por não usar os valores obtidos pela mineração sobre a mortalidade de indivíduos com mpox para definir os parâmetros do modelo epidemiológico, preferindo utilizar valores encontrados na literatura ((RISER et al., 2023)). Surpreendentemente, conseguimos, por meio de nossas análises, chegar a uma taxa de probabilidade de morte semelhante à encontrada em (RISER et al., 2023). No entanto, não podemos afirmar o mesmo em relação ao tempo médio até a morte do indivíduo contaminado, supondo que isso seja resultado da subnotificação e erro humano no processo de levantamento dos casos.

No entanto, conseguimos extrair informações relevantes sobre o tempo de diagnóstico da doença, visto que mais de 95% dos registros continham informações sobre a data de diagnóstico. Os dados indicam que o tempo médio até o diagnóstico do mpox nas Américas é de aproximadamente 6,95 dias, uma informação importante relacionada à eficácia das medidas adotadas pelas instituições de saúde competentes. Ao analisar os parâmetros epidemiológicos derivados do processo de otimização, observamos diferenças marcantes entre os países estudados. Com exceção dos parâmetros considerados constantes, devido à sua estabilidade ou origem confiável, como aqueles obtidos de fontes confiáveis ou provenientes do processo de mineração de dados, tais como a duração de abstinência, tempo até a recuperação, tempo de incubação, tempo até a detecção, probabilidade de detecção, tempo até a morte e probabilidade de morte, notamos uma variabilidade significativa entre cada nação, revelando as nuances específicas de cada surto.

A probabilidade de um indivíduo pertencer ao grupo de alto risco apresentou uma variação considerável, sendo 0,27 no Peru e 0,8 nos Estados Unidos, o país com a maior população modelada. Intrigantemente, apesar de ter uma proporção maior de indivíduos de alto risco em relação à sua população, a taxa de infecção do grupo de alto risco nos Estados Unidos foi relativamente baixa, marcando 0,5, enquanto no Peru foi de 4,0. Esses números indicam uma diferença significativa no contato entre os indivíduos do grupo de alto risco e os portadores do vírus, sendo muito mais intenso nos Estados Unidos.

Na representação do compartimento de abstinência neste estudo, onde somente os

indivíduos do grupo suscetível de alto risco têm a capacidade de adotar medidas de isolamento e abstinência, observa-se que isso não pareceu ter impactos significativos na evolução dos casos da doença na maioria dos cenários estudados. Isso se deve ao fato de que a população de alto risco apenas sofre evasão de indivíduos, sem receber novos integrantes ao longo da epidemia. Portanto, em países modelados com um pequeno percentual de indivíduos fazendo parte do grupo de alto risco, como é o caso da Colômbia, não se observa um benefício expressivo com a adoção de políticas de abstinência para essa população.

No entanto, em países com um número elevado de indivíduos no grupo de alto risco e com uma alta probabilidade de abstinência, como nos Estados Unidos, a estratégia de abstinência se mostra eficaz em conter o número de pessoas infectadas durante a epidemia. Isso ocorre porque os indivíduos que têm maior propensão à infecção não são contaminados ao adotarem as medidas de abstinência. Posteriormente, eles transitam para o grupo de baixo risco após o período de isolamento, conforme observado na gráfico, de distribuição da população ao longo da epidemia nos estados unidos, presente no apêndice B. Esse resultado destaca a relevância da abstinência como uma estratégia eficaz para reduzir a propagação da doença, especialmente em contextos onde há uma concentração significativa de indivíduos de alto risco e uma alta adesão a medidas de isolamento e abstinência.

Quanto à taxa de vacinação, notamos uma tendência de ser mais baixa nos países com maior população. A adesão à vacina foi de 0,0001 para o Brasil, Colômbia e Estados Unidos, destacando um padrão intrigante de comportamento de adesão à vacinação em relação ao tamanho populacional.

Ao examinar as datas e os disseminadores iniciais da epidemia em cada rede, especialmente nos países onde foram identificadas múltiplas ondas, como Peru e México, observamos características intrigantes. No México, as duas ondas começaram simultaneamente em redes distintas, um fenômeno que merece uma análise mais detalhada e que é discutida a seguir. No Peru, uma segunda onda epidêmica surgiu 12 dias após o início da primeira. Além disso, o número de disseminadores iniciais da segunda onda foi menor do que na primeira, o que está em consonância com o fato de que o pico da segunda onda foi inferior ao da primeira, conforme evidenciado na seção 5.4. Essas nuances destacam a complexidade da dinâmica de transmissão do mpox, ressaltando a importância de considerar fatores temporais e comportamentais ao modelar e entender surtos epidêmicos em diferentes contextos nacionais.

Os impactos dos valores encontrados para estes parâmetros são graficamente demonstrados no apêndice B. A determinação do número de ondas sobrepostas representa um dos passos primordiais para alcançar nosso objetivo de incorporar essas múltiplas ondas no modelo epidemiológico, usando redes de contato. A abordagem adotada envolveu a contagem de picos na curva epidemiológica suavizada, como ilustrado na Figura 17. No entanto, os resultados da simulação do modelo SPLICE revelam que essa técnica pode não ser suficiente para identificar o número real de ondas. Isso ocorre porque ela pode não conseguir detectar ondas cujos picos foram suprimidos por ondas maiores. Por exemplo, ao analisarmos o caso do México, de acordo com a análise de picos, a segunda onda do surto de mpox deveria ter atingido seu platô no final do período de simulação, após o dia 200. Contudo, a simulação com SPLICE sugere que o México provavelmente enfrentou uma epidemia de onda única. Isso é evidenciado pela variável de delay, que indica o início da epidemia em cada rede. No caso da segunda onda, o valor é igual a 0, indicando que ela começou no mesmo dia que a primeira. Outra possibilidade é que de fato duas ondas tenham atingido o país, ao mesmo tempo, mas em populações, ou redes, diferentes. De modo que as distintas características das redes, tenha feito com que a dinâmica epidemiológica tenha sido diferente nas duas ondas, o que é denotado por exemplo, pela diferença nos picos das ondas, conforme exibido na figura 20 .

Neste sentido, apresentamos na seção C.1 do apêndice C, um estudo prévio da utilização do trabalho aqui apresentado sob a perspectiva de uma outra abordagem de identificação do número de ondas presentes nos surtos epidemiológicos. A saber, com auxílio de especialistas em epidemiologia, fomos capazes de realizar uma análise qualitativa das características das curvas endemiológicas de cada um dos países estudados, para assim determinar o número de ondas presentes em cada uma delas. Conforme descrito no apêndice, os resultados sugerem que adotar essa abordagem em conjunto com os métodos desenvolvidos no trabalho, fornecem como resultado uma descrição muito mais precisa da dinâmica de transmissão do mpox.

Dessa forma, é fundamental que futuras pesquisas se dediquem a investigar técnicas mais aprofundadas para identificar ondas sobrepostas durante o surto de mpox de 2022. Além disso, é essencial caracterizar essas redes conectadas, compreendendo as particularidades das populações envolvidas. Isso nos permitirá entender os fenômenos que estão por trás das disparidades observadas, justificando por que essas redes enfrentaram a epidemia com intensidades variadas e em momentos distintos. Somente através de uma análise detalhada desses fatores

poderemos obter uma compreensão abrangente da dinâmica do surto de mpox nas Américas durante esse período crítico.

No processo de otimização dos parâmetros, baseado nos valores de referência e nos intervalos de busca definidos na subseção 5.5.1, observamos que alguns desses valores e intervalos não foram capazes de capturar adequadamente a dinâmica da epidemia em determinados cenários. Parte desse efeito já era esperado e previsto para certos parâmetros, como a probabilidade de um indivíduo pertencer ao grupo de alto risco, obtida a partir do estudo conduzido por Bragazzi et al. (2023). No entanto, esse parâmetro, embora derivado de dados de um país americano (Canadá), foi limitado aos casos de mpox ocorridos em um intervalo de apenas 7 dias. Portanto, optamos por não considerar o valor de referência como uma constante, mas definimos um intervalo de busca mais amplo para encontrar o valor ótimo do parâmetro, visando mitigar a limitação dos dados.

Ainda assim, esses parâmetros, juntamente com outros, não se ajustaram de forma adequada aos dados utilizados na simulação, conforme evidenciado na tabela 10, que apresenta os valores dos parâmetros após o processo de otimização. Notamos que, em quatro dos cinco países analisados, a taxa de vacinação dos indivíduos do grupo de alto risco atingiu o valor máximo definido no intervalo. Similarmente, em três países, a taxa de vacinação dos indivíduos do grupo de baixo risco foi igual a 0,0005. Além disso, para a probabilidade de abstinência e isolamento, em dois dos cinco países estudados, os valores extremos foram atingidos. Por fim, no caso da probabilidade de um indivíduo pertencer ao grupo de alto risco, em dois países, o valor otimizado correspondeu ao máximo definido no intervalo. Esses resultados levantam a possibilidade de que os valores reais dessas variáveis possam estar situados fora dos intervalos definidos, indicando que o erro mínimo obtido pode não corresponder ao erro mínimo real que o modelo proposto pode atingir. Esta discrepância sugere a necessidade de uma análise mais aprofundada dos dados e da escolha dos intervalos para garantir uma modelagem mais precisa e representativa da dinâmica da epidemia em diferentes contextos.

6.1 ANÁLISE DOS RESULTADOS DA SIMULAÇÃO

A construção do modelo epidemiológico compartimental com redes de contato, destinado a descrever a dinâmica do mpox nas Américas, representa o cerne do nosso trabalho. No capítulo 4, desenvolvemos e detalhamos o nosso modelo, baseando-nos em informações

epidemiológicas e sociais. Este modelo é uma contribuição original, pois não foi previamente proposto por outros pesquisadores; no entanto, é uma composição que se baseia em preceitos já estabelecidos na literatura.

Uma das pressuposições incorporadas neste modelo é a inclusão de um compartimento, denominado estado de isolamento e abstinência, ao qual apenas a parte da população considerada de alto risco tem acesso. Ademais, o modelo considera a possibilidade de transição da população de alto risco para baixo risco por meio do estado de abstinência. No entanto, a maior contribuição reside na capacidade de representar epidemias como a soma dos resultados de iterações que ocorrem em várias redes, com cada rede sendo representada por uma instância do modelo.

Durante o desenvolvimento deste estudo, conduzimos uma análise dos impactos de diferentes estratégias de vacinação para compreender os efeitos decorrentes de maiores taxas de vacinação. No entanto, é importante ressaltar que algumas simplificações foram adotadas ao longo do processo. Não consideramos a possibilidade de indivíduos vacinados se tornarem suscetíveis à doença novamente. Além disso, não exploramos a interação de indivíduos em estado de abstinência com a vacinação.

Outra simplificação adotada está na definição de uma taxa constante de vacinação durante toda a epidemia. Neste sentido, futuros trabalhos podem se concentrar em representar melhor a dinâmica de estratégias de intervenção, como a vacinação, praticadas pelas autoridades de saúde. Isso envolve considerar que essas estratégias variam ao longo do surto, sendo aplicadas em determinados períodos (intervalos de tempo definidos) e com diferentes taxas de adesão

Para mais detalhes sobre as estratégias adotadas neste estudo e os resultados obtidos com as diferentes taxas de vacinação, a seção C.1 do Apêndice C pode ser consultada. Nessa seção, são apresentados aspectos mais específicos referentes à análise das redes de populações e seus efeitos sob diferentes estratégias de vacinação.

Sugerimos que futuras pesquisas aprofundem a investigação sobre os impactos da vacinação, considerando a possibilidade de reinfeção em indivíduos vacinados e explorando a interação entre a abstinência e a vacinação. Essa investigação pode fornecer informações adicionais sobre como esses fatores podem afetar a dinâmica de transmissão de epidemias, utilizando o modelo SPLICE como base.

Com isso em mente, examinamos os resultados da simulação do mpox nas nações americanas, empregando o modelo SPLICE. Uma observação importante diz respeito ao erro obtido ao comparar a simulação com os dados reais e suavizados, conforme a função de erro descrita na subseção 3.1.4. Analisando os resultados apresentados na tabela 11, notamos que, em todos os casos, o erro relacionado à forma, que avalia a capacidade do modelo em reproduzir a forma da curva epidemiológica da doença, é inferior ao erro de distância, que compara os dados previstos com os dados reais. Essa discrepância está associada à dispersão dos casos reais, que apresentam uma grande variação de valores, como mencionado anteriormente. O erro em relação aos dados suavizados sugere que a estimativa do número de ondas ou dos valores dos parâmetros pode não ter sido precisa o suficiente para calibrar o modelo adequadamente.

Ademais, existem outras características relacionadas ao mpox que não são abrangidas pelo modelo. Mudanças abruptas e espontâneas no comportamento social, por exemplo, podem alterar a forma da onda epidemiológica e são difíceis de representar em modelos compartimentais. Eventos pontuais, como feriados ou festividades, também podem ocorrer durante um surto e impactar o número de novos casos. Tais eventos não podem ser previstos pelo modelo, que pode não ser flexível o bastante para se adaptar a essas variações.

Apesar das limitações mencionadas, conseguimos demonstrar a viabilidade de descrever surtos com múltiplas ondas usando modelos compartimentais, sem recorrer a estruturas de reinfecção, como em (GHOSH; GHOSH, 2022). Mesmo com uma taxa de erro associada, conseguimos reproduzir com um nível razoável de fidelidade os episódios epidêmicos de mpox nas nações americanas, como exemplificado nos resultados da simulação no Peru, que inclui duas redes conectadas, representando duas ondas sobrepostas.

Por fim, para mitigar as limitações decorrente das incertezas em relação aos valores-chave e aos pressupostos essenciais utilizados para determinar os parâmetros epidemiológicos do modelo. e aumentar a confiabilidade das descobertas do nosso estudo, optamos por realizar análises de sensibilidade. Essa abordagem nos permitiu explorar uma ampla gama de valores para os parâmetros epidemiológicos, demonstrando a robustez das nossas descobertas diante das variações nos valores desses parâmetros.

Ao submeter o modelo a diferentes cenários, através da análise de sensibilidade, pudemos observar como as variações nos parâmetros impactaram as previsões do modelo. Essa abordagem possibilitou uma compreensão mais abrangente da resposta do modelo em face das

incertezas dos valores dos parâmetros, fortalecendo a confiabilidade e a robustez das conclusões alcançadas.

Dessa forma, ao evidenciar a consistência das descobertas em uma ampla faixa de valores de parâmetros, a análise de sensibilidade desempenhou um papel fundamental na validação e na confirmação da solidez das nossas conclusões, superando as limitações impostas pela incerteza decorrente do surto de mpox de 2022.

Por fim, percebemos que o modelo desenvolvido possui uma capacidade robusta de ajuste e simulação, sendo apto a considerar uma ampla gama de cenários que envolvem redes interconectadas de populações. Isso inclui a habilidade de encontrar soluções eficazes diante da diversidade de combinações de parâmetros, abrangendo desde o número e tamanho das populações até coeficientes de transmissão (beta) e outras variáveis epidemiológicas relevantes mencionadas no estudo. No decorrer da pesquisa, evidenciamos a adaptabilidade do modelo a diversas redes, apresentando dois cenários potenciais para o surto de mpox no Brasil.

Assim, esse modelo se revela uma ferramenta significativa, fornecendo suporte para pesquisadores na área da saúde e epidemiologia para compreender, do ponto de vista epidemiológico, os eventos ocorridos durante o surto de mpox nas regiões investigadas. Sua flexibilidade e capacidade de ajuste a diferentes contextos possibilitam não apenas a análise retrospectiva, mas também a projeção de cenários futuros, fortalecendo a capacidade de resposta diante de eventuais novos surtos ou epidemias.

7 CONSIDERAÇÕES FINAIS

Neste estudo, apresentamos um inovador modelo epidemiológico compartimental diferencial, integrando redes de populações conectadas, para simular epidemias com múltiplas ondas sobrepostas. Nosso modelo foi elaborado para descrever a dinâmica de transmissão do mpox nas Américas durante o surto de 2022, com base nos dados fornecidos pela OPAS. Por meio desses dados, conseguimos extrair informações relevantes sobre a epidemia global nos países americanos, informações essas que não foram exploradas em estudos anteriores que, em sua maioria, focalizaram análises em outros países. Ao longo deste trabalho, também discutimos questões fundamentais relacionadas ao processo de modelagem e estimativa dos parâmetros epidemiológicos que definem a dinâmica de transmissão de doenças como o mpox.

Contudo, observamos uma limitação importante decorrente da escassez de informações nos dados fornecidos pela OPAS, o que afetou diretamente os resultados do modelo. Apesar dessa limitação, conseguimos validar a eficácia da abordagem de redes conectadas para representar epidemias com múltiplas ondas. No entanto, não consideramos diversas formas de relações sociais e comportamentos individuais, fatores essenciais na propagação de doenças epidêmicas.

Futuros estudos podem se dedicar a encontrar maneiras mais eficazes de mitigar o impacto da falta de dados na estimação dos parâmetros epidemiológicos. Além disso, é vital explorar outras técnicas para detectar o número de redes conectadas em epidemias com múltiplas ondas e expandir o modelo proposto para incorporar diferentes tipos de comportamentos individuais. Esta expansão permitiria uma representação mais abrangente e precisa da dinâmica de transmissão do mpox e de outras doenças epidêmicas, levando em conta as nuances das interações sociais e comportamentais dos indivíduos envolvidos.

REFERÊNCIAS

- AL-RAEEI, M. The study of human monkeypox disease in 2022 using the epidemic models: Herd immunity and the basic reproduction number case. **Annals of Medicine and Surgery**, Wolters Kluwer Health, v. 85, n. 2, p. 316, 2023.
- AL-SHOMRANI, M. M.; MUSA, S. S.; YUSUF, A. Unfolding the transmission dynamics of monkeypox virus: An epidemiological modelling analysis. **Mathematics**, MDPI, v. 11, n. 5, p. 1121, 2023.
- ANDERSON, R. M.; MAY, R. M. **Infectious diseases of humans: dynamics and control**. [S.l.]: Oxford university press, 1991.
- AUGSBURGER, I. B.; GALANTHAY, G. K.; TAROSKY, J. H.; RYCHTÁŘ, J.; TAYLOR, D. Voluntary vaccination may not stop monkeypox outbreak: A game-theoretic model. **PLOS Neglected Tropical Diseases**, Public Library of Science San Francisco, CA USA, v. 16, n. 12, p. e0010970, 2022.
- BERGSTRA, J.; BENGIO, Y. Random search for hyper-parameter optimization. **Journal of machine learning research**, v. 13, n. 2, 2012.
- BETTI, M.; FARRELL, L.; HEFFERNAN, J. M. A pair formation model with recovery: Application to monkeypox. **medRxiv**, Cold Spring Harbor Laboratory Press, p. 2022–08, 2022.
- BRAGAZZI, N. L.; HAN, Q.; IYANIWURA, S. A.; OMAME, A.; SHAUSAN, A.; WANG, X.; WOLDEGERIMA, W. A.; WU, J.; KONG, J. D. Adaptive changes in sexual behavior in the high-risk population in response to human monkeypox transmission in canada can help control the outbreak: insights from a two-group, two-route epidemic model. **Journal of Medical Virology**, Wiley Online Library, 2023.
- BRAUER, F.; CASTILLO-CHAVEZ, C.; FENG, Z. **Mathematical models in epidemiology**. [S.l.]: Springer, 2019. v. 32.
- CANTO, F. J. A.; AVILA-VALES, E. J. Fitting parameters of seir and sird models of covid-19 pandemic in mexico. **Preprint**, p. 1–11, 2020.
- CARRIÓN, E. F. Mathematical models for introduction, spread and early detection on infectious diseases in veterinary epidemiology. Universidad Complutense de Madrid, 2017.
- Center for Disease Control and Prevention. **Technical Report 3: Multi-National Mpox Outbreak, United States, 2022**. 2023. Accessed on Jun 25, 2023. Disponível em: <<https://www.cdc.gov/poxvirus/mpox/cases-data/technical-report/report-3.html>>.
- Centers for Disease Control and Prevention. **Global Distribution of Monkeypox Outbreaks, 2022**. 2022. Accessed on May 1, 2023. Disponível em: <<https://www.cdc.gov/poxvirus/mpox/response/2022/world-map.html>>.
- CNN (Cable News Network). **WHO says mpox is no longer a global health emergency**. 2023. Accessed on May 23, 2023. Disponível em: <<https://edition.cnn.com/2023/05/11/health/mpox-global-health-emergency-who/index.html>>.

CONTROL, C. for D.; PREVENTION. **Key Characteristics for Identifying Mpox**. 2023. Acessado em: 23/10/2023. Disponível em: <<https://www.cdc.gov/poxvirus/mpox/clinicians/clinical-recognition.html>>.

CONTROL, C. for D.; PREVENTION. **What to Do If You Are Sick**. 2023. Acessado em: 23/10/2023. Disponível em: <<https://www.cdc.gov/poxvirus/mpox/if-sick/what-to-do.html>>.

CORTÉS-MARTÍNEZ, K. V.; ESTRADA-ESQUIVEL, H.; MARTÍNEZ-REBOLLAR, A.; HERNÁNDEZ-PÉREZ, Y.; ORTIZ-HERNÁNDEZ, J. The state of the art of data mining algorithms for predicting the covid-19 pandemic. **Axioms**, MDPI, v. 11, n. 5, p. 242, 2022.

DEMONGEOT, J.; GRIETTE, Q.; MAGAL, P. Si epidemic model applied to covid-19 data in mainland china. **Royal Society Open Science**, The Royal Society, v. 7, n. 12, p. 201878, 2020.

DIEKMANN, O.; HEESTERBEEK, J. A. P.; METZ, J. A. On the definition and the computation of the basic reproduction ratio r_0 in models for infectious diseases in heterogeneous populations. **Journal of mathematical biology**, Springer, v. 28, p. 365–382, 1990.

DOU, Y.-M.; YUAN, H.; TIAN, H.-W. Monkeypox virus: past and present. **World Journal of Pediatrics**, Springer, v. 19, n. 3, p. 224–230, 2023.

FINE, P.; JEZEK, Z.; GRAB, B.; DIXON, H. The transmission potential of monkeypox virus in human populations. **International journal of epidemiology**, Oxford University Press, v. 17, n. 3, p. 643–650, 1988.

GAO, J.; ZHOU, C.; LIANG, H.; JIAO, R.; WHEELLOCK, A. M.; JIAO, K.; MA, J.; ZHANG, C.; GUO, Y.; LUO, S. et al. Monkeypox outbreaks in the context of the covid-19 pandemic: Network and clustering analyses of global risks and modified seir prediction of epidemic trends. **Frontiers in public health**, Frontiers, v. 11, p. 41, 2023.

GARNER, M.; HAMILTON, S. et al. Principles of epidemiological modelling. **Revue Scientifique et Technique-OIE**, v. 30, n. 2, p. 407, 2011.

GESSAIN, A.; NAKOUNE, E.; YAZDANPANAHA, Y. Monkeypox. **New England Journal of Medicine**, Mass Medical Soc, v. 387, n. 19, p. 1783–1793, 2022.

GHOSH, K.; GHOSH, A. K. Study of covid-19 epidemiological evolution in india with a multi-wave sir model. **Nonlinear Dynamics**, Springer, v. 109, n. 1, p. 47–55, 2022.

HARRIS, C. R.; MILLMAN, K. J.; WALT, S. J. van der; GOMMERS, R.; VIRTANEN, P.; COURNAPEAU, D.; WIESER, E.; TAYLOR, J.; BERG, S.; SMITH, N. J.; KERN, R.; PICUS, M.; HOYER, S.; KERKWIJK, M. H. van; BRETT, M.; HALDANE, A.; RÍO, J. F. del; WIEBE, M.; PETERSON, P.; GÉRARD-MARCHANT, P.; SHEPPARD, K.; REDDY, T.; WECKESSER, W.; ABBASI, H.; GOHLKE, C.; OLIPHANT, T. E. Array programming with NumPy. **Nature**, Springer Science and Business Media LLC, v. 585, n. 7825, p. 357–362, set. 2020. Disponível em: <<https://doi.org/10.1038/s41586-020-2649-2>>.

HETHCOTE, H. W. The mathematics of infectious diseases. **SIAM review**, SIAM, v. 42, n. 4, p. 599–653, 2000.

HEVNER, A.; CHATTERJEE, S.; HEVNER, A.; CHATTERJEE, S. Design science research in information systems. **Design research in information systems: theory and practice**, Springer, p. 9–22, 2010.

HUNTER, J. D. Matplotlib: A 2d graphics environment. **Computing in Science & Engineering**, IEEE COMPUTER SOC, v. 9, n. 3, p. 90–95, 2007.

KERMACK, W. O.; MCKENDRICK, A. G. A contribution to the mathematical theory of epidemics. **Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character**, The Royal Society London, v. 115, n. 772, p. 700–721, 1927.

KERMACK, W. O.; MCKENDRICK, A. G. Contributions to the mathematical theory of epidemics. ii.—the problem of endemicity. **Proceedings of the Royal Society of London. Series A, containing papers of a mathematical and physical character**, The Royal Society London, v. 138, n. 834, p. 55–83, 1932.

KLUGE, H.; AMMON, A. Monkeypox in europe and beyond—tackling a neglected disease together. **Eurosurveillance**, European Centre for Disease Prevention and Control, v. 27, n. 24, p. 2200482, 2022.

KNIGHT, J.; TAN, D. H.; MISHRA, S. Maximizing the impact of limited vaccine supply under different early epidemic conditions: a 2-city modelling analysis of monkeypox virus transmission among men who have sex with men. **CMAJ**, Can Med Assoc, v. 194, n. 46, p. E1560–E1567, 2022.

KOROSEC, C.; BANUET-MARTINEZ, M.; YANG, Y.; AL. et. Monkeypox: A review of epidemiological modelling studies and how modelling has led to mechanistic insight. **PREPRINT (Version 1) available at Research Square**, February 2023.

KREUTZ, L.; REZENDE, M.; MATÉ, Y. Varíola dos macacos (monkeypox virus-poxviridae): Uma breve revisão. **Ars Veterinaria**, v. 38, n. 3, p. 111–115, 2022.

KUPFERSCHMIDT, K. **Why the Monkeypox Outbreak Is Mostly Affecting Men Who Have Sex with Men**, **Science**. 2022. Accessed on May 1, 2023.

LAUER, S. A.; GRANTZ, K. H.; BI, Q.; JONES, F. K.; ZHENG, Q.; MEREDITH, H. R.; AZMAN, A. S.; REICH, N. G.; LESSLER, J. The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: estimation and application. **Annals of internal medicine**, American College of Physicians, v. 172, n. 9, p. 577–582, 2020.

LIAO, L.-C.; HSU, C.-Y.; CHEN, H.-H.; LAI, C.-C. Estimating the global spread of epidemic human monkeypox with bayesian directed acyclic graphic model. **Vaccines**, MDPI, v. 11, n. 2, p. 468, 2023.

LÓPEZ-FLORES, M. M.; MARCHESIN, D.; MATOS, V.; SCHECTER, S. Differential equation models in epidemiology. 2021.

MAGAL, P.; SEYDI, O.; WEBB, G. Final size of a multi-group sir epidemic model: Irreducible and non-irreducible modes of transmission. **Mathematical biosciences**, Elsevier, v. 301, p. 59–67, 2018.

MAJEE, S.; JANA, S.; KAR, T. Dynamical analysis of monkeypox transmission incorporating optimal vaccination and treatment with cost-effectiveness. **Chaos: An Interdisciplinary Journal of Nonlinear Science**, AIP Publishing LLC, v. 33, n. 4, p. 043103, 2023.

MATHIEU, E.; SPOONER, F.; DATTANI, S.; RITCHIE, H.; ROSER, M. Mpox (monkeypox). **Our World in Data**, 2022. <https://ourworldindata.org/monkeypox>.

MOLLA, J.; SEKKAK, I.; ORTIZ, A. M.; MOYLES, I.; NASRI, B. Mathematical modeling of mpox: a scoping review. **One Health**, Elsevier, p. 100540, 2023.

MOLLISON, D.; ISHAM, V.; GRENFELL, B. Epidemics: models and data. **Journal of the Royal Statistical Society Series A: Statistics in Society**, Oxford University Press, v. 157, n. 1, p. 115–129, 1994.

MURRAY, J. D. **Mathematical biology: I. An introduction**. [S.l.]: Springer, 2002.

NAKOUNE, E.; OLLIARO, P. **Waking up to monkeypox**. [S.l.]: British Medical Journal Publishing Group, 2022.

NGUNGU, M.; ADDAI, E.; ADENIJI, A.; ADAM, U. M.; OSHINUBI, K. Mathematical epidemiological modeling and analysis of monkeypox dynamism with non-pharmaceutical intervention using real data from united kingdom. **Frontiers in Public Health**, Frontiers Media SA, v. 11, 2023.

NORTH, G. Data analysis | time series analysis. In: HOLTON, J. R. (Ed.). **Encyclopedia of Atmospheric Sciences**. Oxford: Academic Press, 2003. p. 621–624. ISBN 978-0-12-227090-1. Disponível em: <<https://www.sciencedirect.com/science/article/pii/B0122270908001317>>.

PASTOR-SATORRAS, R.; CASTELLANO, C.; MIEGHEM, P. V.; VESPIGNANI, A. Epidemic processes in complex networks. **Reviews of modern physics**, APS, v. 87, n. 3, p. 925, 2015.

PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, E. Scikit-learn: Machine learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011.

PEFFERS, K.; TUUNANEN, T.; ROTHENBERGER, M. A.; CHATTERJEE, S. A design science research methodology for information systems research. **Journal of management information systems**, Taylor & Francis, v. 24, n. 3, p. 45–77, 2007.

PERAKIS, G.; SINGHVI, D.; LAMI, O. S.; THAYAPARAN, L. Covid-19: A multiwave sir-based model for learning waves. **Production and Operations Management**, Wiley Online Library, v. 32, n. 5, p. 1471–1489, 2023.

RHODES, T.; LANCASTER, K.; ROSENGARTEN, M. **A model society: maths, models and expertise in viral outbreaks**. [S.l.]: Taylor & Francis, 2020. 253–256 p.

RISER, A. P.; HANLEY, A.; CIMA, M.; LEWIS, L.; SAADEH, K.; ALARCÓN, J.; FINN, L.; KIM, M.; ADAMS, J.; HOLT, D. et al. Epidemiologic and clinical features of mpox-associated deaths—united states, may 10, 2022–march 7, 2023. **Morbidity and Mortality Weekly Report**, Centers for Disease Control and Prevention, v. 72, n. 15, p. 404, 2023.

ROSSUM, G. V.; DRAKE, F. L. **Python 3 Reference Manual**. Scotts Valley, CA: CreateSpace, 2009. ISBN 1441412697.

SALJE, H.; KIEM, C. T.; LEFRANCQ, N.; COURTEJOIE, N.; BOSETTI, P.; PAIREAU, J.; ANDRONICO, A.; HOZÉ, N.; RICHET, J.; DUBOST, C.-L. et al. Estimating the burden of sars-cov-2 in france. **Science**, American Association for the Advancement of Science, v. 369, n. 6500, p. 208–211, 2020.

SCHOENBACH, V. J. et al. **Understanding the fundamentals of epidemiology**. [S.l.]: University of North Carolina at Chapel Hill, 1999.

Scikit-Learn Contributors. **GridSearchCV - Scikit-Learn Documentation**. 2023. Acessado em: 07/10/2023. Disponível em: <https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html>.

SOFONEA, M. T.; CAUCHEMEZ, S.; BOËLLE, P.-Y. Epidemic models: why and how to use them. **Anaesthesia, critical care & pain medicine**, Elsevier, v. 41, n. 2, p. 101048, 2022.

STOCHIȚOIU, R. D.; PETRICA, M.; REBEDEA, T.; POPESCU, I.; LEORDEANU, M. A self-supervised neural-analytic method to predict the evolution of covid-19 in romania. **arXiv preprint arXiv:2006.12926**, 2020.

TEAM, T. pandas development. **pandas-dev/pandas: Pandas**. Zenodo, 2020. Disponível em: <<https://doi.org/10.5281/zenodo.3509134>>.

TITANJI, B. K.; TEGOMOH, B.; NEMATOLLAHI, S.; KONOMOS, M.; KULKARNI, P. A. Monkeypox: a contemporary review for healthcare professionals. In: OXFORD UNIVERSITY PRESS. **Open forum infectious diseases**. [S.l.], 2022. v. 9, n. 7, p. ofac310.

World Health Organization. **Multi-country monkeypox outbreak in non-endemic countries: Update, 2022**. 2022. Accessed on Jun 25, 2023. Disponível em: <<https://www.who.int/emergencies/disease-outbreak-news/item/2022-DON388>>.

World Health Organization. **WHO recommends new name for monkeypox disease, 2022**. 2022. Accessed on May 1, 2023. Disponível em: <<https://www.cdc.gov/poxvirus/mpox/response/2022/world-map.html>>.

WORMSER, G. P.; POURBOHLOUL, B. **Modeling Infectious Diseases in Humans and Animals By Matthew James Keeling and Pejman Rohani Princeton, NJ: Princeton University Press, 2008. 408 pp., Illustrated. \$65.00 (hardcover)**. [S.l.]: The University of Chicago Press, 2008.

WU, J.; DHINGRA, R.; GAMBHIR, M.; REMAIS, J. V. Sensitivity analysis of infectious disease models: methods, advances and their application. **Journal of The Royal Society Interface**, The Royal Society, v. 10, n. 86, p. 20121018, 2013.

XIRIDOU, M.; MIURA, F.; ADAM, P.; COUL, E. Op de; WIT, J. de; WALLINGA, J. The fading of the mpox outbreak among men who have sex with men: a mathematical modelling study. **medRxiv**, Cold Spring Harbor Laboratory Press, p. 2023–01, 2023.

YANG, S.; GUO, X.; ZHAO, Z.; ABUDUNAIBI, B.; ZHAO, Y.; RUI, J.; WANG, Y.; SONG, W.; WEI, H.; CHEN, T. Possibility of mpox viral transmission and control from high-risk to the general population: a modeling study. **BMC Infectious Diseases**, Springer, v. 23, n. 1, p. 119, 2023.

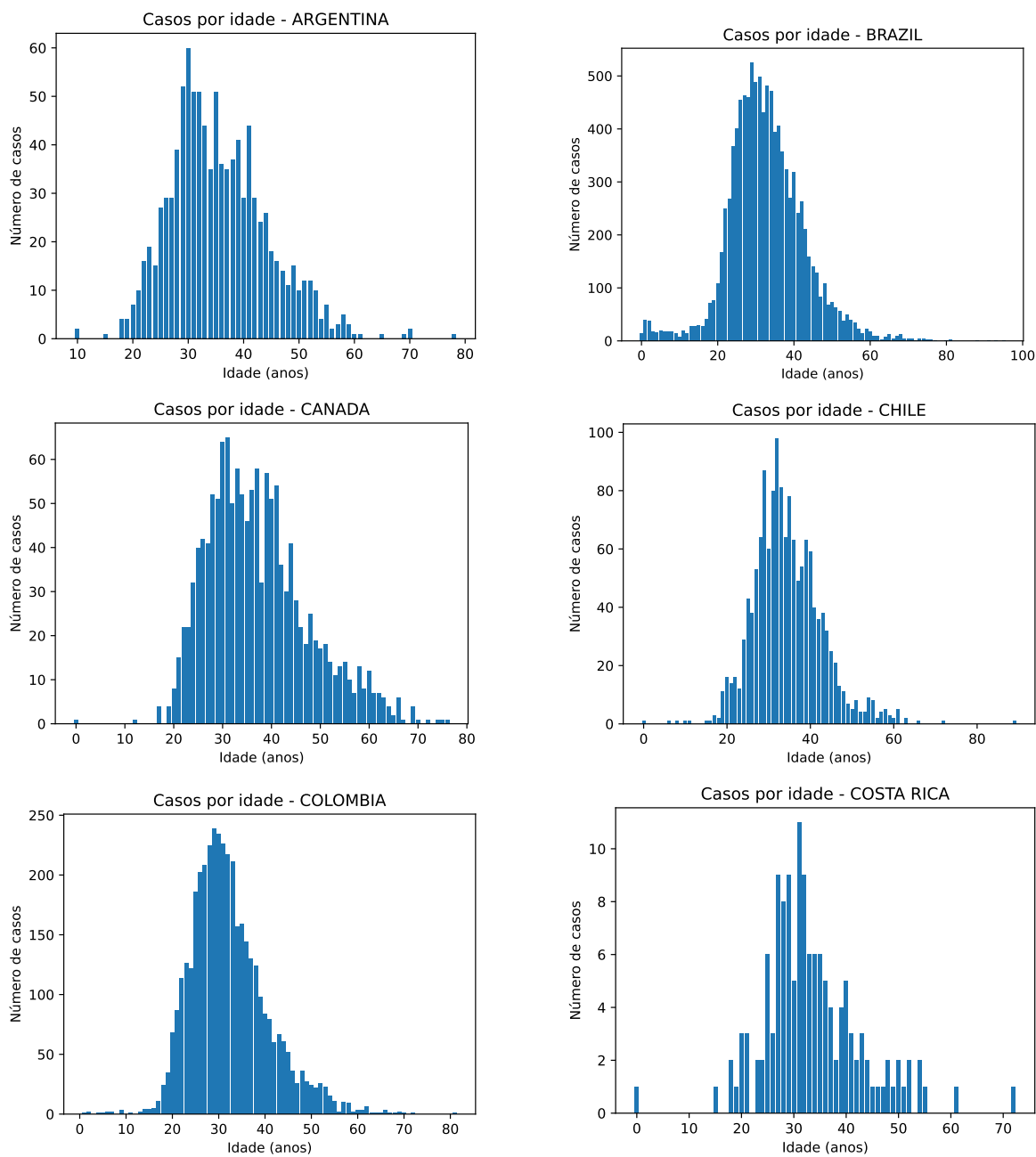
YUAN, P.; TAN, Y.; YANG, L.; ARUFFO, E.; OGDEN, N. H.; BELAIR, J.; ARINO, J.; HEFFERNAN, J. M.; WATMOUGH, J.; CARABIN, H. et al. Modelling vaccination and control strategies of outbreaks of monkeypox at gatherings. **Frontiers in Public Health**, Frontiers, p. 4485, 2022.

YUAN, P.; TAN, Y.; YANG, L.; ARUFFO, E.; OGDEN, N. H.; BELAIR, J.; HEFFERNAN, J.; ARINO, J.; WATMOUGH, J.; CARABIN, H. et al. Assessing transmission risks and control strategy for monkeypox as an emerging zoonosis in a metropolitan area. **Journal of Medical Virology**, Wiley Online Library, v. 95, n. 1, p. e28137, 2023.

APÊNDICES

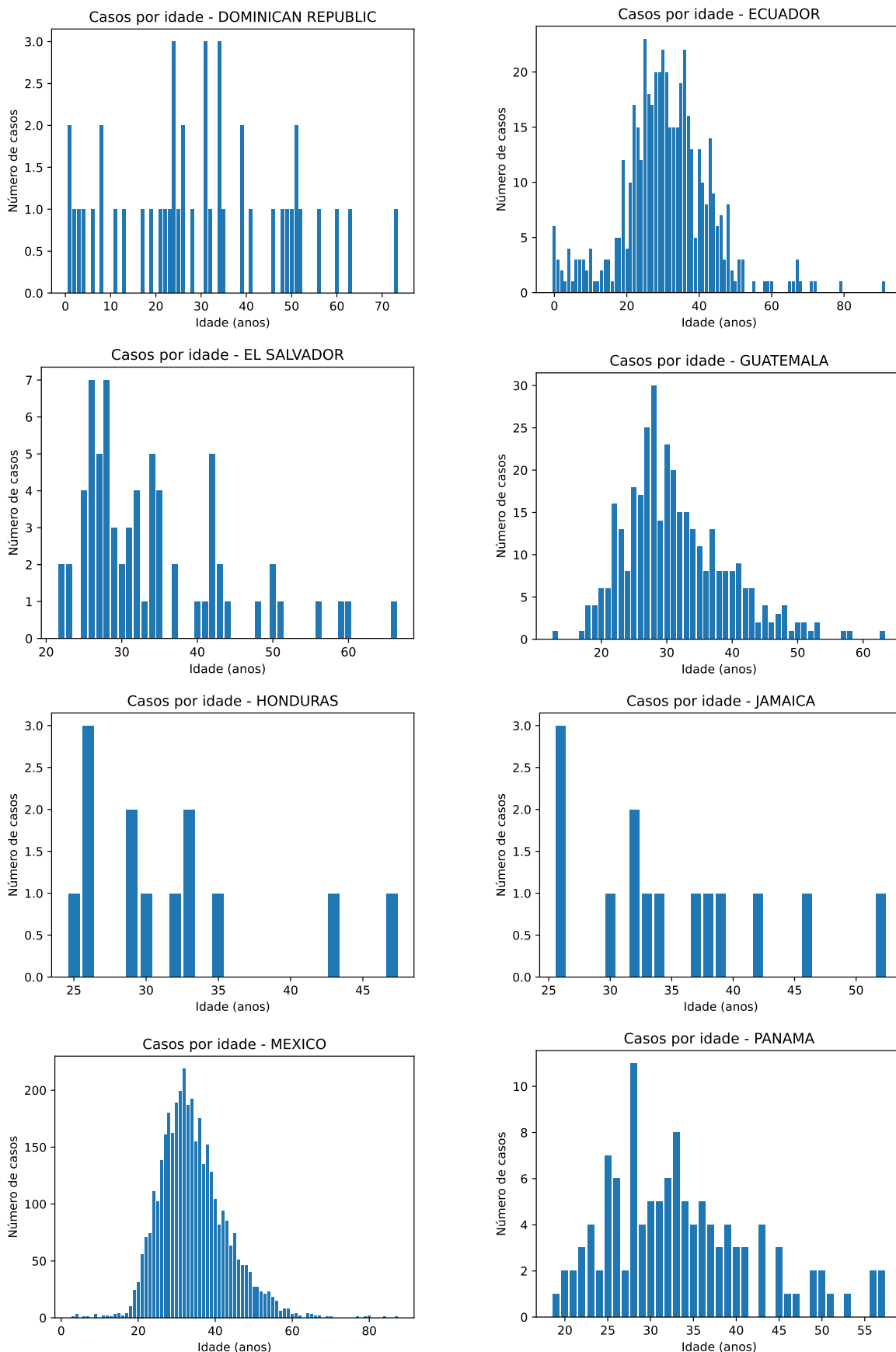
APÊNDICE A – Distribuição de casos por faixa etária nos países das Américas

Figura 24 – Número de casos por idade dos países das Américas.



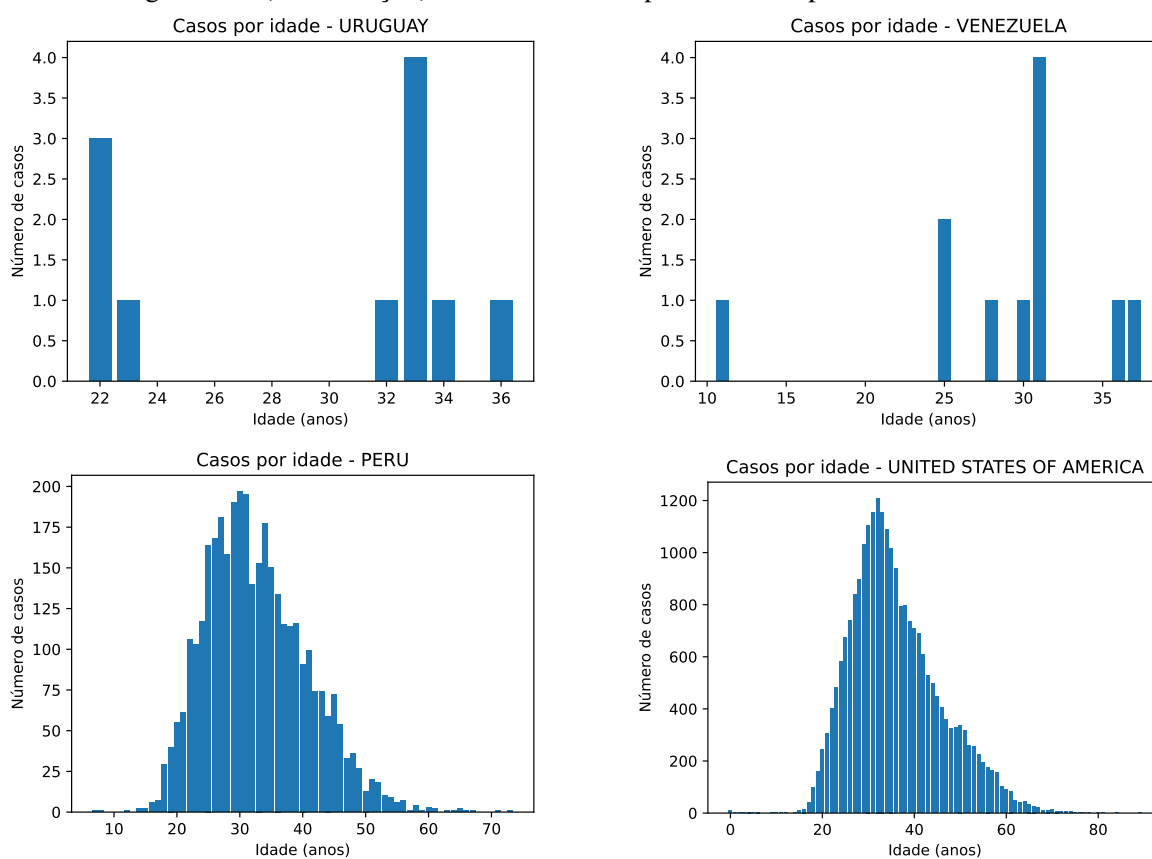
Fonte: O autor

Figura 25 – (Continuação) Número de casos por idade dos países das Américas.



Fonte: O autor

Figura 26 – (Continuação) Número de casos por idade dos países das Américas.

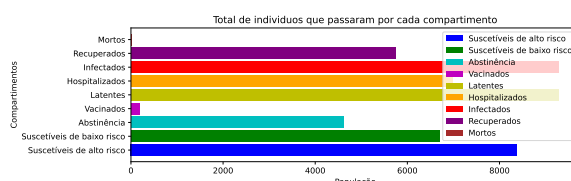
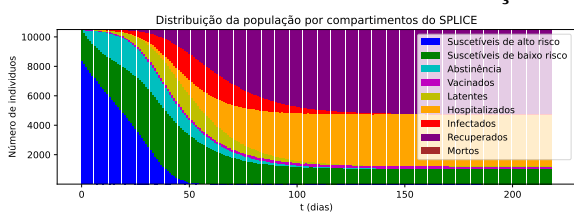


Fonte: O autor

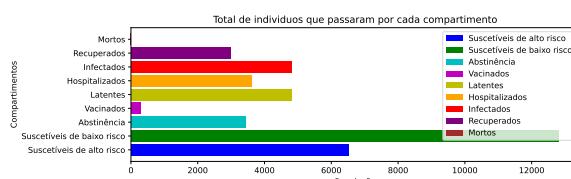
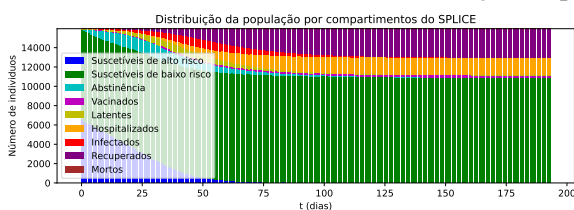
APÊNDICE B – Distribuição da população por compartimentos do SPLICE

As ilustrações abaixo, denotam a evolução da distribuição populacional entre os compartimentos do modelo proposto, ao longo do tempo. E a quantidade de indivíduos total que passou por cada um dos compartimentos para os países estudados. Adiante exibimos a evolução

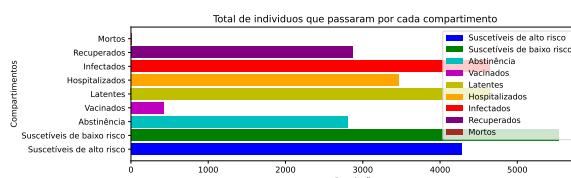
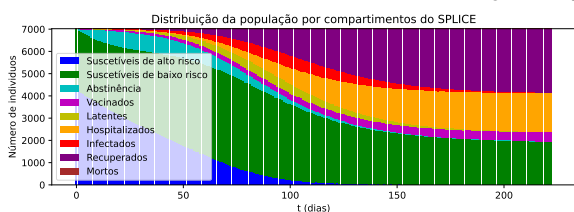
Distribuição da população no Brasil



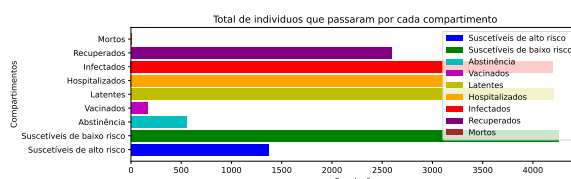
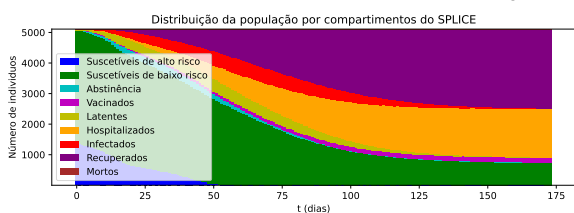
Distribuição da população na Colômbia



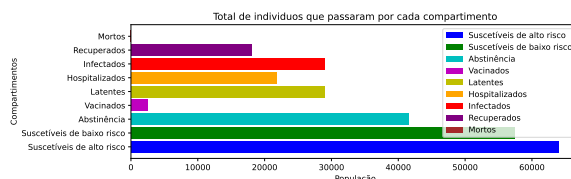
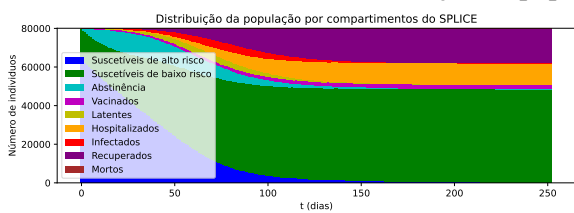
Distribuição da população no México



Distribuição da população no Peru



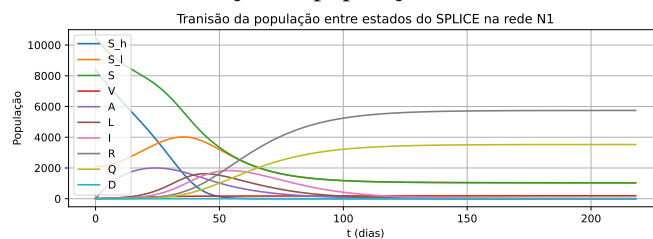
Distribuição da população nos Estados Unidos



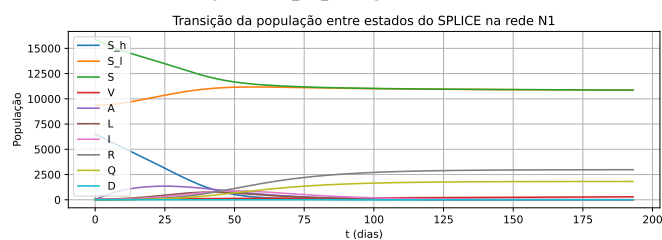
Fonte: O autor

da epidemia sob a perspectiva de cada um dos compartimentos utilizados no modelo, para as redes de cada país.

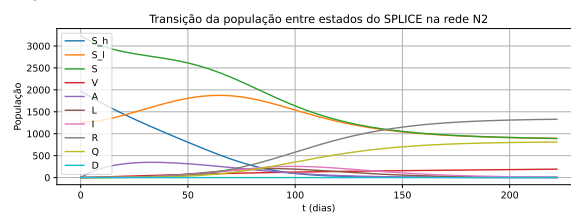
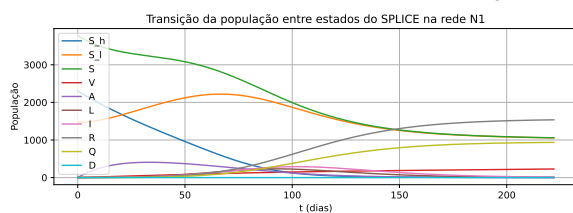
Distribuição da população no Brasil



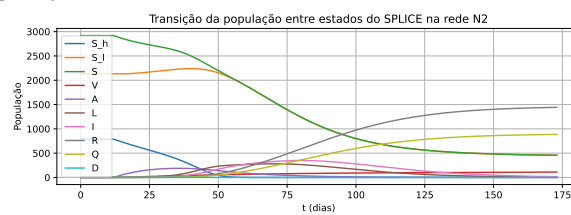
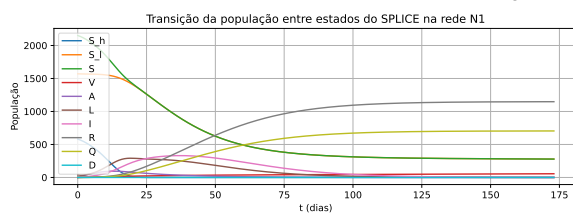
Distribuição da população na Colômbia



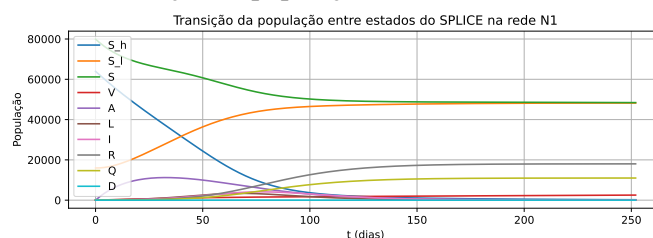
Distribuição da população no México



Distribuição da população no Peru



Distribuição da população nos Estados Unidos



Fonte: O autor

APÊNDICE C – Explorando Novos Cenários:

C.1 AVALIANDO DIFERENTES CENÁRIOS DE REDES CONECTADAS

Durante a análise dos resultados da simulação e da sensibilidade, observamos que, em alguns casos (países), o modelo apresentou uma tendência de subestimar os casos, especialmente no início e/ou final do surto. Isso se evidenciou pela discrepância entre os casos previstos e os casos reais nessas extremidades da curva epidemiológica.

Essa discrepância pode ter decorrido de uma estimativa equivocada do número de redes de populações ou de ondas sobrepostas, como já discutido em seções anteriores. Nesta etapa, apresentamos os resultados considerando diferentes cenários epidêmicos para o Brasil e o Peru. Nosso objetivo foi repetir o processo de descoberta e otimização dos valores dos parâmetros, além da simulação e análise de sensibilidade para esses dois países. Contudo, adotamos uma nova abordagem para determinar o número de ondas do surto.

A identificação do número de ondas na propagação de uma epidemia foi realizada por um especialista humano, por meio de uma análise minuciosa da forma da curva suavizada da epidemia. Esse processo envolveu uma avaliação detalhada da tendência e dos padrões observados na curva temporal, considerando aspectos como a evolução temporal dos casos, variações significativas na taxa de contágio e possíveis períodos de declínio e ressurgimento da doença.

É relevante destacar que essa análise especializada foi conduzida com base em critérios objetivos e conhecimento especializado na área epidemiológica. O especialista considerou diversos indicadores visuais e qualitativos presentes na curva suavizada, evitando subjetividades no processo. Como resultado o novo número de ondas para os dois países, são exibidas na tabela 13 abaixo:

| País | Número de Redes |
|--------|-----------------|
| Brasil | 2 |
| Peru | 4 |

Tabela 13 – cenários das redes de populações conectadas nas Américas

Os novos parâmetros associados a cada rede n do surto nos países é exibido na tabela 14 abaixo:

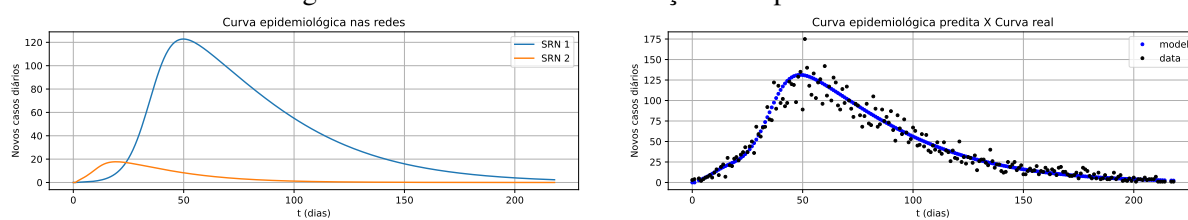
Tabela 14 – Parâmetros vinculados as redes conectadas resultantes

| Parâmetro de cada rede N | Brasil | Peru |
|--------------------------|------------------------------------|--|
| Tamanho da população | N1:28000 N2: 1750 | N1:5110 N2: 4380 N3: 1460 N4: 730 |
| Início do surto | N1:0 N2: 0 | N1:0 N2: 0 N3: 52 N4: 147 |
| Disseminadores iniciais | N1:13 N2: 153 | N1:29 N2: 1 N3: 9 N4: 1 |

Fonte: O autor

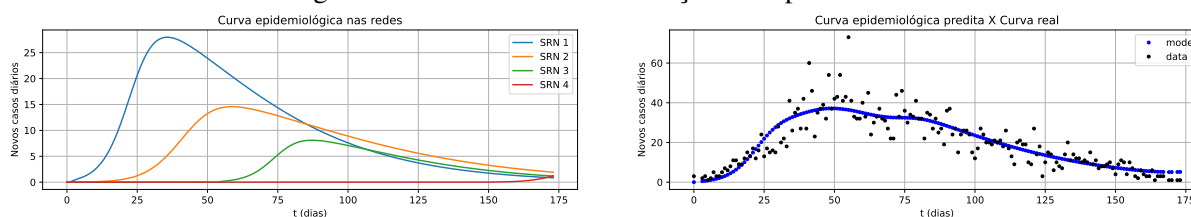
Os resultados das simulações e da análise de sensibilidade, realizadas com o modelo proposto, confirmaram as observações feitas pelo especialista. Ficou evidente que a abordagem utilizada pelo especialista na identificação do número de ondas da epidemia demonstrou ser mais precisa e confiável em comparação ao método anteriormente adotado, baseado na contagem de picos na série temporal suavizada.

Figura 27 – Resultados da simulação do mpox no Brasil



Fonte: O autor

Figura 28 – Resultados da simulação do mpox no Peru

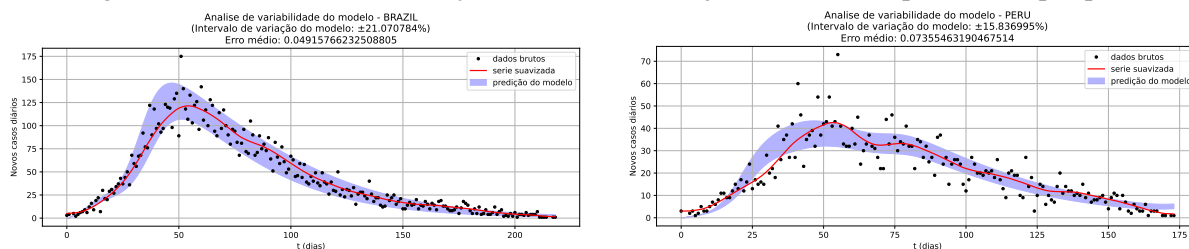


Fonte: O autor

Através da análise de sensibilidade, foi possível observar como o modelo, elaborado

com base nas suposições de especialistas, conseguiu representar de maneira mais precisa a curva epidemiológica. Este modelo, por meio de suas previsões, abarcou um número maior de casos em comparação à abordagem anterior de determinação de ondas. Na figura 29, é evidente que o modelo se ajustou de forma mais acurada, não apresentando discrepâncias significativas no número de casos nas extremidades da onda.

Figura 29 – Resultado da simulação considerando conjunto variado de parâmetros por país



onda epidemiológica resultante da simulação de 20 combinações dos parâmetros ótimos, variando 50% dos parâmetros a uma taxa de $\pm 15\%$

Fonte: O autor

O erro associado a simulação realizada com a melhor combinação de parâmetros (tabela 15), demonstram como de fato, como o modelo foi capaz de gerar previsões mais próximas da realidade. Esta diferença é muito relevante no caso do Brasil, onde o erro decresceu consideravelmente. Para o Peru, por sua vez, não houve melhoria significativa uma vez que a serie temporal de casos brutos do Peru, contem vários pontos dispersos.

Tabela 15 – Taxa de erro da simulação para os países selecionados

| Métrica de Erro | Brasil | Peru |
|---------------------------------------|--------|--------|
| Erro de Forma (ShLoss) | 0,0004 | 0,0010 |
| Erro de Distância (DtLoss) | 0,0007 | 0,0013 |
| Erro Total | 0,0332 | 0,0610 |

Fonte: O autor

Ao comparar os gráficos gerados pelo modelo com os dados reais, constatou-se uma maior proximidade entre as previsões do modelo e a realidade quando a determinação do número de ondas foi realizada com base na análise especializada do especialista. Isso evidenciou uma maior eficácia e precisão dessa abordagem, comparada à abordagem inicialmente adotada no estudo, na identificação do número de ondas e na representação da dinâmica epidêmica.

C.2 AVALIANDO IMPACTOS DE DIFERENTES ESTRATÉGIAS DE VACINAÇÃO

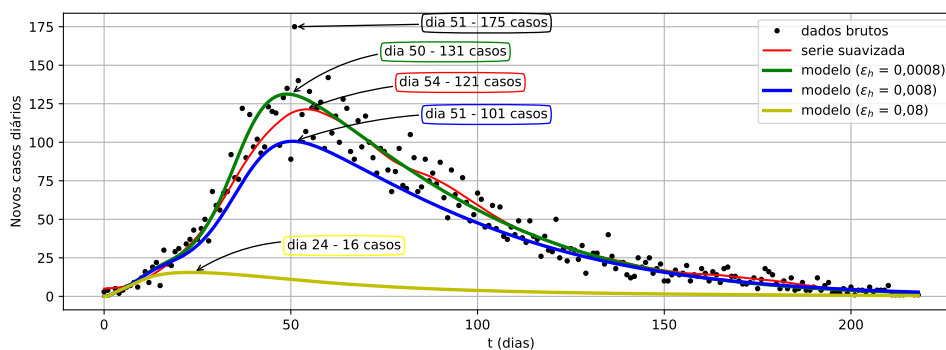
Nesta seção, objetivamos avaliar os impactos resultantes da adoção de diferentes estratégias de vacinação, considerando o modelo proposto.

Conforme discutido no capítulo de discussão de resultados (Capítulo 6), a taxa diária de vacinação no grupo de alto risco (ϵ_h), mesmo atingindo seu valor máximo estipulado (0,0008), não demonstrou impactos significativos no número de indivíduos afetados pela doença.

Agora, exploraremos os impactos decorrentes da aplicação de estratégias de vacinação que aumentem exponencialmente a taxa de vacinação do grupo de alto risco, com uma taxa variando de 10^n vezes, onde $0 \geq n < 3$. Para isso, realizamos novamente a simulação para o Brasil, como estudo de caso, considerando um cenário com duas redes de populações conectadas, o qual se mostrou mais apropriado do que o cenário de apenas uma onda, conforme discutido na seção anterior.

Para avaliar o impacto desse aumento na taxa de vacinação, calculamos a cobertura percentual da epidemia, ou seja, a proporção da população exposta à doença. Isso nos permite compreender os benefícios de incentivar a vacinação como medida de controle de novos surtos. Além disto no gráfico presente na figura 30, podemos observar a evolução diária de casos da epidemia, para cada um dos cenários simulados, considerando as diferentes taxas de vacinação. Além disto, exibimos também o pico da curva epidêmica e os valores a ele associado. Isto significa quando e quanto foi o maior número de casos diário observados em cada cenário, o que também nos permite avaliar os impactos da vacinação, com cada cobertura vacinal imposta.

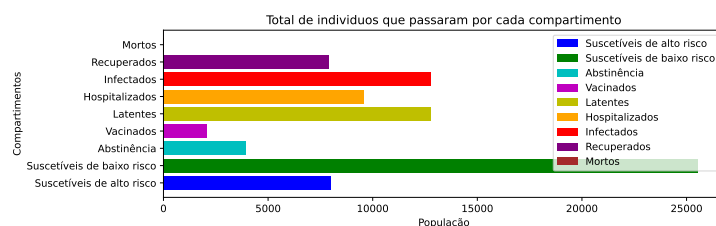
Figura 30 – Resultado das simulações dos diferentes cenários de taxa de vacinação



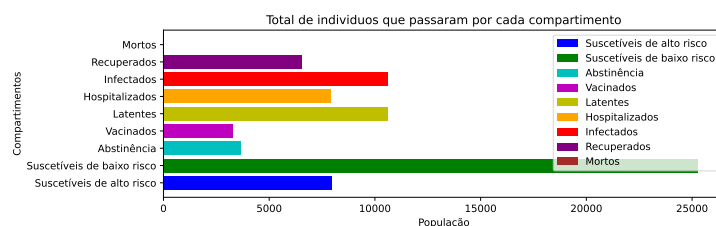
Fonte: O autor

Nos gráficos da Figura 31, observamos a distribuição dos indivíduos em cada com-

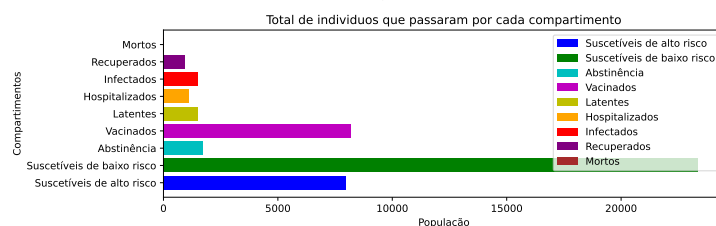
partimento do modelo, considerando a epidemia como um todo, ou seja, a soma dos resultados obtidos para cada uma das duas redes.



(a) Distribuição da população considerando a taxa de vacinação (ϵ_h) = 0,0008



(b) Distribuição da população considerando a taxa de vacinação (ϵ_h) = 0,008



(c) Distribuição da população considerando a taxa de vacinação (ϵ_h) = 0,08

Figura 31 – Distribuição da população entre os compartimentos para cada cenário de vacinação estudado.

Ao analisar esses gráficos, notamos uma redução significativa no número de indivíduos nos estágios latente ou infectado. Entretanto, ao examinar a tabela 16, percebemos que, embora o número total de indivíduos afetados pela epidemia tenha diminuído considerando a população total, essa redução foi proporcionalmente menor na segunda rede populacional.

Tabela 16 – Cobertura da epidemia para cada cenário de vacinação estudado

| ϵ_h \ Redes | N1 | N2 | Total |
|----------------------|--------|--------|--------|
| 0,0008 | 41,80% | 58,52% | 42,78% |
| 0,008 | 34,13% | 56,66% | 35,45% |
| 0,08 | 2,40% | 42,09% | 4,73% |

Fonte: O autor

O fenômeno descrito no parágrafo anterior pode ser compreendido através da análise das características da segunda rede. Como evidenciado na Tabela 14, a Rede N2 apresenta uma proporção de força de infecção-população (E_0/N) superior à da primeira rede. Na primeira rede,

o número inicial de infectados representa 0,05% da população, enquanto na segunda rede, esse número é de 8,74%. Portanto, mesmo aumentando o número de indivíduos vacinados, a redução no número de indivíduos contaminados, nesta rede, não é significativa. Em número absolutos o número total de casos obtidos em cada cenário, e comparados com os casos reais e a série suavizada, são apresentadas na tabela 17, abaixo.

Tabela 17 – Total de casos em cada cenário de vacinação estudado

| Cenário \ Redes | N1 | N2 | Total |
|-----------------------|------|-----|-------|
| Dados brutos | - | - | 9418 |
| Dados Suavizados | - | - | 9421 |
| $\epsilon_h = 0,0008$ | 8768 | 818 | 9586 |
| $\epsilon_h = 0,008$ | 7149 | 795 | 7944 |
| $\epsilon_h = 0,08$ | 501 | 620 | 1127 |

Fonte: O autor

Adicionalmente, é importante destacar que, apesar de termos aumentado progressivamente a taxa de vacinação nesta análise, essa intervenção foi realizada sem a alteração da taxa de vacinação no grupo de baixo risco. Isso permitiu que a doença se propagasse entre os indivíduos desse grupo, embora a uma taxa muito menor. Essa condição também explica a atenuação na curva epidemiológica apresentada na Figura 30. Por exemplo, quando a taxa diária de vacinação no grupo de alto risco foi de 0,08, a epidemia atingiu seu pico antecipadamente em comparação com os outros cenários, amortecendo o avanço da epidemia ao longo dos dias.

APÊNDICE D – Script python de otimização dos parâmetros, simulação e avaliação do modelo

D.0.0.0.1 SPLICE MPX SIMULATOR - PAHO ACADEMIC STUDY

- SPONSORED BY: *PAHO*
- DEVELOPED BY: *DIEGO FRIAS, PhD.*
- CONTRIBUTED BY: *VITOR SANTOS*
- AFFILIATION: *BAHIA STATE UNIVERSITY, UNEB, SALVADOR, BAHIA, BRAZIL*
- DATE OF RELEASE: *SEPTEMBER/2023*

D.0.0.0.2 IMPORTING PYTHON LIBRARIES

```
1 import os
2 import uuid
3 import time
4 import glob
5 import random
6 import copy as cp
7 import numpy as np
8 import pandas as pd
9 import datetime as dt
10 from pandas import read_csv
11 from matplotlib import pyplot as plt
```

D.0.0.0.3 SMOOTHED TIME SERIES PREPROCESSING FUNCTIONS

```
1 def GetInflex(xdata, y, inflexstep, thres):
2
3     step = inflexstep
4
5     N = len(y)
```

```
6
7     InflexTarget = []
8     iInflexTarget = []
9
10    xinflex2 = []
11    for n in range(2*step, N-2*step):
12        aux = (y[n-2*step]-2*y[n-step]+y[n])*(y[n]-2*y[n+
13            step]+y[n+2*step])
14        if aux < 0 and np.abs(aux) > thres:
15            xinflex2.append(n)
16
17    dif = []
18    for i in range(1, len(xinflex2)):
19        dif.append(xinflex2[i]-xinflex2[i-1])
20
21    if len(dif)>0:
22        thres = np.min(dif)
23
24    SeriesMinimumidx = []
25    serie_idx = []
26
27    serie_idx.append(xinflex2[0])
28    for i in range(len(dif)):
29        if dif[i] > thres:
30            SeriesMinimumidx.append(serie_idx)
31            serie_idx = []
32            serie_idx.append(xinflex2[i+1])
33        else:
34            serie_idx.append(xinflex2[i+1])
35
36    if len(serie_idx) > 0:
37        SeriesMinimumidx.append(serie_idx)
```

```

37
38     for j in range(len(SeriesMinimumidx)):
39         iInflexTarget.append(SeriesMinimumidx[j])
40         InflexTarget.append(xdata[SeriesMinimumidx[j]])
41     else:
42         print("DEBB: NO INFLEXION POINT WAS FOUND IN DATA
43             ... IT IS SUGGESTED TO STOP RUNNING TO REVIEW
44             INPUT DATA ...")
45
46     return InflexTarget, iInflexTarget, xinflex2
47
48 # GetInflex

```

```

1 def GetPeak(xdata, y, peakstep, xthres, ythres):
2
3     step = peakstep
4
5     N = len(y)
6     yM = np.max(y)
7
8     PeakTarget = []
9     iPeakTarget = []
10
11     xextreme = []
12     for n in range(step, N-step):
13         aux = (y[n]-y[n-step])*(y[n+step]-y[n])
14         if aux <= 0 and ((y[n] - y[n-step] > xthres) or (y[
15             n] - y[n+step] > xthres)):
16             xextreme.append(n)
17
18     dif = []
19     for i in range(1, len(xextreme)):

```

```
19     dif.append(xextreme[i]-xextreme[i-1])
20
21
22     if len(dif)>0:
23
24         thres = np.min(dif)
25
26         SeriesMinimumidx = []
27         serie_idx = []
28
29         serie_idx.append(xextreme[0])
30         for i in range(len(dif)):
31             if dif[i] > thres:
32                 SeriesMinimumidx.append(serie_idx)
33                 serie_idx = []
34                 serie_idx.append(xextreme[i+1])
35             else:
36                 serie_idx.append(xextreme[i+1])
37
38         if len(serie_idx) > 0:
39             SeriesMinimumidx.append(serie_idx)
40
41
42
43         for i in range(len(SeriesMinimumidx)):
44             if np.max(y[SeriesMinimumidx[i]]) > ythres * yM
45                 :
46                 iPeakTarget.append(SeriesMinimumidx[i])
47                 PeakTarget.append(xdata[SeriesMinimumidx[i]
48                                     ])
49                 break
50         for j in range(i+1, len(SeriesMinimumidx)):
```

```

49         if (SeriesMinimumidx[j][0] > SeriesMinimumidx[j
        -1][-1] + 2 * peakstep) and (np.max(y[
        SeriesMinimumidx[j]]) > ythres * yM):
50             iPeakTarget.append(SeriesMinimumidx[j])
51             PeakTarget.append(xdata[SeriesMinimumidx[j
        ]])
52
53     else:
54         print("DEBB: NO PEAK WAS FOUND IN DATA ... IT IS
        SUGGESTED TO STOP RUNNING TO REVIEW INPUT DATA
        ...")
55
56     return PeakTarget, iPeakTarget, xextreme
57
58 # GetPeak

```

```

1 def GroundTruth(data, inflexstep, peakstep, inflexthres,
2 xpeakthres, ypeakthres, ctry):
3     y = np.array(list(data[ctry]["cases"]))
4     xdata = np.array(list(data[ctry]["epi_days"]))
5
6     InflexTarget, iInflexTarget, xInflex = GetInflex(xdata,
7     y, inflexstep, inflexthres)
8
9     PeakTarget, iPeakTarget, xPeaks = GetPeak(xdata, y,
10    peakstep, xpeakthres, ypeakthres)
11
12    print("PEAK ", iPeakTarget, "\n", xPeaks)
13
14    plt.figure(figsize = (6,3))
15    plt.plot(xdata, y, "b-")

```

```

14     xtrgt = xdata[xInflex]
15     plt.plot(xtrgt, y[xInflex], "k.")
16
17     xtrgt1 = xdata[xPeaks]
18     plt.plot(xtrgt1, y[xPeaks], "r.")
19
20     plt.title("Picos da epidemia para "+ctry)
21     plt.xlabel('Dias')
22     plt.ylabel('Casos')
23     plt.grid()
24
25     plt.show()
26
27     print("Ground Truth Sh Features: \nNumber of Peaks: ",
28           len(iPeakTarget), "\nNumber of Inflexion Regions: ",
29           len(iInflexTarget))
30
31     TotCases = np.sum(np.array(list(data[ctry]["cases"])))
32
33     return InflexTarget, iInflexTarget, PeakTarget,
34           iPeakTarget, TotCases
35
36 # GroundTruth

```

D.0.0.0.4 READING, VISUALIZING AND PLOTTING EPIDEMIC RAW DATA

```

1 path = os.getcwd()
2 csv_files = glob.glob(os.path.join(path, "raw/raw_*.csv"))
3
4 data= {}
5 country=' '
6 age_interval=[]

```

```
7
8 for f in csv_files:
9     fname = f.split("\\")[-1]
10    country = fname.split("_")[2]
11    data[country] = (pd.read_csv(f))
12    strp = fname.split("AGE_")
13    minage = int(strp[-1].split("_SEX")[0].split("-")[0])
14    maxage = int(strp[-1].split("_SEX")[0].split("-")[1])
15    age_interval.append([minage, maxage])
16
17    data[country]['onset_dates'] = pd.to_datetime(data[
18        country]['onset_dates'])
19    consecutive_days=[]
20    consecutive_days.append(0)
21    for i in range(len(data[country])-1):
22        consecutive_days.append(consecutive_days[-1]+\\
23            (pd.to_datetime(data[country]['onset_dates '
24                ][i+1])-pd.to_datetime(data[country]['
25                onset_dates '][i])).days)
26    data[country]["epi_days"] = consecutive_days
27
28    print('country:', country)
29    print('Content:')
30    display(data[country])
31    print()
32    f = plt.figure(figsize=(10,3))
33    plt.plot(data[country]['epi_days'], data[country]['cases
34        '], 'k.')
```

```
35 plt.show()
```

D.0.0.0.5 READING, VISUALIZING AND PLOTTING EPIDEMIC SMOOTHED TIME SERIES

```

1 path = os.getcwd()
2 scsv_files = glob.glob(os.path.join(path, "smoothed/ground_
   *.csv"))
3 sdata={}
4 scountry=''
5 sage_interval=[]
6
7 mdata = []
8
9 ctr=0
10
11 for f in scsv_files:
12     sfname = f.split("\\")[-1]
13     scountry = sfname.split("_")[2]
14     sdata[scountry] = pd.read_csv(f)
15
16     md = cp.deepcopy(sdata[scountry])
17
18     strp = sfname.split("AGE_")
19     minage = int(strp[-1].split("_SEX")[0].split("-")[0])
20     maxage = int(strp[-1].split("_SEX")[0].split("-")[1])
21     sage_interval.append([minage, maxage])
22
23     sdata[scountry]['onset_dates'] = pd.to_datetime(sdata[
   scountry]['onset_dates'])
24     sconsecutive_days=[]
25     sconsecutive_days.append(0)

```

```

26 for i in range(len(sdata[scountry])-1):
27     sconsecutive_days.append(sconsecutive_days[-1]+\
28         (pd.to_datetime(sdata[scountry]['
                onset_dates'][i+1])-pd.to_datetime(sdata
                [scountry]['onset_dates'][i])).days)
29 sdata[scountry]["epi_days"] = sconsecutive_days
30
31 print('Country:', scountry)
32 print('Consecutive days:', len(sconsecutive_days), '
        Smoothed len:', len(sdata[scountry]))
33 print('Content:')
34 display(sdata[scountry])
35 print()
36 f = plt.figure(figsize=(10,3))
37 x = sdata[scountry]['epi_days']
38 y = sdata[scountry]['cases']
39 plt.plot(x,y,'r.-')
40 plt.plot(data[scountry]['epi_days'],data[scountry]['
        cases'],'k.')
41 plt.xticks(rotation='vertical')
42 plt.grid()
43 plt.xlabel('onset dates')
44 plt.ylabel('new daily cases')
45 plt.show()
46 ctr+=1
47
48 # data = sdata #era comentario

```

SIMULATOR

```

1 def SPLICE_ModelEvaluate(Nn, prms, T, verbose, ctry, data):
2
3     x = data[ctry]['epi_days']

```

```

4     y = data[ctry]['cases']
5
6     [N, t0, h_risk, beta_h, beta_l, epsilon_h, epsilon_l,
       Pa, Ta, Trec, Tinc, Tdet, Pdet, Tdeath, Pdeath, E0]
       = prms
7     FeatureNames = ["NetSize", "Delay", "h_risk", "Beta_h",
                      "Beta_l", "Epsilon_h", "Epsilon_l", "Pa", "Ta", "
                      Trec", "Tinc", "Tdet", "Pdet", "Tdeath", "Pdeath", "
                      E0"]
8     printable_prms = [f"Label: {label} - Value: {value}\n"
                        for value, label in zip(prms, FeatureNames)]
9     print("Parameters SPLICE :\n","".join(printable_prms))
10
11     s_h = np.zeros((Nn,T), dtype=float)
12     s_l = np.zeros((Nn,T), dtype=float)
13     a = np.zeros((Nn,T), dtype=float)
14     v = np.zeros((Nn,T), dtype=float)
15     l = np.zeros((Nn,T), dtype=float)
16     h = np.zeros((Nn,T), dtype=float)
17     i = np.zeros((Nn,T), dtype=float)
18     r = np.zeros((Nn,T), dtype=float)
19     d = np.zeros((Nn,T), dtype=float)
20     c = np.zeros((Nn,T), dtype=float)
21
22     S_h = np.zeros(T, dtype=float)
23     S_l = np.zeros(T, dtype=float)
24     A = np.zeros(T, dtype=float)
25     V = np.zeros(T, dtype=float)
26     L = np.zeros(T, dtype=float)
27     H = np.zeros(T, dtype=float)
28     I = np.zeros(T, dtype=float)
29     R = np.zeros(T, dtype=float)

```

```

30 D = np.zeros(T, dtype=float)
31 C = np.zeros(T, dtype=float)
32
33 # Executar simula~o
34 chart_time_label = uuid.uuid4()
35 legend=[]
36 coverage = []
37 detected = []
38
39 for nt in range(Nn):
40
41     for t in range(int(t0[nt])+1):
42         s_h[nt][t] = N[nt] * h_risk[nt]
43         s_l[nt][t] = N[nt] * (1 - h_risk[nt])
44
45     l[nt][int(t0[nt])] = E0[nt]
46     s_h[nt][int(t0[nt])] = ( N[nt] - E0[nt] ) * h_risk[
47         nt]
48     s_l[nt][int(t0[nt])] = ( N[nt] - E0[nt] ) * (1 -
49         h_risk[nt])
50
51     aux1 = 1 - Pdeath[nt]
52     aux2 = 1 - Pdet[nt]
53     lamb_h = 1 * beta_h[nt] / N[nt]
54     lamb_l = 1 * beta_l[nt] / N[nt]
55     alpha = 1 / Tinc[nt]
56     delta = 1 * aux1 * Pdet[nt] / Tdet[nt]
57     gamma_i = 1 * aux1 * aux2 / Trec[nt]
58     mu_i = 1 * Pdeath[nt] * aux2 / Tdeath[nt]
59     theta_h = Pa[nt] / Ta[nt]
60     phi = 1 / Ta[nt]

```

```

60     ok = False
61
62     if Trec[nt] - Tdet[nt] > 0:
63
64         gamma_h = 1 * Pdet[nt] * aux1 / ( Trec[nt] -
65             Tdet[nt] )
66
67         if Tdeath[nt] - Tdet[nt] > 0:
68
69             mu_h = 1 * Pdeath[nt] / ( Tdeath[nt] - Tdet
70                 [nt] )
71
72             ok = True
73
74         if ok:
75
76             for t in range(int(t0[nt])+1,T):
77
78                 new_Sh_2L = (s_h[nt][t-1] * lamb_h) * i[nt
79                     ][t-1]
80                 new_Sl_2L = (s_l[nt][t-1] * lamb_l) * i[nt
81                     ][t-1]
82                 new_A = s_h[nt][t-1] * theta_h
83                 new_Sh_2V = (s_h[nt][t-1] * epsilon_h[nt])
84                 new_Sl_2V = (s_l[nt][t-1] * epsilon_l[nt])
85                 new_I = l[nt][t-1] * alpha
86                 new_I_2R = gamma_i * i[nt][t-1]
87                 new_I_2H = delta * i[nt][t-1]
88                 new_I_2D = mu_i * i[nt][t-1]
89                 new_H_2R = gamma_h * i[nt][t-1]
90                 new_H_2D = mu_h * i[nt][t-1]

```

```

88     new_A_2S1 = phi * a[nt][t-1]
89
90     s_h[nt][t] = s_h[nt][t-1] - new_Sh_2L -
          new_Sh_2V - new_A
91     s_l[nt][t] = s_l[nt][t-1] - new_Sl_2L -
          new_Sl_2V + new_A_2S1
92     l[nt][t] = l[nt][t-1] + new_Sh_2L +
          new_Sl_2L - new_I
93     i[nt][t] = i[nt][t-1] + new_I - new_I_2R -
          new_I_2H - new_I_2D
94     h[nt][t] = h[nt][t-1] + new_I_2H - new_H_2R
          - new_H_2D
95     r[nt][t] = r[nt][t-1] + new_I_2R + new_H_2R
96     d[nt][t] = d[nt][t-1] + new_I_2D + new_H_2D
97     a[nt][t] = a[nt][t-1] + new_A - new_A_2S1
98     v[nt][t] = v[nt][t-1] + new_Sh_2V +
          new_Sl_2V
99     c[nt][t] = new_I_2H
100
101
102     total_infected = (s_h[nt][0] + s_l[nt][0]) - (
          s_h[nt][-1] + s_l[nt][-1] + v[nt][-1] + a[nt
          ][-1])#added follow s_l non infecteds at end
103     coverage.append((total_infected / (s_h[nt][0] +
          s_l[nt][0]))
104     total_notified = np.sum(c[nt])
105     detected.append((total_notified / total_infected
          )
106
107     if verbose:
108         plt.figure(figsize=(10,3))
109         plt.plot(s_h[nt], label='S_h')

```

```

110 plt.plot(s_l[nt], label='S_l')
111 plt.plot([ (a + b) for a, b in zip(s_h[nt],
112     s_l[nt]) ], label='S')
112 plt.plot(v[nt], label='V')
113 plt.plot(a[nt], label='A')
114 plt.plot(l[nt], label='L')
115 plt.plot(i[nt], label='I')
116 plt.plot(r[nt], label='R')
117 plt.plot(h[nt], label='Q')
118 plt.plot(d[nt], label='D')
119
120 # plt.title("SRN "+str(nt+1)+" N = "+str
121     ('{:6d}'.format(int(round(N[nt]))))+ \
122     # " popula~o atingida: "+str('{:1f}'.
123     format(100*coverage[-1]))+\
124     # "% \ncasos detectados: "+str('{:1f}'.
125     format(100*detected[-1]))+"%")
126 plt.title(f"Transi~o da popula~o
127     entre estados do SPLICE na rede N{nt+1}"
128     )
129 plt.legend()
130 plt.xlabel('t (dias)')
131 plt.ylabel('Popula~o')
132 plt.grid()
133 plt.savefig(f"SRN-{nt+1}
134     _splice_compartments_{chart_time_label}.
135     png",bbox_inches='tight')
136 plt.savefig(f"SRN-{nt+1}
137     _splice_compartments_{chart_time_label}.
138     pdf",format='pdf',bbox_inches='tight')
139 plt.show()

```

```

132     else:
133         print("DEBB --- network ",nt," have wrong
           parameters: Trec[nt]<Tdet[nt] | Tdeath[nt]<
           Tdet[nt] ... \n>>>>>>>>> NETWORK VALUES NOT
           UPDATED\n")
134         break
135
136     if ok:
137
138
139         if verbose:
140             plt.figure(figsize=(10,3))
141             for nt in range(Nn):
142                 legend.append('SRN '+str(nt+1))
143                 plt.plot(c[nt])
144             #         legend.append('data')
145             #         plt.plot(x,y,'k.')
146             plt.legend(legend)
147             plt.xlabel('t (dias)')
148             plt.grid()
149             plt.title('Curva epidemiolÃgica nas redes')
150             plt.ylabel('Novos casos diÃrios')
151             plt.savefig(f"splice_srns_{chart_time_label}.
           png",bbox_inches='tight')
152             plt.savefig(f"splice_srns_{chart_time_label}.
           pdf",format='pdf',bbox_inches='tight')
153             plt.show()
154
155             # sum srn results
156
157         for t in range(T):
158             for nt in range(Nn):

```

```

159         S_h[t]+= s_h[nt][t]
160         S_l[t]+= s_l[nt][t]
161         A[t]+= a[nt][t]
162         V[t]+= v[nt][t]
163         L[t]+= l[nt][t]
164         H[t]+= h[nt][t]
165         I[t]+= i[nt][t]
166         R[t]+= r[nt][t]
167         D[t]+= d[nt][t]
168         C[t]+= c[nt][t]
169
170     # plot srns results X real data
171
172     if verbose:
173         plt.figure(figsize=(10,3))
174         plt.plot(x, C[x], "b.", label="model")
175         plt.plot(x, y, "k.", label="data")
176         plt.legend()
177         plt.title('Curva epidemiolÃgica predita X
178                 Curva real')
179         plt.xlabel('t (dias)')
180         plt.grid()
181         plt.ylabel('Novos casos diÃrios')
182         plt.savefig(f"splice_result_{chart_time_label}.
183                 png",bbox_inches='tight')
184         plt.savefig(f"splice_result_{chart_time_label}.
185                 pdf",format='pdf',bbox_inches='tight')
186         plt.show()
187
188     # return C, c, coverage, detected, ok
189     return C, c, coverage, detected, ok, {'V':(V[-1] - V
190     [0]), 'R':(R[-1] - R[0]), 'D':(D[-1] - D[0])}

```

```

187
188 # SPLICE_ModelEvaluate

```

D.0.0.0.6 OPTIMIZER SECTION

A BASIC PLOTTER OF A MODEL PREDICTION AND RAW DATA

```

1 def plotbasefig(data, ctry, Serie):
2
3     x = data[ctry]['epi_days']
4     plt.figure(figsize=(12,4))
5     plt.plot(x, data[ctry]['cases'],'k.')
6     plt.plot(x, Serie[x],'b.')
7     plt.ylabel('Notified(detected) daily cases')
8     plt.xlabel('onset day since epidemic start')
9     plt.grid()
10    legend=["Data","B Prediction"]
11    return legend
12
13 # plotbasefig

```

COMPARATOR (CALCULATES A METRIC OF ERROR FOR A PREDICTION OF THE MODEL)

```

1 def LossFunc(data, ctry, mdY, sdata):
2
3     Loss = 1.0
4     ShLoss = DtLoss = -1234
5
6     gtX = np.array(data[ctry]['epi_days'])
7     gtY = np.array(data[ctry]['cases'])
8     gtD = np.diff(gtX)
9
10    sgtY = np.array(sdata[ctry]['cases'])
11
12    TotCases = np.sum(gtY)

```



```
43     Mmd = np.max(mY)
44
45     if Mmd > 0 and np.min(mY) >= 0:
46
47         nmd = mY / Mmd
48
49         RelShapeDiff = np.abs( w * (nmd - ngt) )
50         MRelShapeDiff = np.max(RelShapeDiff)
51         NormMAE = np.mean(RelShapeDiff)
52
53         AbsDiff = np.abs( w * (mY - gtY) )
54         MAbsDiff = np.max(AbsDiff)
55         MAE = np.mean(AbsDiff)
56
57         ShLoss = 1.0 * np.sum(RelShapeDiff) / totngt
58         DtLoss = 1.0 * np.sum(AbsDiff) / TotCases
59
60         W = ShLoss / DtLoss - 0.5
61
62         if W > 1:
63             W = 1
64         elif W < 0:
65             W = 0
66
67         Loss = 50.0 * ( W * ShLoss + (1 - W) * DtLoss
68             )
69
70         print("shape metric ",ShLoss," distance metric
71             ",DtLoss," weight ",W," Loss: ",Loss)
72
73     else:
```



```

23     m = G[:,0]
24     M = G[:,1]
25     stp = G[:,2]
26
27     PRMS = []
28
29     f = 0
30     prms = []
31     for i in range(Nn):
32         prms = check_add(prms, f, x[f][i], m[f], M[f], stp[
33             f], 0, F)
34
35     PRMS.append(prms)
36
37     f = 1
38     prms = [0]
39     for i in range(1,Nn):
40         prms = check_add(prms, f, x[f][i], m[f], M[f], stp[
41             f], 0, F)
42
43     PRMS.append(prms)
44
45     for f in range(2,11):
46         prms = check_add([], f, x[f], m[f], M[f], stp[f],
47             Nn, F)
48         PRMS.append(prms[0])
49
50     f = 11
51     prms = check_add([], f, x[f], m[f], M[f], stp[f], Nn, F
52         )
53     add = np.array(prms[0]) * PRMS[9][0]
54     PRMS.append(list(add))
55
56     for f in range(12,F-1):

```

```

51     prms = check_add([], f, x[f], m[f], M[f], stp[f],
52                     Nn, F)
53
54     PRMS.append(prms[0])
55
56     f = F-1
57     prms = []
58     for i in range(Nn):
59         prms = check_add(prms, f, x[f][i], m[f], M[f], stp[
60                         f], 0, F)
61     PRMS.append(prms)
62
63     return PRMS
64
65 # getPRMS

```

D.0.0.0.7 GRADIENT UPDATER

FILLING THE NEXT COORDINATE OF A PARAMETER ACCORDING WITH ITS PARTIAL DERIVATIVE AND LEARNING RATE

```

1 def UpdateGradient(LossL, Loss0, LossR, f, nt, x, xgrad,
2                   imax, G, FeatureNames, lr_model, lr_value):
3
4     def GoLeft(X, jump):
5
6         if jump < 1:
7             jump = 1
8
9         if X > 0:
10            if X - jump > 0:
11                Y = X - jump
12                jump = -jump
13            print("GRADIENT RESULT FOR FEATURE ",
14                  FeatureNames[f], \

```

```
13         ": EXITING FROM A LOCAL MAXIMUM TO
14             THE LEFT WITH JUMP ", jump)
15     else:
16         Y = 0
17         jump = -X
18         print("GRADIENT RESULT FOR FEATURE ",
19             FeatureNames[f], \
20                 ": EXITING FROM A LOCAL MAXIMUM UP TO
21                 THE LEFT BOUNDARY WITH JUMP ",
22                 jump)
23     else:
24         Y = X
25         jump = 0
26         print("GRADIENT RESULT FOR FEATURE ",
27             FeatureNames[f], ": STAYING AT LEFT BOUNDARY"
28             )
29
30     return Y, jump
31
32 def GoRight(X, jump, rbdry):
33
34     if jump < 1:
35         jump = 1
36
37     if X < rbdry:
38         if X + jump < rbdry:
39             Y = X + jump
40             print("GRADIENT RESULT FOR FEATURE ",
41                 FeatureNames[f], \
42                     ": EXITING FROM A LOCAL MAXIMUM TO
43                     THE RIGHT WITH JUMP ", jump)
```

```
37         else:
38             Y = rbdry
39             jump = rbdry - X
40             print("GRADIENT RESULT FOR FEATURE ",
41                   FeatureNames[f],\
42                   ": EXITING FROM A LOCAL MAXIMUM UP TO
43                   THE RIGHT BOUNDARY WITH JUMP ",
44                   jump)
45
46         else:
47             Y = X
48             jump = 0
49             print("GRADIENT RESULT FOR FEATURE ",
50                   FeatureNames[f],": STAYING AT RIGHT BOUNDARY
51                   ")
52
53         return Y, jump
54
55     #
56     -----
57
58     rbdry = imax[f]
59
60     if lr_model == 0:
61         if lr_value > 0:
62             LR = 1.0 / lr_value ** Loss0
63         else:
64             LR = 1.0 / 10 ** Loss0
65     else:
66         if lr_value > 0 and lr_value <= 1:
67             LR = lr_value
68         else:
```

```
62         LR = 0.01
63
64     jump = 0
65     if nt < 0:
66         X = x[f]
67         Y = X
68     else:
69         X = x[f][nt]
70         Y = X
71
72     if Loss0 > LossL and Loss0 > LossR:
73
74         dL_L = Loss0 - LossL
75         dL_R = Loss0 - LossR
76
77         if dL_L >= dL_R:
78
79             jump = int(np.floor(LR * Loss0/dL_L))
80
81             Y, jump = GoLeft(X, jump)
82
83         else:
84
85             jump = int(np.floor(LR * Loss0/dL_R))
86
87             Y, jump = GoRight(X, jump, rbdry)
88
89     elif Loss0 < LossL and Loss0 < LossR:
90
91         print("GRADIENT RESULT FOR FEATURE ", FeatureNames[f
92               ], ": STAYING AT A LOCAL MINIMUM ")
```

```

93     else:
94
95         dL_RL = LossR - LossL
96
97         if np.abs(dL_RL)>0:
98
99             jump = int(np.floor(LR * Loss0/np.abs(dL_RL)))
100
101         else:
102
103             jump = 1
104
105         if dL_RL > 0:
106
107             Y, jump = GoLeft(X, jump)
108
109         else:
110
111             Y, jump = GoRight(X, jump, rbdry)
112
113     if nt < 0:
114         xgrad[f] = Y
115     else:
116         xgrad[f][nt] = Y
117
118     return xgrad, jump
119
120 # UpdateGradient

```

D.0.0.0.8 POINT EVALUATOR: SIMULATION OF MODEL AT A GIVEN POINT IN THE PARAMETRIC GRID

CONTROLS THE SOLUTION OF THE DIRECT PROBLEM FOR A GIVEN PARAMETER SET

```

1 def EvalPoint(f, nt, sense, x, G, Coordinates, Loss, Series
  , Waves, legend, Nn, T, BLossPos, BShLossPos, \
2         BDtLossPos, BShLoss, BDtLoss, WLossPos, data,
          ctry, sdata, verbose):
3
4     BLoss = Loss[BLossPos]
5     WLoss = Loss[WLossPos]
6
7     loss = float("nan")
8
9     direction = ["RIGHT","CENTER","LEFT"] #
10
11     F = len(G)
12
13     wx = cp.deepcopy(x)
14
15     wPRMS = getPRMS(wx, G, Nn, F)
16
17     print("\nEXPLORING TO THE ",direction[1-sense]," -
          PARAMETERS: ",wPRMS[f])
18
19     if wx in Coordinates:
20
21         prevpos = Coordinates.index(wx)
22         loss = Loss[prevpos]
23         print("COMBINATION ALREADY EXPLORED ... \nLoss: ",
          loss)
24         ModelY = Series[prevpos]
25         ModelX = np.arange(len(ModelY))
26
27     else:

```

```
28
29     prevpos = -1
30
31     Serie, CollWaves, coverage, detected, ok, _ =
32         SPLICE_ModelEvaluate(Nn, wPRMS, T, False, ctry,
33             data)
34
35     if ok:
36
37         loss, ShLoss, DtLoss = LossFunc(data, ctry,
38             Serie, sdata)
39
40         if verbose:
41
42             if nt > -1:
43                 legend.append("f: "+str(f)+" nt: "+str(
44                     nt+1)+" dir: "+ direction[1-sense])
45             else:
46                 legend.append("f: "+str(f)+" dir: "+
47                     direction[1-sense])
48
49         Coordinates.append(wx)
50         Loss.append(loss)
51         Series.append(Serie)
52         Waves.append(CollWaves)
53
54         Pos = len(Loss)-1
55
56         if Loss[-1] < BLoss:
57             BLoss = Loss[-1]
58             BLossPos = Pos
```

```

55         if ShLoss < BShLoss:
56             BShLoss = ShLoss
57             BShLossPos = Pos
58
59         if DtLoss < BDtLoss:
60             BDtLoss = DtLoss
61             BDtLossPos = Pos
62
63         if Loss[-1] > WLoss:
64             WLoss = Loss[-1]
65             WLossPos = Pos
66
67     return loss, Coordinates, Loss, Series, Waves, legend,
68           BLossPos, BShLossPos, BDtLossPos, \
69           BShLoss, BDtLoss, WLossPos, prevpos
70 # EvalPoint

```

EXPLORING MODEL SOLUTIONS DECREASING AND INCREASING A GIVEN PARAMETER

```

1 def LateralXplorer(x, xgrad, imax, f, nt, G, Coordinates,
2   Loss, Loss0, Series, Waves, legend, Nn, T, \
3     BLossPos, BShLossPos, BDtLossPos,
4     BShLoss, BDtLoss, WLossPos, data,
5     ctry, FeatureNames, lr_model, \
6     lr_value, sdata, verbose):
7
8     BLoss = Loss0
9
10    wx = cp.deepcopy(x)

```

```
11     if nt == -1:
12
13         if wx[f] == 0:
14
15             AtBdry = True
16
17         else:
18
19             wx[f] -= 2
20
21     else:
22
23         if wx[f][nt] == 0:
24
25             AtBdry = True
26
27         else:
28
29             wx[f][nt] -= 2
30
31     if not AtBdry:
32
33         LossL, Coordinates, Loss, Series, Waves, legend,
34         BLossPos, BShLossPos, BDtLossPos, BShLoss,
35         BDtLoss, WLossPos, prevpos = \
36         EvalPoint(f, nt, -1, wx, G, Coordinates, Loss,
37         Series, Waves, legend, Nn, T, BLossPos,
38         BShLossPos, BDtLossPos, \
39         BShLoss, BDtLoss, WLossPos, data, ctry,
40         sdata, verbose)
```

```
38
39     print("\nNOT EXPLORED BECAUSE IT IS AT LEFT
40           BOUNDARY. ASSIGNING Loss =", Loss0)
41     LossL = Loss0
42
43     wx = cp.deepcopy(x)
44
45     rbdry = imax[f]
46
47     AtBdry = False
48
49     if nt == -1:
50
51         if wx[f] == rbdry:
52
53             AtBdry = True
54
55         else:
56
57             wx[f] += 2
58
59     else:
60
61         if wx[f][nt] == rbdry:
62
63             AtBdry = True
64
65         else:
66
67             wx[f][nt] += 2
68
```

```

69     if not AtBdry:
70
71         LossR, Coordinates, Loss, Series, Waves, legend,
              BLossPos, BShLossPos, BDtLossPos, BShLoss,
              BDtLoss, WLossPos, prevpos = \
72     EvalPoint(f, nt, 1, wx, G, Coordinates, Loss,
              Series, Waves, legend, Nn, T, BLossPos,
              BShLossPos, BDtLossPos, \
73             BShLoss, BDtLoss, WLossPos, data, ctry,
              sdata, verbose)
74
75     else:
76
77         print("\nNOT EXPLORED BECAUSE IT IS AT RIGHT
              BOUNDARY. ASSIGNING Loss =", Loss0)
78         LossR = Loss0
79
80     xgrad, jump = UpdateGradient(LossL, Loss0, LossR, f, nt
              , x, xgrad, imax, G, FeatureNames, lr_model,
              lr_value)
81
82     return Coordinates, Loss, legend, BLossPos, BShLossPos,
              BDtLossPos, BShLoss, BDtLoss, WLossPos, Series,
              Waves, xgrad, jump
83
84 # LateralXplorer

```

EXPLORING SOLUTIONS AROUND THE CURRENT POINT IN THE PARAMETER GRID

```

1 def LocalXplorer(data, ctry, Nn, G, imax, Coordinates, Loss
      , Series, Waves, BLossPos, BShLossPos, BDtLossPos, \
2         BShLoss, BDtLoss, WLossPos, T, lr_model,
          lr_value, sdata, verbose):

```

```
3
4     BFeat = -1
5     BNtwrk = -2
6
7     FeatureNames = ["NetSize", "Delay", "h_risk", "Beta_h",
8                     "Beta_l", "Epsilon_h", "Epsilon_l", "Pa", "Ta", "
9                     Trec", "Tinc", "Tdet", "Pdet", "Tdeath", "Pdeath", "
10                    E0"]
11
12     Serie0 = Series[BLossPos]
13     Loss0 = Loss[BLossPos]
14     BLossPos0 = BLossPos
15     WLossPos0 = WLossPos
16
17     legend = []
18
19     F = len(G)
20
21     x = cp.deepcopy(Coordinates[BLossPos])
22     y = cp.deepcopy(Loss[BLossPos])
23
24     xgrad = cp.deepcopy(x)
25     igrad = []
26
27     print("input prm coordinates: ", x)
28
29     PRMS = getPRMS(x, G, Nn, F)
30
31     print("input prms:\n", PRMS)
32
33     f = 0
```

```

32     print("\n- EXPLORING FEATURE ", FeatureNames[f])
33
34     for nt in range(Nn):
35
36         print("NETWORK ", nt+1)
37
38         prevBLossPos = BLossPos
39
40         Coordinates, Loss, legend, BLossPos, BShLossPos,
41         BDtLossPos, BShLoss, BDtLoss, WLossPos, Series,
42         Waves, \
43         xgrad, jump = LateralXplorer(x, xgrad, imax, f, nt,
44         G, Coordinates, Loss, Loss0, Series, Waves,
45         legend, Nn, T, \
46         BLossPos, BShLossPos, BDtLossPos,
47         BShLoss, BDtLoss, WLossPos, data,
48         ctry, FeatureNames, \
49         lr_model, lr_value, sdata, verbose)
50
51         igrad.append(jump)
52     if verbose:
53         plt.figure(figsize=(1+int(len(igrad))/2), 3)
54         plt.bar(np.arange(len(igrad)), igrad)
55         plt.grid()
56         plt.xlabel("features")
57         plt.ylabel("jumps")
58         plt.show()
59
60     if BLossPos > prevBLossPos:
61         BFeat = f
62         BNtwrk = nt

```

```

58     f = 1
59
60     print("\n- EXPLORING FEATURE ", FeatureNames[f])
61
62     for nt in range(1, Nn):
63
64         print("NETWORK ", nt+1)
65
66         prevBLossPos = BLossPos
67
68         Coordinates, Loss, legend, BLossPos, BShLossPos,
69         BDtLossPos, BShLoss, BDtLoss, WLossPos, Series,
70         Waves, \
71         xgrad, jump = LateralXplorer(x, xgrad, imax, f, nt,
72         G, Coordinates, Loss, Loss0, Series, Waves,
73         legend, Nn, T, \
74         BLossPos, BShLossPos, BDtLossPos,
75         BShLoss, BDtLoss, WLossPos, data,
76         ctry, FeatureNames, \
77         lr_model, lr_value, sdata, verbose)
78
79         igrad.append(jump)
80
81         if verbose:
82             plt.figure(figsize=(1+int(len(igrad))/2), 3))
83             plt.bar(np.arange(len(igrad)), igrad)
84             plt.grid()
85             plt.xlabel("features")
86             plt.ylabel("jumps")
87             plt.show()
88
89         if BLossPos > prevBLossPos:
90             BFeat = f

```

```

84         BNtwrk = nt
85
86     nt = -1
87
88     for f in range(2,F-1):
89
90         print("\n- EXPLORING FEATURE ",FeatureNames[f])
91
92         prevBLossPos = BLossPos
93
94         Coordinates , Loss , legend , BLossPos , BShLossPos ,
          BDtLossPos , BShLoss , BDtLoss , WLossPos , Series ,
          Waves , \
95     xgrad , jump = LateralXplorer(x, xgrad , imax , f , nt ,
          G , Coordinates , Loss , Loss0 , Series , Waves ,
          legend , Nn , T , \
96             BLossPos , BShLossPos , BDtLossPos ,
          BShLoss , BDtLoss , WLossPos , data ,
          ctry , FeatureNames , \
97             lr_model , lr_value , sdata , verbose)
98
99     igrad.append(jump)
100    if verbose:
101        plt.figure(figsize=(1+int(len(igrad))/2),3))
102        plt.bar(np.arange(len(igrad)),igrad)
103        plt.grid()
104        plt.xlabel("features")
105        plt.ylabel("jumps")
106        plt.show()
107
108    if BLossPos > prevBLossPos:
109        BFeat = f

```

```

110         BNtwrk = nt
111
112
113     f = F-1 # network's initial infectors
114
115     print("\n- EXPLORING FEATURE ", FeatureNames[f])
116
117     for nt in range(Nn):
118
119         print("\n\n
120             -----\
121             nNETWORK ", nt+1, "\n
122             -----
123
124         ")
125
126         prevBLossPos = BLossPos
127
128         Coordinates, Loss, legend, BLossPos, BShLossPos,
129             BDtLossPos, BShLoss, BDtLoss, WLossPos, Series,
130             Waves, \
131         xgrad, jump = LateralXplorer(x, xgrad, imax, f, nt,
132             G, Coordinates, Loss, Loss0, Series, Waves,
133             legend, Nn, T, \
134                 BLossPos, BShLossPos, BDtLossPos,
135                 BShLoss, BDtLoss, WLossPos, data,
136                 ctry, FeatureNames, \
137                 lr_model, lr_value, sdata, verbose)
138
139         igrad.append(jump)
140
141         if verbose:
142             plt.figure(figsize=(1+int(len(igrad)/2), 3))
143             plt.bar(np.arange(len(igrad)), igrad)

```

```
132         plt.grid()
133         plt.xlabel("features")
134         plt.ylabel("jumps")
135         plt.show()
136
137     if BLossPos > prevBLossPos:
138         BFeat = f
139         BNtwrk = nt
140
141     if np.sum(np.abs(np.array(igrad))) > 0:
142
143         print("EXPLORING THE STRUCTURAL-SUBGRADIENT-DRIVEN
144             POINT")
145
146         strut_subgrad = cp.deepcopy(x)
147
148         chg=0
149
150         f=0
151         for nt in range(Nn):
152             if x[f][nt] != xgrad[f][nt]:
153                 chg += 1
154                 strut_subgrad[f][nt] = xgrad[f][nt]
155
156         f=1
157         for nt in range(Nn):
158             if x[f][nt] != xgrad[f][nt]:
159                 chg += 1
160                 strut_subgrad[f][nt] = xgrad[f][nt]
161
162         f=F-1
163         for nt in range(Nn):
164             if x[f][nt] != xgrad[f][nt]:
165                 chg += 1
```

```
163         strut_subgrad[f][nt] = xgrad[f][nt]
164
165     if chg > 0:
166
167         gradPRMS = getPRMS(strut_subgrad, G, Nn, F)
168
169         print("strut-xgrad ",strut_subgrad)
170         print("strut-gradPRMS ",gradPRMS)
171
172         SerieSGP, CollWaves, coverage, detected, ok, _
           = SPLICE_ModelEvaluate(Nn, gradPRMS, T,
           False, ctry, data)
173
174         if ok:
175
176             lossSGP, SGP_ShLoss, SGP_DtLoss = LossFunc(
           data, ctry, SerieSGP, sdata)
177
178             if verbose:
179                 legend = plotbasefig(data, ctry, Serie0
           )
180                 legend = plotbasefig(data, ctry, Series
           [BLossPos])
181                 plt.plot(data[ctry]['epi_days'],
           SerieSGP[data[ctry]['epi_days']], "m.
           ")
182                 legend.append("structural -gradient "+
           str(lossSGP))
183                 plt.legend(legend)
184                 plt.show()
185
186         Coordinates.append(strut_subgrad)
```

```
187         Loss.append(lossSGP)
188         Series.append(SerieSGP)
189
190         if lossSGP < Loss[BLossPos]:
191             BLossPos = len(Loss) - 1
192
193     else:
194
195         print("NOT EXPLORING THE STRUCTURAL SUB-
196             GRADIENT-DRIVEN POINT BECAUSE ALL JUMPS = 0"
197             )
198
199     print("EXPLORING THE SYSTEMIC-SUBGRADIENT-DRIVEN
200         POINT")
201
202     syst_subgrad = cp.deepcopy(x)
203
204     chg=0
205
206     for f in range(2,F-1):
207         if x[f] != xgrad[f]:
208             chg += 1
209             syst_subgrad[f] = xgrad[f]
210
211     if chg > 0:
212
213         gradPRMS = getPRMS(syst_subgrad, G, Nn, F)
214
215         print("syst-xgrad ",syst_subgrad)
216         print("syst-gradPRMS ",gradPRMS)
```

```
215 SerieTGP, CollWaves, coverage, detected, ok, _
      = SPLICE_ModelEvaluate(Nn, gradPRMS, T,
      False, ctry, data)
216
217 if ok:
218
219     lossTGP, TGP_ShLoss, TGP_DtLoss = LossFunc(
      data, ctry, SerieTGP, sdata)
220
221     if verbose:
222         legend = plotbasefig(data, ctry, Serie0
      )
223         legend = plotbasefig(data, ctry, Series
      [BLossPos])
224         plt.plot(data[ctry]['epi_days'],
      SerieTGP[data[ctry]['epi_days']], "m.
      ")
225         legend.append("systemic-gradient "+str(
      lossTGP))
226         plt.legend(legend)
227         plt.show()
228
229     Coordinates.append(syst_subgrad)
230     Loss.append(lossTGP)
231     Series.append(SerieTGP)
232
233     if lossTGP < Loss[BLossPos]:
234         BLossPos = len(Loss) - 1
235
236 else:
237
```

```

238         print("NOT EXPLORING THE SYSTEMIC SUB-GRADIENT -
                DRIVEN POINT BECAUSE ALL JUMPS = 0")
239
240     else:
241
242         print("NOT EXPLORING THE ANTI-GRADIENT POINTS
                BECAUSE ALL JUMPS = 0")
243
244     return Coordinates, Loss, Series, Waves, BLossPos,
           BShLossPos, BDtLossPos, BShLoss, BDtLoss, WLossPos,
           BFeat, BNtwrk
245
246 # LocalXplorer

```

D.0.0.1 COUNTRY SELECTION & DATA COLLECTION

```

1 # ['ARGENTINA', 'BRAZIL', 'CANADA', 'CHILE', 'COLOMBIA', '
   GUATEMALA', 'MEXICO', 'PERU', 'UNITED STATES OF AMERICA
   ']
2 # Input bellow the country name for simulation and analysis
   process
3 ctry = 'UNITED STATES OF AMERICA'
4
5 print("SIMULATING COUNTRY: ",ctry)
6
7 y = np.array(list(data[ctry]["cases"]))
8 xdata = np.array(list(data[ctry]["epi_days"]))

```

D.0.0.2 DATA VISUALIZATION AND PREPROCESSING

```

1 inflexstep = 4
2 peakstep = 5

```

```

3 inflexthres = 1e-5
4 xpeakthres = 0.1
5 ypeakthres = 0.25
6
7 InflexTarget, iInflexTarget, PeakTarget, iPeakTarget,
   TotCases = GroundTruth(sdata, inflexstep, peakstep, \
8 inflexthres, xpeakthres, ypeakthres, \
9 ctry)

```

D.0.0.2.1 HYPER-PARAMETER CONFIGURATION

```

1 print("HYPER-PARAMS VALUES")
2
3 # Simulation time
4 T = 1 + xdata[-1] # calculated form data
5
6 print("SIMULATING ",T," DAYS")
7
8 # Number of networks - using number of peaks
9 Nn_2_countries = {
10     'ARGENTINA':3,
11     'BRAZIL':1,
12     'CANADA':2,
13     'CHILE':3,
14     'COLOMBIA':1,
15     'GUATEMALA':3,
16     'MEXICO':2,
17     'PERU':2,
18     'UNITED STATES OF AMERICA':1
19 }
20
21 Nn = Nn_2_countries[ctry]

```

```

22
23 print(f"MODELING {Nn} NETWORKS USING {3*Nn-1 + 13}
    FEATURES")
24
25 # Learning rate
26 # 2 learning models available
27 # lr_model = 0 : lr = 1 / lr_value^Loss # lr_rate is an
    inverse potential funcion of Loss
28 # lr_model = 1 : lr = lr_ value # constant
    lr_rate
29
30 lr_model = 0 # choose model
31 lr_value = 0.1 # in ]0,1]
32
33 print(f"LEARNING MODEL {lr_model}: LEARNING PARAMETER {
    lr_value}")
34
35 # Searching strategy parameters
36
37 maxAttempts4GoodInit = 50
38
39 print("MAXIMUM NUMBER OF RANDOM TRIES FOR OBTAINING A GOOD
    CIV (COMBINAÃŒÃŒO - INICIAL - DE - VARIÃŒÃVEIS) ",
    maxAttempts4GoodInit)
40
41 # Loss (error) value accepted as a good CIV
42 MaxLoss = 0.25
43
44 print("ERROR ACCEPTED AS A GOOD CIV ", MaxLoss)
45
46 # Number of re-initializations and optimal solutions to
    find at macro search process for uncertainty analysis

```

```

47 nInits = 40
48
49 print("NUMBER OF RANDOM RE-INITIALIZATIONS OF THE SEARCH
      ALGORITHM ", nInits)
50
51 # Maximum number of iterations during a search for an
      optimal solution at BSO-CIV
52 nWalks = int( 50 *np.sqrt(Nn))
53
54 print("MAXIMUM NUMBER OF STEPS IN THE SEARCH ALGORITHM ",
      nWalks)
55
56 # Uncertainty Check
57
58 UncertaintyThres = 0.8
59
60 print("UNCERTAINTY CHECK THRESHOLD ", UncertaintyThres)

```

PARAMETER GRID DEFINITION

```

1 Ymax = np.max(y)
2
3 # DEFINE GRID
4
5 G = []
6
7 Glevel = 40
8
9 V = 0.2
10 v = 0.001
11
12 # grid for SRN size N (SIZE OF SUSCEPTIBLE POPULATION - TO
      SEARCH)
13

```

```
14 dN = int(10.0 * Ymax)
15 mN = dN
16 MN = int(500.0 * Ymax)
17
18 G.append([mN, MN, dN])
19
20 print("NETWORK SIZE RANGE: [", mN, ",", MN, "] STEP = ", dN)
21
22 # grid for delay (days) of transmission between networks (
    SRNS ACTIVATION DAYS - TO SEARCH)
23
24 dDelay = int( T / Glevel) #5 # 0.01
25 mDelay = 0
26 MDelay = int(0.85 * T)
27
28 G.append([mDelay, MDelay, dDelay])
29
30 print("NETWORK DELAY RANGE: [", mDelay, ",", MDelay, "] STEP =
    ", dDelay)
31
32 # grid for (h_risk) (PERCENT OFF HIGH RISK GROUP - TO
    SEARCH)
33
34 dC = 0.01
35 mC = 0.2
36 MC = 0.8
37
38
39 G.append([mC, MC, dC])
40
41 print("HIGH RISK PERCENTAGE RANGE: [", mC, ",", MC, "] STEP =
    ", dC)
```

```
42
43 # grid for contagiousness (beta_h) (TRANSMISSION
    COEFFICIENT ON HIGH RISK GROUP - TO SEARCH)
44
45 dC = 0.1
46 mC = dC
47 MC = 4.0
48
49 G.append([mC,MC,dC])
50
51 print("TRANSMISSION COEFFICIENT RANGE ON RIGH RISK GROUP (
    beta_h): [", mC,",",MC,"] STEP = ",dC)
52
53 # grid for contagiousness (beta_L) (TRANSMISSION
    COEFFICIENT ON LOW RISK GROUP - TO SEARCH)
54
55 dC = 0.01
56 mC = 0.05
57 MC = 2.0
58
59 G.append([mC,MC,dC])
60
61 print("TRANSMISSION COEFFICIENT RANGE ON LOW RISK GROUP (
    beta_l): [", mC,",",MC,"] STEP = ",dC)
62
63 # grid for (epsilon_h) (VACCINATION COEFFICIENT ON RIGH
    RISK GROUP - TO SEARCH)
64
65 dC = 0.0001
66 mC = 0.0005
67 MC = 0.0008
68
```

```
69 G.append([mC,MC,dC])
70
71 print("VACCINATION COEFFICIENT ON HIGH RISK GROUP (
      epsilon_h): [", mC,",",",MC,"] STEP = ",dC)
72
73 # grid for (epsilon_l) (VACCINATION COEFFICIENT ON LOW RISK
      GROUP - TO SEARCH)
74
75 dC = 0.0001
76 mC = 0.0001
77 MC = 0.0005
78
79 G.append([mC,MC,dC])
80
81 print("VACCINATION COEFFICIENT ON LOW RISK GROUP (epsilon_l
      ): [", mC,",",",MC,"] STEP = ",dC)
82
83 # grid for (P_a) (ABSTINENCE PROBABILITY - HERE CONSTANT)
84
85 dC = 0.01
86 mC = 0.3
87 MC = 0.5
88
89 G.append([mC,MC,dC])
90
91 print("ABSTINENCE PROBABILITY RANGE (Pa): [", mC,",",",MC,"]
      STEP = ",dC)
92
93 # grid for (T_a) (ABSTINENCE TIME - HERE CONSTANT)
94
95 refTa = 21.0
96 VTa = v
```

```
97 mTa = (1.0 - VTa)* refTa
98 MTa= (1.0 + VTa) * refTa
99 dTa = (MTa - mTa) / 2
100
101 G.append([mTa,MTa,dTa])
102
103 print("ABSTINENCE TIME RANGE (Ta): [" ,mTa," ,",MTa," ] STEP =
      ",dTa)
104
105 # grid for Trec - (RECOVERY TIME - HERE CONSTANT)
106
107 refTrec = 21.0
108 VTrec = v
109 mTrec = (1.0 - VTrec)* refTrec
110 MTrec = (1.0 + VTrec)* refTrec
111 dTrec = (MTrec - mTrec) / 2
112
113 G.append([mTrec,MTrec,dTrec])
114
115 print("RECOVERY TIME RANGE (Trec): [" , r'{:2f}'.format(
      mTrec)," ,",r'{:2f}'.format(MTrec)," ] STEP = ",r'{:2f}'.
      format(dTrec))
116
117 # grid for Tinc (HERE CONSTANT)
118
119 refTinc = 8
120 VTinc = v
121 mTinc = (1.0 - VTinc)* refTinc
122 MTinc = (1.0 + VTinc)* refTinc
123 dTinc = (MTinc - mTinc) / 2
124
125 G.append([mTinc,MTinc,dTinc])
```

```
126
127 print("INCUBATION TIME RANGE (Tinc): [" , r'{:2f}'.format(
      mTinc),",",r'{:2f}'.format(MTinc),"] STEP = ",r'{:2f}'.
      format(dTinc))
128
129 # grid for Tdet (HERE CONSTANT)
130
131 # we work with % of Trec
132 # Tdet = fraction of Trec, such that the Time2det = Tdet *
      Trec
133
134 ref_tdet = 6.95
135 refTdet = ref_tdet / refTrec
136 VTdet = v
137 mTdet = (1.0 - VTdet)* refTdet
138 MTdet = (1.0 + VTdet)* refTdet
139 dTdet = (MTdet - mTdet) / 2
140
141 G.append([mTdet,MTdet,dTdet])
142
143 print("DET_TIME/REC_TIME RATIO RANGE (Tdet): [" , r'{:2f}'.
      format(mTdet),",",r'{:2f}'.format(MTdet),"] STEP = ",r'
      {:2f}'.format(dTdet))
144
145 # grid for Pdet (HERE CONSTANT)
146
147 ref_Pdet = 0.5
148 VPdet = v
149 mPdet = (1 - VPdet) * ref_Pdet
150 MPdet = (1 + VPdet) * ref_Pdet
151 dPdet = (MPdet - mPdet) / 2
152
```

```
153 G.append([mPdet,MPdet,dPdet])
154
155 print("PROBABILITY OF DETECTION RANGE (Pdet): [" , r'{:2f}'.format(
      format(mPdet),"",r'{:2f}'.format(MPdet)," ] STEP = ",r'
      {:2f}'.format(dPdet))
156
157 # grid for Tdeath (HERE CONSTANT)
158
159 refTdeath = 68
160 VTdeath = v
161 mTdeath = (1.0 - VTdeath)* refTdeath
162 MTdeath = (1.0 + VTdeath)* refTdeath
163 dTdeath = (MTdeath - mTdeath) / 2 / Glevel
164 G.append([mTdeath,MTdeath,dTdeath])
165
166 print("TIME TO DEATH RANGE (Tdeath): [" , r'{:2f}'.format(
      mTdeath),"",r'{:2f}'.format(MTdeath)," ] STEP = ",r'{:2f}
      }'.format(dTdeath))
167
168 # grid for Pdeath (HERE CONSTANT)
169
170 refPdeath = 0.0013
171 V = v
172 mPdeath = (1.0 - V)* refPdeath
173 MPdeath = (1.0 + V)* refPdeath
174 dPdeath = (MPdeath - mPdeath) / 2
175 G.append([mPdeath,MPdeath,dPdeath])
176
177 print("PROBABILITY OF DEATH RANGE (Pd): [" , r'{:2f}'.format(
      (mPdeath),"",r'{:2f}'.format(MPdeath)," ] STEP = ",r'{:2
      f}'.format(dPdeath))
178
```

```

179 # grid for E0 (FORCE OF INFECTION - TO SEARCH)
180
181 dE0 = 2
182 mE0 = 1
183 ME0 = int(0.05 * MN) # 1000.0
184
185 G.append([mE0,ME0,dE0])
186
187 print("NUMBER OF INITIAL INFECTORS' RANGE (E0): [", mE0,","
      ,ME0,"] STEP = ",dE0)
188
189 # NUMBER OF TYPES OF MODEL PARAMETERS
190
191 F = len(G)
192
193 # CALCULATING THE LIMITS OF THE PARAMETRIC GRID
194
195 imax = []
196 for f in range(F):
197     imax.append(1 + int((G[f][1]-G[f][0])/G[f][2]))

```

INITIALIZATION MODULE

```

1 def Initialize(inits, maxAttempts4GoodInit, Nn, imax, G, F,
2     verbose):
3     Attempts4GoodInit = 0
4
5     print("INITIALIZATION STEP: ", inits)
6
7     BLoss = 1e10

```

```
8
9  while BLoss >= MaxLoss and Attempts4GoodInit <
10      maxAttempts4GoodInit:
11
12      Attempts4GoodInit += 1
13
14      iN = [0] * Nn
15
16      for nt in range(Nn):
17
18          iN[nt] = random.randrange(1, imax[0] - 1)
19
20      iDelay = [0] * Nn
21
22      for nt in range(1, Nn):
23
24          iDelay[nt] = iDelay[nt-1] + int(imax[1] *
25              random.uniform(0.01, 1.1*nt/Nn))
26
27      iHRisk = random.randrange(int(imax[2]/2), imax[2])
28
29      iBetaH = random.randrange(int(imax[2]/2), imax[2])
30
31      iBetaL = random.randrange(int(imax[2]/2), imax[2])
32
33      iEpsilonH = random.randrange(int(imax[2]/2), imax
34          [2])
35
36      iEpsilonL = random.randrange(int(imax[2]/2), imax
37          [2])
38
39      iPa = random.randrange(int(imax[2]/2), imax[2])
```

```
36
37     iTa = random.randrange(int(imax[2]/2), imax[2])
38
39     iTrec = random.randrange(1, imax[3])
40
41     iTinc = random.randrange(1, imax[4])
42
43     iTdet = random.randrange(1, imax[5])
44
45     iPdet = random.randrange(1, imax[6])
46
47     iTdeath = random.randrange(1, imax[7])
48
49     iPdeath = random.randrange(1, imax[8])
50
51     iE0 = [0] * Nn # initial infected
52
53     for nt in range(Nn):
54
55         iE0[nt] = random.randrange(1, imax[9])
56
57     iprms = [iN, iDelay, iHRisk, iBetaH, iBetaL,
58             iEpsilonH, iEpsilonL, iPa, iTa, iTrec, iTinc,
59             iTdet, iPdet, iTdeath, iPdeath, iE0]
60
61     prms = getPRMS(iprms, G, Nn, F)
62
63     print("\nLOOKING FOR INITIAL PRMS ITERATION: ",
64           Attempts4GoodInit, "/", maxAttempts4GoodInit)
65
66     Serie, CollWaves, coverage, detected, ok, _ =
67         SPLICE_ModelEvaluate(Nn, prms, T, False, ctry,
```

```
        data)
64
65     if ok:
66         loss, ShLoss, DtLoss = LossFunc(data, ctry,
67             Serie, sdata)
68
69         Loss.append(loss)
70         Coordinates.append(iprms)
71         Series.append(Serie)
72         Waves.append(CollWaves)
73
74         print('Loss: ', loss)
75
76         if loss < BLoss:
77             BPrms = iprms
78             BLoss = loss
79             BLossPos = len(Loss) - 1
80             print("\nNew Prms: ", prms)
81
82     else:
83         print("INITIALIZATION FAILED - WRONG PARAMETER
84             COMBINATION .....")
85
86         Attempts4GoodInit -= 1
87
88     bestCoord = Coordinates[BLossPos]
89     bestSerie = Series[BLossPos]
90     bestWaves = Waves[BLossPos]
91
92     return BPrms, BLoss, BLossPos, bestCoord, bestSerie,
93         bestWaves
```

D.0.1 BSO-CIV:

```

1 def Searcher(inits, nWalks, data, ctry, Nn, G, imax,
    Coordinates, Loss, Series, Waves, BLossPos, BShLossPos,
    \
2         BDtLossPos, BShLoss, BDtLoss, WLossPos, T,
        lr_model, lr_value, sdata, verbose):
3
4     walk = 0
5
6     Converged = 0
7
8     while Converged < 1 and walk < nWalks:
9
10        walk += 1
11
12        print("\nINIT STEP: ", inits, " WALK ", walk, " ON GRID
            \n")
13
14
15        Coordinates, Loss, Series, Waves, BLossPos,
            BShLossPos, BDtLossPos, BShLoss, BDtLoss,
            WLossPos, BFeat, BNtwrk = \
16        LocalXplorer(data, ctry, Nn, G, imax, Coordinates,
            Loss, Series, Waves, BLossPos, BShLossPos,
            BDtLossPos, \
17                BShLoss, BDtLoss, WLossPos, T,
                    lr_model, lr_value, sdata, verbose)
18
19        plt.plot(Loss)
20        plt.grid()
21        plt.ylabel("Loss value")
22        plt.xlabel("number of configurations")

```

```
23     plt.show()
24
25     legend = plotbasefig(data, ctry, Series[BLossPos])
26     plt.legend(legend)
27     plt.grid()
28     plt.show()
29
30     print("INITIAL LOSS: ", InitLoss[-1], " BEST LOSS: ",
31           Loss[BLossPos])
32
33     if BFeat == -1 and BNtwrk == -2 and BLossPos ==
34         prevLossPos:
35         print(" A GLOBAL MINIMUM FOUND ...")
36         Converged = 1
37     else:
38         print(">>> BEST LOSS AT CURRENT SEARCHING STEP:
39             ", Loss[BLossPos], " <<<")
40         if Loss[BLossPos] == Loss[-2]:
41             print("BEST POINT FOUND FOLLOWING THE LOCAL
42                 STRUCTURAL GRADIENT ...")
43         elif Loss[BLossPos] == Loss[-1]:
44             print("BEST POINT FOUND FOLLOWING THE LOCAL
45                 SYSTEMIC GRADIENT ...")
46         else:
47             print("BEST POINT FOUND IN THE NEIGHBORHOOD
48                 ...")
49
50     prevLossPos = BLossPos
51
52     if walk == nWalks:
53         print(" A GLOBAL MINIMUM APPROACHED ...")
```

```

49     return Coordinates, Loss, Series, Waves, BLossPos,
        BShLossPos, BDtLossPos, BShLoss, BDtLoss, WLossPos,
        BFeat, BNtwrk

```

D.0.2 REPETITIONS OF SEARCHING FOR OPTIMAL SOLUTIONS

```

1  verbose = True
2
3  InitTime = time.time()
4
5  clr = ['b-', 'r-', 'y-', 'g-', 'm-', 'k-', 'k.'] # patterns to
        plot
6
7  InitLoss = []
8  InitCoordinates = []
9
10 out = 0
11 inits = 0
12
13 CoordinatesMinimum = []
14 LossMinimum = []
15 SeriesMinimum = []
16
17 while out == 0 and inits < nInits:
18
19     Coordinates = []
20     Loss = []
21     Series = []
22     Waves = []
23
24     inits += 1

```

```
25 BPrms, BLoss, BLossPos, bestCoord, bestSerie, bestWaves
    = Initialize(inits, maxAttempts4GoodInit, Nn, imax,
    G, F, True)
26
27 Coordinates = []
28 Coordinates.append(bestCoord)
29 Series = []
30 Series.append(bestSerie)
31 Waves = []
32 Waves.append(bestWaves)
33
34 BLossPos = 0
35
36 loss, ShLoss, DtLoss = LossFunc(data, ctry, Series[0],
    sdata)
37
38 BLoss = loss
39
40 Loss = []
41 Loss.append(BLoss)
42
43 print("\nINITIAL GUESS #", inits, " LOSS: ", BLoss)
44
45 legend = plotbasefig(data, ctry, Series[BLossPos])
46
47 InitLoss.append(BLoss)
48 InitCoordinates.append(BPrms)
49 WLossPos = BLossPos
50
51 prevLossPos = BLossPos
52 BShLossPos = BLossPos
53 BDtLossPos = BLossPos
```

```

54     BShLoss = BLoss
55     BDtLoss = BLoss
56
57     Coordinates, Loss, Series, Waves, BLossPos, BShLossPos,
        BDtLossPos, BShLoss, BDtLoss, WLossPos, BFeat,
        BNtwrk = \
58     Searcher(inits, nWalks, data, ctry, Nn, G, imax,
        Coordinates, Loss, Series, Waves, BLossPos,
        BShLossPos, \
59             BDtLossPos, BShLoss, BDtLoss, WLossPos, T,
                lr_model, lr_value, sdata, verbose)
60
61     LossMinimum.append(Loss[BLossPos])
62     CoordinatesMinimum.append(Coordinates[BLossPos])
63     SeriesMinimum.append(Series[BLossPos])
64
65     b = getPRMS(CoordinatesMinimum[-1], G, Nn, F)
66
67     print("BEST PARAMETERS ON THE INITIALIZATION STEP: ",
        inits, "\n", b)
68
69     legend = plotbasefig(data, ctry, SeriesMinimum[-1])
70     plt.legend(legend)
71     plt.show()
72
73
74     toc = time.time()
75     print("END... InitTime: ", InitTime, " Ending Time: ", time.
        time(), " Elapsed Time : ", (toc-InitTime)/60/60, " hours")

```

D.0.3 UNCERTAINTY ANALYSIS SECTION

```

1 BNmb = int (0.2 * nInits)
2
3 if BNmb < 2:
4     BNmb = 2
5
6 idx = np.argsort(LossMinimum)
7
8
9 print("LOOKING FOR PARAMETER VARIATION IN THE BEST ",BNmb,"
    SOLUTIONS:\n")
10
11 bPRMS =[]
12
13 for i in range(BNmb):
14
15     print("best regression: ",i+1," / ",BNmb," LOSS = ",
        LossMinimum[idx[i]])
16
17     bPRMS.append(getPRMS(CoordinatesMinimum[idx[i]], G, Nn,
        F))
18     FeatureNames = ["NetSize", "Delay", "h_risk", "Beta_h",
        "Beta_l", "Epsilon_h", "Epsilon_l", "Pa", "Ta", "
        Trec", "Tinc", "Tdet", "Pdet", "Tdeath", "Pdeath", "
        E0"]
19     printable_prms = [f"Value: {value}, Label: {label}" for
        value, label in zip(bPRMS[-1], FeatureNames)]
20     print("Parameters: ",i+1,":\n","".join(printable_prms))
21
22 print("Interval of variation of parameters in country ",
        ctry,":\n")
23 bmin = np.min(bPRMS,axis=0)
24 print("Minimum parameters \n",bmin)

```

```

25 bmax = np.max(bPRMS,axis=0)
26 print("Maximum parameters \n",bmax)
27 bmean = (np.array(bmin)+np.array(bmax))/2
28 print("Mean parameters \n",bmean)
29 bstdev = (np.array(bmax)-np.array(bmin))/2
30 print("Std. dev. of parameters \n",bstdev)
31 # relvar = np.array(bstdev)/np.array(bmean)
32 relvar = np.nan_to_num(np.array(bstdev) / np.array(bmean))
33 print("VariaÃ§Ã£o relativa dos parÃ¢metros (%)\n", 100 *
      relvar)
34
35 if np.max(relvar) >= UncertaintyThres:
36     print("WARNING: UNCERTAINTY CHECK FAILED WITH THRESHOLD
      ",UncertaintyThres, " - RESULTS ARE <NOT> RELIABLE
      !!! ",np.max(relvar))
37 else:
38     print("UNCERTAINTY CHECK OK WITH THRESHOLD ",
      UncertaintyThres," - RESULTS ARE RELIABLE !!! ",np.
      max(relvar))

```

D.0.4 SENSITIVITY ANALYSIS SECTION

```

1  '''
2      PLOT BEST REGRESSIONS AND VARIABILITY OF RESULTS WITH
3      PARAMETER VARIATION
4      CALCULATES PERCENTAGE OF SUSCEPTIBLES INFECTED (
5      COVERAGE) AND PERCENT OF INFECTED DETECTED
6  '''
7
8  nVariants = 20
9
10 stdev = 0.15

```

```
9
10 x = data[ctry]['epi_days']
11
12 Coverage = []
13 Detected = []
14 label = 0
15 for i in range(BNmb):
16     label +=1
17     print("best regression: ",i+1," / ",BNmb," LOSS = ",
18           LossMinimum[idx[i]])
19     FeatureNames = ["NetSize", "Delay", "h_risk", "Beta_h",
20                     "Beta_l", "Epsilon_h", "Epsilon_l", "Pa", "Ta", "
21                     Trec", "Tinc", "Tdet", "Pdet", "Tdeath", "Pdeath", "
22                     E0"]
23     printable_prms = [f"Value: {value}, Label: {label}" for
24                       value, label in zip(bPRMS[-1], FeatureNames)]
25     print("Parameters: ",i+1,":\n","".join(printable_prms))
26
27
28     LossArr =[]
29     SerieArr = []
30     WaveArr = []
31     Coverage = []
32     Detected = []
33
34     var = 0
35
36     while var < nVariants:
37
38         var += 1
39
40         print("variant: ",var," / ",nVariants)
```



```
63
64         for n in range(1, Nn):
65
66             newPRMS[f][n] = newPRMS[f][0]
67
68         if f == F-1:
69
70             for n in range(Nn):
71
72                 if np.random.uniform(0,1) < 0.5:
73
74                     newPRMS[f][n] = newPRMS[f][n] *
75                         ( 1 + stdev * (1- np.random
76                             .uniform(0,2)))
77
78 Serie, CollWaves, coverage, detected, ok, _ =
79     SPLICE_ModelEvaluate(Nn, newPRMS, T, False, ctry
80     , data) #era True
81
82 if ok:
83
84     SerieArr.append(Serie)
85     WaveArr.append(CollWaves)
86     Coverage.append(coverage)
87     Detected.append(detected)
88
89     loss, ShLoss, DtLoss = LossFunc(data, ctry,
90     Serie, sdata)
91
92     LossArr.append(loss)
93
94 else:
```

```

90
91     var -= 1
92     print("INVALID PARAMETER COMBINATION ... TRYING
93           AGAIN")
94
95     std = np.std(SerieArr,axis=0)
96     mean = np.mean(SerieArr,axis=0)
97
98     stdcov = np.std(Coverage,axis=0)
99     meancov = np.mean(Coverage,axis=0)
100
101     stddet = np.std(Detected,axis=0)
102     meandet = np.mean(Detected,axis=0)
103
104     plt.figure(figsize=(12,4))
105     plt.plot(x, data[ctry]['cases'],'k.',label="data points
106            ")
107     plt.plot(x, sdata[ctry]['cases'],'r-',label="smoothed
108            time series")
109     plt.fill_between(np.arange(len(std)), (mean-std), (mean
110            +std), color='b', alpha=.3,label="model interval")
111     plt.title(str(nVariants)+ " conjuntos de parâmetros
112            variados aleatoriamente considerados" + \
113            "\nfaixa de variação dos parâmetros: "+r'$
114            \pm$'+str('{:0f}'.format(100*stdev))+ "% "
115            + \
116            "Intervalo de variação do modelo: "+r'$\pm$
117            '+str('{:0f}'.format(100*np.mean(std)/np.
118            mean(mean)))+ "% " + \
119            "\npopulação atingida: : "+str(100*np.array(
120            meancov))+ "% "+r'$\pm$'+str(100*np.array(

```

```

        stdcov))+"% "+ \
112     "casos detectados : "+str(100*np.array(
        meandet))+"% "+r'$\pm$'+str(100*np.array(
        stddet))+"% "+\
113     "\nErro mÃdio: "+str(np.mean(LossArr)))
114 plt.grid()
115 plt.legend(["dados brutos","serie suavizada","
        prediÃço do modelo"])
116 plt.savefig(f"splice_sensitivity_{label}.png",
        bbox_inches='tight')
117 plt.savefig(f"splice_sensitivity_{label}.pdf",format='
        pdf',bbox_inches='tight')
118 plt.show()

```

D.0.4.1 Show results of simulation with best params

```

1 newPRMS = cp.deepcopy(bPRMS[0])
2 Serie, CollWaves, coverage, detected, ok, results =
        SPLICE_ModelEvaluate(Nn, newPRMS, T, True, ctry, data)
3 print(CollWaves)
4 print(results)
5 print(ctry)
6 Loss, ShLoss, DtLoss = LossFunc(data, ctry, Serie, sdata)
7 print('Loss ',Loss, 'ShLoss ',ShLoss, 'DtLoss ',DtLoss)

```