

# Inferência de Parâmetros em um Modelo Epidemiológico em Rede para Simulação da Propagação da Dengue em Contexto Intramunicipal

Welber Paraizo Ferreira<sup>1</sup>, Antônio Tadeu Azevedo Gomes<sup>1</sup>, Márcia Ito<sup>2</sup>,  
Caroline de Oliveira Costa Souza Rosa<sup>3</sup>

<sup>1</sup> Laboratório Nacional de Computação Científica (LNCC),

<sup>2</sup> Centro Paula Souza (CPS),

<sup>3</sup> Universidade Federal Fluminense (UFF)

wparaizo@posgrad.lncc.br, atagomes@lncc.br

marcia.ito@cpspos.sp.gov.br, carolineocsr@id.uff.br

**Abstract.** *Network-based modeling of disease spread allows understanding spatiotemporal dissemination patterns and supporting epidemiological surveillance strategies. A recurring challenge in this type of modeling is estimating the probability of infection spreading from one node to another. To address this, we propose an approach based on the Generalized Inverse Infection Model (GIIM), applied to heterogeneous real-world data, integrating structural attributes (potential flow of people moving between regions), epidemiological attributes (local incidence), and environmental attributes (climatic variables) into a directed graph representing the main connectivity pathways between regions. The probability of transmission between regions is modeled through a logistic function parameterized over these attributes, with coefficients estimated via Genetic Algorithms, aiming at predicting the onset timing of local outbreaks. The process includes stochastic simulations using the Susceptible-Infected (SI) model and evaluation through the Monte Carlo method. The contribution of this paper lies in the adaptation and application of a reproducible framework for calibrating epidemiological models on a real intramunicipal-scale network, combining heterogeneous data integration, evolutionary optimization, and rigorous quantitative evaluation, with potential application in public health decision-support systems. The results indicate a high predictive capability of the model in identifying the onset of local outbreaks and suggest that epidemiological and environmental factors exert a more significant influence on dengue spread than potential mobility flows between regions.*

**Resumo.** *A modelagem em redes da propagação de doenças permite compreender padrões espaço-temporais de disseminação e apoiar estratégias de vigilância epidemiológica. Um desafio recorrente neste tipo de modelagem é estimar a probabilidade de espalhamento da infecção de um nó para outro da rede. Para tanto, propomos uma abordagem baseada no Modelo de Infecção Inversa Generalizado (GIIM) aplicada a dados reais heterogêneos, integrando atributos estruturais (fluxo potencial de pessoas se movimentando entre regiões), epidemiológicos (incidência local) e ambientais (variáveis climáticas) em um grafo*

*dirigido representando as vias principais de ligação entre regiões. A probabilidade de transmissão entre regiões é modelada por meio de uma função logística parametrizada sobre esses atributos, com coeficientes estimados por meio de Algoritmos Genéticos visando prever o momento de início de surtos locais. O processo inclui simulações estocásticas via modelo Suscetível-Infected (SI) e avaliação pelo método de Monte Carlo. A contribuição do artigo reside na adaptação e emprego de um arcabouço reprodutível para calibração de modelos epidemiológicos em uma rede real de escala intramunicipal, combinando integração de dados heterogêneos, otimização evolutiva e avaliação quantitativa rigorosa, com potencial aplicação em sistemas de apoio à decisão em saúde pública. Os resultados obtidos indicam elevada capacidade preditiva do modelo na identificação do início de surtos locais e sugerem que fatores epidemiológicos e ambientais exercem influência mais significativa na propagação da dengue do que os fluxos potenciais de mobilidade entre regiões.*

## **1. Introdução**

As arboviroses (como dengue, zika e chikungunya) são um problema de saúde pública global e trazem um complicador na modelagem epidemiológica tradicional devido à interação entre humanos, vetores (mosquitos) e condições ambientais, e às conexões que permitem a exportação do vírus entre localidades infectadas e suscetíveis. Nesse sentido, a modelagem em redes da propagação de doenças infecciosas permite capturar a heterogeneidade de contatos, a mobilidade urbana e outras dependências espaço-temporais. Diversos trabalhos exploram essa alternativa de modelagem, apresentando modelos que se propõem a prever números de casos [Liu et al. 2020], prover alertas [Sood et al. 2023], ou apoiar a tomada de decisão quanto a medidas preventivas [Reyna-Lara et al. 2022]. No entanto, a maioria dos modelos de propagação exige, como entrada, probabilidades reais de infecção nas arestas da rede, probabilidades estas que, no geral, não estão disponíveis.

Para superar a lacuna supracitada, o Modelo de Infecção Inversa Generalizado (*Generalized Inverse Infection Model – GIIM*) [Bóta and Gardner 2023] surge como uma abordagem inovadora, capaz de estimar essas probabilidades a partir de observações reais de surtos epidemiológicos. Estudos anteriores utilizando o GIIM demonstraram sucesso ao identificar como fatores determinantes o tamanho da população e a umidade absoluta em epidemias de H1N1 na Suécia [Bóta et al. 2021], ou o Produto Interno Bruto (PIB) e a adequação do habitat do vetor em epidemias de zika nas Américas [Gardner et al. 2018].

Neste artigo, aplicamos o GIIM para analisar a propagação da dengue no município de São José do Rio Preto – SP, representando cada região como um nó em uma rede cujas arestas indicam rotas potenciais de transmissão. Devido à escassez de dados abertos de mobilidade, essas rotas foram baseadas em fluxos de transporte público e na proximidade entre regiões, sendo calculadas por um modelo gravitacional. No GIIM, as probabilidades de infecção são formuladas como uma função logística aplicada a uma soma ponderada de atributos conhecidos e normalizados, permitindo quantificar a contribuição de cada atributo. Neste trabalho, usamos como atributos:

- Fatores Ambientais: consideramos a temperatura média, a precipitação e a umidade relativa, fatores que regulam o ciclo de vida dos mosquitos;
- Fatores Demográficos: utilizamos a densidade populacional como um indicador crítico da disponibilidade de hospedeiros humanos;

- Fatores Epidemiológicos: utilizamos as taxas de incidência da dengue nas regiões;
- Mobilidade: representamos a mobilidade pela quantidade e periodicidade de ônibus disponíveis entre as regiões.

A otimização dos parâmetros do modelo que representam a ponderação dos atributos acima é realizada por um Algoritmo Genético que busca minimizar o erro entre simulações estocásticas (via método de Monte Carlo) baseadas no modelo Suscetível-Infectado (SI) e dados reais de início do surto em cada região. O erro é medido por meio da área sob a curva Característica de Operação do Receptor (*Receiver Operating Characteristic* – ROC). O objetivo central deste estudo é avaliar a adaptação e o emprego do GIIM para calibração de modelos epidemiológicos em uma rede real de escala intramunicipal. Os resultados obtidos corroboram a literatura associada ao GIIM [Gardner et al. 2018, Bóta et al. 2021], apontando contribuição moderada a reduzida da mobilidade urbana frente aos demais atributos. Entendemos que a evolução desse modelo, incorporando novos atributos, pode consolidar-se como ferramenta de suporte à decisão para a alocação estratégica de recursos de vigilância e intervenções de controle, mitigando o impacto da dengue e outras arboviroses em populações vulneráveis.

O restante do artigo está estruturado como se segue. Na Seção 2 apresentamos os trabalhos relacionados e identificamos a contribuição deste artigo frente aos mesmos. Na Seção 3 apresentamos o GIIM e sua adaptação ao contexto intramunicipal de São José do Rio Preto – SP. Na Seção 4 apresentamos os dados reais disponíveis e como eles são aplicados no GIIM. Na Seção 5 detalhamos as simulações executadas necessárias ao GIIM, e na Seção 6 os resultados obtidos. Finalmente, na Seção 7 apresentamos nossas considerações finais e perspectivas de trabalhos futuros.

## 2. Trabalhos Relacionados

A modelagem da propagação da dengue tem sido muito investigada sob diferentes perspectivas, envolvendo abordagens estatísticas, modelos mecânicos e técnicas baseadas em aprendizado de máquina. Diversos estudos têm buscado compreender o papel relativo de fatores ambientais, demográficos e de mobilidade urbana na dinâmica de transmissão da doença [Antonio et al. 2017, Xavier et al. 2021, Bomfim et al. 2020, Kiang et al. 2021, Chen and Moraga 2025].

Em contextos de análises epidemiológicas em larga escala, [Antonio et al. 2017] investigaram a propagação da dengue em todos os municípios brasileiros durante um período de 12 anos. Os autores demonstraram que a distribuição espacial de casos segue leis de potência ao longo do tempo e que a incidência da doença escala aproximadamente de forma linear com o tamanho da população do município. O estudo sugere que a dispersão da doença em longas distâncias está associada principalmente à circulação de indivíduos infectados, em vez da mobilidade limitada do vetor.

A influência de fatores ambientais na incidência de dengue tem sido bastante estudada. [Xavier et al. 2021] investigaram a região metropolitana do Rio de Janeiro utilizando dados climáticos obtidos por sensoriamento remoto, vindo de satélites. Integrando essas informações em modelos ARMAX, os autores mostraram que variáveis como temperatura e precipitação podem ser utilizadas para prever a incidência de dengue com horizonte de até um ano, destacando o potencial de dados climáticos derivados de satélite como ferramenta de apoio à vigilância epidemiológica em regiões tropicais.

Abordagens baseadas em redes neurais também têm sido utilizadas para modelar a propagação da doença. [Chen and Moraga 2025] propuseram um modelo baseado em redes neurais recorrentes para prever casos de dengue em diferentes cidades brasileiras. O estudo integra múltiplos modais de transporte, incluindo aéreo, rodoviário e hidroviário, permitindo estimar o efeito de casos importados entre regiões. Os resultados indicam que modelos que ignoram o fluxo populacional tendem a subestimar o risco de surtos, enquanto a inclusão de informações de mobilidade melhora a capacidade preditiva.

A mobilidade urbana também tem sido estudada em escalas espaciais menores. [Bomfim et al. 2020] analisaram a dinâmica da dengue na cidade de Fortaleza utilizando dados de transporte público para estimar o fluxo de pessoas entre 119 bairros. O estudo comparou modelos mecânicos baseados em compartimentos epidemiológicos, como SEIR, com abordagens baseadas em redes neurais, demonstrando que a incorporação de dados de mobilidade urbana melhora a capacidade de prever tanto o tempo de ocorrência quanto a intensidade dos picos epidêmicos.

Outra fonte de dados para estimar padrões de mobilidade urbana são os registros de telefonia celular. [Kiang et al. 2021] analisaram registros de chamadas de aproximadamente 11 milhões de usuários na Tailândia para estimar a conectividade entre províncias. Os resultados mostraram que regiões geograficamente distantes, mas fortemente conectadas por deslocamentos populacionais, apresentam padrões de incidência de dengue altamente correlacionados. Além disso, os autores propuseram um sistema de previsão baseado em um “mosaico adaptativo”, no qual diferentes modelos estatísticos são utilizados dinamicamente de acordo com o horizonte temporal e o contexto espacial analisado.

Embora diversos estudos apontem a mobilidade urbana como um fator importante na propagação da doença, outros trabalhos indicam que sua influência pode variar dependendo do contexto epidemiológico. [Leandro et al. 2024] investigaram a dinâmica da dengue em Foz do Iguaçu durante o período de restrições de mobilidade impostas pela pandemia de COVID-19. Utilizando Mapas de Autovetores de Moran (MEM) para analisar a estrutura espacial da incidência, os autores observaram que a transmissão manteve forte estruturação espacial local mesmo em cenários de mobilidade, sugerindo que fatores ambientais e estruturais podem desempenhar papel dominante em determinadas situações.

Paralelamente a essas abordagens, métodos baseados em redes têm sido propostos para modelar explicitamente o processo de propagação da doença entre regiões conectadas. Nesse contexto, o *Generalized Inverse Infection Model* (GIIM) [Bóta and Gardner 2023] propõe um método geral para estimar probabilidades de transmissão em arestas de uma rede a partir de dados observados de propagação. O modelo formula o problema como uma tarefa de otimização inversa, na qual os parâmetros que definem as probabilidades de infecção são ajustados de forma a reproduzir os padrões observados de disseminação da doença.

Em uma perspectiva mais geral de processos de contágio em redes complexas, [Brockmann and Helbing 2013] demonstraram que padrões espaço-temporais aparentemente irregulares de epidemias podem ser descritos de forma mais simples quando a distância geográfica é substituída por uma distância efetiva derivada dos fluxos de mobilidade entre regiões. Esse resultado reforça a relevância de representar sistemas epidemiológicos por redes ponderadas por deslocamento humano.

Diferentemente de abordagens puramente estatísticas ou baseadas em aprendizado profundo, o GIIM incorpora explicitamente atributos associados às arestas da rede, possibilitando analisar o impacto relativo de diferentes fatores – como mobilidade, condições ambientais e características demográficas – na dinâmica de propagação. Estudos anteriores que utilizaram o GIIM concentraram-se em escalas espaciais amplas, como redes nacionais [Bóta et al. 2021] ou internacionais [Gardner et al. 2018], nas quais os nós representam grandes regiões geográficas conectadas por fluxos populacionais agregados.

Neste trabalho, exploramos o uso do GIIM em uma escala intraurbana. Neste caso, modelamos a propagação da dengue entre regiões de uma mesma área urbana utilizando dados de mobilidade derivados da rede de transporte público disponível. Essa perspectiva permite investigar como padrões de deslocamento dentro da cidade interagem com fatores ambientais e demográficos para influenciar a disseminação da doença, oferecendo uma análise complementar às abordagens em larga escala predominantes na literatura.

### 3. Metodologia

#### 3.1. Construção da Rede Espacial

A propagação espacial da doença foi modelada sobre um grafo direcionado e ponderado  $G = (V, E)$ , no qual cada vértice  $v \in V$  representa um bairro (segundo classificação do Instituto Brasileiro de Geografia e Estatística – IBGE) do município de São José do Rio Preto – SP, enquanto cada aresta direcionada  $(u, v) \in E$  representa um fluxo potencial de mobilidade entre os respectivos bairros. O peso estrutural associado a cada aresta  $(u, v)$  é denotado por  $F_{uv}$  e representa a intensidade relativa de deslocamento entre os bairros. O fluxo foi estimado a partir de um modelo gravitacional (vide Seção 4), utilizado para aproximar padrões de mobilidade urbana.

Cada vértice da rede possui atributos epidemiológicos e demográficos. Em particular, cada bairro  $u$  é associado a uma série temporal de incidência epidemiológica  $I_u(t)$ , enquanto a população residente  $pop_v$  é associada ao bairro de destino  $v$ . Adicionalmente, o modelo incorpora variáveis ambientais globais dependentes do tempo  $t$ , incluindo temperatura média  $T(t)$ , umidade relativa  $U(t)$  e precipitação acumulada  $P(t)$ . O horizonte temporal considerado é  $t = 0, \dots, 52$ , correspondente às 53 semanas epidemiológicas analisadas no ano de 2022.

#### 3.2. Modelo Epidemiológico

A dinâmica de propagação foi modelada por meio de um processo estocástico do tipo Suscetível–Infectado (SI), no qual cada região pode assumir os estados suscetível ou infectado. Uma vez infectado, o bairro permanece nesse estado até o final da simulação.

A escolha desse modelo simplificado se justifica por duas razões. A primeira delas diz respeito ao objetivo do presente trabalho, que é analisar a dinâmica de propagação espacial da doença entre regiões conectadas, e não modelar detalhadamente o ciclo completo de recuperação. Cumpre destacar que, no contexto do GIIM, o modelo SI é recorrentemente utilizado por sua simplicidade estrutural e por permitir representar de forma direta o processo de difusão da infecção em uma rede ao longo do tempo, sem introduzir parâmetros adicionais de recuperação ou imunidade, sendo uma abordagem validada em estudos anteriores [Bóta and Gardner 2023].

Além disso, os dados epidemiológicos utilizados neste estudo correspondem a séries temporais agregadas de incidência semanal, nas quais a distinção explícita entre estados de recuperação ou suscetibilidade renovada não é diretamente observável. Nesse cenário, o modelo SI fornece uma representação consistente para inferir padrões de transmissão espacial a partir das observações disponíveis.

A probabilidade de transmissão da infecção de um bairro  $u$  para um bairro  $v$  no tempo  $t$  é definida por  $\sigma(w_{uv}(t))$ , com:

$$w_{uv}(t) = A + \alpha F_{uv} + \beta I_u(t) + \gamma T(t) + \delta U(t) + \eta P(t) + \zeta pop_v \quad (1)$$

onde  $\sigma(\cdot)$  representa a função sigmoide  $\sigma(x) = \frac{1}{1+e^{-x}}$ . O termo  $A$  corresponde ao intercepto do modelo, enquanto  $\alpha, \beta, \gamma, \delta, \eta$  e  $\zeta$  representam os coeficientes associados aos atributos considerados. Todos esses parâmetros são estimados durante o processo de otimização. Em conjunto, os coeficientes indicam a contribuição relativa dos diferentes fatores incluídos no modelo, como mobilidade urbana, condições ambientais e características demográficas, na determinação da probabilidade de transmissão entre regiões conectadas na rede. Essa formulação permite integrar simultaneamente diferentes tipos de fatores na estimativa das probabilidades de transmissão entre bairros.

A probabilidade de um bairro suscetível  $v$  tornar-se infectado em  $t$  é dada por:

$$P_v(t) = 1 - \prod_{u \in \mathcal{N}_v^I(t)} (1 - \sigma(w_{uv}(t))), \quad (2)$$

onde  $\mathcal{N}_v^I(t)$  representa o conjunto de bairros vizinhos de  $v$  no estado infectado no tempo  $t$ . Essa formulação assume independência entre as tentativas de transmissão provenientes de diferentes bairros vizinhos infectados.

#### 4. Base de Dados

Os dados utilizados neste estudo referem-se à incidência semanal de casos de dengue registrados no município de São José do Rio Preto – SP ao longo do ano de 2022. As séries temporais de incidência foram agregadas por bairro da cidade, permitindo representar cada bairro como um nó na rede de propagação da doença.

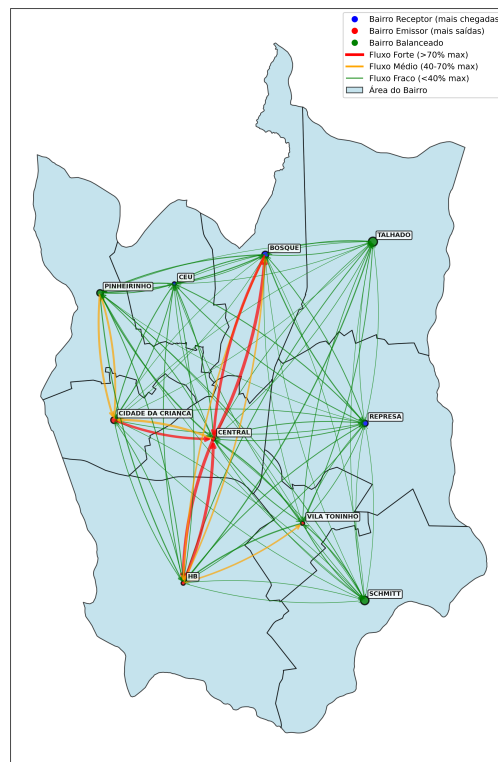
Informações demográficas foram obtidas a partir do Censo Demográfico de 2022 do IBGE [Brasil 2023]. A população residente em cada bairro foi utilizada como um indicador do potencial de suscetibilidade local, refletindo a disponibilidade de hospedeiros humanos para a transmissão da doença.

Além disso, o modelo incorpora variáveis ambientais associadas à dinâmica de transmissão da dengue, incluindo temperatura média, precipitação acumulada e umidade relativa. Essas variáveis são conhecidas por influenciar diretamente o ciclo de vida do mosquito *Aedes aegypti*, afetando parâmetros como sobrevivência, reprodução e taxa de desenvolvimento do vetor.

Por fim, foram considerados indicadores de mobilidade urbana entre os bairros da cidade. Devido à ausência de dados diretos de deslocamento populacional, os fluxos potenciais de mobilidade foram aproximados utilizando informações da rede de transporte

público urbano. Para isso, foram analisadas as rotas das linhas de ônibus, representadas como seqüências de bairros atendidos ao longo de cada percurso. A partir dessas seqüências, foram identificadas transições entre bairros consecutivos, permitindo estimar a conectividade entre diferentes regiões da cidade.

Com base nessas transições, foi construída uma medida de custo de deslocamento entre bairros, definida de forma inversamente proporcional à frequência de conexões observadas nas rotas do transporte público. Esses elementos foram então utilizados na construção de uma rede de interação entre bairros por meio de um modelo gravitacional [Capoani 2023], no qual a intensidade potencial de fluxo entre dois bairros é proporcional à oferta de transporte nas regiões de origem e destino e inversamente proporcional ao custo estimado de deslocamento entre elas. A adoção dessa abordagem está alinhada a estudos anteriores que demonstraram a aplicabilidade de modelos gravitacionais para estimar mobilidade intraurbana a partir de redes de ônibus, especialmente em cidades nas quais esse sistema representa a principal estrutura de deslocamento urbano [Hong and Jung 2016]. A Figura 1 ilustra a rede de interação obtida para São José do Rio Preto – SP a partir dessa abordagem.



**Figura 1. Rede de interação entre bairros de São José do Rio Preto – SP, com indicação dos bairros mais influentes em termos de graus de chegada e de saída dos vértices e intensidade potencial de fluxos nas arestas.**

## 5. Experimentos Computacionais

Antes da execução do modelo, todos os atributos utilizados foram normalizados para o intervalo  $[0, 1]$  por meio de normalização min-max. Esse procedimento foi adotado devido à heterogeneidade de escala entre as variáveis consideradas, como incidência epide-

miológica, população e variáveis ambientais, evitando que atributos com maior magnitude numérica dominem o processo de calibração dos parâmetros do modelo. Embora a normalização min-max tenha sido adotada pela sua simplicidade e compatibilidade com o processo de otimização, reconhece-se sua sensibilidade a valores extremos. A investigação de *outliers* e a comparação com métodos robustos de normalização serão consideradas em trabalhos futuros.

A observação real utilizada no processo de calibração foi construída a partir das séries temporais de incidência epidemiológica. Para cada região da rede, foi definida uma série binária de surtos baseada na detecção de anomalias na incidência semanal. Especificamente, um surto é identificado quando a incidência observada excede a média histórica local acrescida de duas vezes o desvio padrão, ou seja,

$$I_v(t) > \mu_v(t) + 2\phi_v(t)$$

onde  $\mu_v(t)$  e  $\phi_v(t)$  são calculados a partir de uma janela móvel de oito semanas. A escolha do multiplicador de duas vezes o desvio padrão busca identificar aumentos estatisticamente significativos na incidência. A janela móvel de oito semanas foi adotada para capturar o padrão epidemiológico local de curto prazo, fornecendo uma estimativa estável da média e da variação sem levar em consideração períodos de tempo muito grandes.

Adicionalmente, impõe-se um critério de persistência mínima de três semanas consecutivas, no qual a incidência deve permanecer acima da média histórica da janela considerada. Esse critério tem o objetivo de reduzir a detecção de picos isolados, garantindo que apenas aumentos sustentados na incidência sejam classificados como surtos. Uma vez detectado o início do surto, o evento permanece em surto até o final da série temporal analisada. Esse processo resulta em uma matriz binária  $Y_v(t)$  indicando a ocorrência de surtos em cada região ao longo do tempo.

A dinâmica de propagação da doença na rede é simulada utilizando um processo estocástico baseado no modelo epidemiológico descrito anteriormente. No tempo inicial, o estado de cada região é amostrado a partir de uma probabilidade proporcional à incidência observada na semana inicial, normalizada pela população local. A evolução temporal do processo é então simulada de acordo com as probabilidades de transmissão associadas às arestas da rede.

Devido à natureza probabilística do modelo, a simulação é realizada por meio do método de Monte Carlo. Para cada conjunto de parâmetros do modelo são executadas 10 simulações independentes da dinâmica de propagação. Em cada execução, registramos o estado de cada região ao longo do tempo. A probabilidade estimada de ocorrência de surto em cada região e instante temporal é então obtida como a frequência relativa de infecção observada nas simulações.

A estimativa dos parâmetros do modelo é realizada por meio de um Algoritmo Genético, implementado utilizando a biblioteca PyGAD [Gad 2024]. Nesse processo, cada indivíduo da população representa um conjunto de parâmetros do modelo, incluindo o termo de viés (intercepto) e os coeficientes associados aos diferentes atributos considerados na função de transmissão. A qualidade de cada solução é avaliada comparando-se as probabilidades de surto simuladas com a matriz binária de surtos observados.<sup>1</sup>

---

<sup>1</sup>Detalhes acerca desses experimentos estão disponíveis para consulta em: <https://github.com/>

Como função objetivo, utiliza-se a área sob a curva ROC (*Area Under Curve – AUC*), que mede a capacidade do modelo em distinguir corretamente eventos de surto e não-surto ao longo da rede e do tempo. O processo evolutivo utiliza seleção por torneio, cruzamento de ponto único, mutação aleatória e elitismo, permitindo explorar o espaço de parâmetros em busca da configuração que maximiza a concordância entre as simulações do modelo e os padrões epidemiológicos observados.

## 6. Resultados e discussões

Para avaliar o impacto relativo dos diferentes fatores presentes no modelo, foram realizados dois experimentos computacionais distintos. No primeiro experimento, o modelo foi ajustado utilizando os valores reais de mobilidade entre as regiões, conforme descrito na Seção 4. No segundo experimento, foram utilizados valores sintéticos de mobilidade. Esses valores sintéticos foram gerados a partir da transformação  $1 - F_{uv}$ , aplicada aos valores normalizados de mobilidade. Essa estratégia foi adotada porque, após o processo de normalização, a maioria dos valores de mobilidade apresentou magnitude próxima de zero. Dessa forma, a transformação permitiu avaliar o comportamento do modelo em um cenário onde os fluxos de mobilidade assumem valores elevados.

Em ambos os experimentos, a otimização dos parâmetros foi realizada com um algoritmo genético executado por 500 gerações. No experimento utilizando os dados reais de mobilidade, o processo de otimização alcançou um valor de AUC igual a 0.99986965. Já no experimento utilizando os dados sintéticos de mobilidade, o valor de AUC obtido foi 0.99993483. Esses valores indicam que, em ambos os cenários, o modelo apresentou alta capacidade de discriminar corretamente os períodos de ocorrência de surtos.

A função peso utilizada para estimar a probabilidade de transmissão em cada aresta do grafo é definida a partir de uma combinação linear dos atributos considerados no modelo, seguida da aplicação de uma função sigmoide. No experimento com dados reais de mobilidade, os coeficientes estimados pelo processo de otimização foram:

$$w_{uv}(t) = 0.365 + 0.475 \cdot F_{uv} + 0.105 \cdot I_u(t) - 0.030 \cdot T(t) - 0.373 \cdot U(t) - 0.397 \cdot P(t) + 0.089 \cdot pop_v$$

A análise das partes lineares (produtos dos coeficientes pelos valores normalizados dos atributos) indica que os fatores ambientais desempenham papel relevante na dinâmica de transmissão modelada. Em particular, as partes lineares associadas à umidade e à precipitação apresentaram as maiores contribuições relativas, com frações de variância de 30.9% e 30.2%, respectivamente. A parte linear associada à mobilidade apresentou fração de variância de 21.9%, indicando participação moderada na formação da probabilidade de transmissão. A parte linear associada à incidência na região de origem também apresentou contribuição moderada, refletindo a influência da dinâmica local na propagação da doença. Já a parte linear associada à temperatura apresentou contribuição bastante reduzida, com fração inferior a 0.2% da variância total.

No experimento com mobilidade sintética, os coeficientes estimados pelo algoritmo genético resultaram na seguinte função peso:

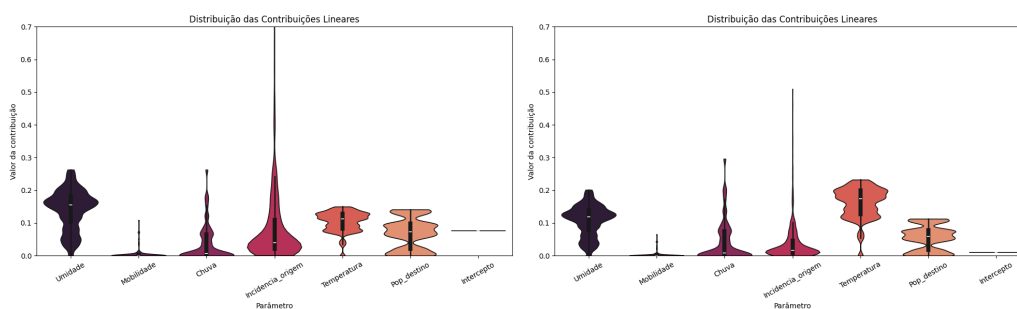
$$w_{uv}(t) = 0.196 + 0.142 \cdot F_{uv} - 0.14 \cdot I_u(t) + 0.26 \cdot T(t) - 0.484 \cdot U(t) - 0.104 \cdot P(t) + 0.148 \cdot pop_v$$

Nesse cenário, observa-se uma mudança significativa na importância relativa das variáveis. A parte linear associada à umidade continuou tendo a maior contribuição para a variância, com 51.8%. Em seguida, destacam-se a incidência na região de origem e a temperatura, cujas partes lineares apresentam frações de variância de 23.3% e 11.2%, respectivamente. A mobilidade, apesar de ter sido artificialmente amplificada no experimento, apresentou contribuição relativamente pequena para a variância total da parte linear, correspondendo a aproximadamente 1.9%.

Outro aspecto observado foi a inversão de sinal nos coeficientes das variáveis incidência  $I_u(t)$  e temperatura  $T(t)$  entre os dois experimentos. Esse comportamento pode estar relacionado à redistribuição da influência relativa dos atributos durante o processo de otimização, especialmente quando há alterações na estrutura dos fluxos de mobilidade. Como a função de peso é definida por uma combinação linear dos fatores considerados, diferentes configurações dos dados podem levar o algoritmo a encontrar soluções com desempenho preditivo semelhante, ainda que com sinais distintos para alguns coeficientes. Além disso, possíveis correlações entre variáveis ambientais e a dinâmica da incidência podem gerar efeitos compensatórios durante o ajuste dos parâmetros.

Esses resultados mostram que, mesmo com os fluxos de mobilidade artificialmente elevados, os fatores ambientais continuam desempenhando papel dominante na dinâmica de propagação capturada pelo modelo. Esse comportamento pode indicar que a estrutura epidemiológica representada pelo modelo é mais sensível às condições ambientais e à dinâmica local da incidência do que à intensidade dos fluxos de mobilidade entre regiões.

A Figura 2 apresenta a distribuição dos valores das partes lineares, em módulo, associadas a cada atributo ao longo das 53 semanas analisadas para os dois experimentos.



**Figura 2. Distribuição dos valores das partes lineares ao longo das 53 semanas para os dois experimentos. À esquerda são apresentados os resultados utilizando os dados de fluxo potencial calculado a partir do modelo gravitacional, enquanto à direita são apresentados os resultados utilizando os dados sintéticos de fluxo.**

A análise conjunta desses resultados indica que o modelo é capaz de capturar de forma consistente os padrões de propagação presentes nos dados, mantendo elevada capacidade preditiva mesmo sob diferentes configurações dos fluxos de mobilidade. Contudo, os experimentos sugerem que os fatores ambientais e a dinâmica epidemiológica local apresentam influência mais significativa na determinação das probabilidades de transmissão estimadas pelo modelo.

Cabe destacar que os experimentos conduzidos nesta etapa tiveram como objetivo investigar a sensibilidade do modelo em relação aos fluxos de mobilidade entre regiões,

por meio da comparação entre cenários com mobilidade observada e mobilidade sintética. Assim, não foi realizada uma análise sistemática da contribuição individual dos demais atributos considerados, como temperatura, precipitação, umidade, densidade populacional e incidência local, o que constitui uma direção relevante para trabalhos futuros. Ligado a isso, a incorporação de procedimentos formais de análise de incerteza dos parâmetros estimados permanece como perspectiva para investigações futuras.

## 7. Conclusões

Este trabalho apresentou um modelo de propagação da dengue baseado em redes, no qual diferentes fatores associados à dinâmica epidemiológica foram incorporados para estimar a probabilidade de transmissão entre regiões de um município. O trabalho integrou informações de mobilidade urbana, variáveis ambientais e dados de incidência semanal da doença, permitindo representar a dinâmica espacial da transmissão com um grafo.

Os experimentos demonstraram que o modelo apresenta elevada capacidade de reproduzir os surtos observados nos dados reais, alcançando valores elevados da métrica AUC nos cenários analisados. A análise das contribuições dos atributos indicou que fatores ambientais, especialmente umidade e precipitação, exercem influência significativa na variação da função de transmissão estimada pelo modelo, enquanto a mobilidade apresentou participação de moderada a reduzida dependendo da configuração considerada.

De forma geral, os resultados sugerem que a dinâmica local da incidência e as condições ambientais desempenham papel relevante na propagação da doença dentro do sistema modelado. A estrutura proposta mostra-se promissora como ferramenta para análise da disseminação espacial da dengue e pode contribuir para o entendimento dos fatores que influenciam a ocorrência de surtos.

Como limitação do presente estudo, destaca-se que a avaliação foi conduzida com dados referentes a um único município e a um período epidemiológico específico. Assim, a capacidade de generalização da abordagem para outros contextos urbanos e para diferentes horizontes temporais ainda deve ser investigada.

Como trabalhos futuros, pretende-se incorporar dados mais detalhados de mobilidade populacional, bem como explorar outras variáveis ambientais e socioeconômicas que possam influenciar a dinâmica de transmissão da doença. Além disso, a aplicação do modelo em diferentes municípios e a validação em anos subsequentes poderão contribuir para avaliar sua robustez preditiva em distintos contextos epidemiológicos.

## Agradecimentos

Este trabalho contou com apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Processo nº 2023/10080-3, à qual os autores agradecem pelo financiamento.

## Referências

- Antonio, F. J., Itami, A. S., de Picoli, S., Teixeira, J. J. V., and Mendes, R. d. S. (2017). Spatial patterns of dengue cases in Brazil. *PloS one*, 12(7):e0180715.
- Bomfim, R., Pei, S., Shaman, J., Yamana, T., Makse, H. A., Andrade, J. S., Lima Neto, A. S., and Furtado, V. (2020). Predicting dengue outbreaks at neighbourhood level using human mobility in urban areas. *Journal of the Royal Society Interface*, 17(171).

- Bóta, A. and Gardner, L. (2023). A generalized framework for the estimation of edge infection probabilities. *Algorithms*, 16(8):390.
- Bóta, A., Holmberg, M., Gardner, L., and Rosvall, M. (2021). Socioeconomic and environmental patterns behind H1N1 spreading in Sweden. *Scientific Reports*, 11(1):22512.
- Brasil, I. (2023). Censo demográfico 2022. *Dados nacionais. Fundação Instituto Brasileiro de Geografia e Estatística. Brasil*.
- Brockmann, D. and Helbing, D. (2013). The hidden geometry of complex, network-driven contagion phenomena. *science*, 342(6164):1337–1342.
- Capoani, L. (2023). Review of the gravity model: origins and critical analysis of its theoretical development. *SN Business & Economics*, 3(5):95.
- Chen, X. and Moraga, P. (2025). Dengue forecasting and outbreak detection in Brazil using LSTM: integrating human mobility and climate factors. *Infectious Disease Modelling*.
- Gad, A. F. (2024). Pygad: An intuitive genetic algorithm python library. *Multimedia tools and applications*, 83(20):58029–58042.
- Gardner, L. M., Bota, A., Gangavarapu, K., Kraemer, M. U., and Grubaugh, N. D. (2018). Inferring the risk factors behind the geographical spread and transmission of zika in the Americas. *PLoS neglected tropical diseases*, 12(1):e0006194.
- Hong, I. and Jung, W.-S. (2016). Application of gravity model on the korean urban bus network. *Physica A: Statistical Mechanics and its Applications*, 462:48–55.
- Kiang, M. V., Santillana, M., Chen, J. T., Onnela, J.-P., Krieger, N., Engø-Monsen, K., Ekapirat, N., Areechokchai, D., Prempre, P., Maude, R. J., et al. (2021). Incorporating human mobility data improves forecasts of dengue fever in Thailand. *Scientific reports*, 11(1):923.
- Leandro, A. S., Chiba de Castro, W. A., Garey, M. V., and Maciel-de Freitas, R. (2024). Spatial analysis of dengue transmission in an endemic city in Brazil reveals high spatial structuring on local dengue transmission dynamics. *Scientific reports*, 14(1):8930.
- Liu, K., Zhang, M., Xi, G., Deng, A., Song, T., Liid, Q., Kang, M., and Yin, L. (2020). Enhancing fine-grained intra-urban dengue forecasting by integrating spatial interactions of human movements between urban regions. *PLoS Neglected Tropical Diseases*, 14(12):1–22.
- Reyna-Lara, A., Soriano-Paños, D., Arias-Castro, J. H., Martínez, H. J., and Gómez-Gardeñes, J. (2022). A metapopulation approach to identify targets for Wolbachia-based dengue control. *Chaos*, 32(4).
- Sood, S. K., Sood, V., Mahajan, I., and Sahil (2023). An intelligent healthcare system for predicting and preventing dengue virus infection. *Computing*, 105(3):617–655.
- Xavier, L. L., Honório, N. A., Pessanha, J. F. M., and Peiter, P. C. (2021). Analysis of climate factors and dengue incidence in the metropolitan region of Rio de Janeiro, Brazil. *Plos one*, 16(5):e0251403.