

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA CIVIL E AMBIENTAL**

**AVALIAÇÃO E MODELAGEM DA RELAÇÃO ENTRE
DENGUE, CLIMA E SANEAMENTO EM MUNICÍPIOS DO
ESTADO DO RIO DE JANEIRO**

THIAGO TEIXEIRA MENDONÇA

ORIENTADOR: CARLOS HENRIQUE RIBEIRO LIMA

**DISSERTAÇÃO DE MESTRADO EM TECNOLOGIA AMBIENTAL E
RECURSOS HÍDRICOS**

BRASÍLIA/DF: SETEMBRO – 2025

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA CIVIL E AMBIENTAL**

**AVALIAÇÃO E MODELAGEM DA RELAÇÃO ENTRE DENGUE,
CLIMA E SANEAMENTO EM MUNICÍPIOS DO ESTADO DO RIO
DE JANEIRO**

THIAGO TEIXEIRA MENDONÇA

**DISSERTAÇÃO SUBMETIDA AO DEPARTAMENTO DE ENGENHARIA CIVIL
E AMBIENTAL DA FACULDADE DE TECNOLOGIA DA UNIVERSIDADE DE
BRASÍLIA COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A
OBTENÇÃO DO GRAU DE MESTRE EM TECNOLOGIA AMBIENTAL E
RECURSOS HÍDRICOS.**

APROVADA POR:

**Prof CARLOS HENRIQUE RIBEIRO LIMA , PhD (ENC-UnB)
(Orientador)**

**Prof. RAQUEL MORAES SOARES, PhD (ENC-UnB)
(Examinador Interno)**

**Prof. CELSO AUGUSTO GUIMARÃES SANTOS, PhD (UFPB/CT/DECA)
(Examinador Externo)**

BRASÍLIA/DF, 26 DE SETEMBRO DE 2025

FICHA CATALOGRÁFICA

MENDONÇA, THIAGO TEIXEIRA

Avaliação e modelagem da relação entre Dengue, clima e saneamento em municípios do estado do Rio de Janeiro [Distrito Federal] 2025.

x, 117p., 210 x 297 mm (ENC/FT/UnB, Mestre, Tecnologia Ambiental e Recursos Hídricos, 2025).

Dissertação de Mestrado – Universidade de Brasília. Faculdade de Tecnologia.

Departamento de Engenharia Civil e Ambiental.

1. Dengue

2. Saneamento Básico

3. Saúde Coletiva

4. Modelagem Matemática

5. Sistema Nacional de Informações em Saneamento Básico

I. ENC/FT/UnB

II. Título (série)

REFERÊNCIA BIBLIOGRÁFICA

THIAGO, T.M. (2025). Estudo de relações entre Dengue, clima e saneamento em municípios do estado do Rio de Janeiro. Dissertação de Mestrado em Tecnologia Ambiental e Recursos Hídricos. Publicação MTARH.DM-17A/25. Departamento de Engenharia Civil e Ambiental, Universidade de Brasília, Brasília, DF, 117p.

CESSÃO DE DIREITOS

AUTOR: Thiago Teixeira Mendonça

TÍTULO DA DISSERTAÇÃO DE MESTRADO: Estudo de relações entre Dengue, clima e saneamento em municípios do estado do Rio de Janeiro

GRAU: Mestre em Tecnologia Ambiental e Recursos Hídricos ANO: 2025

É concedida à Universidade de Brasília a permissão para reproduzir cópias desta monografia de Projeto Final e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte desta monografia de Projeto Final pode ser reproduzida sem a autorização por escrito do autor.

Thiago Teixeira Mendonça
SQN 406 Bloco K, Asa Norte.
70847-110 Brasília – DF – Brasil.

*"Rompi tratados
Traí os ritos
Quebrei a lança
Lancei no espaço
Um grito, um desabafo
E o que me importa
É não estar vencido
Minha vida, meus mortos
Meus caminhos tortos
Meu Sangue Latino
Minh'alma cativa."
Paulinho Mendonça*

Dedicatória

Dedico este trabalho a toda a população brasileira, que contribuiu indiretamente na elaboração deste trabalho através dos impostos que financiaram a minha pesquisa e de diversos outros colegas universitários. Espero que os debates e resultados apresentados aqui possam colaborar com ao menos um pingô no enorme mar do conhecimento da ciência, contribuindo para um maior bem-estar de todas as pessoas.

Agradecimentos

Agradeço a todas as pessoas que estiveram comigo, direta ou indiretamente, em todos os processos de elaboração deste trabalho. Das menores às maiores ajudas, só cheguei aqui por conta dos meus círculos de apoio.

Agradeço ao Programa de Pós-Graduação em Tecnologia Ambiental e Recursos Hídricos pelas imensas oportunidades e portas abertas. O que eu aprendi nas disciplinas, trocas com colegas, professores, congressos, e mais, é inexplicável em palavras. Agradeço especialmente ao meu orientador, professor Carlos Henrique Ribeiro Lima, que, além das excelentes contribuições com meu trabalho ao longo de toda sua construção, também me deu “puxões de orelha” necessários para me tornar um profissional mais capacitado. O professor Carlos me permitiu voar pela UnB, conhecendo disciplinas de diversos programas diferentes que contribuíram muito com meu aprendizado multidisciplinar. Também faço um agradecimento especial à professora Raquel pelas diversas conversas e apoios, técnicos ou não, durante toda minha estadia no programa.

Agradeço à minha família e demais pessoas que eu amo. Sem seus apoios, definitivamente eu nunca conseguiria chegar onde estou atualmente: feliz, saudável, realizado profissionalmente e contente com a vida. Monike, Márcio, Isis, Alice, Flaviane, Carol, Daniel e Vera: família; Flora, Luísa, Luiza, Luzia e demais Lulu's da minha vida, Rô, Carol, Juliana, Nico, Aud, Tai, Daniel e demais amizades-amores da vida, do Rio a Floripa; Steps, Guilherme, Samy, Felipe, Tina, João, Lua e Carlota: colegas candangas de quarto, forró, trilhas e muito mais. Por fim, também agradeço aos amigos dos Camarões, que sempre me acompanharam e me apoiaram onde quer que eu estivesse. Amores desde a infância até o fim.

Por fim, agradeço à população brasileira por investir em mim. Sem a bolsa que recebi, distribuída pela FINATEC, eu não conseguiria ter me mantido em Brasília durante esse período.

RESUMO

As epidemias de dengue vêm representando uma crescente ameaça à saúde global. Sua transmissão, modulada por uma ampla rede de fatores climáticos e socioambientais, cresceu até ultrapassar fronteiras geográficas onde sua presença não era historicamente frequente. Embora haja na literatura debates qualitativos sobre a influência do saneamento básico na sua disseminação, há uma lacuna quanto à incorporação destes parâmetros em análises quantitativas, mais especificamente em modelos mecanísticos.

Partindo da hipótese de que o saneamento básico pode atuar como modulador dos efeitos do clima na dinâmica da doença, o presente trabalho objetivou avaliar a influência de determinantes de saneamento na epidemiologia da dengue e propor uma abordagem para sua incorporação em modelos de transmissão da doença. Duas abordagens metodológicas foram aplicadas no contexto do estado do Rio de Janeiro. A primeira consistiu na aplicação de análises estatísticas para explorar relações entre dados epidemiológicos de dengue, saneamento básico e variáveis sociodemográficas. Surto epidêmicos ocorridos foram analisados espaço-temporalmente, dados de dengue e saneamento foram correlacionados e uma análise multivariada foi aplicada para identificar e classificar variáveis relacionadas com o saneamento básico que possam ter correlações com a dengue. A segunda abordagem foi a implementação de um modelo mecanístico do tipo SIR-SI-SI (humanos, mosquitos, ovos), que leva em consideração parâmetros biológicos e variáveis externas. O modelo foi calibrado e testado (período 2014–2022) em seis municípios do estado com distintas características epidemiológicas, urbanas e socioculturais.

Como principais resultados obtidos, a análise multivariada categorizou com sucesso os indicadores de saneamento básico em conjuntos com menores dimensões, que foram interpretados e nomeados, a exemplo de “Cobertura dos Serviços”, “Qualidade da Água” e “Eficiência da Rede”. O modelo mecanístico foi capaz de reproduzir a sazonalidade da doença com resultados satisfatórios e os parâmetros “Capacidade de Carga” e “Taxa de Sobrevivência Aquática” foram identificados como fundamentais para a modelagem via incorporação de covariáveis relacionadas às condições de saneamento básico. As análises de correlação revelaram aparentes paradoxos, em que indicadores de melhores serviços de saneamento se associaram a maiores incidências de dengue. Com isso, incutiu-se que a urbanização pode atuar como fator de confusão relevante neste contexto.

Dessa forma, o trabalho contribui para a construção de um arcabouço de integração de determinantes de saneamento básico na modelagem da dengue, possibilitando a quantificação de impactos de políticas públicas de saneamento na saúde humana.

ABSTRACT

Dengue epidemics have been posing a growing threat to global health. Its transmission, modulated by a wide network of climatic and socio-environmental factors, has grown to exceed geographical boundaries where its presence was not historically found. Although there are qualitative debates in the literature about the influence of basic sanitation on its spread, there is a gap in the incorporation of these parameters into qualitative analyses, more specifically, in mechanistic models.

Based on the hypothesis that basic sanitation can act as a modulator of the climate's effects on the dynamics of the disease, this study aimed to address this gap by developing an approach for incorporating sanitation determinants into mechanistic models. Two methodological approaches were applied in the context of the state of Rio de Janeiro, Brazil. The first consisted of applying statistical analyses to explore relationships between dengue, sanitation, and sociodemographic data. Epidemic outbreaks were analyzed spatio-temporally, dengue and sanitation data were correlated, and a multivariate analysis was applied to identify and classify variables related to sanitation that may be correlated to dengue. The second approach was the development and calibration of a mechanistic SIR-SI-SI model (humans, mosquitoes, eggs), which incorporated biological parameters and external variables. It was calibrated and tested (2014–2022) in six municipalities of the state with distinct epidemiological, urban, and sociocultural characteristics.

For the main results, the multivariate analysis successfully categorized sanitation indicators into smaller-dimension sets, which were interpreted and named, for example, “Service Coverage”, “Water Quality”, and “Network Efficiency”. The mechanistic model was able to reproduce the disease's seasonality with satisfactory results. The “Carrying Capacity” and “Aquatic Survival Rate” parameters were identified as fundamental for modeling through the incorporation of covariates related to basic sanitation conditions. The correlation analyses revealed apparent paradoxes, where indicators of better sanitation services were associated with higher incidences of dengue. From this, it was inferred that urbanization may act as a relevant confounding factor in this context.

Thus, this work contributes to the construction of a framework for the integration of sanitation determinants into dengue modeling, enabling the quantification of the impacts of sanitation policies on human health.

SUMÁRIO

1. INTRODUÇÃO.....	1
2. OBJETIVO GERAL.....	4
2.1 OBJETIVOS ESPECÍFICOS.....	4
3. REFERENCIAL TEÓRICO.....	5
3.1 A DENGUE.....	5
3.1.1 História e Geografia da Dengue.....	7
3.1.2 Estratégias de controle da Dengue.....	9
3.1.3 Clima, Saneamento, Sociedade e Dengue.....	10
3.2 A CIÊNCIA DA EPIDEMIOLOGIA.....	13
3.2.1 Erros e vieses em estudos epidemiológicos.....	15
3.3 MÉTODOS E MEDIDAS MATEMÁTICOS APLICADOS EM ESTUDOS DE SAÚDE PÚBLICA.....	17
3.4 MODELOS MATEMÁTICOS UTILIZADOS NA SAÚDE PÚBLICA.....	20
3.5 O ESTADO DO SANEAMENTO BÁSICO NO BRASIL.....	24
4. REVISÃO DE LITERATURA.....	26
4.1 MODELOS MECANÍSTICOS DE DENGUE.....	26
4.2 ABORDAGENS E PARÂMETROS UTILIZADOS EM MODELOS “SIR”.....	29
4.2.1 PARÂMETROS DOS MODELOS.....	29
4.2.2 CALIBRAÇÃO, TREINAMENTO E VALIDAÇÃO.....	33
5. MATERIAIS E MÉTODOS.....	35
5.1 OBTENÇÃO DE DADOS.....	35
5.1.1 Dados de dengue.....	36
5.1.2 Dados meteorológicos.....	37
5.1.3 Dados de saneamento básico.....	38
5.2 RECORTE ESPACIAL E TEMPORAL.....	39
5.2.1 Campos dos Goytacazes.....	42
5.2.2 Duque de Caxias.....	42
5.2.3 Macaé.....	42
5.2.4 Paraty.....	42
5.2.5 Resende.....	43
5.2.6 Rio de Janeiro.....	43
5.3 ANÁLISE EXPLORATÓRIA DA DENGUE NOS MUNICÍPIOS DO RIO DE JANEIRO..	43
5.3.1 Contagem de meses epidêmicos.....	44
5.3.2 Análises de distribuição espacial e temporal.....	44
5.4 CORRELAÇÃO ENTRE DENGUE E INDICADORES DE SANEAMENTO.....	45
5.4.1 Variável Resposta.....	45
5.4.2 Análise de correlação.....	46
5.5 SELEÇÃO E AGRUPAMENTO DE INDICADORES DE SANEAMENTO BÁSICO.....	47
5.5.1 Pré-processamento e seleção de variáveis do SINISA.....	47
5.5.2 Análise de Componentes Principais (PCA).....	47
5.6 MODELO MECANÍSTICO DE DENGUE.....	48
5.6.1 Condições iniciais do modelo.....	48

5.6.2	Parâmetros utilizados do modelo.....	50
5.6.3	Equações do modelo final SIR-SI-SI.....	52
5.6.4	Análise exploratória manual dos parâmetros.....	54
5.6.5	Processo de calibração do modelo.....	55
6.	RESULTADOS E DISCUSSÃO.....	57
6.1	ANÁLISE EXPLORATÓRIA.....	57
6.1.1	Epidemias no estado do Rio de Janeiro.....	63
6.2	CORRELAÇÕES ENTRE DENGUE E INDICADORES DE SANEAMENTO.....	67
6.3	ANÁLISE DE COMPONENTES PRINCIPAIS.....	72
6.3.1	Seleção de indicadores do SINISA.....	72
6.3.2	Componentes Principais dos eixos do SINISA -.....	72
6.3.3	Associação entre os Componentes Principais e a Dengue.....	74
6.4	MODELO SIR-SI-SI.....	76
6.4.1	Resultados da calibração do modelo.....	76
6.4.2	Ajuste do modelo aos dados observados.....	79
6.4.3	Perspectivas de incorporação de parâmetros de saneamento básico no modelo.....	86
7.	CONSIDERAÇÕES FINAIS E RECOMENDAÇÕES.....	88
8.	CONCLUSÕES.....	92
	REFERÊNCIAS BIBLIOGRÁFICAS.....	94
	APÊNDICES.....	101
	APÊNDICE A - DIAGRAMA DETALHADO DO MODELO SIR-SI-SI.....	102
	APÊNDICE B - PARÂMETROS SELECIONADOS PARA ANÁLISE DE COMPONENTES PRINCIPAIS.....	103
	APÊNDICE C - RESULTADOS DAS ANÁLISES DE CORRELAÇÃO ENTRE VARIÁVEIS DE SANEAMENTO BÁSICO E O DESFECHO: “INCIDÊNCIA MÉDIA ANUAL” EM 82 MUNICÍPIOS DO RIO DE JANEIRO.....	106
	APÊNDICE D - RESULTADOS DETALHADOS DAS ANÁLISES DE COMPONENTES PRINCIPAIS DOS INDICADORES DO SINISA.....	110
	APÊNDICE E - RESULTADOS DA ANÁLISE BIBLIOMÉTRICA.....	114
	APÊNDICE F - LISTA DE SUGESTÕES (ROTEIRO) PARA PESQUISAS FUTURAS.....	117

Lista de Quadros

Quadro 4.1 - Referências de valores fixos ou intervalos adotados para parâmetros na literatura	32
Quadro 5.1 - Fontes utilizadas para obtenção de dados do trabalho.....	36
Quadro 5.2 - Estações meteorológicas utilizadas para obtenção de dados.....	37
Quadro 5.3 - Opções utilizadas para download de dados do SNIS.....	39
Quadro 5.4 - Condições iniciais utilizadas no modelo mecanístico e equações utilizadas para seus cálculos	49
Quadro 5.5 - Valores e cálculos adotados para as variáveis utilizadas nas equações do modelo SIR-SI-SI	51
Quadro 5.6 - Descrição das variáveis utilizadas nas equações do modelo SIR-SI-SI.....	54
Quadro 6.1 - Quantidade de indicadores selecionados, de cada eixo do SINISA.....	72
Quadro 6.2 - Características dos componentes principais com maior explicação de variância.....	74
Quadro 6.3 - Resultado da calibração de parâmetros do modelo nos 6 municípios testados, em comparação com intervalos observados na literatura.....	75
Quadro 6.4 - Quadro 6.4: Resultado da calibração de parâmetros do modelo nos 6 municípios testados, em comparação com intervalos observados na literatura. (A) Valores após calibração semanal. (B) Valores após calibração mensal.....	76

Quadros nos apêndices

Quadro C.1 - Valores de correlação entre incidência média mensal e variáveis do SINISA.....	APÊNDICE C
Quadro D.1 - Sumário dos resultados da PCA - Eixo Resíduos Sólidos.....	APÊNDICE D
Quadro D.2 - Componentes principais do eixo Resíduos Sólidos do SINISA.....	APÊNDICE D
Quadro D.3 - Sumário dos resultados da PCA - Eixo Água e Esgoto.....	APÊNDICE D
Quadro D.4 - Componentes principais do eixo Água e Esgoto do SINISA.....	APÊNDICE D
Quadro D.5 - Sumário dos resultados da PCA - Eixo Águas Pluviais.....	APÊNDICE D
Quadro D.6 - Componentes principais do eixo Águas Pluviais do SINISA.....	APÊNDICE D

Lista de Figuras

Figura 3.1 - Canal endêmico da dengue, subdividido por semanas epidemiológicas, no Equador. Adaptado de Gutiérrez, Vallejo e Rivas (2021).....	15
Figura 3.2 - esquema de variáveis confundidoras em estudos epidemiológicos.....	17
Figura 3.3 - esquema de um modelo SIR.....	22
Figura 4.1 - Compartimentos utilizados no modelo de Campos et al. (2021).....	27
Figura 4.2 - Compartimentos utilizados nos modelos de Chanprasopchai (E) e Alves (D).26	
Figura 5.1 - Fluxograma da metodologia.....	35
Figura 5.2 - Mapa de localização e delimitação de municípios selecionados para o modelo SIR.....	40
Figura 5.3 - Temperatura média anual dos municípios selecionados para o modelo mecanístico.....	41
Figura 5.4 - Índice de Gini e Precipitação anual média nas estações selecionadas para cada município para o modelo mecanístico.....	41
Figura 5.5 - Compartimentos utilizados no modelo SIR-SI-SI.....	52
Figura 6.1 - Casos mensais de dengue em seis municípios do estado do Rio de Janeiro ao longo da série temporal.....	58
Figura 6.2 - Incidência mensal em seis municípios do estado do Rio de Janeiro ao longo da série temporal.....	59
Figura 6.3 - Incidência média mensal em seis municípios do estado do Rio de Janeiro.....	61
Figura 6.4 - Comparação entre incidência, temperatura e precipitação médias mensais na cidade do Rio de Janeiro.....	62
Figura 6.5 - Quantidades de epidemias observadas em cada mês do ano, nos municípios do estado do Rio de Janeiro.....	63
Figura 6.6 - Mapa apresentando a espacialização do total de meses epidêmicos de dengue, entre 2014 e 2021, no estado do Rio de Janeiro.....	65
Figura 6.7 - Mapa apresentando a espacialização da incidência da dengue, entre 2014 e 2021, no estado do Rio de Janeiro.....	66
Figura 6.8 - Gráfico de dispersão entre incidência anual mediana de dengue nos municípios do Rio de Janeiro com três indicadores do SINISA. Painel superior: IN014_AE (Consumo micromedido por economia); Painel central: IN016_RS (Taxa de cobertura regular do serviço de coleta de RDO em relação à população urbana); Painel inferior: IN040_AP (Parcela de Domicílios em Situação de Risco de Inundação).....	67

Figura 6.9: Gráfico de correlações entre as 10 variáveis de saneamento do SINISA mais correlacionadas com a incidência de dengue em 82 municípios do Rio de Janeiro.....	68
Figura 6.10: Gráfico de correlações entre as 15 variáveis de saneamento do IBGE mais correlacionadas com a incidência de dengue em 92 municípios do Rio de Janeiro.....	70
Figura 6.11 - Gráfico de Scree representando o acumulado do percentual de variância explicada por cada componente principal (resultados da PCA).....	73
Figura 6.12 - Comparação entre calibração do modelo com agregamento semanal (esquerda) e mensal (direita) para o município de Duque de Caxias nas primeiras 100 semanas epidemiológicas.....	80
Figura 6.13 - Comparação entre calibração do modelo com agregamento semanal (esquerda) e mensal (direita) para o município de Paraty nas primeiras 100 semanas epidemiológicas.....	82
Figura 6.14 - Modelo calibrado com agregamento semanal para o município de Macaé. Período de treinamento (2014-2017).....	83
Figura 6.15 - Modelo calibrado com agregamento semanal para o município de Resende. Comparação entre período de treinamento (2014-2017) e período de validação (2018-2022).....	84
Figura 6.16 - Modelo calibrado com agregamento semanal para o município de Campos dos Goytacazes. Comparação entre período de treinamento (2014-2017) e período de validação (2018-2022).....	85

Lista de Símbolos, Nomenclatura e Abreviações

Agência Nacional de Águas e Saneamento Básico (ANA)

Análise de Componentes Principais (PCA)

Direito Humano à água e ao esgotamento sanitário (DHAES)

Instituto Brasileiro de Geografia e Estatística (IBGE)

Organização das Nações Unidas (ONU)

Organização Mundial da Saúde (OMS)

Modelo Autoregressivo com Defasagens Distribuídas (ARDL)

Pesquisa de Informações Básicas Municipais (MUNIC)

Pesquisa Nacional por Amostra de Domicílios (PNAD)

Programa Nacional de Saneamento Rural (PNSR)

Sistema Nacional de Informações em Saneamento Básico (SINISA)

Modelo mecanístico com humanos suscetíveis, infectados e recuperados (SIR)

Modelo mecanístico com humanos suscetíveis, infectados e recuperados, vetores suscetíveis e recuperados e ovos suscetíveis e recuperados (SIR-SI-SI)

1. INTRODUÇÃO

A dengue representa uma das mais graves e crescentes ameaças à saúde pública global, sendo classificada pela Organização Mundial da Saúde (OMS) como um dos dez principais riscos à saúde humana (WHO, 2019). Originalmente concentrada em regiões tropicais, suas fronteiras geográficas têm sido expandidas para novas áreas, em um fenômeno atribuído às mudanças climáticas, que alteram os padrões de temperatura e precipitação, promovendo condições suscetíveis à proliferação de mosquitos *Aedes aegypti* (Messina *et al.*, 2014, 2019). Este fenômeno tem intensificado a atenção da comunidade científica sobre a doença e seus determinantes (Aguiar *et al.*, 2022; Ogunlade *et al.*, 2023).

A transmissão da dengue é um processo complexo influenciado por uma ampla rede de fatores. Além de variáveis climáticas, a dinâmica da doença está profundamente ligada às características socioeconômicas, comportamentais da população suscetível e ao acesso a serviços de saneamento básico (Andrade *et al.*, 2024; de Vasconcelos; de Lima; Cardoso, 2023; Pinho; Mendonça; Lima, 2024; Queiroz; Silva, 2024). A proliferação do vetor *Aedes aegypti* está diretamente associada à disponibilidade de água parada, frequentemente resultante de deficiências em estruturas de saneamento básico. O vínculo entre saneamento e saúde é tão significativo que a OMS estima que a cada dólar investido em saneamento, pode-se gerar uma economia de quatro dólares em gastos com saúde global (ONU, 2014). Os resultados obtidos por Fernandes e Silva (2020) corroboram tal afirmação ao observar que municípios do Rio de Janeiro com melhores condições de saneamento têm gastos inferiores com enfermidades relacionadas a saneamento inadequado.

Apesar da reconhecida importância desses múltiplos fatores, as modelagens matemáticas para análise da disseminação da dengue frequentemente restringem a incorporação de parâmetros às variáveis meteorológicas, como temperatura, precipitação e umidade (Aguiar *et al.*, 2022). A quantificação de relações existentes entre dengue e clima é um campo amplamente debatido no campo da biomatemática, com uma gama de equações desenvolvidas e aplicadas na literatura para representar numericamente aspectos biológicos, como taxa de sobrevivência de ovos e taxa de picada dos mosquitos (Mordecai *et al.*, 2017; Nuraini *et al.*, 2021).

No entanto, modelos que se baseiam apenas em variáveis meteorológicas possuem limitações, uma vez que não conseguem justificar a grande heterogeneidade existente na incidência da doença observada em locais com climas semelhantes. Para além de aspectos

culturais, comportamentais, e mesmo da estocasticidade, uma hipótese é levantada: as condições de saneamento podem atuar como um modulador da influência do clima nas dinâmicas da doença. Durante um evento de chuva, por exemplo, um município com piores condições de drenagem urbana e resíduos sólidos oferece condições propícias para a instituição de criadouros para os mosquitos através de água parada.

Embora diversos autores tratem qualitativamente dos efeitos da falta de saneamento na incidência da doença (Heller, 2016; Oliveira *et al.*, 2023; Queiroz; Silva; Heller, 2020; Ribeiro *et al.*, 2021), há uma notável carência de estudos quantitativos. Esta lacuna no corpo científico é confirmada na revisão sistemática de Ogunlade *et al.* (2023): enquanto estratégias de controle químico e biológico são amplamente modeladas em estudos científicos (representando 31% e 60% do total dos artigos analisados, respectivamente), o controle ambiental foi tema de apenas 9% dos estudos.

Essa lacuna de investigações de uma área tão importante pode ocorrer devido a uma dificuldade de se mensurar as múltiplas dimensões do saneamento de forma padronizada. Nesse sentido, o banco de dados do Sistema Nacional de Informações em Saneamento Básico (SINISA) oferece uma oportunidade para superar esse desafio (Borges *et al.*, 2022). A disponibilidade de uma vasta gama de indicadores operacionais, financeiros e de cobertura, para milhares de municípios brasileiros, pode permitir a realização de análises robustas com fins de avaliar sua influência na dinâmica da dengue.

É neste contexto que a presente pesquisa se insere, propondo-se a preencher parte dessa lacuna ao investigar as complexas interações entre clima, saneamento e a incidência da dengue. Utilizando um recorte espacial nos municípios do estado do Rio de Janeiro (RJ), uma unidade federativa com grande heterogeneidade socioambiental, boa disponibilidade de dados e com surtos expressivos de dengue nos anos recentes, este trabalho emprega uma abordagem metodológica com duas frentes complementares: uma estatística e uma mecanística.

Inicialmente, foram aplicados métodos estatísticos, como análises espaciais, de correlação e de componentes principais, para explorar padrões e associações entre indicadores de saneamento básico do SINISA e a ocorrência da doença. Na sequência, um modelo epidemiológico mecanístico compartimental (SIR-SI-SI) foi desenvolvido para simular as dinâmicas de transmissão da doença em diferentes níveis (mosquitos, vetores adultos e vetores em fase de ovos), avaliando o impacto das variáveis climáticas nos parâmetros biológicos. A construção do modelo e os cálculos dos parâmetros que os compõem foram realizados com base em trabalhos recentes e relevantes da literatura

(Alves, 2021; de Vasconcelos; de Lima; Cardoso, 2023; Mordecai *et al.*, 2017; Nuraini *et al.*, 2021), com o objetivo final de explorar possibilidades de incorporação de variáveis de saneamento básico na estrutura de modelos como este.

O trabalho está estruturado em seis partes, sendo a atual uma breve introdução à temática. Na parte seguinte apresenta-se o objetivo geral e os específicos, seguidos pelo referencial teórico e revisão bibliográfica, realizados com materiais acadêmicos relevantes à pesquisa, com interesse especial em materiais mais recentes. As metodologias são apresentadas de maneira detalhada e, por fim, são evidenciados e debatidos os resultados.

2. OBJETIVO GERAL

Avaliar e modelar a influência de indicadores de saneamento básico e de variáveis climáticas na epidemiologia da dengue em municípios do estado do Rio de Janeiro.

2.1 OBJETIVOS ESPECÍFICOS

- Caracterizar padrões espaço-temporais da distribuição da dengue em municípios do estado do Rio de Janeiro;
- Identificar agrupamentos nos dados de saneamento básico do SINISA que tenham associação à ocorrência da dengue;
- Desenvolver um modelo mecanístico de dengue capaz de avaliar as dinâmicas de disseminação da doença a partir de dados climáticos.
- Avaliar a incorporação de variáveis de saneamento básico na estrutura do modelo mecanístico de dengue.

3. REFERENCIAL TEÓRICO

Neste capítulo são apresentados conceitos e documentos fundamentais para a concepção e construção do trabalho como um todo. Os conceitos levantados neste capítulo são utilizados para embasar e justificar as metodologias e a análise dos resultados do estudo.

Os subcapítulos seguem uma sequência lógica que compreende aspectos essenciais para a compreensão da dengue. Inicialmente, processos biológicos do vetor *Aedes aegypti* e da transmissão do vírus são pontuados. Na sequência, apresenta-se um histórico da disseminação, do controle e da compreensão dos fatores determinantes da doença no mundo e no Brasil. Destaca-se a importância das estratégias de controle e das abordagens mais recentes, que adotam uma visão holística ao integrar fatores sociais, econômicos, ambientais, de saneamento básico e outros para enfrentá-la de maneira mais eficiente. A ciência da epidemiologia, seus limites e os métodos comumente empregados em estudos epidemiológicos são também abordados. Por fim, discute-se os dados de saneamento básico no Brasil. O capítulo é concluído com um gancho que conecta as informações debatidas à revisão bibliográfica (Capítulo 4), que reúne estudos recentes e relevantes utilizados como referência neste trabalho.

3.1 A DENGUE

A dengue é uma arbovirose urbana de grande relevância no cenário da saúde pública mundial. Sua classificação como “arbovirose” é dada devido às características da sua disseminação: do inglês *arthropod borne virus*, arboviroses são doenças causadas por vírus transmitidos por artrópodes. No caso da dengue, seus principais vetores são os mosquitos fêmeas do gênero *Aedes*, infectados pelo vírus da dengue (DENV). Destacam-se as espécies *Aedes aegypti* e *Aedes albopictus* como vetores primários e secundários, respectivamente. Os mosquitos da espécie *Aedes aegypti* podem transmitir também as doenças da Febre Amarela, Zika e Chikungunya, se infectados pelos vírus correspondentes, o que demonstra a relevância do estudo desse vetor para a saúde pública.

Quatro sorotipos distintos são definidos para a dengue: DENV-1, DENV-2, DENV-3 e DENV-4. Uma vez infectado por um sorotipo, um indivíduo desenvolve imunidade duradoura a ele. No entanto, mantém-se suscetível ao restante. Nesses casos, ocorrem as denominadas infecções secundárias, que têm maior probabilidade de aparição de sintomas mais intensos. Isso ocorre devido ao efeito denominado “aprimoramento dependente de anticorpos” (sigla *ADE* em inglês), onde os anticorpos de uma infecção

anterior não são capazes de neutralizar um novo sorotipo, paradoxalmente facilitando sua entrada e replicação viral (Katzelnick *et al.*, 2017).

A transmissão ocorre através de um ciclo que envolve mosquitos (vetores) e humanos (hospedeiros). Quando uma fêmea infectada pica uma pessoa, o vírus é inoculado e fica em incubação por um período médio de 4 a 10 dias. Se infectado, este indivíduo pode desenvolver sintomas, que geralmente persistem por 2 a 7 dias, e tornar-se infeccioso, podendo transmitir o vírus a outros mosquitos suscetíveis.

Quando um mosquito não infectado pica uma pessoa infecciosa, ele pode adquirir o vírus. Dá-se início ao período de incubação extrínseco, que dura de 8 a 10 dias, ao fim do qual ele também se tornará infeccioso. Durante esse intervalo de tempo, o vírus se replica e se dissemina pelo organismo do mosquito até alcançar suas glândulas salivares. Uma vez infeccioso, o mosquito permanece assim por toda a sua vida, podendo transmitir a doença a cada nova picada.

Os sintomas mais comuns da doença são febre, dor de cabeça, náuseas e dores intensas nos ossos e articulações, motivo pelo qual é popularmente referenciado em inglês pelo nome “*breakbone fever*” (febre quebra-ossos). Para as infecções secundárias, podem ser observados sintomas de febre hemorrágica que podem culminar, inclusive, na morte. Essas formas mais graves representam um fardo considerável para os sistemas de saúde, especialmente em países da Ásia e da América Latina, resultando em hospitalizações e óbitos (WHO, 2023).

A existência dos quatro diferentes sorotipos virais é um fator que complexifica a dinâmica epidemiológica da dengue: uma pessoa recuperada da infecção por um dos sorotipos se torna resistente a ele mas mantém-se suscetível à infecção por outros sorotipos. Pode ocorrer de múltiplos sorotipos circularem conjuntamente em uma mesma região, embora, em geral, apenas um ocorra em predominância em relação aos outros, sendo o responsável pela maioria dos casos durante tais períodos. Mudanças na predominância de sorotipos podem proporcionar grandes surtos de casos, levando a colapsos no sistema de saúde.

O *Aedes aegypti* é um vetor que possui afinidade a ambientes urbanos e grande eficiência na disseminação da doença. Os ovos do mosquito são encontrados em recipientes com água parada (denominados “criadouros”), como poças d’água, pneus, vasos de planta, caixas d’água mal vedadas e outros. É nas cidades que o mosquito encontra melhores condições de sobrevivência, com faixas de temperatura mais quentes e disponibilidade de alimentos, em especial em ambientes mais adensados e que possuem

irregularidades na coleta de resíduos (que disponibilizam criadouros artificiais) e no abastecimento de água (que implicam no armazenamento doméstico de água) (Messina *et al.*, 2014; Queiroz; Silva, 2024). Essa característica, junto à preferência do mosquito por sangue humano e ao seu hábito de viver próximo a edificações, justificam a forte associação existente entre a doença e a urbanização (Almeida; Cota; Rodrigues, 2020; Reinhold; Lazzari; Lahondère, 2018). Ainda, o mosquito possui capacidade de picar múltiplas pessoas durante um mesmo ciclo de alimentação, fundamentando o alto potencial de propagação da doença (Gubler, 1998 apud Alves, 2021).

Toda essa complexidade levantada demonstra a importância da elaboração e aprimoramento de estudos epidemiológicos contínuos para o acompanhamento dos casos da doença e prevenção de surtos, que, além de causar dor e sofrimento, geram gastos para o sistema de saúde e diminuição da produtividade da população.

3.1.1 História e Geografia da Dengue

Registros históricos sugerem a ocorrência de doenças com sintomas semelhantes aos da dengue há séculos. Apesar disso, foi apenas em meados do século XX que cientistas isolaram e descreveram os dois primeiros sorotipos do vírus da dengue em laboratório, o que trouxe maior compreensão sobre sua epidemiologia (Messina *et al.*, 2014). A partir desse momento, já com maior conhecimento sobre a doença, foi observado um rápido crescimento na notificação de casos, justificado pela propagação do vetor *Aedes aegypti*, ao redor do mundo, fato intensificado pelo rápido crescimento da urbanização, do adensamento populacional e da mobilidade urbana (Messina *et al.*, 2014). Até 1970, grandes epidemias eram concentradas em poucos países, mas atualmente mais de 100 nações são consideradas endêmicas para a dengue, estimando-se que metade da população mundial esteja em risco de infecção (WHO, 2023).

No Brasil, os primeiros casos formalmente notificados de dengue datam de 1946, embora haja estimativas de que a doença tenha sido introduzida no país ainda no século XIX através de navios negreiros (Silva; Mariano; Scopel, 2008). Andrade e Lopes (2021) e Senhoras (2020) incutem em seus trabalhos sobre a complexa interação entre a história das epidemias da dengue e de outras doenças infecciosas no Brasil e no mundo, vulnerabilidades sociais associadas e o significativo efeito de doenças infecciosas no desenvolvimento da humanidade.

Levantamentos globais mais recentes apontam para um crescimento acelerado no número de casos, com um aumento de aproximadamente dez vezes, passando de cerca de

500 mil para 5,2 milhões de casos anuais entre 2000 e 2019 (WHO, 2023). Ainda assim, apesar dos números alarmantes, autores argumentam que a verdadeira incidência da dengue pode estar subestimada nos registros oficiais (Bhatt *et al.*, 2013), chegando à ordem de 390 milhões. Isso ocorre porque uma grande proporção dos casos é assintomática ou apresenta sintomas leves e inespecíficos, semelhantes aos de outras doenças febris, levando a uma subnotificação significativa.

A doença vem se consolidando como um persistente problema de saúde pública em diversos países da América Latina ao longo dos anos (Badii *et al.*, 2007; Ibarrarán *et al.*, 2023), sendo considerada pelo Ministério da Saúde do Brasil (Brasil, 2023) como a arbovirose urbana de maior relevância epidemiológica nas Américas e no próprio país. O Brasil registrou a maior incidência da doença em 2022 entre todos os países do Mercosul (COVIGSAL, 2023), o que proporcionou uma ampla gama de impactos socioeconômicos e na saúde. O país tem enfrentado picos epidêmicos de grande magnitude nos anos recentes e, somente no primeiro trimestre de 2024, foram estimados mais de 2,6 milhões de casos prováveis nas unidades federativas brasileiras, acompanhados de um elevado número de óbitos (Brasil, 2024; Westin, 2024; Souza *et al.*, 2024).

O histórico da dengue no Brasil foi marcado por uma complexa trajetória de vigilância epidemiológica e controle vetorial. No contexto da luta contra a Febre Amarela, no início do século XX, algumas ações de controle foram instituídas direcionadas à redução da população de mosquitos *Aedes* (Salles *et al.*, 2018). Essa campanha nacional obteve grande sucesso, culminando na erradicação do vetor *Aedes aegypti* do território nacional, com reconhecimento oficial na década de 1940 pela Organização Pan-Americana da Saúde (OPAS) (Silva; Mariano; Scopel, 2008; Teixeira; Barreto, 1996). Essa vitória foi temporária: o vetor foi gradualmente reintroduzido no país em focos isolados a partir das décadas seguintes, até que, nos anos 1980, uma reintrodução em larga escala deu origem à primeira epidemia moderna documentada, em Roraima, estabelecendo o cenário endêmico que persiste até hoje (Teixeira; Barreto, 1996). Esse ciclo de erradicação e reintrodução evidencia não apenas a necessidade de vigilância contínua, mas também a importância de se compreender o repertório de ferramentas de controle disponíveis. A análise das estratégias empregadas ao longo dessa história, de suas limitações e das inovações que surgiram em resposta, é fundamental para o enfrentamento atual da doença, conforme detalhado a seguir.

3.1.2 Estratégias de controle da Dengue

Múltiplas estratégias de controle já foram implementadas globalmente para mitigar o impacto de doenças transmitidas por vetores. As técnicas de prevenção da dengue mais tradicionalmente empregadas baseiam-se em três mecanismos principais de controle vetorial: controle mecânico, controle biológico e controle químico (Zara *et al.*, 2016). O controle mecânico visa eliminar ou reduzir os criadouros do vetor e minimizar o contato entre mosquitos e hospedeiros por meio de ações como a remoção de água acumulada em recipientes e o uso de telas. O controle biológico consiste na introdução de organismos que atuam como predadores ou patógenos para as larvas ou mosquitos adultos, como o uso de peixes larvófagos ou microrganismos que liberam toxinas nos criadouros. Por fim, o controle químico envolve a aplicação de compostos, muitos deles neurotóxicos, para eliminar os mosquitos adultos e suas formas imaturas. Seu uso inadequado pode acarretar prejuízos ao meio ambiente e à saúde humana.

O uso dessas medidas tradicionais pode ter eficiência limitada e os resultados demorarem a aparecer. Para os controles mecânicos e biológicos, há dependência na adesão da comunidade, o que nem sempre é alcançado de maneiras satisfatórias. Quanto ao controle químico, os mosquitos podem desenvolver resistência aos inseticidas químicos. Em resposta a esses desafios, métodos inovadores foram desenvolvidos em anos mais recentes (Wong *et al.*, 2022). Uma dessas abordagens é a liberação de mosquitos geneticamente modificados, que podem atuar na redução de fertilidade dos vetores ou na resistência do mosquito ao desenvolvimento viral (BRASIL, 2024). Outra frente promissora é o desenvolvimento de vacinas, que buscam conferir imunidade aos hospedeiros. Um marco importante ocorreu em 2021, quando o Comitê Consultivo sobre Práticas de Imunização (ACIP) dos Estados Unidos recomendou pela primeira vez a vacinação contra a dengue no país, influenciando a tomada de decisões semelhantes em outros países. O uso da vacina foi permitido recentemente no Brasil, diante das crises sanitárias ocorridas em anos recentes, e parcelas populacionais mais suscetíveis receberam doses iniciais do imunizante como medida para conter os impactos da doença (Taylor, 2024).

A trajetória da disseminação da dengue no Brasil e no mundo, especialmente ao longo das últimas décadas (Souza *et al.*, 2024), em conjunto com a evolução das estratégias de controle e prevenção, demonstra a importância crucial da ciência para a melhoria da qualidade de vida. Se, no passado, o conhecimento sobre o agente etiológico da dengue era incipiente, os avanços científicos das últimas décadas permitiram um profundo

entendimento dos aspectos biológicos do vírus e de seu vetor. Esse conhecimento foi fundamental para o desenvolvimento de métodos de controle progressivamente mais eficientes, desde as abordagens mecânicas predominantes no passado, passando pelos controles biológicos e químicos, até as recentes tecnologias de mosquitos geneticamente modificados e vacinas. Apesar desses notáveis progressos, o monitoramento contínuo da dengue permanece essencial, visto que a doença afeta de forma desproporcional as populações mais vulneráveis aos seus determinantes sociais e ambientais, as quais frequentemente enfrentam barreiras no acesso a essas novas tecnologias e métodos de prevenção (de Sousa *et al.*, 2021). Esses aspectos relacionados à vulnerabilidade serão explorados com maior detalhe em seções subsequentes deste trabalho.

3.1.3 Clima, Saneamento, Sociedade e Dengue

O fenômeno de transmissão da dengue é profundamente modulado por um conjunto de variáveis ambientais, climáticas e sociais. Interações desses fatores em diferentes níveis estabelecem condições propícias para a proliferação e desenvolvimento dos mosquitos *Aedes aegypti*, viabilizando a ocorrência de surtos e epidemias. Neste subcapítulo, detalhes de alguns dos determinantes relacionados com a endemia da doença são apresentados.

Variáveis meteorológicas, como temperatura, precipitação e umidade, são apontadas como fatores primários que influenciam a ecologia, o desenvolvimento e a sobrevivência dos mosquitos *Aedes*. Em faixas de temperaturas mais elevadas, o ciclo de vida do vetor é acelerado, reduzindo o tempo entre a fase de ovo e a de mosquito adulto, ao mesmo tempo em que o período de incubação extrínseco do vírus é encurtado (Alkhalidy, 2017; Ogunlade *et al.*, 2023; Sebastianelli *et al.*, 2024).

Estas relações não são lineares, todavia. Temperaturas excessivamente altas ou baixas afetam negativamente os processos de sobrevivência e transmissão da doença: a faixa de temperatura viável para a transmissão é entre 18 e 34°C, com maior eficiência entre 26 e 29°C (Reinhold; Lazzari; Lahondère, 2018). Isso é justificado pela complexa interação que ocorre entre processos, como a taxa de oviposição, a probabilidade de sobrevivência dos ovos, o tempo de vida médio de um mosquito adulto, dentre outros. Cada um desses processos possui faixas de temperatura ótimas e a simulação desses fatores pode ser feita através de equações da biomatemática, como as funções de Brière e Quadrática (Jin; Quinn; Shi, 2022; Mordecai *et al.*, 2017)

As chuvas, por sua vez, aumentam a disponibilidade de criadouros potenciais, favorecendo a abundância da população de vetores. As relações entre temperatura e

umidade são complexas: segundo Alto e Juliano (2001), o impacto da precipitação na produção de mosquitos adultos é fortemente dependente da temperatura. Em condições mais quentes, na faixa de 26 a 30°C, a disponibilidade de água em abundância favorece uma maior produção de mosquitos. Nesta mesma faixa de temperatura, condições que promovem a secagem completa de criadouros reduzem a população de larvas e, conseqüentemente, a produção de mosquitos. Por outro lado, a mesma variabilidade da chuva tem efeitos menos pronunciados em temperaturas mais amenas (próximas de 22°C).

A quantificação dessa relação entre precipitação e incidência da dengue é um objeto de estudo na literatura, com diferentes abordagens sendo aplicadas. No contexto de João Pessoa (Paraíba), métodos estatísticos avançados foram utilizados para identificar defasagens temporais entre ocorrência de chuva e aumentos nos casos de dengue (Santos *et al.*, 2019). Os autores encontraram um efeito mais forte da precipitação sobre a doença com *lag* temporal na ordem de 90 dias. Neste mesmo contexto, os autores indicam que a temperatura não se mostrou como um preditor útil para a doença. Esta conclusão se diferencia de outras literaturas levantadas para as condições do Rio de Janeiro, o que demonstra como o efeito de variáveis climáticas sobre a doença pode variar regionalmente. É possível que locais que mantêm-se numa faixa de temperatura ótima para a doença sejam limitados pelas condições de precipitação, enquanto em locais como o Rio de Janeiro, que constantemente ultrapassa os limites de temperatura ótima para sobrevivência dos mosquitos *Aedes* e suas fases de desenvolvimento, a doença tenha maior correlação com a temperatura.

Observar essas relações entre temperatura e umidade demonstra o quão complexas são as dinâmicas populacionais do vetor e da transmissão de dengue, que sofrem influência, também, de diversas outras variáveis, em diferentes intensidades. Ainda nesse contexto, as mudanças climáticas representam uma ameaça crescente à saúde global, fato exemplificado pela expansão geográfica de atuação do *Aedes aegypti* para regiões onde a doença historicamente não era endêmica, como na Europa e na América do Norte (Messina *et al.*, 2019).

Condições de saneamento e de infraestrutura urbana são determinantes cruciais da doença, em escala local. A urbanização acelerada, muitas vezes desacompanhada de planejamento e investimentos adequados, cria ambientes propícios para a proliferação do vetor. A falta de acesso à rede de água potável contínua faz com que indivíduos armazenem água em recipientes muitas vezes irregulares, enquanto a coleta inadequada de resíduos pode aumentar também a oferta de criadouros (Maciel-de-Freitas; Valle, 2014).

Uma evidência quantitativa dessa relação foi demonstrada por Pinho (2023), que encontrou correlações significativas entre incidência de dengue e índices de tratamento de esgoto (negativa moderada, $r = -0.63$) e de abastecimento de água (negativa fraca, $r = -0.34$), em municípios de Goiânia.

Relações entre saneamento e saúde são reconhecidas, como no estudo da OMS que estima que investimentos em saneamento geram economias com gastos em saúde (WHO, 2018). A importância desses fatores é tão significativa que pesquisadores estimam que até 95% do ônus da Dengue e da Encefalite Japonesa, no Japão, pode ser atribuído às condições inadequadas de água e saneamento (Carlton *et al.*, 2012).

Esta análise é corroborada por Léo Heller, especialista em saneamento e antigo relator sobre direito humano à água potável e ao saneamento da ONU. Em uma entrevista, o pesquisador retrata haver um forte vínculo entre acesso deficiente a serviços de saneamento e a ocorrência de arboviroses, como a dengue (Heller, 2016). Esta relação é, porém, mais complexa do que apenas o acesso ou falta de acesso a saneamento básico: de acordo com Queiroz, Silva e Heller (2020), embora as deficiências nos serviços de saneamento sejam frequentemente associadas a maiores taxas de incidência de doenças, as medidas aplicadas para solucionar tais problemas frequentemente falham em obter bons resultados em questão de saúde pública. Segundo os autores, isto ocorre pela orientação tecnocêntrica e economicista utilizada na construção de soluções, que utilizam soluções padronizadas e desconsideram toda uma complexidade socio-cultural presente nos territórios.

Uma abordagem mais eficaz para a universalização do saneamento ocorreria, portanto, através da construção de soluções com participação da comunidade, de forma que os indivíduos que a compõem se apropriem das tecnologias no contexto em que vivem. Como exemplo de alternativa à prática comum, uma oficina de reciclagem, que trata da importância da separação de resíduos, seria mais eficiente caso buscasse se encaixar na realidade da comunidade, considerando as dificuldades e aptidões locais, para além de apenas uma explicação generalizada do direcionamento dos materiais.

Essa perspectiva é alinhada ao referencial do Direito Humano à água e ao esgotamento sanitário (DHAES), um conceito aprovado e reconhecido pela Assembleia Geral da ONU em 2010. A visão a partir do DHAES desloca o paradigma do saneamento como um mero serviço para o de um direito fundamental, indissociável do direito à vida, à saúde e à dignidade. Uma abordagem com base no conceito do DHAES ultrapassa a simples análise de cobertura de redes de água e esgoto, trazendo enfoque às populações

mais vulneráveis e mais particularidades culturais, reconhecendo os indivíduos como sujeitos de direitos.

O uso do DHAES como referencial teórico é de grande valor para esta dissertação, por trazer enfoque às dimensões da desigualdade na distribuição da dengue. Diversos estudos apontam que pessoas em situação de maior vulnerabilidade possuem risco relativo superior aos efeitos da doença em comparação com pessoas com condições melhores de vida (do Carmo *et al.*, 2020; Salles *et al.*, 2018).

Apesar disso, a relação pode ser paradoxal, quando os dados são analisados em dimensões mais amplas. Nos documentos publicados por Andrade *et al.* (2024) e Oliveira *et al.* (2023) os autores encontraram resultados indicando que municípios com melhores condições sociais e acesso a serviços de saneamento básico possuíam, também, as maiores incidências de dengue dentre os municípios analisados. A justificativa dada pelos autores é relacionada a fenômenos de urbanização, que, apesar de trazerem melhorias em indicadores de saneamento, também disponibilizam criadouros e proporcionam aumentos na densidade populacional, fornecendo condições ideais para proliferação dos mosquitos e do vírus. Outra justificativa dada é devida à perda de serviços ecossistêmicos de controle da dengue fornecidos pela ecologia local causados pelo desmatamento.

Ainda, de forma a complexificar o debate, Novaes *et al.* (2022) demonstram como as características climáticas e de densidade populacional de cada local afetam significativamente as correlações entre saneamento e dengue. Isso reforça a ideia de que não existe uma solução única para a problemática da dengue e que, para que estudos cheguem a resultados satisfatórios, contextualizações devem ser feitas.

Os pontos levantados neste capítulo evidenciam a grande complexidade inerente à epidemiologia da dengue. As dinâmicas da doença são influenciadas por interações entre a biologia do vetor, condições climáticas, grau de urbanização, capacidade dos sistemas de saúde, a qualidade e a equidade dos serviços de saneamento, o perfil socioeconômico e comportamental da população. É com base nessa natureza multifatorial que o presente trabalho justifica a abordagem adotada, ao buscar compreender quantitativamente os efeitos de variáveis climáticas e de saneamento básico na ocorrência de dengue a partir de um modelo mecanístico.

3.2 A CIÊNCIA DA EPIDEMIOLOGIA

A epidemiologia, do grego: *epi* (sobre) + *demos* (povo) + *logos* (estudo), é a ciência que constitui a base da saúde pública. Uma das definições clássicas é apresentada

por Last (2001) como “o estudo da distribuição e dos determinantes de estados ou eventos relacionados à saúde em populações específicas, e sua aplicação na prevenção e controle dos problemas de saúde” (apud Bonita; Beaglehole; Kjellström, 2010). Os resultados obtidos em estudos epidemiológicos são fundamentais para orientar a tomada de decisões e para contribuir com o desenvolvimento e a avaliação de intervenções voltadas ao controle e à prevenção desses problemas.

Quatro dimensões-chave são utilizadas na epidemiologia para compreender desfechos em saúde: indivíduos/grupos afetados (quem), vias de transmissão (como), variação temporal e espacial (quando/onde) e determinantes relacionados (por quê). Os estudos epidemiológicos são categorizados em abordagens “descritivas” e “analíticas”. A primeira abordagem é utilizada para caracterizar a ocorrência das doenças, organizar e resumir dados segundo as dimensões citadas. Com ela, hipóteses são formuladas, mas não testadas. Já na abordagem analítica, as hipóteses são testadas, utilizando-se de grupos-controle para avaliar associações entre fatores de risco e o desenvolvimento de doenças, o que permite a comparabilidade entre resultados (Bonita; Beaglehole; Kjellström, 2010).

As abordagens mencionadas no parágrafo anterior são subdivididas entre diferentes estudos, cada um com diferentes utilidades e aplicações. O presente trabalho utiliza metodologias aplicadas nos denominados “estudos ecológicos”. Nesses estudos, as unidades de análise não são indivíduos, mas sim grupos populacionais delimitados por recortes espaço-temporais. Estudos ecológicos são úteis para analisar grandes volumes de dados e identificar correlações em larga escala, como os efeitos de variáveis ambientais, de saneamento e socioeconômicas no desfecho, a exemplo da ocorrência de dengue.

Outra aplicação da epidemiologia descritiva é no monitoramento de padrões temporais da casuística de doenças endêmicas, isto é, que possuem certa estabilidade na casuística ao longo do tempo. As doenças infecciosas, como a dengue, comumente exibem variações sazonais de casos, com flutuações previsíveis ao longo do ano. O acompanhamento dessas flutuações permite a construção de canais endêmicos, que são ferramentas fundamentais utilizadas na vigilância da saúde e representam as incidências esperadas de uma doença em cada período do ano, em um determinado local. É com base nos canais endêmicos que limites são estabelecidos e indicam quando picos de casos são considerados normais ou surtos epidêmicos. As instituições de vigilância de saúde fazem acompanhamento contínuo desses dados para promover ações apropriadas de prevenção e controle dos casos caso sejam observadas irregularidades.

Os limites de um canal endêmico delimitam diferentes zonas, representadas pelas áreas coloridas na figura 3.1, abaixo. São elas: zona de êxito, zona de segurança, zona de alerta e zona de epidemia. A delimitação dos limites dos canais é calculada com base em quartis ou desvios-padrão a partir do valor da média histórica da incidência para o mesmo período, a depender da literatura utilizada.

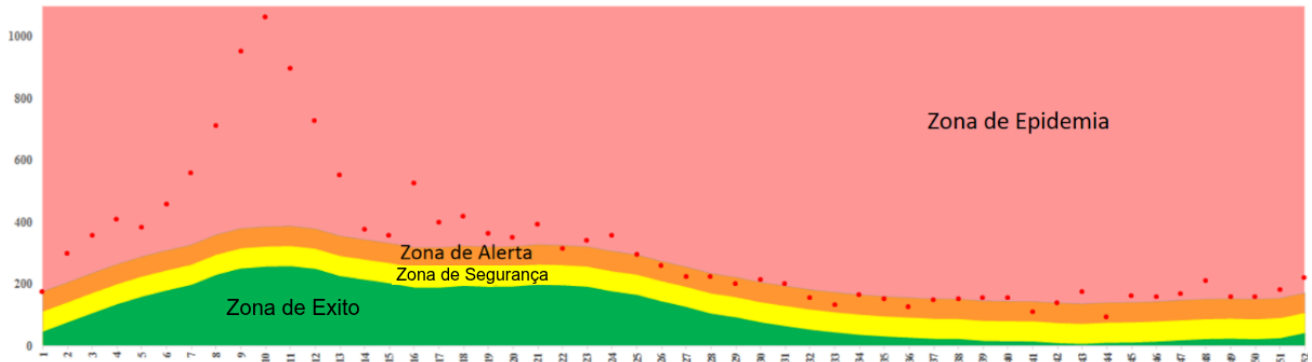


Figura 3.1: Canal endêmico da dengue, subdividido por semanas epidemiológicas, no Equador. Adaptado de Gutiérrez, Vallejo e Rivas (2021)

Quando a incidência de uma doença, em determinado período e local, ultrapassa o limite superior esperado desse canal, e chega à zona de epidemia, configura-se uma situação de epidemia. Se essa ocorrência for geograficamente restrita, denomina-se um surto epidêmico. Se a epidemia se disseminar para mais de um país ou continente, passa a ser considerada uma pandemia.

A compreensão desses conceitos, bem como o reconhecimento de erros e vieses e a caracterização de medidas de frequência e de mortalidade são essenciais para os estudos epidemiológicos.

3.2.1 Erros e vieses em estudos epidemiológicos

Neste subcapítulo são apresentados os tipos de erro mais comuns em estudos epidemiológicos, bem como as técnicas usadas para redução dos seus efeitos, com enfoque especial em exemplos relevantes para o presente estudo. Amplas quantidades de variáveis são analisadas em estudos epidemiológicos na busca de quantificar suas reais influências sobre a saúde humana. Espera-se que os estudos sejam válidos e precisos, isto é, que os resultados se aproximem da realidade e sejam reprodutíveis. Dois tipos de erros são comumente observados em estudos desse tipo e que podem comprometer a qualidade dos resultados: erros aleatórios e erros sistemáticos.

Os erros aleatórios são causados por acasos amostrais e afetam a precisão dos resultados, resultando em intervalos de confiança amplos. Já os erros sistemáticos, ou

vieses, são causados por falhas nos desenhos e metodologias aplicadas nos estudos. Estes erros podem distorcer e tendenciar resultados para alguma direção específica, afetando a validade do estudo. Exemplos de erros sistemáticos são: viés de seleção, viés de informação, viés de ambiguidade temporal e viés de confusão. O último viés é o de maior interesse em estudos ecológicos e receberá o enfoque nesta seção.

Os vieses de confusão ocorrem quando associações entre variáveis e desfechos são distorcidas pela presença de outras variáveis, denominadas “variáveis de confusão”. Para ser caracterizada como tal, ela deve ao mesmo tempo ser associada à exposição, não ser mediadora entre a exposição e o desfecho e possuir associação ao desfecho independentemente da exposição. Um erro por viés de confusão ocorre quando variáveis relevantes para o estudo são erroneamente excluídas das análises.

Um exemplo prático, no universo do presente trabalho, está diretamente ligado à conceituação essencial de Determinantes Sociais de Saúde (DSS) (Buss; Pellegrini Filho, 2007). É sabido que a presença de ambientes propícios para reprodução do vetor e a falta de acesso a serviços de saneamento básico (exposições) possuem efeitos diretos sobre a incidência da dengue (desfecho). Tratam-se de exposições diretamente ligadas ao desfecho. Por outro lado, classe social, escolaridade, raça e gênero, são fortemente correlacionadas às variáveis anteriores, não são causadoras da dengue. Estas variáveis atuam, portanto, como variáveis confundidoras: pessoas em situações de maior vulnerabilidade social geralmente têm piores condições de moradia, menos acesso a serviços de saúde e de saneamento, além de menor disponibilidade de tempo para prevenção e tratamento, o que as torna desproporcionalmente mais afetadas pela doença.

Diferentes estratégias podem ser adotadas durante o planejamento dos estudos para mitigar os efeitos de vieses de confundimento. Para estudos analíticos, a randomização, o pareamento ou a aplicação de restrições são estratégias comuns que possuem bons resultados. Já em estudos descritivos, é comum o uso de técnicas de estratificação e de análises multivariadas, que permitem compreender e ajustar matematicamente os efeitos das variáveis de confusão, reduzindo os efeitos delas sobre o desfecho.

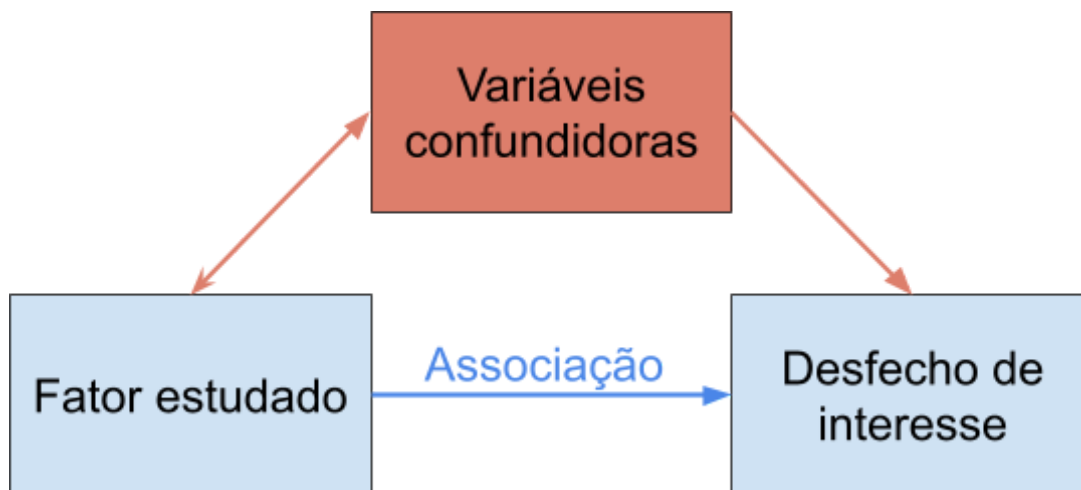


Figura 3.2: esquema de variáveis confundidoras em estudos epidemiológicos.

3.3 MÉTODOS E MEDIDAS MATEMÁTICOS APLICADOS EM ESTUDOS DE SAÚDE PÚBLICA

No capítulo anterior, conceitou-se e debateu-se sobre a importância da epidemiologia para a saúde pública. Os pilares desta ciência são a avaliação e o monitoramento do estado da saúde das populações, que possibilitam a implantação eficiente de técnicas de controle e prevenção (Bonita; Beaglehole; Kjellström, 2010). As análises quantitativas são úteis para apoiar no monitoramento de doenças, embasando tomadas de decisão; a Matemática e a Estatística oferecem ferramentas úteis para tal. Neste capítulo, conceitos e cálculos utilizados em estudos epidemiológicos mais relevantes para o presente estudo são explicitados.

As duas principais medidas para quantificar a ocorrência de uma doença são a incidência e a prevalência. Ambas são medidas para espaços de tempo padronizados, em geral meses ou semanas epidêmicas. A incidência mede o número de novos casos que ocorrem em um determinado espaço de tempo, em uma população. Esta medida é útil para doenças que possuem tempo de recuperação mais curto, como gripe e dengue. A prevalência, por outro lado, mede a quantidade de indivíduos atualmente infectados em determinado ponto no tempo. Esta medida é utilizada para acompanhamento de doenças crônicas, como HIV e câncer (Bonita; Beaglehole; Kjellström, 2010).

As medidas de quantificação utilizam valores brutos ou padronizados, proporcionais à população a que estão sendo comparadas. Valores brutos são usados quando o interesse da pesquisa é conhecer o total de casos em uma população, como por exemplo para a compra de vacinas. Para que sejam feitas comparações justas entre diferentes espaços ou períodos, os dados brutos são padronizados em proporções populacionais: uma medida comum de incidência é em casos por 100.000 habitantes. Isto

permite a observação de quais regiões sofrem maiores impactos por determinada doença. Uma aplicação é para delimitar quais municípios, ou bairros, possuem maior incidência da doença, buscar compreender os motivos para tal. Um município menor pode possuir uma quantidade bruta de casos inferior a outro, mais populoso, mas sua incidência ser maior, justificada por piores condições de saneamento, acesso a serviços de saúde, climáticas, culturais, dentre outros. Tanto valores brutos de dados quanto incidências serão utilizados neste trabalho e mencionados no decorrer do texto.

Para determinar se um aumento de casos de determinada doença é caracterizado como “normal” ou uma anomalia estatística, utiliza-se a ferramenta do canal endêmico. O cálculo dos limites pode ser feito de diferentes maneiras, utilizando-se a média ou a mediana como base. O Ministério da Saúde indica, em norma técnica, o uso de mediana e quartis, onde o primeiro quartil indica o limite inferior do canal, enquanto o limite superior é dado pelo terceiro quartil (BRASIL, 2023). Também é indicado o uso de média e desvios-padrão. Um exemplo de cálculo está na equação abaixo:

$$\text{Limite superior do canal endêmico} = \underline{x} + 1,96 * s, \quad (3.1)$$

onde:

\underline{x} é a média de casos de dengue em cada mês, em cada município.

s é o desvio padrão de casos de dengue em cada mês, em cada município.

A literatura indica diferentes valores para a multiplicação do desvio padrão. Os valores mais comuns são 1,96 e 2, assemelhando-se a um teste de hipóteses estatísticas com intervalo de confiança na ordem de 95%, onde um valor superior a este ponto é considerado um evento estatisticamente raro, com baixa probabilidade de ocorrer. O estabelecimento de um limiar de significância é uma ferramenta simples e fundamental para análises estatísticas e está revisitado no contexto da análise de correlação (capítulo 5.4.2) deste trabalho.

Além de analisar variações temporais, pode ser de interesse de um estudo epidemiológico observar os padrões espaciais de distribuição de uma doença. A epidemiologia espacial é uma subdivisão da epidemiologia que analisa a espacialidade das doenças.

Análises espaciais buscam determinar objetivamente se a distribuição de eventos de saúde, como a incidência de dengue, é determinada por padrões geográficos ou é aleatória. A hipótese inicial em testes estatísticos espaciais indica que há aleatoriedade espacial em uma distribuição, isto é, que um cenário observado em um município, por exemplo, não possui relações com os valores da vizinhança. Esta hipótese pode ser rejeitada em favor da hipótese de que há autocorrelação espacial. A autocorrelação espacial pode ser categorizada entre a autocorrelação positiva e a negativa.

A autocorrelação espacial positiva aponta que a distribuição espacial não é aleatória e que locais com valores semelhantes, como por exemplo municípios com valores elevados de incidência, estão agrupados. A autocorrelação negativa, por outro lado, também indica que a distribuição não é aleatória, mas que há dissimilaridade entre municípios. Isto é, na autocorrelação negativa, municípios com valores elevados de incidência estão próximos de municípios com valores mais baixos.

O método do I de Moran é um dos métodos mais utilizados para o cálculo da dependência espacial (Gimond, 2023). O método entrega como resultado a intensidade da correlação e a significância estatística do cálculo. Valores significativos positivos indicam autocorrelação positiva e vice-versa. Já para os casos de não significância, não há evidências para rejeitar a hipótese nula de aleatoriedade espacial. O I de Moran é uma estatística global, que resume o padrão de toda a região de estudo em um único valor, não possuindo capacidade de delimitar a localização dos agrupamentos. O método denominado “Indicadores locais de associação espacial” (LISA, em inglês) foi elaborado para superar esta limitação (Anselin, 1995).

Além de padrões temporais e espaciais, outra ferramenta útil aplicada em estudos de saúde pública são as análises de associação entre variáveis. As análises de correlação permitem quantificar relações existentes entre variáveis e desfechos na saúde, como a incidência da dengue. Diferentes coeficientes existem para esse fim e sua escolha deve ser feita dependendo das características dos dados estudados.

O coeficiente de Pearson (r) é uma das medidas mais simples e reconhecidas para análises de correlação. Neste método, a relação linear entre duas variáveis é quantificada, sendo útil para dados paramétricos. Muitas vezes, dados epidemiológicos e ambientais não seguem distribuições estatísticas normais, o que faz com que o coeficiente de Pearson seja ineficiente quando aplicado nesses contextos (Oliveira *et al.*, 2023). Mesmo assim, este método continua sendo aplicado na literatura, como observado por Abdullah *et al.* (2022).

No contexto de dados de saúde, o uso de coeficientes não paramétricos, como o de Spearman (ρ), é mais recomendado do que o r . Na pesquisa de Abdullah et al., os autores observaram que a maior parte dos estudos analisados utilizaram ρ para compreender relações entre variáveis. Um exemplo desta aplicação está na pesquisa de Ribeiro *et al.* (2021), que aplicou correlação de Spearman para avaliar relações entre indicadores de infestação de *Aedes aegypti* e taxas de incidência da dengue em municípios do Rio de Janeiro. Os autores encontraram correlações positivas entre as variáveis e identificaram fragilidades no método utilizado para cálculo dos indicadores, a partir do qual sugeriram melhorias.

Os conceitos abordados neste capítulo estão explorados nos capítulos 5.3 e 5.4 com o intuito de reconhecer correlações entre variáveis e padrões de dependência espacial da dengue no contexto dos municípios estudados.

3.4 MODELOS MATEMÁTICOS UTILIZADOS NA SAÚDE PÚBLICA

Um “modelo” é caracterizado como uma representação matemática simplificada da realidade, que busca descrever comportamentos dentro de um sistema. No contexto do presente trabalho, modelos matemáticos se apresentam como ferramentas úteis para testagem de hipóteses, avaliação de cenários e apoio na tomada de decisões. Expressados através de conjuntos de equações matemáticas, os modelos podem ser tão complexos quanto necessários para um estudo: “todos os modelos estão errados, mas alguns são úteis” foi um aforismo elaborado pelo estatístico George Box ainda na década de 1970.

O objetivo de um modelo não é, portanto, simular perfeitamente a realidade, mas sim construir uma abstração que capture a essência de determinado fenômeno estudado. Um dos desafios durante a elaboração de modelos eficientes é a manutenção do equilíbrio entre a complexidade do modelo e sua aproximação da realidade. Uma das métricas para a qualificação deste equilíbrio é o Critério de Informação de Akaike (AIC, em inglês). Este método compara a aderência e a simplicidade de diferentes modelos, favorecendo os que possuem maior aproximação e menor complexidade.

No campo da epidemiologia, os modelos podem ser classificados de diferentes formas. Uma distinção fundamental está entre as abordagens mecânicas e as estatísticas. Os modelos estatísticos identificam padrões e correlações nos dados, simulando implicitamente os processos biológicos relacionados à disseminação da doença. Já os modelos mecânicos buscam representar explicitamente os mecanismos que governam a transmissão da doença. Nestes, são considerados aspectos como as taxas de sobrevivência,

de oviposição e de transmissão vertical, probabilidade de transmissão de vetor para hospedeiro e hospedeiro para vetor, dentre outros parâmetros.

Outra distinção importante é na abordagem dos efeitos da aleatoriedade nos resultados. Em modelos estocásticos, elementos de aleatoriedade são incorporados nos cálculos para representar os efeitos do acaso nas mecânicas de disseminação da doença, sendo uma abordagem valiosa para sistemas onde a variabilidade e o acaso são componentes críticos da transmissão, como na introdução inicial de um patógeno (ou um novo sorotipo de doença) em uma população suscetível, ou em simulações em que as populações de hospedeiros e vetores são muito pequenas. Por outro lado, a abordagem determinística não adota elementos de aleatoriedade, o que faz com que cálculos que utilizam condições iniciais idênticas sempre cheguem aos mesmos resultados finais. Esta abordagem é adequada para análises quando há um prévio entendimento dos mecanismos físicos e biológicos que governam o sistema. A abordagem utilizada neste trabalho foi a determinística, uma vez que foram incorporados conhecimentos consolidados na literatura sobre o ciclo de vida do vetor e as dinâmicas de disseminação da doença, para a construção do modelo mecanístico.

O pilar da modelagem mecanística de doenças infecciosas foi construído ainda no início do século XX, proposto originalmente por Ross, aprimorado por McDonald e, depois, por Kermack e McKendrick (Brauer; Castillo-Chavez; Feng, 2019). Nasce, com esses autores, o modelo compartimental denominado “SIR”, que divide a população de hospedeiros entre os compartimentos “suscetível”, “infectado” e “recuperado”. Neste método, um hospedeiro pode estar em apenas um compartimento, em determinado ponto no tempo, e a transição entre compartimentos é dada através de taxas de infecção (λ) e de recuperação (γ).

Um indivíduo é denominado “suscetível” se estiver passível de contrair a doença. Um indivíduo é considerado “infectado” se estiver atualmente com a doença e tiver capacidade de transmiti-la. Já os indivíduos “recuperados” são aqueles que já passaram pela doença e adquiriram imunidade. A taxa de infecção é dada com base na frequência de contato entre suscetíveis e infectados e a probabilidade da transmissão do patógeno. A taxa de recuperação é relacionada ao tempo em que um indivíduo infectado leva para se livrar da infecção.

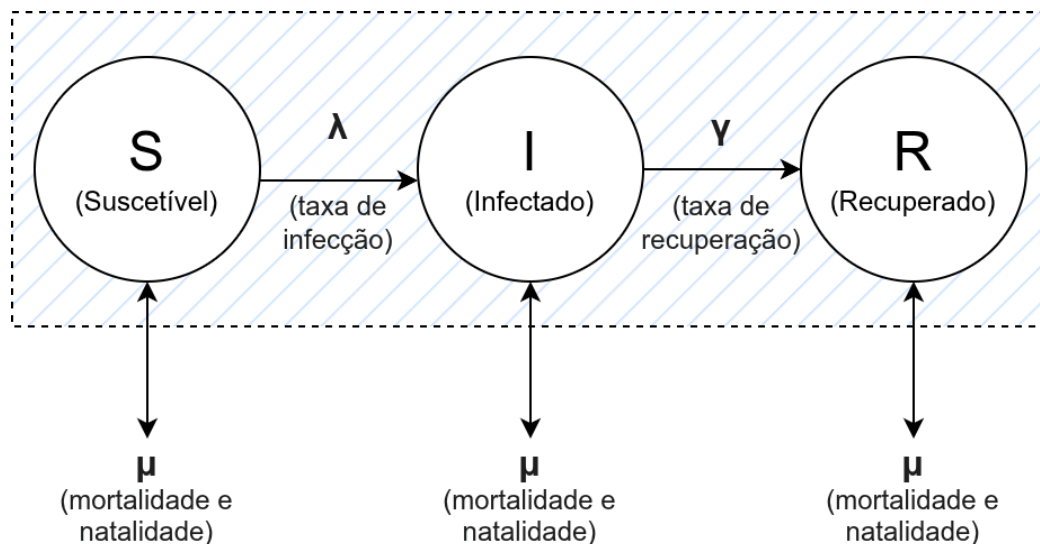


Figura 3.3: esquema de um modelo SIR.

A elaboração dos primeiros modelos SIR foi um marco importante para os estudos epidemiológicos. Contudo, estes modelos possuem limitações em escopo e aplicação quando consideram, por exemplo, que um indivíduo recuperado adquire imunidade contra a doença para todo o restante da sua vida ou então de que uma doença é transmitida entre dois indivíduos. Isto motivou a elaboração de modelos mais complexos com o passar do tempo, que consideram outros compartimentos e taxas de transição para explicar a dinâmica infecciosa das doenças. No contexto da dengue, é comum a adoção de compartimentos para os vetores, de forma a simular interações vetor-hóspede.

Além dos modelos mecanísticos, expostos acima, também são recorrentes, na literatura, estudos que utilizam modelos estatísticos para investigar a epidemiologia da dengue. Diferente da abordagem mecanística, que simula explicitamente os aspectos da transmissão da doença, nos modelos estatísticos essa simulação é implícita e o enfoque é feito em encontrar (e quantificar) as associações existentes entre as variáveis de interesse e o desfecho, neste contexto, a incidência da doença.

Dada a natureza complexa dos dados epidemiológicos e ambientais utilizados em estudos como esses, é comum o uso de testes não-paramétricos, que são recomendados para dados que não seguem uma distribuição normal. Utilizando dados do SINISA, no contexto da COVID-19, Lappicy *et al.* (2023) elaboraram e compararam diferentes modelos lineares generalizados para compreender as relações entre a doença e os dados de saneamento básico, encontrando relações significativas entre as variáveis. O SINISA é o banco de dados utilizado no presente estudo e está apresentado no capítulo seguinte.

Abordagens quantitativas ainda mais avançadas integram diferentes dimensões nas suas análises. Os trabalhos de Rachel Lowe, da *London School of Hygiene and Tropical Medicine*, são exemplos de aplicações assim. Uma de suas linhas de trabalho é a utilização de modelos espaço-temporais não lineares. Em um trabalho de 2021, para o Brasil, a autora encontrou defasagens temporais entre temporadas de chuva e de seca e aumentos na razão de risco da dengue na ordem de 0–3 meses e 3–5 meses, respectivamente (Lowe *et al.*, 2021; Sebastianelli *et al.*, 2024). Esta abordagem é útil pois considera não apenas o tempo, mas a espacialidade das relações entre as variáveis analisadas. Outro trabalho co-autorado por Lowe analisa relações entre mudanças climáticas, anomalias climáticas e a casuística da dengue no Brasil. Neste estudo foram aplicados métodos lineares de regressão e técnicas de mineração de dados para analisar grandes quantidades de dados brutos, indicando as variáveis “circulação prévia do vírus”, “urbanização” e “ocorrência de anomalias de temperatura por períodos prolongados” como os principais fatores que levaram ao aumento da incidência de dengue na região central do Brasil (Barcellos *et al.*, 2024).

As análises numéricas da epidemiologia da dengue podem ser feitas através de uma multitude de métodos e a escolha da abordagem depende dos objetivos de cada pesquisa. Conforme pontuado na introdução, lacunas significativas existem quanto à utilização de variáveis de saneamento básico para estudar a dengue, o que expõe uma oportunidade para o presente trabalho. Da mesma forma que um modelo mecanístico que considera apenas compartimentos de humanos é insuficiente para estudar arboviroses, instiga-se que a inserção de parâmetros de saneamento básico nos modelos pode frutificar em bons resultados. Desta forma, este trabalho opta por uma estrutura mecanística mais detalhada, que contempla compartimentos para humanos, mosquitos e fase aquática do vetor. Isso possibilita avaliar mais detalhadamente os efeitos do acesso a serviços de saneamento básico nos determinantes ambientais, uma área pouco investigada.

Para realizar tal avaliação, é imprescindível dispor de dados de qualidade que caracterizam as condições de saneamento básico no país. O capítulo a seguir explicita as fontes utilizadas e debate aspectos e limitações do uso de dados secundários no estudo, com enfoque no Sistema Nacional de Informações de Saneamento básico (SINISA) (Borges *et al.*, 2022).

3.5 O ESTADO DO SANEAMENTO BÁSICO NO BRASIL

O SINISA, o banco de dados utilizado neste trabalho, foi estabelecido em 1997. Sua função é fornecer informações sobre saneamento para membros da sociedade e tomadores de decisão, sendo uma ferramenta essencial para o acompanhamento da universalização do saneamento no país (Borges *et al.*, 2022).

Apesar da dimensão e importância do SINISA, é importante ressaltar que o panorama de informações sobre saneamento no Brasil é diverso e fragmentado. Diferentes agências, programas e sistemas, cada um com diferentes metodologias e recortes, já lançaram diversos materiais que podem ter, inclusive, informações conflitantes entre si. De acordo com o Instituto Água e Saneamento, essa fragmentação dificulta uma visualização completa da real dimensão e complexidade do setor (Instituto Água e Saneamento, 2024).

Como exemplos de outras fontes para úteis para compreensão das condições de saneamento no país, para além do SINISA, que foca na prestação dos serviços, tem-se: Censo Demográfico e a Pesquisa Nacional por Amostra de Domicílios (PNAD), elaborado pelo IBGE, que oferece uma perspectiva a partir da residência, detalhando as soluções utilizadas pela população; Pesquisa de Informações Básicas Municipais (MUNIC), também do IBGE, que foca no planejamento e na capacidade de gestão das prefeituras; Sistema de Informação de Vigilância da Qualidade da Água para Consumo Humano (Sisagua), vinculado ao Programa Nacional de Vigilância da Qualidade da Água para Consumo Humano (Vigiagua) do Ministério da Saúde, que foca nas características e qualidade das estruturas de abastecimento de água.

Além desses, há uma ampla classe de bancos de dados brasileiros com diferentes enfoques e elaborados a partir de dados primários e secundários, tendo o SINISA como uma das principais fontes para sua elaboração. A Agência Nacional de Águas e Saneamento Básico (ANA), por exemplo, publica os diagnósticos técnicos “Atlas Esgotos” e “Atlas Água”, ambos de grande relevância. O Programa Nacional de Saneamento Rural (PNSR) foca nas condições e planejamento para melhorias no meio rural. Já o Plano Nacional de Saneamento Básico (PNSB) é outra fonte de informações que possui análises pertinentes, mas seu último lançamento foi feito apenas em 2014, estando em processo de revisão desde então.

A diversidade de fontes exposta nos parágrafos anteriores evidencia a complexidade do setor e reforça a capacidade e relevância do SINISA como um sistema abrangente e padronizado para avaliação das condições de saneamento em escala nacional. Apesar disso, assim como os outros, este sistema não é perfeito e considerações devem ser

feitas para sua aplicação. Borges et al. apontam que uma das dificuldades no sistema é quanto à qualidade das informações, que são autodeclaradas pelos representantes municipais dos serviços de saneamento (Borges *et al.*, 2022).

Uma das limitações estruturais de maior significância está na decisão metodológica do sistema de disponibilizar informações apenas de serviços centralizados de água e esgoto. Soluções individuais, como fossas sépticas e poços, não são computadas, embora sejam consideradas pelo PLANSAB. Esta escolha é justificada devido à dificuldade de se garantir que os dados obtidos sejam de qualidade satisfatória. Isso pode implicar na subnotificação da cobertura real de saneamento e representar vieses em estudos que utilizam esses dados, um fator que está considerado nos resultados do presente trabalho.

A complexidade do SINISA, com mais de cem variáveis em cada eixo, também é um desafio metodológico. A escolha e o uso de variáveis é complexa e propensa a erros, visto que diversas das variáveis possuem correlações entre si, o que pode proporcionar multicolinearidade e enviesar resultados. Em uma meta-análise que buscou entender como o banco de dados tem sido utilizado na ciência, Oliveira, Lima e Cardoso (2023) observaram que o SINISA já foi utilizado em mais de 300 artigos. Não foram encontrados estudos dedicados ao agrupamento dos indicadores de saneamento. Uma das contribuições desta dissertação está em abordar esta lacuna e aplicar métodos de agrupamento para utilização dos dados no modelo epidemiológico proposto.

Este capítulo buscou contextualizar as oportunidades e lacunas para utilização de dados do SINISA na presente dissertação. Esta compreensão crítica do banco de dados, em conjunto com os debates realizados sobre a dengue e a epidemiologia, estabelece a base que fundamenta as escolhas metodológicas deste estudo. Deste modo, espera-se que o trabalho contribua para a expansão dos limites do conhecimento científico ao propor novas formas de análise da epidemiologia da dengue que consideram a influência do saneamento básico em suas dinâmicas.

4. REVISÃO DE LITERATURA

4.1 MODELOS MECANÍSTICOS DE DENGUE

Diante do interesse do presente trabalho de explorar relações entre variáveis e disseminação da dengue de forma explícita, neste capítulo são apresentadas variações do modelo de Ross e MacDonald aplicados em tempos mais recentes.

Como ponto de início, tem-se a sistematização realizada por Aguiar et al. (2022), que traz uma varredura de diversos tipos de modelos matemáticos aplicados à epidemiologia da dengue. Esta é uma referência essencial para o presente trabalho, pois além de apresentar as metodologias mais usadas e lacunas na ciência, também contém tabelas com valores adotados para os parâmetros dos modelos, em especial com abordagem vetor-hospedeiro, como o aplicado neste trabalho. Os modelos analisados pelos autores são categorizados a partir das abordagens, uso de controles, sazonalidade, e outras características.

As abordagens são divididas entre as simulações de relações hospedeiro-hospedeiro (focam explicitamente nos humanos, não modelando as dinâmicas dos vetores), vetor-vetor (apenas relações entre vetores; consideram aspectos como infectividade viral e resposta imune), e vetor-hospedeiro (tratam tanto da biologia dos vírus quanto das relações humanas).

As simulações vetor-hospedeiro são feitas com diferentes níveis de complexidade. São encontrados estudos que inserem diferentes compartimentos em seus modelos, detalhando etapas de ovos, larvas e pupas para além dos vetores adultos e hóspedes, além de estudos que incorporam (tornando explícitos seus cálculos) processos dinâmicos da doença, como os efeitos da vacinação, período de incubação em humanos e vetores e a co-circulação de sorotipos. Os conjuntos de equações podem ser expandidos conforme necessário, sendo limitados pela busca de equilibrar a complexidade do modelo e sua capacidade de ajuste. Um ajuste excessivamente estrito, denominado *overfitting*, diminui a capacidade do modelo de compreender diferentes realidades. O Critério de Akaike (AIC) é uma ferramenta estatística frequentemente utilizada para avaliar esse equilíbrio, penalizando modelos excessivamente complexos que não tenham ganhos substanciais na qualidade do ajuste.

A disseminação da dengue ocorre através dos vetores, que são os mosquitos. Não considerar seus aspectos biológicos nos modelos mecanísticos significaria modelar implicitamente seus efeitos, diminuindo a sensibilidade dos modelos. Isso não impede que

resultados interessantes sejam obtidos, entretanto. Por exemplo, Campos *et al.* (2024) utilizaram apenas os três compartimentos originais (humanos suscetíveis, infectados e recuperados - SIR), simulando as taxas de transição implicitamente utilizando apenas duas variáveis: β e γ , encontrando um ajuste aproximado entre os valores observados e os simulados. No entanto, o modelo proposto pelos autores teve limitações quanto ao período de treinamento e na capacidade de aproximação do valor observado de pico de casos, possivelmente devido à maior simplificação adotada.

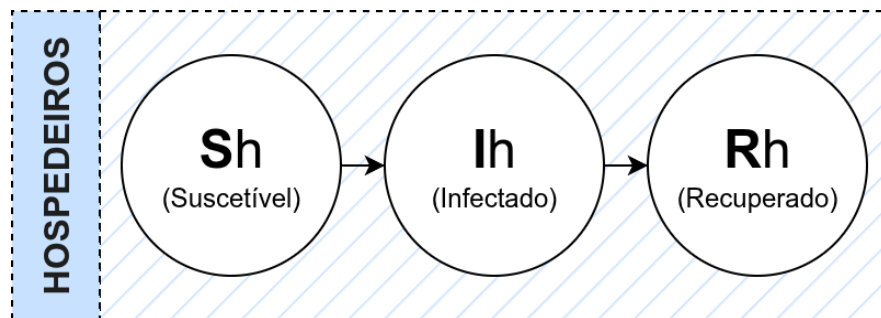


Figura 4.1: Compartimentos utilizados no modelo de Campos *et al.* (2021)

Trabalhos como o de Nuraini *et al.* (2021) e Chanprasopchai *et al.* (2018) aumentam a complexidade ao agregar compartimentos para vetores, criando os modelos “SIR-SI”. Nesses exemplos, o compartimento “R” não é utilizado para vetores devido ao seu curto tempo de vida: um mosquito infectado mantém-se infeccioso até sua morte.

Para representar o período de incubação da doença nos hospedeiros e vetores, Alves (2021) considera também o período de incubação da doença, adicionando um compartimento “E” (exposto) ao modelo, construindo o modelo “SEIR-SEI-SI”. Enquanto isso, De Vasconcelos *et al.* (2023) simula populações sintomáticas e assintomáticas em um modelo “SEMIR”.

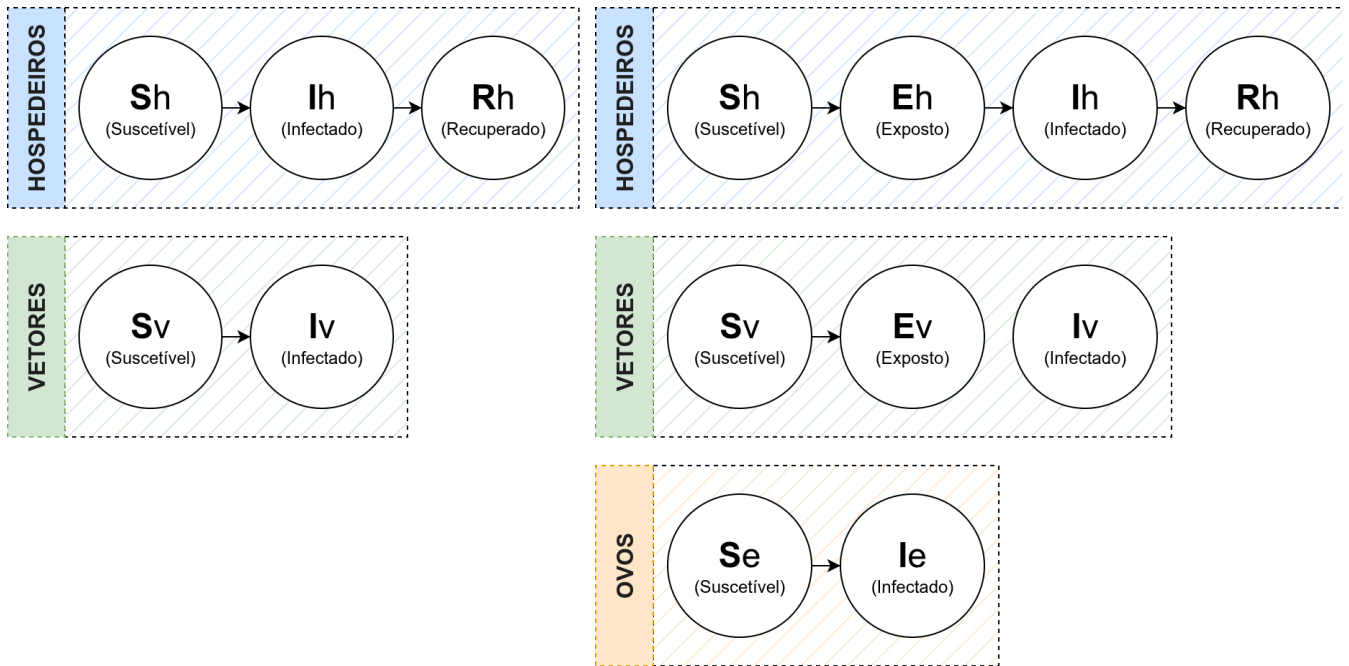


Figura 4.2: Compartimentos utilizados nos modelos de Chanprasopchai (esquerda) e Alves (direita)

Modelos mais simples, como o de Campos *et al.* (2021) e de Nuraini *et al.* (2021), utilizam menos compartimentos e, conseqüentemente, têm menores custos computacionais. Esses dois trabalhos foram elaborados para a realidade do Brasil e da Indonésia, respectivamente, e obtiveram resultados aproximados dos valores reais, sendo, portanto, considerados modelos úteis, apesar de simularem implicitamente os parâmetros e taxas relevantes. Por outro lado, modelos mais complexos têm capacidade de simular explicitamente as taxas e transições biológicas (mortalidade vetorial, período de incubação, oviposição, dentre outras), ao passo em que a solução dos conjuntos de equações que os compõem é mais computacionalmente custosa e lenta. A otimização de processos computacionais complexos é especialmente importante para a vigilância epidemiológica quando precisa, por exemplo, estimar a eficiência de diferentes estratégias de controle, ou estimar crescimento da doença em diferentes regiões, ao passo em que modelos mais complexos podem levar na ordem de dias a semanas para iterar.

A visão geral apresentada por Aguiar *et al.* (2022) complementa os conhecimentos específicos obtidos com as leituras dos artigos e teses expostos neste capítulo, de forma que todos os trabalhos apresentados nesta seção são de grande relevância para a construção teórica e metodológica do presente documento. Dessa forma, com base nas informações levantadas no presente capítulo, foi possível elaborar um modelo determinístico estruturado em compartimentos adequados para os objetivos do trabalho (um modelo

“SIR-SI-SI”), que inclui parâmetros relevantes no debate sobre quantificação das relações entre dengue e saneamento e tenha um equilíbrio razoável entre complexidade, performance e custo computacional. Mais detalhes sobre a construção do modelo estão apresentados no capítulo 5.6.

4.2 ABORDAGENS E PARÂMETROS UTILIZADOS EM MODELOS “SIR”

Para a construção de um modelo mecanístico, uma série de decisões devem ser tomadas para além da estrutura das equações. Nesta seção são apresentados valores fixos e equações utilizadas para cálculo de parâmetros e outras escolhas metodológicas adotadas por diferentes autores, contextualizando e justificando as decisões feitas na metodologia da presente dissertação.

Uma escolha inicial de grande relevância é na maneira de lidar com dados de casos das doenças. A depender do objetivo do estudo, é possível escolher entre o uso de dados proporcionais às respectivas populações (incidência), como feito por Nuraini (2021), ou utilizando contagem bruta de casos (ocorrência), como aplicado por de Vasconcelos *et al.* (2023). Dentre os estudos que utilizam contagem, há alguns que optam por agregar os dados em períodos maiores, como semanais, mensais, ou mesmo de 8 dias, como em Alves (2021). Essa escolha pode ser feita para reduzir custos computacionais e suavizar flutuações inerentes aos dados de saúde.

Para dados de população de humanos, especialmente em estudos que utilizam períodos maiores, condições como taxa de mortalidade e de nascimento ou de migração podem ter efeitos importantes nos compartimentos populacionais. Apesar disso, todos os trabalhos analisados optaram por simplificar esse aspecto e considerar a população constante ao longo de todo o período de simulação.

4.2.1 PARÂMETROS DOS MODELOS

As dinâmicas de transmissão das doenças são influenciadas por diversas variáveis, como exploradas nos capítulos anteriores. Uma vez que a modelagem mecanística busca representar numericamente e de forma explícita a realidade, a definição dos parâmetros que compõem os modelos é uma etapa essencial para a obtenção de bons resultados. As duas maneiras empregadas para definição desses parâmetros são: fixando ou calculando dinamicamente seus valores.

Em geral, os parâmetros que recebem valores fixos são aqueles mais explorados pela ciência, com boa fundamentação teórica sobre os aspectos biológicos, e parâmetros menos sensíveis a fatores ambientais. Como exemplos de parâmetros fixos, tem-se a taxa

de mortalidade de humanos (em geral utiliza-se o inverso da expectativa de vida) e a taxa de recuperação da doença (inverso do tempo necessário para um indivíduo infectado tornar-se saudável novamente).

Por outro lado, os parâmetros mais sensíveis às condições ambientais, que não possuem consenso na literatura, ou que são objetos de análise dos modelos, são frequentemente calculados utilizando funções dinâmicas. Essas funções são elaboradas em relação a variáveis que possuem relações diretas com a mecânica dos parâmetros, como temperatura e umidade relativa. As equações utilizadas variam desde funções lineares simples até equações mais complexas, como as de Brière e quadráticas (Mordecai *et al.*, 2017). As funções polinomiais abaixo foram utilizadas por Nuraini (2021) para quantificar a mortalidade de mosquitos ($\mu\nu$) e sua taxa de picada (b), em relação à temperatura (T):

$$\mu\nu = 0,8692 - 0,1590 * T + 0,01116 * T^2 - 3,408 * 10^{-4} * T^3 + 3,809 * 10^{-6} * T^4 \quad (4.1)$$

$$b = 0,0043 * T + 0,0943 \quad (4.2)$$

Grande parte da literatura analisada corrobora que múltiplas variáveis climáticas, e mesmo de saneamento, podem ter efeitos na disseminação da dengue, conforme explicitado anteriormente. Apesar disso, a maior parte dos modelos de dengue calcula os parâmetros utilizando apenas valores de temperatura. Uma das justificativas para isso é devido à baixa correlação estatística direta existente entre valores de umidade e de casos de dengue, que implica dificuldades para quantificação mecânica dos seus efeitos e a consequente calibração dos modelos.

Para enfrentar essa dificuldade, de Vasconcelos *et al.* (2023) elaboraram um método simples e eficiente para considerar a umidade em seu modelo. No método, uma taxa de evaporação diária é calculada a partir dos valores de umidade e de temperatura. Segundo os autores, essa taxa tem efeitos na quantidade de água disponível nos criadouros dos mosquitos. Com isso, a capacidade de carga do ambiente (K), que descreve a população máxima de mosquitos em determinado momento, é quantificada de forma dinâmica.

Outra abordagem que incorporou variáveis climáticas para além da temperatura foi feita por Nuraini *et al.* (2021). A maior parte das variáveis que compõem o modelo (“SIR-SI”) elaborado pelos autores foram calculadas com base na temperatura. Apenas a

taxa de infecção (βh), um dos parâmetros mais sensíveis do modelo, segundo os autores, foi calculada separadamente. O cálculo de βh foi feito utilizando um modelo estatístico autorregressivo com defasagens distribuídas (ARDL, em inglês), que computa o valor da taxa de infecção no presente considerando os efeitos causados pela temperatura, precipitação, umidade e seu próprio valor no passado.

Essa abordagem estatística para o parâmetro βh gerou um resultado importante para o presente trabalho. Os autores testaram efeitos de temperatura, precipitação e umidade, com defasagem temporal, na incidência da doença, encontrando correlações estatísticas significativas apenas entre incidência de dengue, precipitação e umidade. Isto é, não encontraram correlações entre a doença e a temperatura. Isso complexifica o debate sobre o uso de variáveis meteorológicas em modelos de dengue ao desconsiderar uma variável amplamente utilizada na literatura devido a um achado estatístico. A abordagem adotada por Nuraini *et al.*, construindo um modelo composto por variáveis mecânicas em conjunto com uma variável estatística (βh), é um exemplo de grande relevância para o presente trabalho. Os autores demonstraram como análises estatísticas podem ser úteis para a validação de premissas de um modelo mecanístico e a obtenção de bons resultados.

Além da parametrização de variáveis climáticas, outro desafio presente na modelagem da dengue é conseguir representar o ciclo de vida do vetor sem tornar o modelo excessivamente complexo. Alguns dos estudos integram apenas compartimentos de humanos ou de humanos e vetores adultos em seus modelos (modelos “SIR” e “SIR-SI”, respectivamente) (Aguiar *et al.*, 2022). Modelos como esses consideram de forma implícita as relações existentes entre variáveis climáticas, de saneamento, entre outras, com as fases iniciais do ciclo de vida dos *Aedes aegypti*, sendo incapazes de analisá-las profundamente.

Nesse sentido, Alves (2021) oferece uma solução simples e eficiente para solucionar tal problema, tornando seu modelo sensível para quantificar as relações entre as fases iniciais do mosquito e as condições climáticas. Seu modelo utiliza compartimentos para as fases de ovos, mas não de larva e pupa, e considera uma taxa de sobrevivência aquática (sa), que descreve implicitamente a transição da fase de ovos à fase adulta. Na sua aplicação, o valor de sa é calculado utilizando uma função quadrática que considera apenas temperatura, com valores definidos por Mordecai (2017). Apesar de não ter sido explorado nesse sentido, esse parâmetro também pode ser influenciado por variáveis como umidade (que afeta a dessecação de criadouros) e condições de saneamento (disponibilidade de recipientes para armazenamento de água).

O quadro 4.1, a seguir, contém valores e intervalos adotados para parâmetros relevantes para o presente estudo, em diferentes modelos de dengue na literatura. Na quarta coluna (“Valor fixo adotado ou Intervalo (referência)”), os valores fixos correspondem a valores fixados nos estudos analisados, enquanto os intervalos correspondem a valores plausíveis de referência para checagem de parâmetros calculados. A mesma coluna contém as referências numéricas, entre parênteses, que indicam as referências bibliográficas que forneceram as respectivas informações.

Quadro 4.1: Referências de valores fixos ou intervalos adotados para parâmetros na literatura

Categoria	Símbolo	Descrição	Valor fixo adotado ou Intervalo (referência na literatura)	Unidade
Hospedeiros	Nh	População total	--	Pessoas
	Sh	Proporção de suscetíveis em relação à população total	0.04 a 1 (2) 0.569 (3); 0.3 a 1 (5);	Pessoas
	μ h	Natalidade e Mortalidade	expectativa de vida ⁻¹ (3; 4)	ano ⁻¹
	γ	Taxa de recuperação	3 ⁻¹ a 12 ⁻¹ (1); 30 ⁻¹ (3); 4 ⁻¹ a 32 ⁻¹ (5)	dia ⁻¹
Vetor	μ v	Natalidade e Mortalidade	30 ⁻¹ a 7,7 ⁻¹ (1); 0,026 a 0,043 (2); 12 ⁻¹ (4)	Vetores/dia
	K	Capacidade de carga	1 a 5 (2); 3 (3); 0,25 a 4 (5)	Vetores / Pessoa
	σ	Razão de sexo feminino/masculino	0 a 1 (2); 0.5 a 1,5 (5)	--
	b	Taxa de picada	0,26 a 0,67 (1);	dia ⁻¹
	Pv	Probabilidade (Taxa) de transmissão de hospedeiro para vetor	0 a 0,75 (1); 0,5 a 1 (2); 0,25 a 1 (5)	%
	Ph	Probabilidade (Taxa) de transmissão de vetor para hospedeiro	0 a 0,5 (1); 0,1 a 0,75 (2); 0,25 a 1 (5)	%
Ovos	μ e	Mortalidade	0,01 a 0,2616 (2); 0,01 a 0,125 (5)	dia ⁻¹
	or	Taxa de Oviposição	0,3548 a 9,57 (2); 0,25 a 4 (5)	dia ⁻¹
	dr	Taxa de desenvolvimento ou período de incubação	0,169 (1); 0,011 a 0,158 (2); 0,063 a 0,1 (5)	dia ⁻¹

Categoria	Símbolo	Descrição	Valor fixo adotado ou Intervalo (referência na literatura)	Unidade
	vt	Taxa de transmissão vertical	0,1 (5)	--
	sa	Taxa de sobrevivência aquática	-0,001 a -0,005 (5)	--

Referências no quadro acima: (1) Aguiar *et al.* (2022); (2) de Vasconcelos *et al.* (2023); (3) Nuraini *et al.* (2021); (4) Chanprasopchai *et al.* (2018); (5) Alves (2021).

Cada um dos estudos analisados utiliza diferentes estratégias para determinação ou estimação de parâmetros. Nuraini *et al.* (2021), por exemplo, fixaram metade dos parâmetros em seus modelos, buscando compreender mais a fundo as variações na taxa de infecção. Outros estudos, como o de Alves (2021) e de de Vasconcelos *et al.* (2023), objetivaram compreender o conjunto total de variáveis, estimando conjuntos de mais de 10 variáveis cada.

Uma vez que os modelos devem ser úteis, e não perfeitos, não é possível que exista um consenso ou abordagem única e essencial para a parametrização de modelos de dengue: as escolhas devem ser feitas buscando solucionar da melhor forma os objetivos individuais de cada estudo. Como explicitado e defendido no presente capítulo, as decisões, definições e cálculos devem ser feitos com base na literatura, de forma a garantir que o modelo corresponda à realidade biológica. Isso distingue um modelo mecanístico de um modelo elaborado por aprendizado de máquina, que gera equações em um exercício de matemática e estatística aplicadas, mas que não possibilitam reconhecer a fundo a realidade biológica que governa a disseminação de uma doença.

4.2.2 CALIBRAÇÃO, TREINAMENTO E VALIDAÇÃO

Outro fator relevante na modelagem epidemiológica é a calibração, método usado para estimar os parâmetros. Em geral, a calibração é feita usando dois subconjuntos de dados, nos denominados períodos de “treinamento” e de “validação”. No primeiro período, as variáveis do modelo são alteradas computacionalmente, em um espaço de busca, procurando aproximar os casos estimados com os reais. A validação é a etapa seguinte, quando o modelo já treinado é aplicado em um novo período para conferir sua aproximação.

Diferentes modelos possuem diferentes finalidades. Um modelo com caráter preditivo tem como enfoque gerar previsões aproximadas da realidade. Já modelos com caráter de extração de informação, ou de inferência, são usados para extrair informações

sobre associações entre variáveis-resposta e de entrada, isto é, reconhecer efeitos dos parâmetros que governam as dinâmicas do sistema (Breiman, 2001).

Como exemplos de aplicação das duas categorias, enquanto um modelo preditivo pode ser usado para estimar curvas de casos de dengue em meses futuros, um modelo com enfoque em inferência pode ser usado para quantificar parâmetros, como taxa de sobrevivência e de infecção de mosquitos, em condições passadas.

Para estudos epidemiológicos, é comum a aplicação de modelos com caráter inferencial, uma vez que o foco principal é compreender a dinâmica da doença. Esses modelos utilizam períodos de treinamento mais longos de forma a captar flutuações nos cenários, como sazonalidade, surtos epidêmicos, entre outras flutuações. Essa abordagem é vista em trabalhos como o de De Vasconcelos *et al.* (2023) e Alves (2021), que treinaram seus modelos usando a maior parte dos períodos estudados. Da mesma forma, Nuraini *et al.* (2021) treinaram seu modelo por mais de 5 anos, com período de validação de apenas 4 meses. Ainda, Campos *et al.* (2024), que além de elaborarem um modelo SIR também utilizaram redes neurais, treinaram seus modelos no equivalente a 90% de um ano e validaram-no nos 10% restantes.

Diante do exposto nesta revisão bibliográfica, fica clara a evolução e variedade de modelos mecanísticos construídos para a dengue, que utilizam das estruturas mais simples às mais complexas, incorporando, principalmente, variáveis climáticas em sua estrutura. Em conjunto com os debates realizados anteriormente, a revisão fortalece a argumentação inicial do trabalho ao pontuar uma carência de trabalhos que incorporem explicitamente os determinantes de saneamento básico em seus modelos. Assim, no capítulo seguinte são detalhadas as metodologias adotadas no presente trabalho, buscando preencher esta lacuna de conhecimento. As análises quantitativas e qualitativas entre dengue e saneamento são apresentadas, em conjunto com o processo de construção e calibração de um modelo-base que buscou integrar explicitamente variáveis de saneamento em sua estrutura.

5. MATERIAIS E MÉTODOS

Neste capítulo são apresentadas as fontes dos dados utilizados no presente trabalho, sua obtenção, organização e posterior análise. Os dados foram armazenados e manipulados através da linguagem de programação *R* versão 4.4 e com a interface *RStudio* no sistema operacional *Linux*. Os diversos pacotes que foram utilizados no decorrer do trabalho estão mencionados ao longo do texto ou nas referências bibliográficas.

A metodologia do trabalho é dividida em 4 partes principais. Inicia-se com uma análise exploratória dos dados de dengue e de saneamento nos municípios do estado. Na sequência são apresentadas as correlações entre variáveis de dengue e de saneamento nos respectivos municípios e os métodos de agrupamento das variáveis de saneamento. Por fim, um modelo matemático de casos de dengue é exposto. O fluxograma da figura 5.1 apresenta uma visualização das metodologias aplicadas.

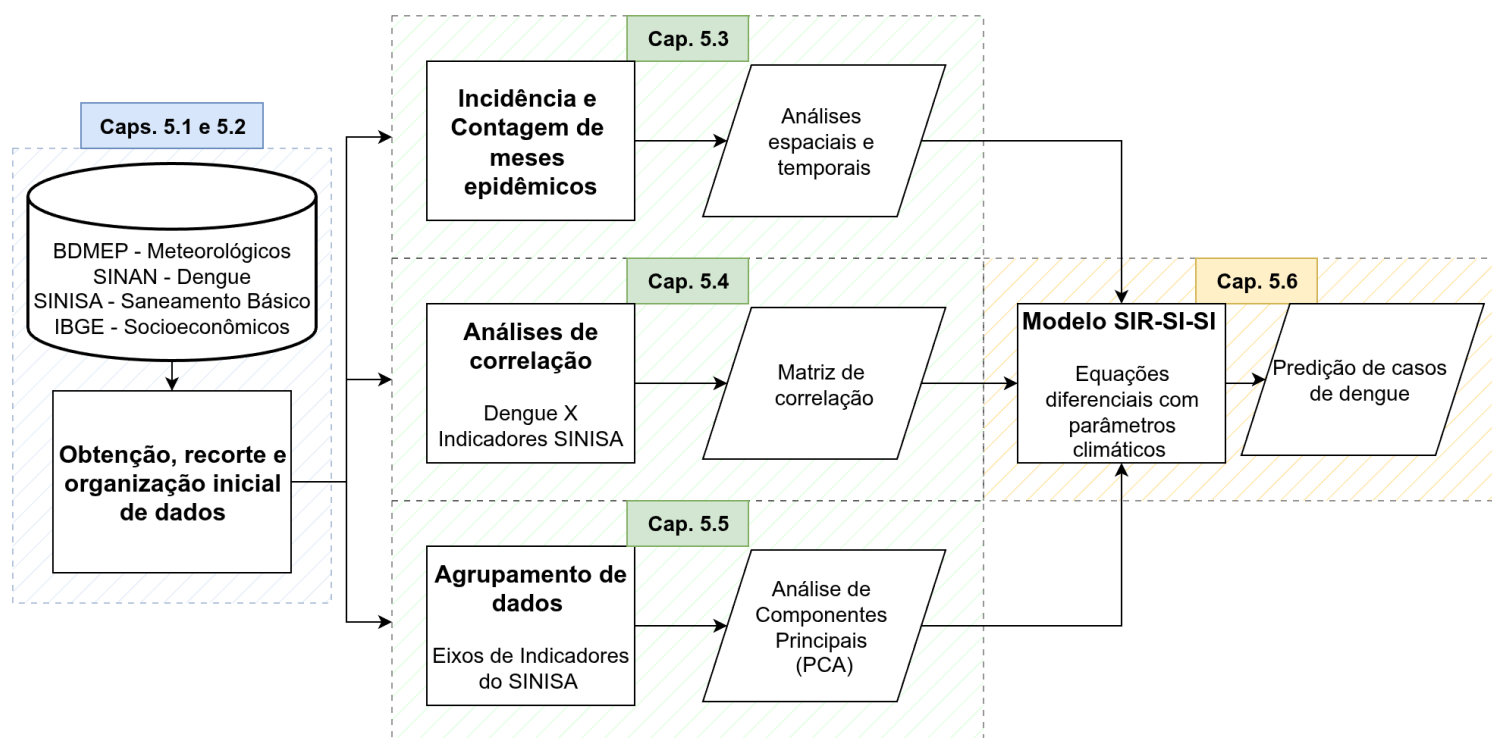


Figura 5.1: Fluxograma da metodologia.

5.1 OBTENÇÃO DE DADOS

Os dados utilizados no presente estudo são separados em três tipos: casos de dengue, dados meteorológicos e dados de saneamento básico. As informações sobre casos de dengue foram obtidas através do Sistema de Informação de Agravos de Notificação

Online (SINAN Online). Os dados meteorológicos originaram-se do Banco de Dados Meteorológicos (BDMEP) do Instituto Nacional de Meteorologia (INMET). Por fim, o SINISA foi utilizado para obter os dados de saneamento básico. Os dados foram organizados para cada município estudado. Detalhamentos sobre a escolha e a utilização das variáveis estão apresentados nos subcapítulos seguintes. O quadro abaixo apresenta um resumo das fontes de dados utilizadas neste estudo.

Quadro 5.1: Fontes utilizadas para obtenção de dados do trabalho

Variável	Descrição	Fonte	Período dados
Dados climatológicos	Temperatura, umidade e precipitação	BDMEP	2014–2022
Dados serviços de Águas Pluviais	Dados diversos de águas pluviais	SNIS-AP	2021
Dados serviços de Água e Esgoto	Dados diversos de água e esgoto	SNIS-AE	2021
Dados serviços de Resíduos Sólidos	Dados diversos de resíduos sólidos	SNIS-RS	2021
Casos dengue	Casos prováveis diários	SINAN-DENGUE	2014–2022
População por município	Estimativas DataSUS	DataSUS	2000–2022*

* Observação: devido ao lançamento tardio de informações do último CENSO, a população de 2022 foi considerada igual à de 2021, conforme Nota Técnica nº 41/2024-CGIAE/DAENT/SVSA/MS

5.1.1 Dados de dengue

O SINAN é um banco de dados de notificações compulsórias, isto é, comunicações obrigatórias de serem feitas por parte dos profissionais de saúde durante o atendimento, em casos de suspeita ou confirmação de doenças, agravo ou outros eventos de saúde pública. É, portanto, um sistema de grande legitimidade, comumente utilizado para realização de estudos epidemiológicos e útil para a compreensão da incidência da dengue. A obtenção dos dados do SINAN foi feita através do TABNET, um site oficial do DATASUS, e do aplicativo *Tabwin* v4.15, ambos recomendados pelo Ministério da Saúde para acesso aos dados do sistema (Brasil, [s. d.]). Além disso, a biblioteca *microdatasus* (Saldanha; Bastos; Barcellos, 2019) da linguagem R também foi utilizada para manuseio e exploração destes dados.

Os dados de dengue foram obtidos a nível federal, recortados e reorganizados para cada município do presente estudo. Os locais de residência (e não os locais onde a notificação foi feita) foram utilizados para indicar a qual município cada caso correspondia. Além disso, foram consideradas como “casos prováveis” todas as notificações, com exceção dos casos descartados, que foram removidos do estudo. Essas escolhas correspondem às recomendações do SINAN. Os dados foram organizados em total de casos diários e em incidência (casos/100.000 habitantes) para cada município.

5.1.2 Dados meteorológicos

Em relação aos dados meteorológicos, o BDMEP foi a fonte de informações escolhida. Esta plataforma oferece dados meteorológicos diários de centenas de estações meteorológicas presentes no Brasil (INMET, 2025). Optou-se por esta fonte, em detrimento a métodos mais complexos, como dados de satélites, para simplificar o estudo. Alves (2021), por exemplo, usou dados de satélite em seu estudo, extraindo apenas o valor do pixel mais próximo do centroide de cada município. Ambas as simplificações são semelhantes, apesar de partirem de origens diferentes. Desta forma, compreende-se que tal simplificação foi suficiente para as análises realizadas.

Os dados foram obtidos para as estações meteorológicas presentes nos municípios utilizados no modelo mecanístico aplicado na metodologia. A seleção de municípios foi feita considerando a disponibilidade de dados e as características geográficas e sociodemográficas, conforme apresentado no subcapítulo 5.2. O quadro a seguir apresenta os códigos das estações que foram utilizadas na obtenção dos dados.

Quadro 5.2: Estações meteorológicas utilizadas para obtenção de dados

Município	código estação	código IBGE
Campos dos Goytacazes	A607	3301009
Duque de Caxias	A603	3301702
Macaé	A608	3302403
Paraty	A619	3303807
Resende	A609	3304201
Rio de Janeiro - Forte de Copacabana	A652	3304557
Rio de Janeiro - Vila Militar	A621	3304557

A data de início e fim dos dados foi delimitada entre 01/01/2014 e 31/12/2022 para corresponder aos dados disponíveis de ocorrência da doença. Foram obtidos dados diários de estações automáticas e as variáveis selecionadas foram as seguintes:

1. Temperatura máxima diária
2. Temperatura média diária
3. Temperatura mínima diária
4. Precipitação total diária
5. Umidade relativa do ar média diária

Os dados climatológicos disponíveis possuem imperfeições em determinados recortes. Em mais de uma estação meteorológica foram observados dados faltantes em sequências de dias até semanas. Nesses casos, optou-se por substituir os dados faltantes por uma média móvel de 7 dias. Mais informações sobre as escolhas e a utilização dos dados meteorológicos estão apresentadas no subcapítulo do modelo mecanístico (capítulo 5.6).

5.1.3 Dados de saneamento básico

Considerando-se a problemática explicitada anteriormente quanto à qualidade dos dados de saneamento básico, optou-se por utilizar os dados do SINISA. O sistema nacional de saneamento básico possui grande relevância e reconhecimento no Brasil e parte do interesse deste estudo foi compreender se esses dados são relevantes para utilização em estudos epidemiológicos. De forma a mitigar as adversidades citadas, optou-se por usar apenas os dados de indicadores que são processados pela equipe do SINISA, o que garante maior confiabilidade nos dados (Borges *et al.*, 2022).

Foram obtidos os dados desagregados disponíveis dos 4 eixos de saneamento básico em todos os municípios do Rio de Janeiro. Como não era objetivado realizar uma análise longitudinal, com as alterações temporais nos serviços de saneamento básico, escolheu-se utilizar os dados de 2022, que foram disponibilizados em 2023 e eram os mais recentes quando o presente estudo teve início. Estes dados foram utilizados em parte com dados de todos os municípios do estado e em parte no recorte apenas dos municípios utilizados neste estudo.

Quadro 5.3: Opções utilizadas para download de dados do SNIS

Tipo	SNIS - AE	SNIS - AP	SNIS - RS
Agregação de dados	Desagregados	--	--
Natureza jurídica	--	--	Todas
Ano Referência	2014–2022	2017–2022	2010–2022
Estado	Rio de Janeiro	Rio de Janeiro	Rio de Janeiro
Faixa populacional	Todas	Todas	--
Municípios	Todos	Todos	Todos
Informações obtidas	Informações gerais, Informações de água, de esgoto e de qualidade; Indicadores operacionais de água, de esgoto e de qualidade	Informações gerais, Cobrança; Financeiro; Infraestrutura; Operacional; Gestão de riscos	Informações gerais e indicadores sobre catadores; Despesa e Trabalhadores; Coleta domiciliar e pública; Coleta de resíduos de serviços de saúde; Varrição, capina e roçada; Serviços de construção civil.

Fonte: autoral, com informações do SNIS (BRASIL. Ministério das Cidades, 2025)

5.2 RECORTE ESPACIAL E TEMPORAL

O recorte espacial do estudo foi feito em municípios situados no estado do Rio de Janeiro (RJ). O estado do RJ é uma unidade federativa com boa qualidade, disponibilidade e variedade de dados climáticos, havendo regiões com temperaturas inferiores e chuvas intensas, como a região serrana, e regiões com temperaturas mais elevadas próximas à orla (Silva, 2014). O estado possui grande variedade econômica e social, havendo municípios vizinhos com condições climáticas semelhantes, mas diferentes condições de manutenção de infraestrutura, saneamento e acesso a serviços de saúde. Todos os municípios do estado foram considerados para as análises estatísticas descritivas e espaciais - subcapítulo 4.3 da metodologia. O recorte temporal dos dados foi feito entre 1º de janeiro de 2014 e 31 de dezembro de 2022, totalizando 9 anos de dados.

Devido à maior complexidade do modelo mecanístico da doença, optou-se por reduzir a quantidade de municípios para comparação entre si. Esta escolha tomou como base as características sócio-geográficas, climáticas e da disponibilidade e qualidade de dados disponíveis para utilização em cada município. A necessidade de uma série temporal longa (9 anos) e com dados consistentes foi um limitante inicial para a escolha de municípios, já que a maior parte deles não atendia este pré-requisito. Desta forma, dentre os municípios com boa disponibilidade de dados, a seleção final foi realizada buscando capturar a heterogeneidade socioespacial e climática no estado. Os municípios selecionados para aplicação do modelo foram, portanto: Campos dos Goytacazes

(referenciado como “Campos” ao longo do texto), Duque de Caxias (referenciado apenas por “Caxias” no texto), Macaé, Paraty, Resende e Rio de Janeiro, a capital do estado. A escolha final de municípios abrangeu diferentes perfis de urbanização, densidade populacional e regimes climáticos, enriquecendo as análises do modelo.

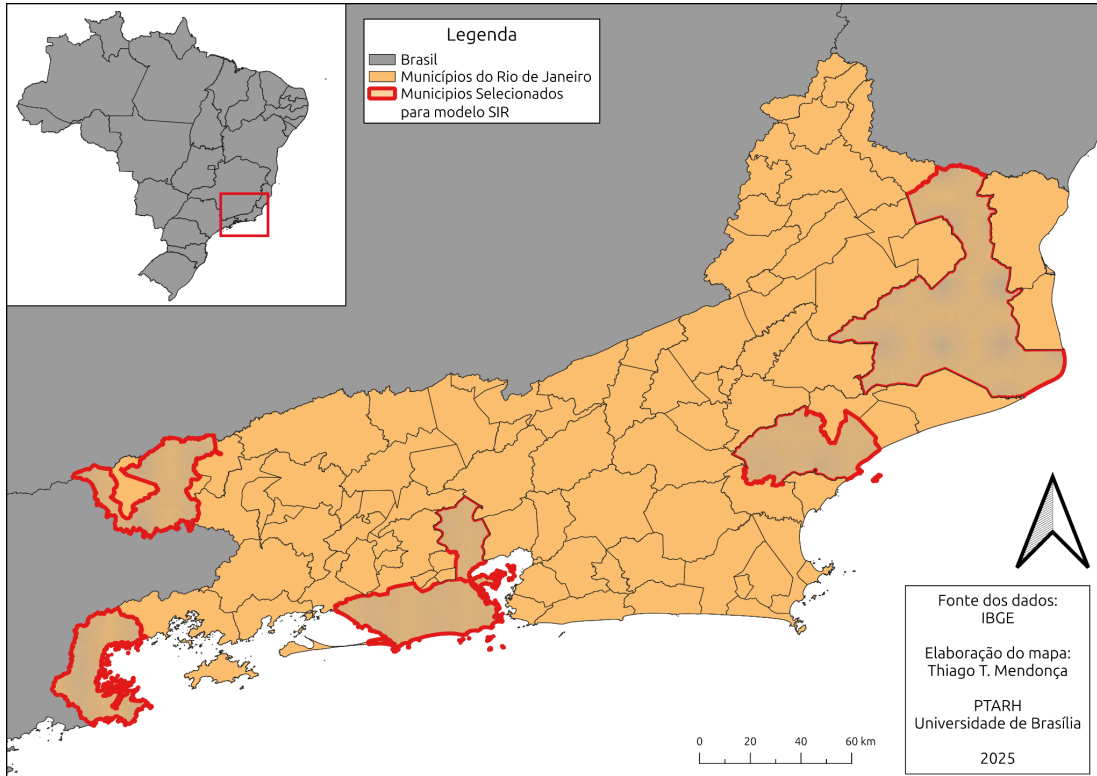


Figura 5.2: Mapa de localização e delimitação de municípios selecionados para o modelo SIR.

Os subcapítulos na sequência contém uma breve descrição dos municípios e apresentam informações climatológicas e sociais dos municípios selecionados para aplicação no modelo matemático. Os gráficos abaixo possuem dados de desigualdade econômica (IBGE, 2010), precipitação anual média e temperatura média para todos os municípios, entre os anos de 2014 e 2022. Os dados climáticos foram calculados utilizando medidas diárias de estações climatológicas mais próximas de cada município, disponibilizadas pelo Banco de Dados Meteorológicos para Ensino e Pesquisa (BDMEP) do Instituto Nacional de Meteorologia (INMET).

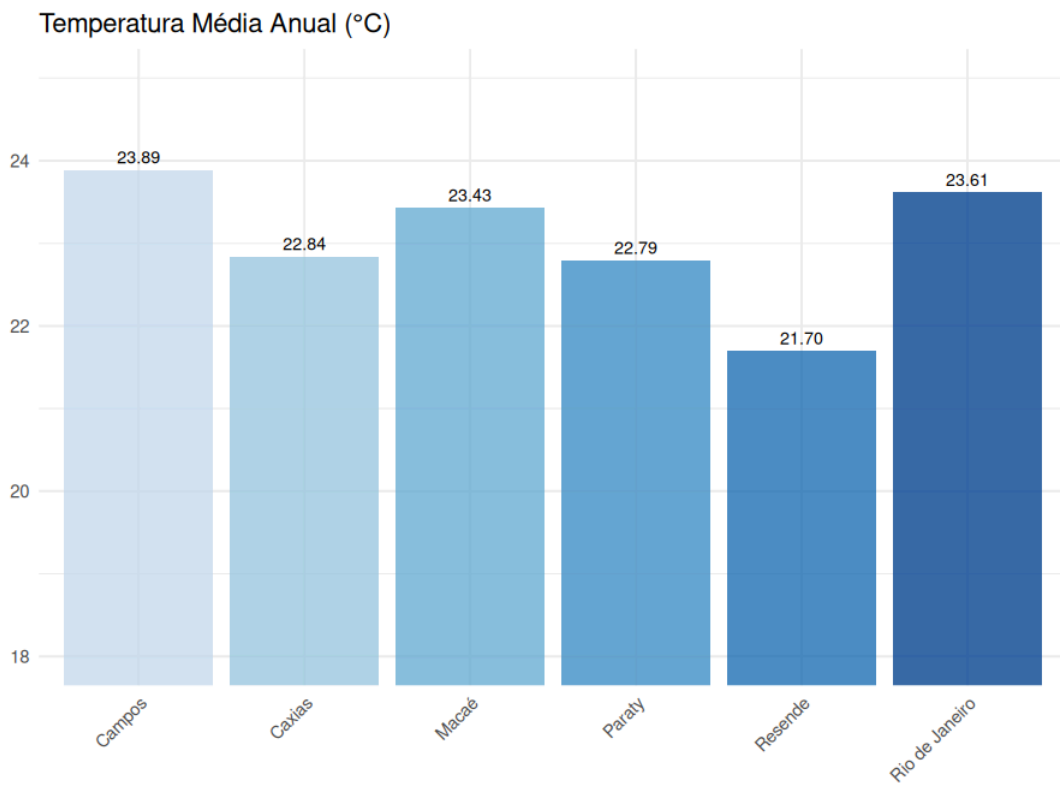


Figura 5.3: Temperatura média anual dos municípios selecionados para o modelo mecânico.

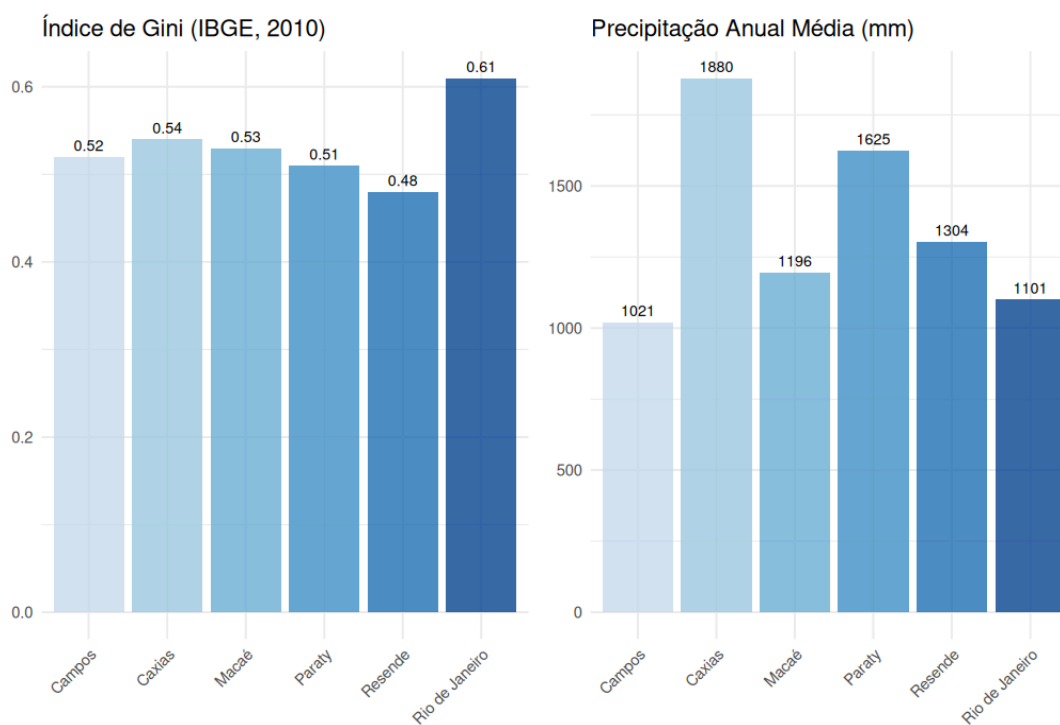


Figura 5.4: Índice de Gini (desigualdade de renda) e Precipitação anual média nas estações selecionadas para cada município para o modelo mecânico.

5.2.1 Campos dos Goytacazes

A cidade de Campos dos Goytacazes fica situada na Região Norte Fluminense do estado do Rio de Janeiro. Possui uma população de 483.551 habitantes (IBGE, 2022) e Produto Interno Bruto (PIB) de R\$ 37,2 bilhões (IBGE, 2021). No último censo que disponibilizou dados de Índice de Gini discriminados por municípios, em 2010, Campos dos Goytacazes apresentava um Índice de 0,52 (IBGE, 2010), o que indica uma situação com considerável desigualdade.

Entre os anos de 2014 e 2022, a temperatura média anual observada foi de aproximadamente 23,9 °C e precipitação anual média em torno de 1.020 mm. O período mais chuvoso é concentrado nos meses de verão (INMET, BDMEP)

5.2.2 Duque de Caxias

Duque de Caxias fica situada ao norte da região metropolitana da cidade do Rio de Janeiro, possuindo condições geográficas semelhantes às da capital. Apesar da proximidade, há uma grande disparidade social entre os dois municípios. Sua população é de 808.161 habitantes (IBGE, 2022) e PIB de R\$ 48,6 bilhões (IBGE, 2021). O índice de Gini era de 0,54 (IBGE, 2010), o que reflete também uma acentuada desigualdade social. Sua temperatura média anual é de 22,8 °C e precipitação anual mais elevada, na ordem de 1880 mm, com período chuvoso concentrado no verão.

5.2.3 Macaé

Macaé é uma cidade costeira reconhecida pelo seu grande polo da indústria petrolífera e, assim como Campos, também está situada ao norte do estado do Rio de Janeiro. Sua população é de 246.391 habitantes (IBGE, 2022) e PIB R\$ 33,7 bilhões (IBGE, 2021), impulsionado pelo setor de petróleo. Apesar do PIB per capita mais elevado em relação aos municípios anteriores, o Índice de Gini em 2010 era de apenas 0,53 (IBGE, 2010), o que indica condição de desigualdade. Em relação ao clima, o município possui temperatura média anual de aproximadamente 23,4 °C e precipitação anual média de 1200 mm, com épocas de chuvas concentradas no verão.

5.2.4 Paraty

Paraty é uma cidade histórica e costeira situada na Região da Costa Verde, ao sul do estado do Rio de Janeiro, fazendo divisa com o estado de São Paulo. É o município com menor população dentre os selecionados para aplicação no modelo: apenas 44.872

habitantes (IBGE, 2022). O PIB de Paraty é de R\$ 1,5 bilhão (IBGE, 2021) e o Índice de Gini de Paraty em 2010 era de 0,51 (IBGE, 2010). O clima é caracterizado por temperaturas médias anuais em torno de 22,8 °C e precipitação ultrapassando 1.600 mm, com chuvas bem distribuídas ao longo do ano, mas com picos durante o verão.

5.2.5 Resende

Localizada na Região do Vale do Paraíba Fluminense, no sul do estado, Resende possui um importante polo industrial e militar. Sua população é de 129.620 habitantes (IBGE, 2022) e PIB de R\$ 9,7 bilhões (IBGE, 2021). O Índice de Gini do município era de 0,48 (IBGE, 2010), o que indica menor desigualdade em comparação com os outros municípios selecionados, mas ainda em um nível expressivo. A precipitação anual média gira em torno de 1.300 mm e a temperatura média anual é de aproximadamente 21,7 °C. As chuvas concentram-se principalmente no período de verão, sendo o clima influenciado pela altitude e relevo do município.

5.2.6 Rio de Janeiro

O Rio de Janeiro é a capital do estado homônimo e a segunda maior metrópole do Brasil. Possui a maior população entre os municípios analisados, com 6.211.223 habitantes (IBGE, 2022), e área total de aproximadamente 1.200 km². Seu PIB é o segundo maior do país, totalizando R\$ 359,6 bilhões (IBGE, 2021). Seu Índice de Gini era 0,61 em 2010 (IBGE, 2010), o que indica uma elevada desigualdade de renda. Seus aspectos climatológicos indicam média anual de temperatura de 23,6 °C e precipitação de 1100 mm. O regime de chuvas ocorre ao longo de todo o ano, com predominância de pancadas de chuva durante o verão.

5.3 ANÁLISE EXPLORATÓRIA DA DENGUE NOS MUNICÍPIOS DO RIO DE JANEIRO

Após a obtenção e organização dos dados, a etapa seguinte da metodologia se refere à aplicação de diferentes análises estatísticas a fim de compreender as características e a distribuição da doença no estado do Rio de Janeiro. Os resultados obtidos são úteis para apoiar na delimitação das etapas seguintes.

Todos os dados das amostras analisadas são considerados discretos, isto é, há um valor finito de dados disponíveis e que podem ser enumerados. Desta forma, as análises estatísticas utilizadas neste subcapítulo são referentes a conjuntos de dados discretos. Os

subcapítulos a seguir compreendem o processamento inicial dos dados e a aplicação de modelos estatísticos espaciais para compreender as correlações existentes.

5.3.1 Contagem de meses epidêmicos

Com base no conceito de canal endêmico (Capítulo 3.2) e na equação 3.1, realizou-se uma comparação entre os casos registrados em cada mês, ano e município e os limites superiores do respectivo canal endêmico. Cada mês foi categorizado como “mês epidêmico” quando o número de casos observados ultrapassou o limite do canal, ou como “mês não epidêmico” quando permaneceu abaixo desse valor. A lógica aplicada segue o esquema apresentado a seguir:

Mês epidêmico = TRUE se número de casos > Limite superior do canal endêmico

Mês epidêmico = FALSE se número de casos ≤ Limite superior do canal endêmico

Este método é uma adaptação do trabalho de Drumond (2020). Os meses categorizados foram contabilizados e analisados através de estatísticas exploratórias. O próximo subcapítulo apresenta as análises espaciais e temporais elaboradas com esses dados.

5.3.2 Análises de distribuição espacial e temporal

O método de I de Moran global foi utilizado para a análise espacial da distribuição da incidência e da contagem de meses epidêmicos nos municípios. Este método fornece informações sobre a dependência espacial dos parâmetros estudados. Define-se um parâmetro como espacialmente dependente quando sua distribuição espacial não é aleatória (Anselin, 1995).

Um teste de hipóteses foi feito com a hipótese nula H_0 afirmando que a distribuição do parâmetro estudado é aleatória. A hipótese alternativa H_1 afirma que a distribuição espacial segue algum padrão. O valor de significância para rejeição das hipóteses foi delimitado em 1%, isto é, p-valores acima de 0,01 são considerados insuficientes para rejeitar a hipótese nula. Nas situações em que há algum padrão de distribuição, o valor de I de Moran é analisado conforme a seguinte lógica: valores de I mais próximos de 1 indicam agrupamento das variáveis, enquanto valores próximos de -1 indicam distanciamento. Os valores mais próximos a zero indicam aleatoriedade na distribuição (Gimond, 2023).

O cálculo de I de Moran foi feito com dois valores distintos de entrada. Na primeira aplicação, utilizou-se o número total de meses epidêmicos entre 2014 e 2022, por

município. Na segunda, utilizaram-se os dados de incidência por 100.000 habitantes em cada município, no mês de fevereiro, em cada um dos anos de estudo. Optou-se pelo mês de fevereiro já que este foi o mês com maior pico de epidemias ao longo dos anos.

Os cálculos foram facilitados pelo uso do pacote *spdep* (versão 1.3-13) na linguagem de programação *R*, em conjunto com o software *GeoDa* (versão 1.22.0.8) (Anselin; Li; Koschinsky, 2022). Para a análise de vizinhança e criação da matriz de pesos espaciais, necessária nos cálculos da estatística *I* de Moran, adotou-se o critério de contiguidade “rainha” (*Queen*) de primeira ordem. Neste método, dois municípios são considerados “vizinhos” quando compartilham qualquer extensão de fronteira, até mesmo um único ponto. Essa escolha foi feita buscando elaborar relações de vizinhança coerentes com a disseminação de doenças, que são complexas e muitas vezes não limitadas por fronteiras geográficas.

As variações temporais da epidemiologia da doença foram realizadas através de análises visuais simplificadas da distribuição espacial da incidência e dos picos de casos ao longo do tempo.

5.4 CORRELAÇÃO ENTRE DENGUE E INDICADORES DE SANEAMENTO

Após o reconhecimento da situação epidemiológica da dengue nos municípios do Rio de Janeiro, nesta etapa o enfoque é dado em analisar suas ligações com as variáveis de saneamento básico. Neste capítulo, são apresentadas as metodologias utilizadas para as análises de correlação entre a incidência de dengue e os indicadores de saneamento básico do SINISA e do IBGE. O objetivo desta análise é obter um panorama das relações existentes e identificar quais variáveis individuais apresentam correlações mais fortes com a dengue, antes de proceder com as técnicas de agrupamento e o modelo SIR.

5.4.1 Variável Resposta

Enquanto o enfoque do estudo é analisar a dengue, foi necessário especificar uma variável resposta para aplicar nas análises de correlação. A opção inicial de utilizar o total bruto de casos foi rejeitada por não permitir realizar comparações diretas entre municípios com populações diferentes. Assim, a variável resposta escolhida foi incidência por 100.000 habitantes para cada município. Mais especificamente, para obter um valor único que representasse a dengue em cada município, adotou-se a mediana das incidências anuais, por ser uma medida de tendência central menos sensível a valores extremos, que foram observados em alguns dos anos em que houve epidemias de dengue.

O cálculo da mediana da incidência anual iniciou-se agregando os dados brutos diários, obtidos do dataSUS, em totais anuais. Esse valor foi dividido pela população estimada do respectivo município naquele ano e multiplicado por 100.000. Obteve-se, assim, a incidência para cada município, para cada ano. Na sequência, calculou-se a mediana dessas taxas anuais para todo o período de estudo, em cada município, obtendo-se apenas um valor de incidência para cada município.

5.4.2 Análise de correlação

O cálculo de correlações foi feito utilizando coeficientes de correlação de Spearman, Pearson e de Kendall. A análise foi feita na linguagem de programação R, utilizando as funções *cor* para quantificar as correlações e o pacote *corrplot* (versão 0.95) para geração de figuras com matrizes de correlação. Para lidar com variáveis faltantes no cálculo de correlações, as abordagens *complete.obs* e *pairwise.complete.obs* foram testadas. Na primeira abordagem, municípios que possuem ao menos um dado faltante são excluídos do cálculo de correlações.

Portanto, devido às características do preenchimento de dados do SINISA, em que diversos municípios deixam de preencher dados, em especial os municípios menores, optou-se por utilizar a segunda abordagem (*pairwise.complete.obs*). Essa técnica consegue aproveitar ao máximo os dados disponíveis, uma vez que ela não descarta as variáveis que não estejam preenchidas em 100% dos municípios, calculando correlações com base nos dados restantes e disponíveis.

A escolha de analisar três diferentes coeficientes de correlação foi dada para possibilitar comparações entre os resultados e para aproveitar a disponibilidade dos métodos da função computacional. O coeficiente de Pearson (r) é um método amplamente aplicado na literatura, que mede a relação linear entre variáveis e é um modelo paramétrico. Os coeficientes de Spearman (ρ) e de Kendall (τ) são métodos de correlação não paramétrica utilizados para avaliar associações monotônicas, que podem ser lineares ou não lineares, entre variáveis.

Além dos cálculos dos coeficientes, também avaliou-se sua significância estatística. O mesmo princípio discutido no capítulo 3.3, para detecção de desvios estatisticamente raros, foi aplicado, considerando significativas as correlações cujos valores absolutos ultrapassaram o limiar de $\pm 2/\sqrt{n}$, equivalente a uma confiança de 95%.

O foco principal dessa análise foi obter a matriz de correlação entre dengue e saneamento. Os resultados obtidos apoiaram na tomada de decisões dos capítulos

seguintes, servindo como base para a escolha das variáveis no agrupamento e a interpretação das componentes principais calculadas.

5.5 SELEÇÃO E AGRUPAMENTO DE INDICADORES DE SANEAMENTO BÁSICO

O objetivo desta etapa é reduzir a dimensionalidade dos dados de saneamento, gerando novas variáveis que tenham uma representação mais geral das condições de saneamento, para serem então aplicadas no modelo mecanístico.

5.5.1 Pré-processamento e seleção de variáveis do SINISA

Como discutido no referencial teórico, o conjunto de dados do SINISA é muito valioso, mas também complexo. O banco de dados bruto possui uma grande quantidade de dados que possuem multicolinearidade inerente entre si, muitos deles com elevado número de dados faltantes no recorte espacial. Por isso, optou-se por realizar um processamento prévio dos dados antes do agrupamento. A abordagem adotada foi semelhante à aplicada por Borges (2024), que lidou com o mesmo banco de dados, com enfoque nos dados de águas pluviais.

A primeira etapa do pré-processamento consistiu em um filtro onde foram selecionadas apenas as variáveis (dos 3 eixos do SINISA) que apresentavam um preenchimento mínimo de 80% nos municípios do estado do Rio de Janeiro. Na sequência, realizou-se outra filtragem qualitativa: os indicadores com maior relevância foram escolhidos com base nos debates expostos no referencial teórico, nas referências bibliográficas e nos conhecimentos do próprio autor. Foram priorizadas as variáveis que contêm relações mais diretas na proliferação do vetor, retirando-se as variáveis que poderiam ter relações indiretas. Além disso, em situações de haver duas variáveis muito semelhantes (com potencial de multicolinearidade), optou-se por manter apenas uma delas. Um total de 41 indicadores foram filtrados durante este processo e estão detalhados no apêndice B deste trabalho.

5.5.2 Análise de Componentes Principais (PCA)

Após a redução na quantidade de dados obtida com o pré-processamento, a próxima etapa foi a aplicação de uma Análise de Componentes Principais (PCA). A PCA é um método da estatística multivariada utilizado para redução de dimensionalidade em conjuntos de dados. Seu princípio funciona a partir da combinação linear das variáveis originais em conjuntos menores, denominados Componentes Principais (PC, em inglês).

Esses PCs não possuem correlação entre si, o que garante essa redução de dimensionalidade, uma vez que os novos conjuntos de variáveis, em geral, conseguem explicar boa parte da variabilidade dos dados originais (Wilks, 2011 apud Borges, 2024).

O método da PCA foi aplicado no conjunto de dados do SINISA para os municípios do Rio de Janeiro. Isso foi feito com apoio do pacote *stats* e a função *prcomp*, na linguagem de computação *R*. A análise foi feita para cada um dos eixos do SINISA, buscando criar componentes que possuem boa capacidade de exprimir as características de cada área do saneamento. O método também foi aplicado em um conjunto com todas as variáveis, compreendendo os três eixos, para analisar a viabilidade dessa técnica em um banco de dados tão amplo e variado.

A etapa final da PCA foi a interpretação dos componentes. As cargas (pesos) de cada variável na composição de cada PC foram analisadas, buscando-se encontrar padrões em uma adaptação da abordagem aplicada por Borges (2024). A autora construiu quatro categorias de tipologias relacionadas à vulnerabilidade urbana a alagamentos, a partir das quais analisou as respectivas componentes com base nas tipologias com cargas mais altas que compunham cada PC. No presente estudo, a caracterização foi feita de baixo para cima, observando as variáveis individuais com maiores cargas para definir os temas de cada componente principal. Essa técnica permitiu que os próprios dados revelassem padrões existentes nos indicadores de saneamento, sem a imposição de categorias pré-definidas.

5.6 MODELO MECANÍSTICO DE DENGUE

Os conjuntos de equações diferenciais e de compartimentos que compõem o modelo matemático de dengue foram elaborados com base nos trabalhos levantados na revisão de literatura (capítulo 4.2). Nos seguintes subcapítulos, inicialmente são apresentados os dados de entrada, os parâmetros utilizados e o conjunto de equações que compõem o modelo final elaborado. No último subcapítulo há um debate sobre o processo de calibração do modelo, concluindo a metodologia do trabalho.

5.6.1 Condições iniciais do modelo

Para o cálculo do modelo mecanístico, dois conjuntos de dados de entrada são necessários: condições iniciais e parâmetros do modelo. O conjunto de condições iniciais corresponde ao número de indivíduos em cada compartimento, no início da simulação (Sh , Sv , Rh , etc. em $t = 0$).

Foram utilizadas um total de 10 condições iniciais para cada município. Para os compartimentos de hospedeiros, delimitou-se a população total (N_h) como a população de cada município, no ano de 2022. Quanto à proporção da população total suscetível à doença (α_{Sh}), os valores adotados na literatura variam entre 4% e 100% (ver quadro 4.1). Este valor foi calculado no processo de calibração do modelo, com valor inicial de 60%, que é um valor coerente com o utilizado por Nuraini (56,9%). O valor inicial para o compartimento de hospedeiros infectados foi definido como o número de casos de dengue no dia 01/01/2014. Por fim, a quantidade de indivíduos recuperados foi calculada subtraindo o total de indivíduos suscetíveis e infectados da população total do município.

Quadro 5.4: Condições iniciais utilizadas no modelo mecanístico e equações utilizadas para seus cálculos

Símbolo	Descrição compartimentos	Unidade	Valor inicial
Sh	População de hospedeiros suscetíveis	Humano	$\alpha_{Sh} * \text{população total}$
Ih	População de hospedeiros infectados	Humano	<i>Número de casos de dengue em 01/01/2014. Quando = 0, adotou-se 1</i>
Rh	População de hospedeiros recuperados	Humano	$N_h - (I_h + S_h)$
Nh	População total de hospedeiros	Humano	<i>População do município</i>
Sv	População de vetores suscetíveis	Mosquito	$K * (S_h + R_h)$
Iv	População de vetores infectados	Mosquito	$K * I_h$
Nv	População total de vetores	Mosquito	$S_v + I_v$
Se	População de ovos suscetíveis	Ovo	$K * (S_h + R_h)$
Ie	População de ovos infectados	Ovo	$K * I_h$
Ne	População total de ovos	Ovo	$S_e + I_e$

Para a população de vetores, as condições iniciais foram calculadas com base no parâmetro de capacidade de carga (K), que indica a população máxima que um meio consegue sustentar, proporcionalmente à quantidade de humanos (unidade: mosquitos/humano). O valor deste parâmetro foi estimado matematicamente no processo de calibração (descrito no capítulo 5.6.5). Sendo assim, a população de vetores suscetíveis foi calculada multiplicando a soma da população de humanos suscetíveis e recuperados por K. Da mesma forma, a população de vetores infectados foi calculada multiplicando a população de hospedeiros inicialmente infectados por K.

Por fim, os compartimentos populacionais dos ovos também foram calculados usando a capacidade de carga. Os cálculos utilizaram equações idênticas às dos vetores.

Sendo assim, as populações iniciais de ovos foram consideradas idênticas às dos ovos, para os respectivos compartimentos.

5.6.2 Parâmetros utilizados do modelo

Os parâmetros do modelo são os valores que multiplicam as equações e governam as dinâmicas biológicas e de transmissão. No contexto do trabalho, estão divididos em duas categorias: parâmetros calibrados e parâmetros não-calibrados.

Os parâmetros calibrados são aqueles cujos valores são incertos ou altamente dependentes de condições locais de cada município. Seus valores finais são estimados através de um processo de otimização (descrito na seção 5.6.5), em um ajuste do modelo aos dados observados. Neste grupo, estão a capacidade de carga (K), a proporção de humanos suscetíveis à doença (α_{Sh}) e mais três parâmetros potencialmente modulados pelas condições de saneamento e que descrevem a biologia do vetor em função da temperatura (equações 5.1 e 5.2).

Os parâmetros não-calibrados são aqueles cujos valores foram pré-definidos com base na literatura e não foram ajustados durante o processo de calibração. Dentre eles, há parâmetros com valores que se mantêm fixos durante toda a simulação, independente das condições ambientais às quais estão expostos, como taxa de mortalidade de humanos (μ_h) e taxa de recuperação da dengue (γ) e parâmetros calculados: a taxa de picada (b) e a taxa de oviposição (or), relacionados respectivamente à quantidade de picadas e de ovos colocados por mosquitos, por dia. Os valores para esses parâmetros foram definidos com base na literatura analisada, com destaque especial ao quadro 4.1.

Conforme exposto no capítulo 4.3.2, na literatura há uma gama de maneiras para cálculo dos parâmetros. Diferentes equações foram testadas buscando simplicidade, velocidade de cálculo e obtenção de resultados coerentes com a realidade. A seleção final de equações aplicadas para o cálculo de parâmetros do modelo foi a seguinte:

Equação de Brière (Mordecai *et al.* 2017):

$$parâmetro = c * T * (T - T_0) * \sqrt{T_m - T} \quad (5.1)$$

Equação Quadrática (Mordecai *et al.* 2017):

$$parâmetro = c * (T - T_0) * (T - T_m) \quad (5.2)$$

Equação linear para cálculo de b (Nuraini, 2017)

$$b = 0,0043 * T + 0,0943 \quad (5.3)$$

Onde:

T é a temperatura medida em cada tempo;

T_0 é a temperatura mínima para transmissão;

T_m é a temperatura máxima para transmissão e

c é uma taxa constante, calculada originalmente por Mordecai *et al.* Este valor foi mantido fixo para o cálculo de or e alterado computacionalmente para as demais variáveis durante a calibração do modelo (capítulo 4.6.5).

O quadro a seguir sintetiza os parâmetros utilizados. Os valores fixos e os valores iniciais para os parâmetros calibráveis foram adotados da literatura (quadro 4.1), servindo como ponto de partida para o processo de calibração (capítulo 5.6.5).

Quadro 5.5: Valores e cálculos adotados para as variáveis utilizadas nas equações do modelo SIR-SI-SI

Categoria	Símbolo	Descrição	Valor	Unidade
Parâmetros não calibrados	μ_e	Mortalidade de ovos	0.004	1/dia
	μ_h	Mortalidade de humanos	$1/(365*65)$	1/dia
	γ	Taxa de recuperação do hospedeiro	1/8	1/dia
	ν_t	Taxa de transmissão vertical	0,1	--
	P_v	Probabilidade de transmissão de hospedeiro para vetor	0,75	--
	P_h	Probabilidade de transmissão de vetor para hospedeiro	0,75	--
	b	Taxa de picada	$0,0043 * T + 0,0943$	1/dia
	or	Taxa de oviposição	$cT * (T - 14,58) * (34,61 - T)^{(1/2)}$	1/dia
Parâmetros calibrados	α_{Sh}	Proporção da população de hospedeiros suscetíveis	0,6	--
	K	Capacidade de carga	0,6	%

Categoria	Símbolo	Descrição	Valor	Unidade
	μ_v	Mortalidade de mosquitos	$c * (T - 9,16) * (T - 37,73)$	1/dia
	dr	Taxa de desenvolvimento	$c * (T - 11,36) * (T - 39,17)$	1/dia
	sa	Taxa de sobrevivência aquática	$c * (T - 13,56) * (T - 38,29)$	1/dia

5.6.3 Equações do modelo final *SIR-SI-SI*

Para compreender explicitamente os efeitos das condições de saneamento na disseminação da doença, optou-se por incluir etapas de desenvolvimento dos mosquitos, que são mais sensíveis a tal. Dessa forma, o modelo final utiliza os compartimentos “*SIR*” para humanos, “*SI*” para mosquitos adultos e “*SI*” para mosquitos na fase de ovos, sendo, portanto, denominado um modelo *SIR-SI-SI*. O esquema simplificado abaixo apresenta os compartimentos utilizados. As setas vermelhas indicam relações existentes entre os compartimentos. Uma versão mais detalhada é apresentada no apêndice do trabalho.

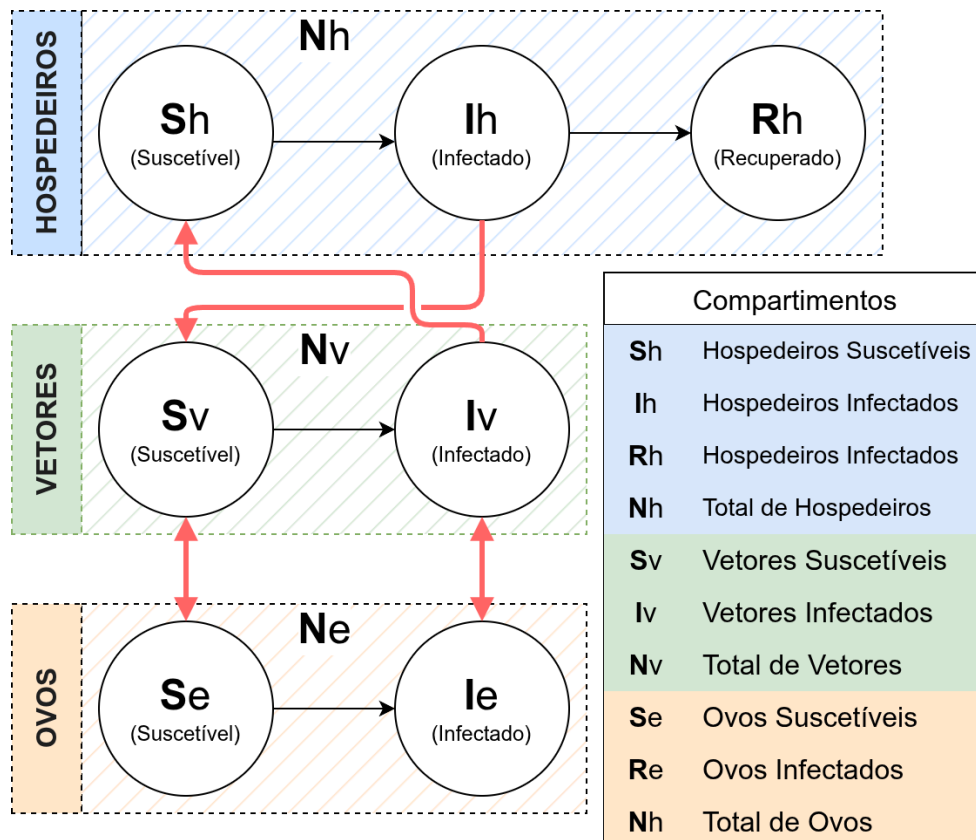


Figura 5.5: Compartimentos utilizados no modelo SIR-SI-SI.

O modelo mecânico é composto por uma série de Equações Diferenciais (EDs), uma para cada compartimento, que contém as diferentes taxas e parâmetros que descrevem as variações na quantidade de indivíduos em cada compartimento ao longo do tempo (dx/dt). Cada uma das taxas e parâmetros utilizados foram obtidos com base na literatura e estão descritos na sequência do texto. Este modelo utiliza uma abordagem determinística, isto é, não incorpora o efeito da aleatoriedade na disseminação da doença. A resolução do modelo foi feita com a linguagem de programação *R*, com apoio do pacote *deSolve*, um pacote especializado em resolução de conjuntos de equações diferenciais. As equações utilizadas no modelo são as seguintes:

Equações relacionadas aos compartimentos de hospedeiros (humanos)

$$\frac{dSh}{dt} = \mu h * (Nh - Sh) - \frac{b*Ph*Iv*Sh}{Nh} \quad (5.4)$$

$$\frac{dIh}{dt} = \frac{b*Ph*Iv*Sh}{Nh} - (\gamma + \mu h) * Ih \quad (5.5)$$

$$\frac{dRh}{dt} = \gamma * Ih - \mu h * Rh \quad (5.6)$$

Equações relacionadas aos compartimentos de vetores (mosquitos) na fase adulta:

$$\begin{aligned} \frac{dSv}{dt} &= dr * 0.5 * sa * Se * (1 - \frac{Nv}{K*Nh}) \\ &- (b * Pv * \frac{Ih}{Nh} + \mu v) * Sv \end{aligned} \quad (5.7)$$

$$\begin{aligned} \frac{dIv}{dt} &= dr * 0.5 * sa * Ie * (1 - \frac{Nv}{K*Nh}) + (b * Pv * \frac{Ih}{Nh}) * Sv \\ &- (\mu v * Iv) \end{aligned} \quad (5.8)$$

Equações relacionadas aos compartimentos de vetores durante a fase de ovos:

$$\frac{dSe}{dt} = or * Sv + (1 - vt) * or * Iv - (\mu e + dr) * Se \quad (5.9)$$

$$\frac{dIe}{dt} = vt * or * Iv - (\mu e + dr) * Ie \quad (5.10)$$

Todas as variáveis utilizadas nas equações foram apresentadas com mais detalhes nos dois capítulos anteriores e os valores fixos e equações adotados para elas estão no quadro 5.1. Os compartimentos referem-se às populações de humanos (hóspedes, compartimentos “h”),

de mosquitos (vetores, compartimentos “v”) e de ovos (compartimentos “e”), conforme a figura 5.5. O quadro a seguir apresenta os nomes de cada variável utilizada no modelo.

Quadro 5.6: Descrição das variáveis utilizadas nas equações do modelo SIR-SI-SI

Símbolo	Descrição
α_{Sh}	Proporção da população de hospedeiros suscetíveis
μ_e	Mortalidade de ovos
μ_h	Mortalidade de humanos
γ	Taxa de recuperação do hospedeiro
ν_t	Taxa de transmissão vertical
P_v	Probabilidade de transmissão de hospedeiro para vetor
P_h	Probabilidade de transmissão de vetor para hospedeiro
μ_v	Mortalidade de mosquitos
d_r	Taxa de desenvolvimento
s_a	Taxa de sobrevivência aquática
b	Taxa de picada
o_r	Taxa de oviposição
K	Capacidade de carga

5.6.4 Análise exploratória manual dos parâmetros

Antes da calibração computacional, uma análise exploratória dos dados foi aplicada buscando compreender qualitativamente os impactos de alterações nos parâmetros, nas dinâmicas do modelo. Alguns dos parâmetros foram alterados individualmente e seus efeitos nas curvas de casos foram observados e registrados em um quadro. Esta análise foi importante para apoiar na decisão de equações a serem adotadas para cálculo de cada parâmetro, inclusive delimitando que alguns dos parâmetros adotassem valores fixos.

Quadro 5.7: Observações do processo de análise exploratória manual dos parâmetros

Parâmetro	Observações
Taxa de transmissão (beta_h)	Reduzir beta_h desacelera a disseminação, diminui o pico de casos e prolonga a epidemia.
Taxa de recuperação (gamma)	Aumentar gamma diminui os picos. Gamma maior significa que uma pessoa fica infectada se cura mais rápido, há menos chance de

	disseminar a doença para mosquitos suscetíveis
Probabilidade de infecção vetor para humano (Pv)	Quanto maior Pv, maior e mais rápido o pico. Picada de mosquitos mais infecciosos é mais provável de transmitir a doença
Probabilidade de infecção humano para vetor (Ph)	Quanto maior Ph, maior e mais rápido o pico (igual a Pv). Quanto mais fácil transmitir, mais mosquitos se tornam infectados
Taxa de mortalidade do vetor (mu_v)	Com maior mortalidade os picos são menores. Diminui a população de vetores e o tempo útil que ele tem para transmitir a doença
Taxa de mortalidade e natalidade dos humanos (mu_h)	Quase não há diferença. O efeito é negligenciável na escala de tempo de um surto epidêmico.
Taxa de picada (b)	Aumentar b acelera a transmissão, resultando em um pico mais alto, mais rápido. Mosquitos que picam mais, transmitem mais

5.6.5 Processo de calibração do modelo

A calibração de um modelo corresponde à etapa de estimativa dos seus parâmetros. Para cada um dos seis municípios do recorte espacial, os parâmetros foram estimados através de um processo de otimização computacional, onde o objetivo foi minimizar a diferença entre os casos de dengue simulados em comparação aos reais.

A função objetivo utilizada para medir o ajuste do modelo foi a Soma dos Quadrados dos Erros (SQE). Este método calcula a diferença entre os valores simulados e observados em cada ponto no tempo, sendo uma métrica comumente utilizada em problemas de ajustes de curvas devido à sua eficiência e simplicidade. A SQE foi calculada com a equação 5.11:

$$SQE = \sum_{i=1}^n (y_{calculado}(i) - y_{observado}(i))^2 \quad (5.11)$$

Onde:

n é o número total de observações no período;

$y_{calculado}(i)$ é o número de casos simulado para o período i

$y_{observado}(i)$ é o número real de casos ocorridos no período i

O processo de minimização da SQE foi feito na linguagem de programação R , utilizando o pacote *DEoptim*, que implementa um algoritmo de otimização global por evolução diferencial (Mullen *et al.*, 2011). Optou-se por utilizar um método de otimização global, em contraponto a otimizadores mais simples, devido à alta complexidade do modelo, que possui 13 variáveis e múltiplas relações não lineares. Este otimizador é

robusto para aplicação em espaços de busca complexos, possuindo maior capacidade de evitar mínimos locais.

Durante os primeiros testes, a calibração de todos os 13 parâmetros, em todo o período, se demonstrou inviável computacionalmente. Diante disso, três estratégias foram instituídas para garantir a eficiência do processo. A primeira consistiu em reduzir a quantidade total de parâmetros a serem otimizados. Apenas os cinco considerados mais sensíveis às condições de saneamento foram mantidos, mantendo os oito restantes fixados, conforme exposto no quadro 5.5. Os parâmetros dependentes de temperatura foram calibrados alterando apenas os valores da constante de proporcionalidade c , enquanto os parâmetros K e α_{Sh} foram calibrados alterando diretamente seus valores.

A segunda estratégia adotada foi a agregação temporal dos dados. Tanto os dados reais quanto os simulados foram consolidados em totais semanais antes do cálculo da SQE. Com isso, além de reduzir o tempo de processamento, essa abordagem também atenuou instabilidades na calibração causadas pela presença de dias com zero casos nos municípios com menores populações. Para explorar o balanço entre resolução temporal de dados e a capacidade de ajuste do modelo, a calibração também foi aplicada utilizando uma agregação mensal. A metodologia aplicada foi semelhante, mas essencialmente diferente à aplicada por Alves (2021): o autor, ao invés de calcular dados diários e agregar em períodos maiores, transformou os parâmetros de seu modelo para uma escala de 8 dias.

A última estratégia foi a divisão da série temporal entre os períodos de treinamento e de validação. Conforme discutido no 4.3.2, a escolha dos respectivos tempos depende do caráter preditivo ou inferencial objetivado no trabalho, tendo influência direta no tempo necessário para resolver o modelo. Optou-se por realizar o treinamento durante os 4 primeiros anos de dados (2014–2017), validando-o no período restante de 5 anos (2018–2022). Essa escolha teve como objetivo reduzir o custo computacional, explorar a capacidade do modelo para captar os mecanismos de transmissão da doença e projetá-lo em um período futuro, diferente daquele utilizado para estimação dos parâmetros.

6. RESULTADOS E DISCUSSÃO

A apresentação de resultados e sua discussão foram formuladas seguindo a mesma ordem lógica utilizada na metodologia, de forma que o conhecimento é construído e agregado conforme evolução dos subcapítulos. Em primeiro momento, é apresentada uma visão geral da epidemiologia de dengue no estado do Rio de Janeiro. Na sequência são apresentadas comparações realizadas entre dados da dengue e de saneamento básico. O resultado da redução de dimensionalidade dos dados de saneamento através da PCA é apresentado e então procede-se à última etapa, onde é aplicado o modelo mecanístico.

6.1 ANÁLISE EXPLORATÓRIA

Os municípios no estado do Rio de Janeiro passaram por múltiplas epidemias de dengue ao longo da série temporal. A figura abaixo apresenta uma visualização da casuística de dengue ao longo do período total de nove anos nos seis municípios selecionados.

É evidente como o município do Rio de Janeiro, o maior do estado, possui um número de casos expressivamente superior em comparação aos outros. Essa grande disparidade dos picos de casos ocorre devido às diferenças populacionais. Em geral, os maiores municípios possuem populações suscetíveis à doença em maior quantidade (figura 6.1).

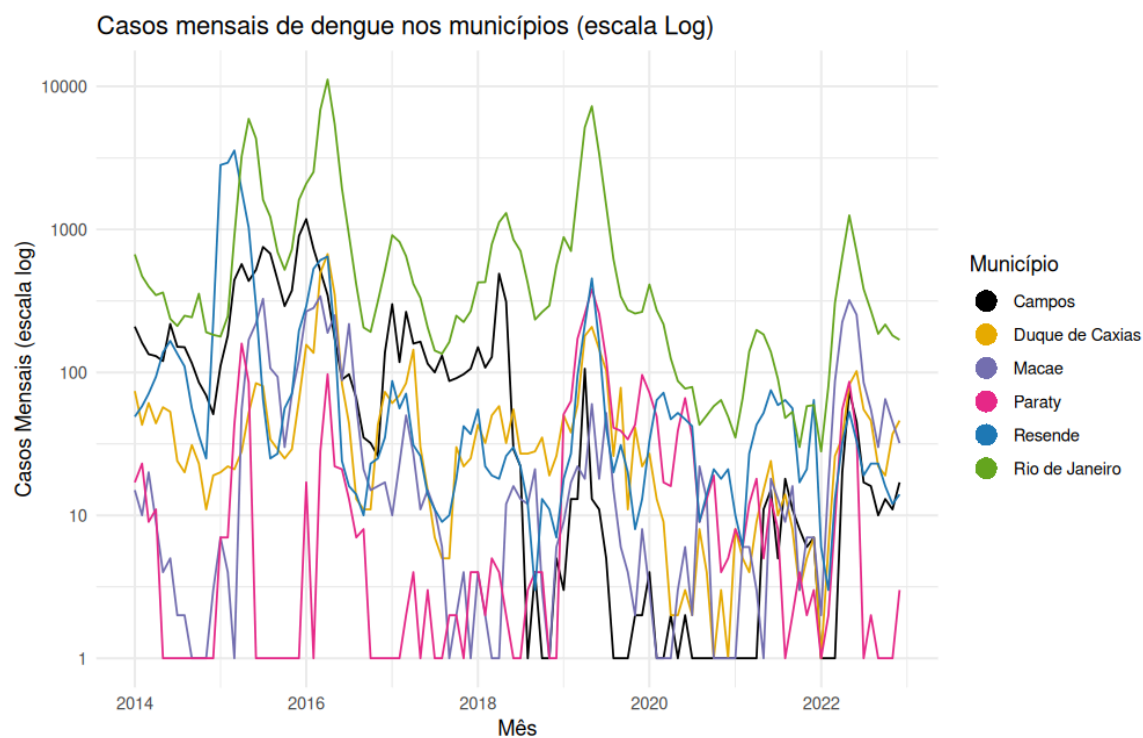
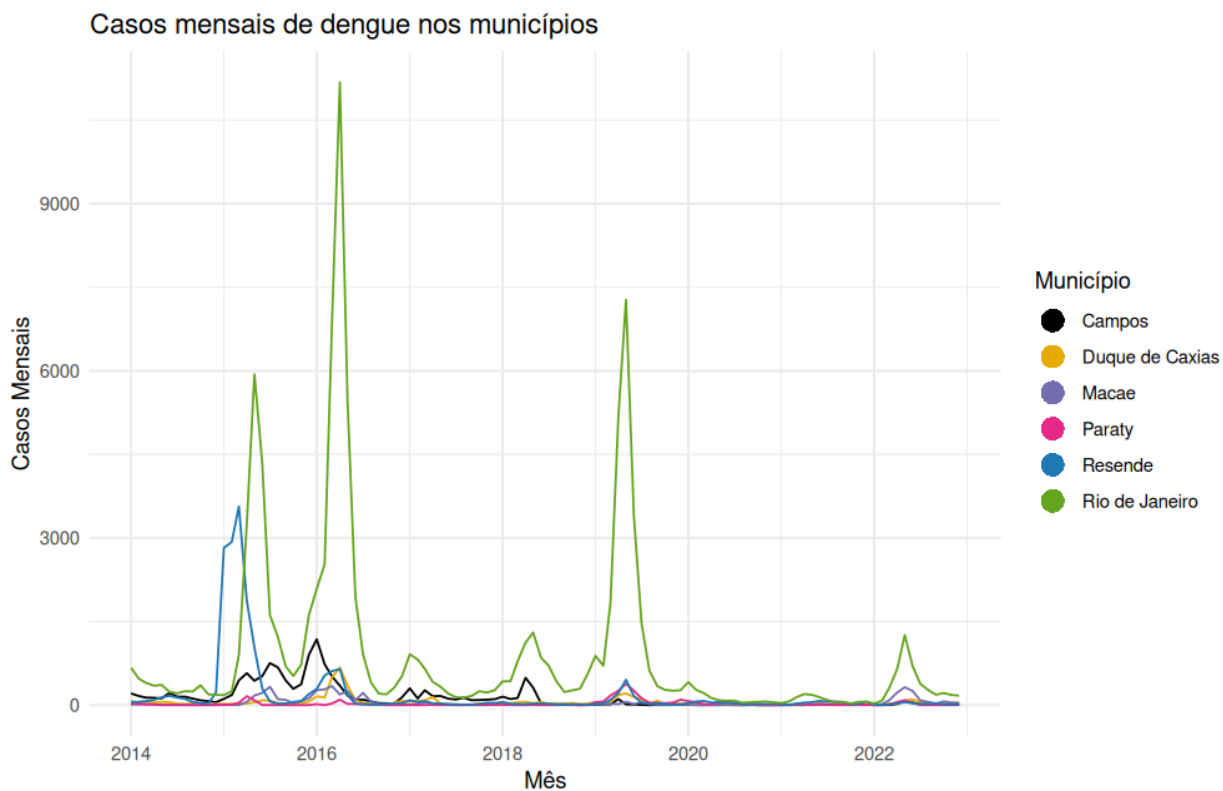


Figura 6.1: Casos mensais de dengue em seis municípios do estado do Rio de Janeiro ao longo da série temporal. Painel A: Escala linear. Painel B: Escala Logarítmica.

Uma análise de casos brutos, porém, não possibilita a realização de comparações diretas, uma vez que as populações dos municípios são variadas. A incidência é uma medida mais competente para casos como esse, conforme apresentado na figura 6.2.

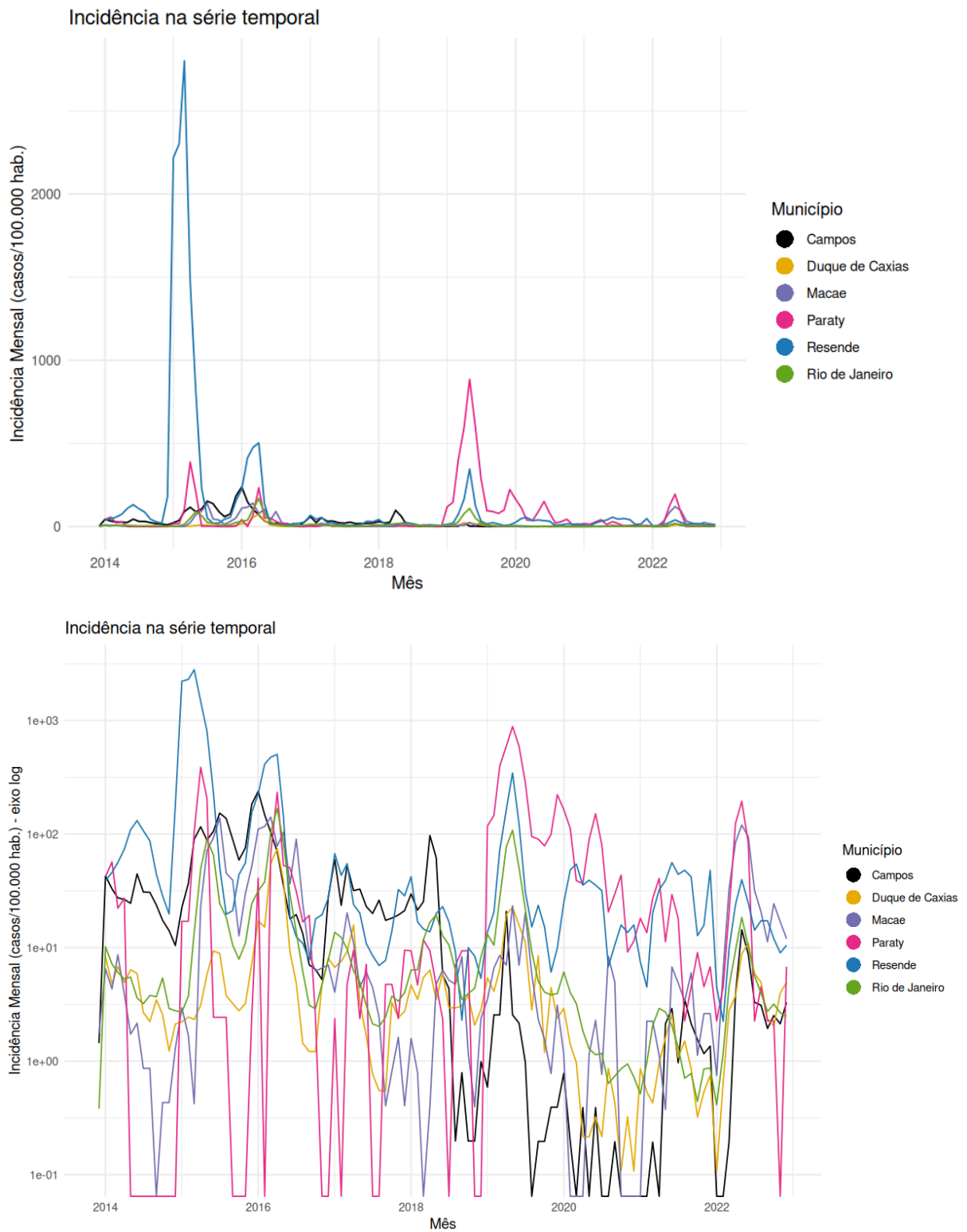


Figura 6.2: Incidência mensal em seis municípios do estado do Rio de Janeiro ao longo da série temporal. Painel A: escala linear. Painel B: escala logarítmica.

Esse novo gráfico mostra uma nova realidade: a cidade do Rio de Janeiro, que na figura 6.1 tem os maiores picos de casos, quando comparada aos outros municípios sob a ótica da incidência, possui uma epidemiologia da dengue mais controlada. Os dois municípios com maiores picos proporcionais são Resende e Paraty, que possuem as menores populações dentre os seis observados. Isso os torna mais sensíveis a flutuações nos valores totais de casos, que podem ser causadas por diversos motivos.

Um exemplo dessa flutuação pode ocorrer devido a movimentos pendulares de entrada e saída de pessoas infectadas, que ocorrem em cidades turísticas, como Paraty, ou por fluxos de ida e vinda de trabalhadores. Se, por um lado, a chegada de pessoas infectadas de outras regiões pode incitar surtos epidêmicos nas populações locais suscetíveis, o oposto também se aplica. Um turista, trabalhador ou outro indivíduo pode se infectar na cidade e levar a doença para sua cidade de origem (Alves, 2021; de Sousa *et al.*, 2021).

Esse fator pode ser relevante em especial quando ocorrem tendências de agrupamento espacial de surtos da doença entre diferentes municípios, o que é mais explorado no subcapítulo seguinte. A magnitude dos efeitos da migração de pessoas foi parametrizada por Alves (2021) em seu estudo.

Outra visualização dos dados é através da média mensal da incidência, o que permite verificar a sazonalidade da doença (Fig. 6.3).

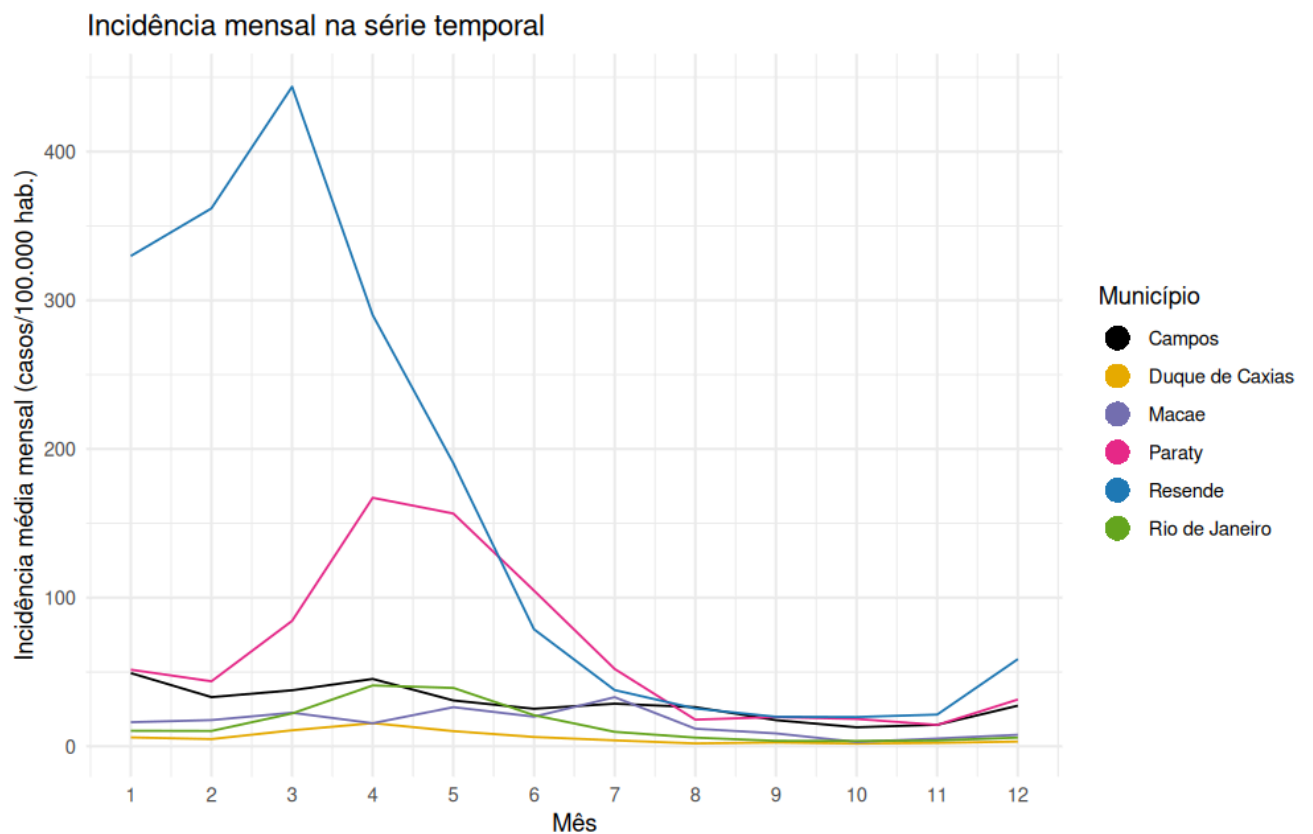


Figura 6.3: Incidência média mensal (2014–2022) em seis municípios do estado do Rio de Janeiro.

Como explicitado no referencial teórico, a dengue é uma doença que possui picos de casos nos meses mais quentes e chuvosos. A figura anterior mostra os efeitos da sazonalidade na casuística da doença, com aumento de casos durante os meses de verão e redução durante o inverno. Outro ponto interessante de se observar nessa mesma figura é a comparação entre incidências nos municípios. Mais uma vez Resende e Paraty se apresentam como os municípios com as curvas mais acentuadas. Além das justificativas anteriores, isso pode ter ocorrido também devido à grande epidemia de 2015, que promoveu um aumento significativo nos casos totais neste ano, aumentando a média de casos, e conseqüentemente da incidência, com efeitos mais expressivos nesses dois municípios menos populosos.

Ainda, Resende é o único município que possui pico de casos em março, e não em abril, como o restante. Isso pode ocorrer devido a diferentes dinâmicas existentes no local. Uma justificativa se dá devido à característica fortemente industrial do município, que atrai um grande volume de trabalhadores de outras cidades (movimento pendular), implicando em uma constante introdução de casos. O efeito da migração em Resende e em Paraty podem tornar ambos os município mais suscetíveis ao crescimento de casos da doença,

quando há condições climáticas favoráveis, do que os outros municípios menos afetados pelo fenômeno da migração.

A figura 6.4, a seguir, reitera o aspecto da sazonalidade e demonstra a presença de efeitos tardios (*time lags*) das variáveis climáticas sobre a incidência da dengue: o pico na curva de incidência ocorre não durante os meses de verão (quando é mais quente e chuvoso), mas sim nas semanas e meses seguintes. Esta observação é consistente com a literatura, como por exemplo nos achados de Nuraini *et al.* (2021). Os autores identificaram correlações entre precipitação e umidade na incidência de dengue com atrasos na ordem de 18 a 50 dias.

A justificativa para o atraso observado pode ser explicada através da mecânica do ciclo de vida do vetor e da disseminação do vírus. Aumentos na chuva e temperatura desencadeiam uma cascata de eventos biológicos que levam tempo para surtir efeitos (Nuraini *et al.* 2021). Exemplificando-se, os tempos necessários para criação de novos criadouros, eclosão de ovos dormentes, desenvolvimento de larvas e pupas, incubação do vírus, dentre outros, podem somar, juntos, um período de semanas até a ascensão dos casos.

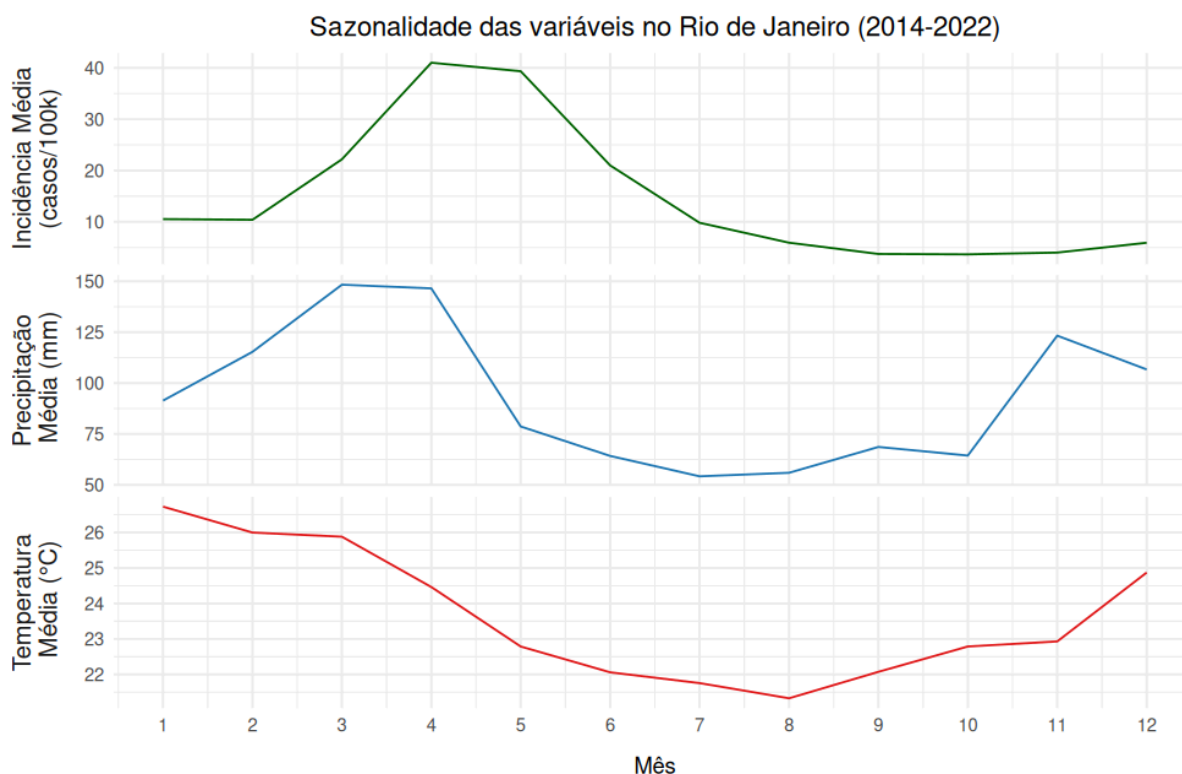


Figura 6.4: Comparação entre incidência, temperatura e precipitação médias mensais na cidade do Rio de Janeiro.

6.1.1 Epidemias no estado do Rio de Janeiro

A seguir, apresentam-se análises realizadas utilizando dados de meses epidêmicos que ocorreram em todos os municípios do estado. Ao longo dos 9 anos analisados no estudo, entre os 92 municípios do Rio de Janeiro, um total de 793 meses foram considerados “epidêmicos”. Todos os municípios analisados passaram por ao menos uma epidemia e a maior quantidade de meses epidêmicos observados em um município foi igual a 12. Apenas 29 municípios foram responsáveis por 40% do total de meses epidêmicos observados.

Os meses de verão se destacaram numa comparação entre os meses com maior frequência de epidemias ao longo de todo o período. Este é um resultado relevante, uma vez que esses são os meses com as maiores médias históricas de casos, o que torna seus limites do canal endêmico mais elevados. O fato de que as anomalias estatísticas ocorrem justamente nesses meses indica que as condições disponíveis durante os meses de primavera e verão (com temperaturas e umidade maiores, alterações em comportamentos da população, dentre outros), criam ambientes ideais para surtos epidêmicos com crescimentos significativos. Isso corrobora o que vem sendo debatido até aqui.

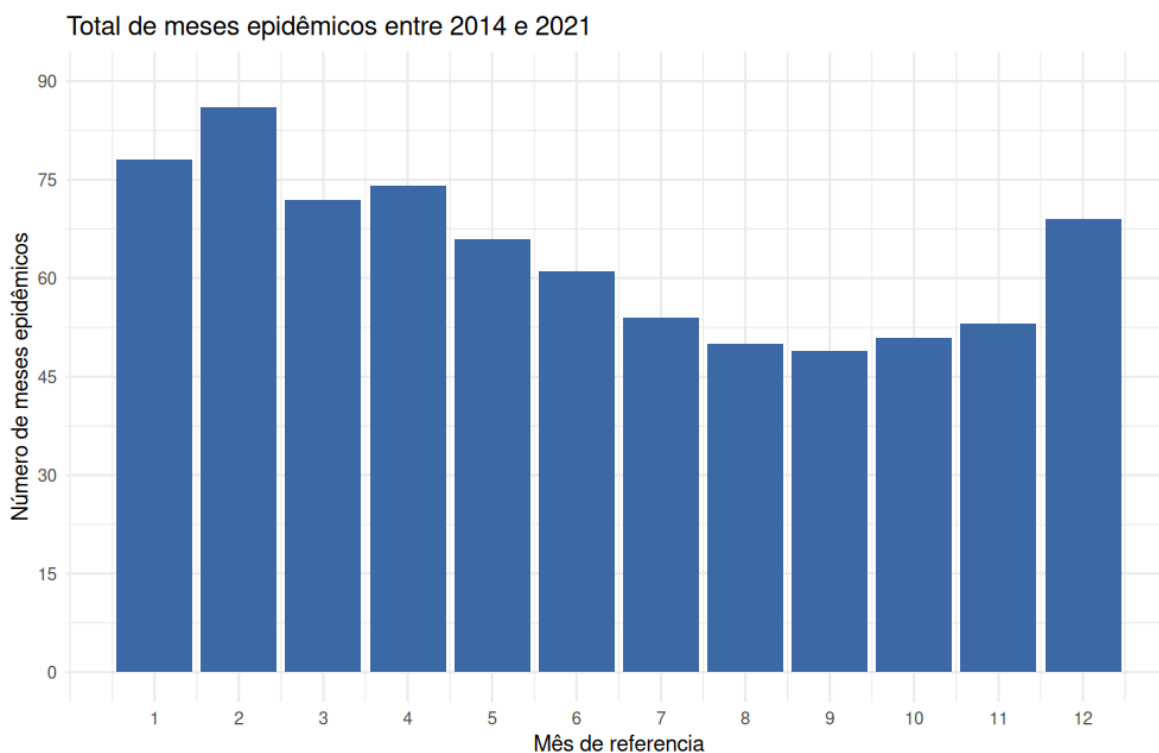


Figura 6.5: Quantidades de epidemias observadas em cada mês do ano, nos municípios do estado do Rio de Janeiro.

Para compreender a distribuição espacial da dengue ao longo do tempo, foram realizadas análises com o método I de Moran. A hipótese nula (H_0) testada afirma que a distribuição da epidemiologia da doença é espacialmente aleatória nos municípios analisados. Quando a hipótese nula é rejeitada, indica-se que pode haver autocorrelação espacial, manifestada de duas maneiras distintas: autocorrelação espacial positiva ou negativa. Uma autocorrelação espacial positiva (denominada por “agrupamento” ou “*cluster*” espacial) ocorre quando municípios com determinadas condições epidêmicas tendem a estar próximos de outros municípios com condições semelhantes. Um exemplo de agrupamento ocorre quando altas taxas de incidência ocorrem em municípios vizinhos em determinado ano.

A autocorrelação negativa (denominada por “dispersão” espacial) descreve um padrão em que municípios com condições epidêmicas se aproximam de outros municípios com condições distintas. Um exemplo é quando um município com alta incidência está próximo de um outro com incidência mais baixa. O método foi aplicado sobre duas perspectivas: para os dados de meses epidêmicos e para a incidência média em cada mês.

A análise espacial do total de meses epidêmicos por município não revelou um padrão espacialmente dependente: o resultado (p -valor = 0,1403 e estatística $I = 0,064$) não possibilita a rejeição da hipótese nula, considerando um nível de significância de 0,01. Isso sugere que as epidemias em todo o período não foram agrupadas em municípios específicos, mas sim que houve variação na localização dos surtos ao longo do período analisado.

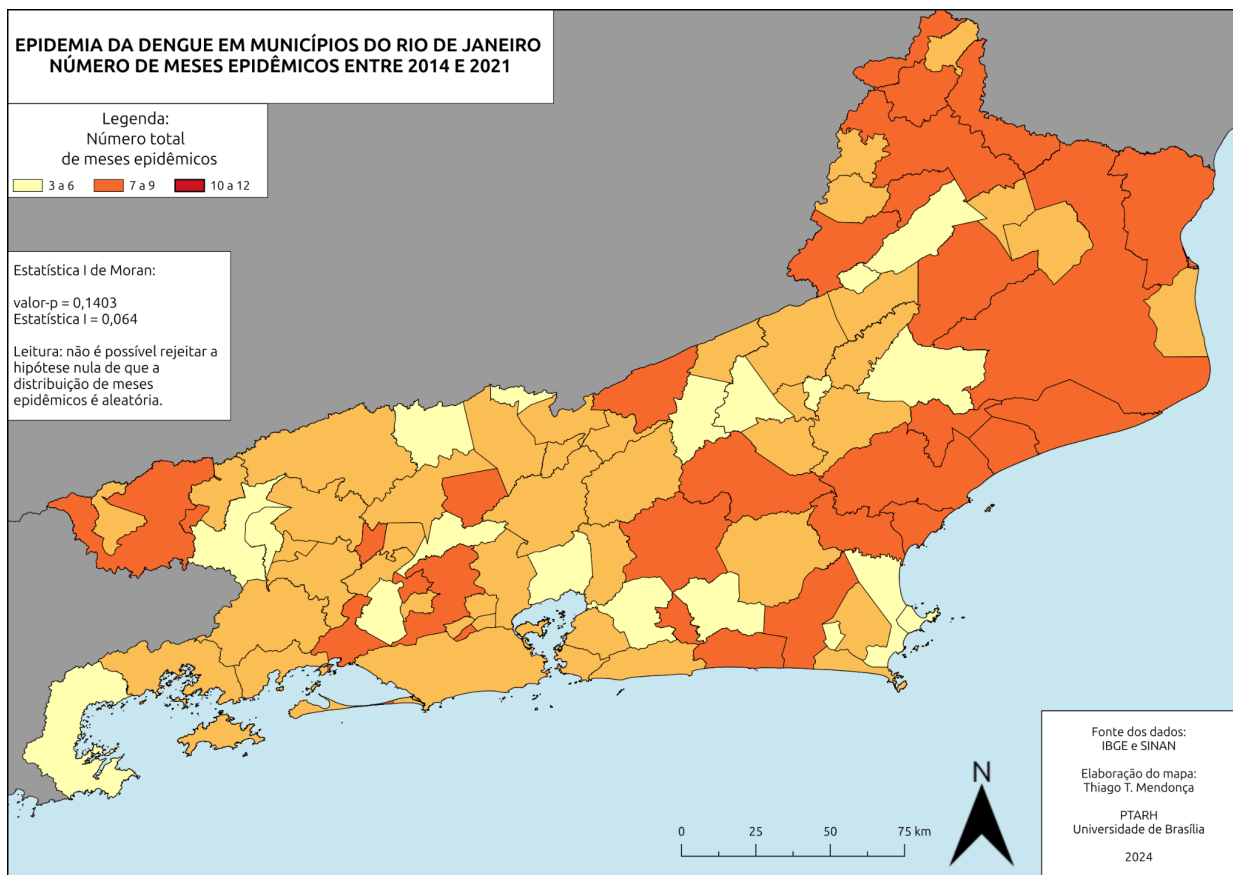


Figura 6.6: Mapa apresentando a espacialização do total de meses epidêmicos de dengue, entre 2014 e 2021, no estado do Rio de Janeiro.

Analisando-se o I de Moran a partir da incidência da doença, o cenário encontrado é diferente. Nessa perspectiva, é possível encontrar anos em que a distribuição da doença é caracterizada como não aleatória. Entre os anos de 2014 e 2021, em 5 deles foram observadas distribuições com valor-p igual ou inferior a 0,01, isto é, às quais se rejeita a hipótese nula de aleatoriedade da distribuição. Considerando a complexidade de relações existentes por trás das dinâmicas de doenças como a dengue, pode-se dizer que os valores de correlações, apesar de não serem considerados altos em uma perspectiva unicamente estatística, são, sim, relevantes sob uma análise de saúde pública: em um dos anos o valor de I foi superior a 0,5, o que indica a existência de uma distribuição agrupada. Nos outros anos, o valor de I flutuou mais próximo de 0, com um ano com $I = 0,30$ e outro $I = 0,27$.

O resultado da análise de distribuição espacial mostra que houve tendência de agrupamento de casos em determinados dos anos estudados. Além disso, a distribuição dos municípios mais afetados pela doença foi diferente com o passar dos anos, conforme observado na figura 6.7.

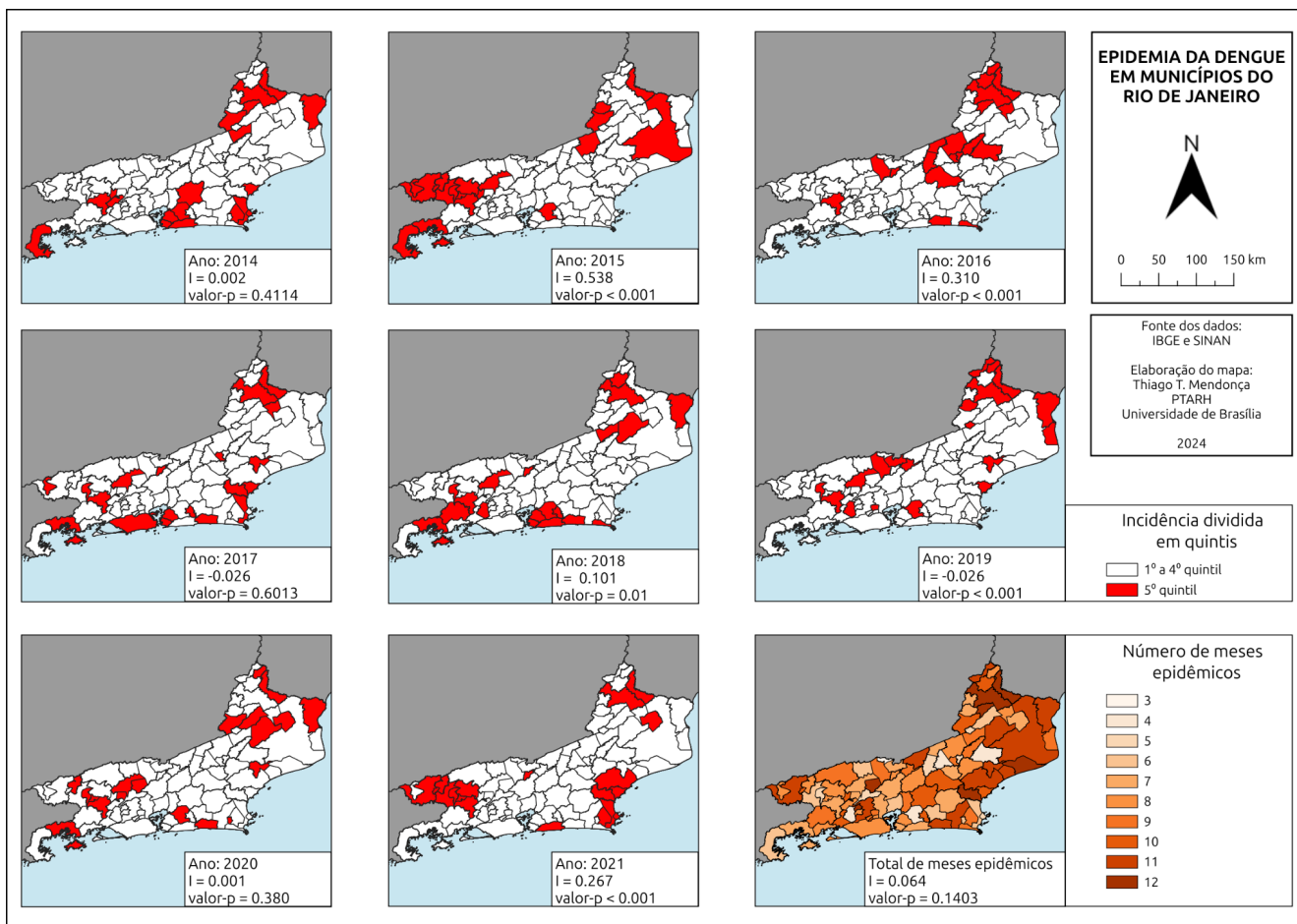


Figura 6.7: Mapa apresentando a espacialização da incidência da dengue, entre 2014 e 2021, no estado do Rio de Janeiro.

Os resultados de agrupamentos observados no presente capítulo podem ser justificados por diversos fatores. Como primeiro ponto, a escala das análises (a nível intermunicipal) pode ser pouco sensível para a compreensão das características da doença, que é, em geral, analisada em escalas maiores (a nível intramunicipal). Estudos como o de Drumond *et al.* (2020) e Pinho *et al.* (2024), foram ambos aplicados no contexto do Distrito Federal e encontraram resultados com significância estatística ainda mais forte.

Como hipótese para outro fator que afeta a distribuição da doença, em especial quanto aos dados agregados ao longo de todo o período, tem-se a natureza reativa de políticas públicas de saúde. Um município, ou região, que sofre uma epidemia em um determinado ano pode direcionar recursos para ações de redução dos impactos dessa doença durante os anos seguintes. Isso implica uma redução de casos neste município, enquanto outros municípios, que estiverem sob menor nível de alerta para controle e

recepção da doença, tornam-se os novos focos de disseminação. Dessa forma, o padrão de agrupamento se altera ao longo dos anos, conforme observado nas figuras 6.6 e 6.7.

6.2 CORRELAÇÕES ENTRE DENGUE E INDICADORES DE SANEAMENTO

Na primeira análise de associações entre variáveis de saneamento básico e o desfecho da dengue, a fonte de dados de saneamento foi o SINISA. Um total de 41 indicadores foram selecionados, conforme apêndice B. Dez municípios foram retirados, por não possuírem nenhum dado preenchido no SINISA, restando 82 ao todo. Com essas informações, ambos os coeficientes de Pearson (relações lineares) e de Spearman (relações não lineares) foram calculados. Buscou-se comparar os resultados de ambas as análises para compreender as associações a partir de abordagens diferentes. A figura 6.8, na sequência, apresenta um gráfico de dispersão entre a incidência mediana de dengue e um indicador representativo de cada eixo do SINISA: IN014_AE (Água e Esgoto), IN016_RS (Resíduos Sólidos) e IN040_AP (Águas Pluviais).

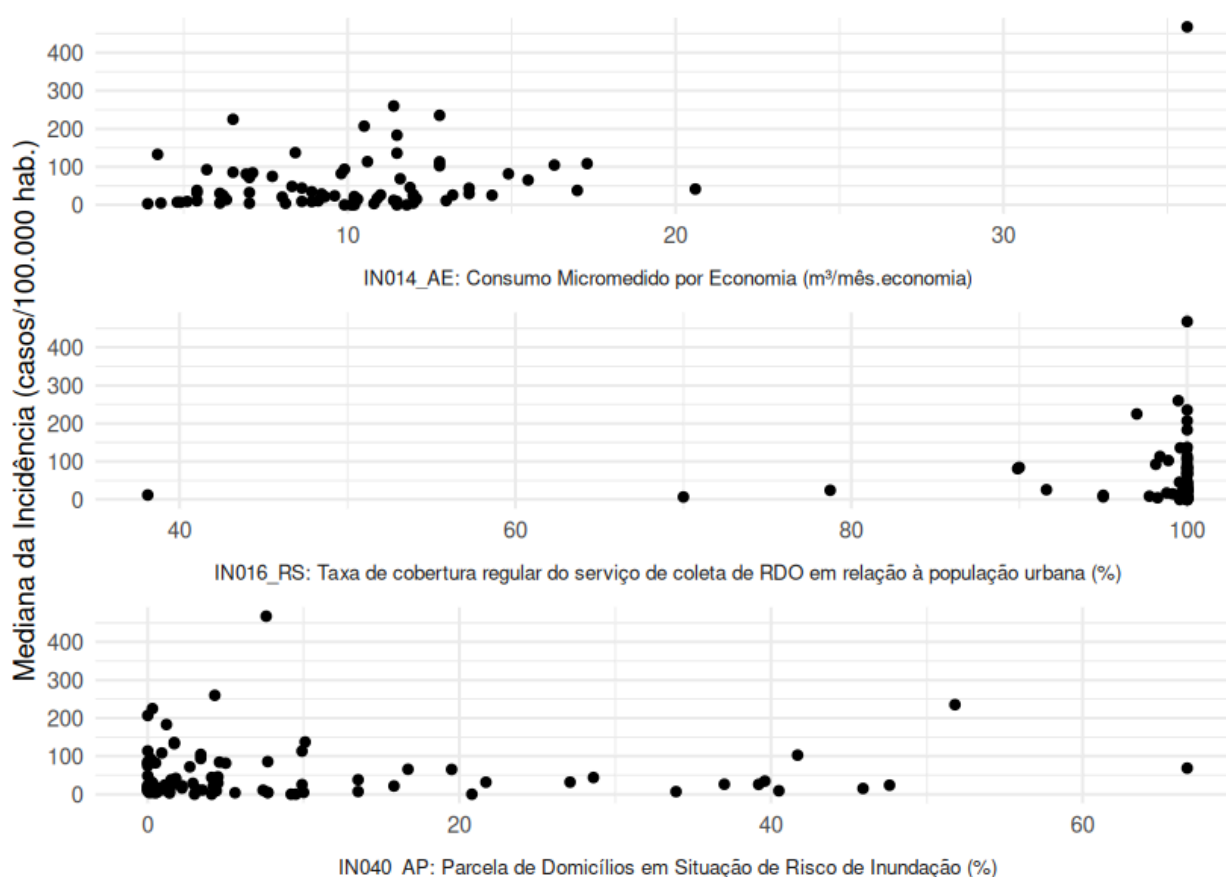


Figura 6.8: Gráfico de dispersão entre incidência anual mediana de dengue nos municípios do Rio de Janeiro com três indicadores do SINISA.

Observando a figura acima, é notável a ausência de padrões visuais que indiquem correlações diretas entre as variáveis. Dentre as três, apenas a variável IN014_AE (Consumo micromedido por economia), do eixo de água e esgoto, possui alguma indicação visual de correlação. Mesmo assim, devido ao ponto fora da curva, o coeficiente de Pearson se demonstrou excessivamente alto ($r = +0,49$). Uma das fragilidades desse coeficiente é sua sensibilidade a *outliers*, o que demonstra a necessidade de se utilizar métodos mais robustos para as análises de correlação, como recomendado pela literatura (Abdullah *et al.*, 2022; Oliveira *et al.*, 2023). Essa realidade se transpõe para todas as outras variáveis, como observado nas figuras e quadros presentes no apêndice C.

As análises utilizando o coeficiente de Spearman conseguiram captar um retrato mais coerente da realidade: as correlações observadas são ainda mais fracas do que as calculadas com r . Mesmo assim, alguns pontos interessantes podem ser levantados. A figura 6.9 apresenta os dez indicadores do SINISA mais correlacionados com a incidência da dengue. A linha vermelha indica o limite de significância estatística calculado, o valor crítico (82 municípios) $\approx \pm 0,22$. O valor crítico foi calculado com $n = 82$ devido à disponibilidade de dados do SINISA: 10 dos municípios não possuíam dados e foram desconsiderados nesta análise.

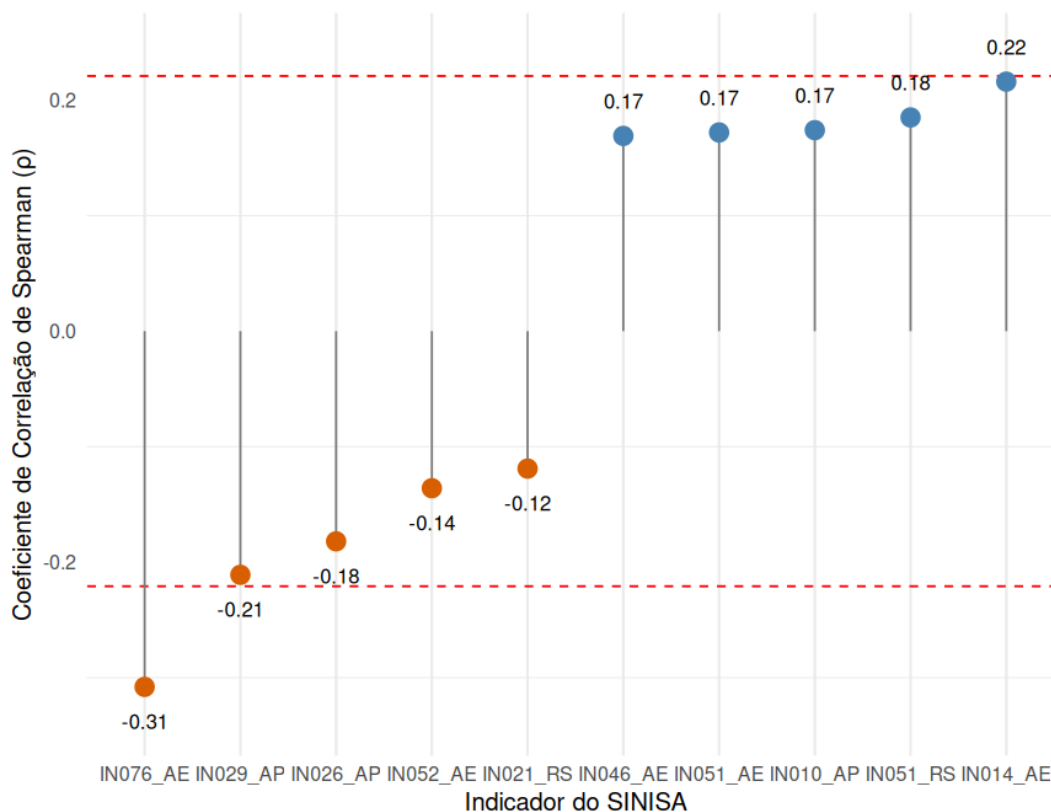


Figura 6.9: Gráfico de correlações entre as 10 variáveis de saneamento do SINISA mais correlacionadas com a incidência de dengue em 82 municípios do Rio de Janeiro.

Dentre todas as associações testadas, apenas dois indicadores tiveram associações fortes o suficiente para ultrapassar o valor crítico que delimita a significância dos parâmetros. A correlação negativa mais forte encontrada foi entre a incidência de dengue e a variável IN076_AE (Incidência das análises de turbidez fora do padrão), com $\rho = -0,31$. Este valor pode sugerir que municípios com melhores condições de qualidade de água tendem a apresentar maior incidência de dengue. Já a correlação positiva mais forte existente foi observada com o indicador IN014_AE (o mesmo que foi acentuado por r), com $\rho = +0,22$. O indicador IN029_AP (Parcela de Cursos d'Água Naturais Perenes com Diques), com $\rho = +0,21$, situou-se logo abaixo do limiar da significância, sendo considerado insignificante, junto a todas as outras variáveis testadas.

O coeficiente de Kendall também foi calculado e seus resultados se aproximaram da análise de Spearman, não trazendo novas interpretações para a discussão. Os resultados obtidos apresentam, de forma geral, um padrão aparentemente paradoxal, em que indicadores de desenvolvimento (como maior consumo de água, melhor qualidade de água e maior percentual de população com esgoto tratado) se associam a uma maior incidência de dengue. Isso pode indicar a existência de um fator de confusão relevante: a urbanização. Municípios mais urbanizados podem possuir, ao mesmo tempo, melhores serviços de saneamento básico e oferecerem condições propícias para proliferação dos *Aedes aegypti*.

Além de maior adensamento populacional, desigualdade social, dentre outros aspectos presentes em municípios de maior porte, outro viés relevante a ser considerado é o de notificação. Municípios mais desenvolvidos frequentemente têm maior capacidade de acompanhamento de endemias através de sistemas de vigilância epidemiológica, ao passo em que sua população possui maior acesso a serviços de diagnóstico. Isso pode levar a uma inflação artificial na incidência da doença quando comparada com municípios que não detêm as mesmas capacidades e subnotificam os casos das doenças.

De forma a complementar a análise, o mesmo método foi também aplicado para os dados de condições de domicílios do Censo de 2022 (IBGE, 2022), que refletem a realidade em um nível mais aproximado, na realidade domiciliar. Os resultados, apresentados na figura 6.10, apresentam uma maior quantidade de parâmetros significantes estatisticamente, em comparação à análise anterior, ao passo em que corroboram os debates realizados a partir dos achados do SINISA. O quadro completo de correlações está inserido no apêndice C. O valor crítico para os dados de saneamento do IBGE foi calculado

considerando todos os 92 municípios do estado devido, uma vez que há disponibilidade de dados em todos eles. Sendo assim, valor crítico ($n = 92$) $\approx \pm 0,21$)

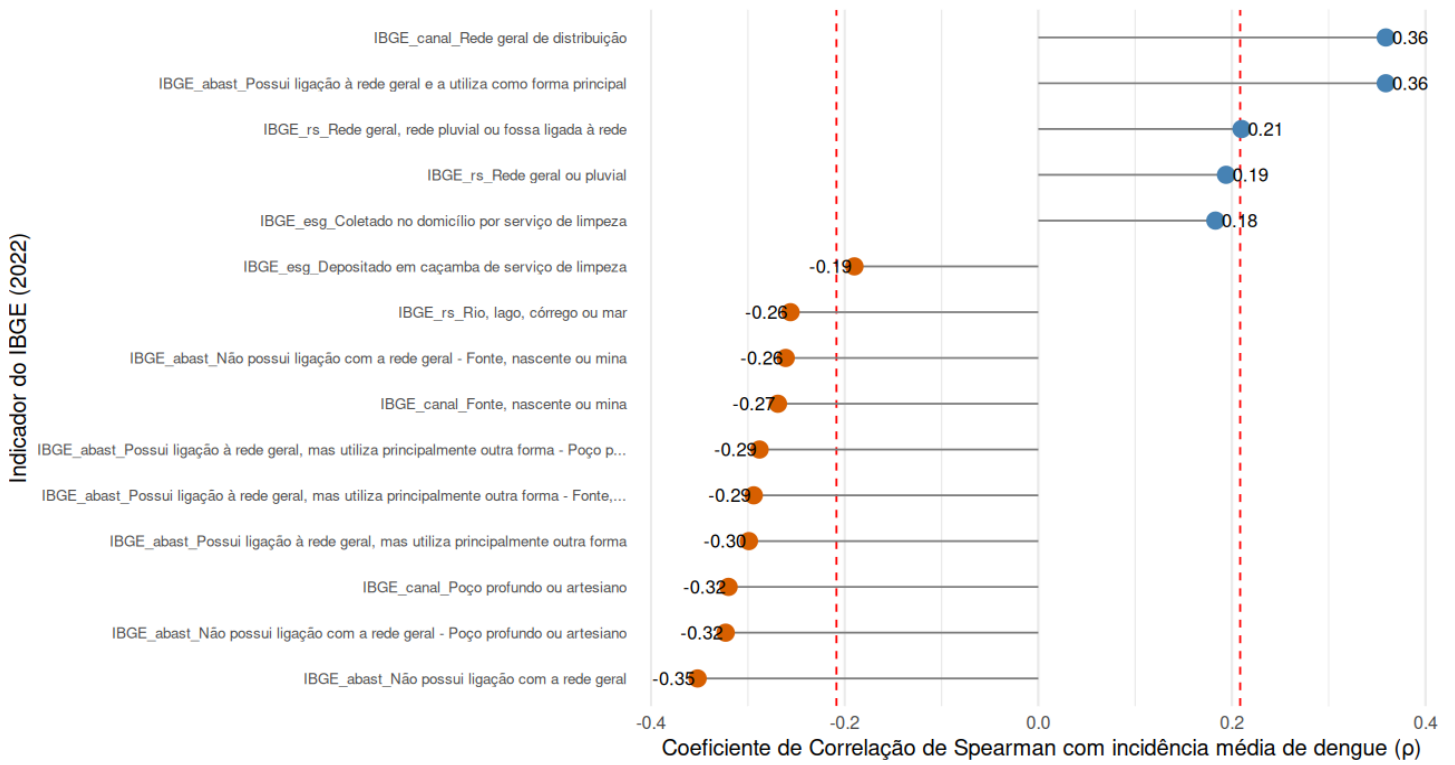


Figura 6.10: Gráfico de correlações entre as 15 variáveis de saneamento do IBGE mais correlacionadas com a incidência de dengue em 92 municípios do Rio de Janeiro.

Conforme discutido nas análises anteriores do presente capítulo, o coeficiente de Spearman (ρ) obteve os resultados mais coerentes para as análises de correlação entre dados de saneamento (SINISA) e de dengue. Sendo assim, a análise de correlação entre dados de saneamento do IBGE e a incidência de dengue utilizou o mesmo coeficiente.

A análise dos dados do IBGE revela associações robustas entre as variáveis e a incidência da doença nos municípios do RJ. Enquanto no SINISA apenas duas variáveis ultrapassaram o limiar de significância, com os dados do IBGE, 12 variáveis foram consideradas significativas, o que pode ser justificado por diferenças metodológicas nas duas fontes de dados, como a completude dos dados e a natureza de coleta domiciliar do Censo, que captura as informações de forma mais direta do que os dados de gestão do SINISA. A correlação positiva mais forte foi observada com o indicador “Possui ligação à rede geral e a utiliza como forma principal” (Eixo IBGE abastecimento), com $\rho = 0,36$, acima do limiar. Por outro lado, as correlações negativas mais fortes foram com as

variáveis “Não possui ligação com a rede geral” (Eixo IBGE abastecimento), com $\rho = -0,35$ e “Não possui ligação com a rede geral - Poço profundo ou artesiano” (Eixo IBGE abastecimento), com $\rho = -0,32$.

De forma geral, as correlações observadas nos dados do IBGE são coerentes com o observado nos dados do SINISA, apesar da diferença na amplitude dos dados. Indicadores que representam acesso a serviços de saneamento formais e centralizados, comuns em áreas urbanizadas, apresentaram correlações positivas e significativas com a dengue. De outra forma, os indicadores que descrevem soluções descentralizadas, individuais ou com ausência de serviços (como o indicador “Não possui ligação com a rede geral - Poço profundo ou artesiano”), mais comuns em áreas menos urbanizadas, mostraram correlações significativas e negativas.

As análises de correlações realizadas entre incidência de dengue e dados de saneamento, com duas perspectivas diferentes (SINISA e IBGE), ambos com resultados semelhantes, evidenciam como as relações entre saneamento e dengue são multifatoriais. As correlações aparentam ser paradoxais, visto que indicadores que descrevem melhores condições de saneamento são positivamente correlacionados com a incidência da dengue, enquanto indicadores que apontam piores condições de saneamento têm correlação negativa com a doença.

Todo esse debate fortalece uma hipótese de que a urbanização pode ser um fator de confusão relevante neste contexto: indicadores de saneamento podem indicar condições de maior urbanização, ao passo em que ambientes urbanizados oferecem condições propícias (adensamento populacional, disponibilidade de criadouros artificiais, etc.) para a disseminação da doença. Além disso, este efeito pode estar sendo potencializado por um viés de notificação: municípios menores, menos urbanizados e com menos investimentos em sistemas centralizados de saneamento básico podem ter, também, menos recursos para manutenção de sistemas de vigilância eficientes, ocasionando uma subnotificação de casos

Diante dessa complexidade, fica evidente que fazer análises isoladas de indicadores é insuficiente para compreender o cenário epidemiológico de forma adequada. Desta forma, buscando-se obter uma visão mais robusta dos dados, uma análise multivariada foi aplicada no capítulo seguinte. A aplicação da análise de componentes principais (PCA), com o intuito de reduzir a dimensionalidade dos dados, agregando múltiplos indicadores em conjuntos menores e mantendo seu poder de explicação, é apresentada na sequência.

6.3 ANÁLISE DE COMPONENTES PRINCIPAIS

Conforme contextualizado ao longo do trabalho e debatido no capítulo anterior, os indicadores do SINISA são numerosos e frequentemente correlacionados entre si, o que dificulta suas análises a nível individual. Uma análise de componentes principais (PCA) foi aplicada para reduzir a dimensão dos dados e sua multicolinearidade. Para isso, uma pré-seleção das variáveis foi realizada, conforme descrita a seguir.

6.3.1 Seleção de indicadores do SINISA

De um universo total de 195 indicadores para os três eixos, apenas 41 restaram ao final, após o pré-tratamento. O primeiro passo desta etapa foi a aplicação de um filtro, que retirou quaisquer variáveis que não estivessem preenchidas em ao menos 80% dos municípios do estado do Rio de Janeiro. Na sequência, a seleção qualitativa foi feita com base na literatura e em conhecimentos do autor sobre a epidemiologia da dengue e suas possíveis relações com saneamento básico. Foram priorizados os indicadores que podem ter efeitos mais diretos sobre a proliferação do vetor. A lista completa dos 41 indicadores está no apêndice B e o quadro 6.1, a seguir, apresenta o resultado do processo de seleção.

Quadro 6.1: Quantidade de indicadores selecionados, de cada eixo do SINISA

Eixo de saneamento	Total	Quant. indicadores inválidos (filtro 80%)	Quant. indicadores válidos descartados (seleção qualitativa)	Quant. indicadores selecionados
Resíduos sólidos	47	28	8	11
Água e Esgoto	123	63	42	18
Águas pluviais	25	8	5	12

6.3.2 Componentes Principais dos eixos do SINISA -

A análise de componentes principais foi aplicada quatro vezes: uma para cada eixo de saneamento básico e uma com todos os dados reunidos. Os resultados detalhados estão apresentados no apêndice D. A figura 6.10 apresenta um gráfico de Scree representando a variância proporcional de cada componente principal (barras) e a variância acumulada (curvas). Uma linha de referência foi adicionada em 75% para indicar um limite a partir do qual a variância explicada é considerada suficiente.

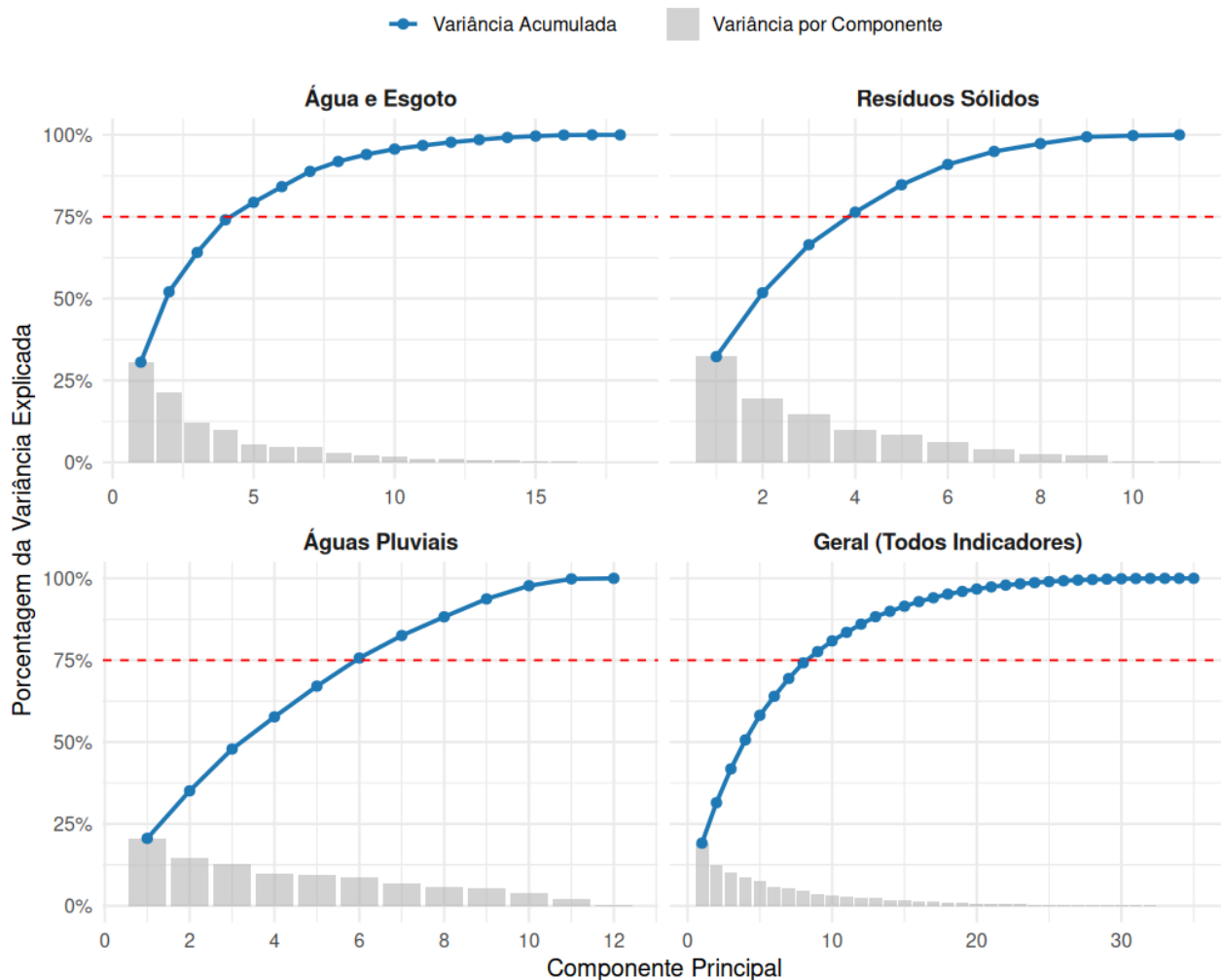


Figura 6.11: Gráfico de Scree representando o acumulado do percentual de variância explicada por cada componente principal (resultados da PCA).

Os gráficos de Scree mostram uma característica relevante dos dados selecionados do SINISA: a intensidade da correlação existente entre as variáveis. Nos eixos de Água e Esgoto e Resíduos Sólidos, as quatro primeiras componentes principais foram capazes de explicar 74,1% e 76,4% da variância total, respectivamente. Isso indica que os indicadores que compõem os respectivos eixos têm grande correlação entre si.

Em contrapartida, para o eixo de Águas Pluviais são necessários seis componentes principais para alcançar o mesmo poder de explicação. O resultado da análise feita usando todos os 41 indicadores é ainda mais amplo, o que é esperado: são necessários nove componentes para ultrapassar a linha de referência.

Além das análises quantitativas, as componentes principais também fornecem informações valiosas em aspectos qualitativos. Em todos os eixos analisados, a observação de indicadores que compõem cada PC resultou em análises relevantes. O quadro 6.2 apresenta um resumo dos resultados encontrados.

Quadro 6.2: Características dos componentes principais com maior explicação de variância

Eixo de Saneamento	Componente	Descrição do Componente	Principais Indicadores	Variância Explicada
Água e Esgoto	PC1_AE	Eficiência (índices de perdas) da Rede de Água	IN009, IN013, IN025, IN050, IN051	30,6%
	PC2_AE	Consumo de água + perdas	IN014, IN022, IN049, IN052, IN053	21,5%
	PC3_AE	Cobertura dos Serviços	IN001, IN023, IN055	12,0%
	PC4_AE	Parâmetros de qualidade da água	IN076 IN084	9,9%
Resíduos Sólidos	PC1_RS	Mão de Obra Operacional	IN001, IN019, IN045, IN051	32,3%
	PC2_RS	Geração e Coleta de Resíduos	IN016, IN021, IN028	19,5%
	PC3_RS	Cobertura da Coleta	IN014, IN015, IN016	14,6%
	PC4_RS	Gastos com serviços	IN002, IN006, IN015	10,0%
Águas Pluviais	PC1_AP	Densidade Urbana e cursos d'água	IN043, IN044, IN026, IN029	20,6%
	PC2_AP	Investimentos com serviços de Drenagem Urbana	IN010, IN048	14,5%
	PC3_AP	Vulnerabilidade a eventos hidrológicos e condições de cursos d'água	IN 021, IN027, IN029, IN041	12,8%
	PC4_AP	Impermeabilização e redes de Drenagem Urbana	IN020, IN021, IN029	9,8%

As análises dos componentes principais trouxeram à luz resultados interessantes tanto no sentido quantitativo, quanto no sentido qualitativo, de forma que a aplicação da PCA cumpriu seu objetivo. O número de indicadores de saneamento foi reduzido de 41 para aproximadamente 14 (quatro para os eixos de AE e RS, respectivamente, e seis para o eixo AP) sem perdas expressivas.

6.3.3 Associação entre os Componentes Principais e a Dengue

Para validar as variáveis sintéticas construídas no subcapítulo anterior, mais uma análise de correlação de Spearman foi realizada. A incidência média da dengue foi comparada com os escores dos dois componentes principais com maior capacidade de explicação, para cada eixo do SINISA. Os resultados estão expostos no quadro 6.3.

Quadro 6.3: Resultado da calibração de parâmetros do modelo nos 6 municípios testados, em comparação com intervalos observados na literatura.

Componente Principal	Interpretação da Dimensão	Coefficiente de Spearman (ρ)
PC1_AE	Eficiência (índices de perdas) da Rede de Água	0.014
PC2_AE	Consumo de Água + perdas	0.002
PC1_RS	Mão de Obra Operacional	-0.078
PC2_RS	Geração e Coleta de Resíduos	-0.162
PC1_AP	Densidade Urbana e cursos d'água	0.19
PC2_AP	Impermeabilização e redes de Drenagem Urbana	0.136

O objetivo dessa análise foi verificar se as novas variáveis sintéticas, que representam de forma mais ampla as dimensões de saneamento básico, possuíam associações mais fortes com a incidência da dengue, em comparação às análises com variáveis individuais realizadas anteriormente. Os resultados apresentados no quadro 6.3 indicam, porém, que nenhum dos componentes principais demonstrou uma correlação estatisticamente significativa com a doença. A associação mais forte foi com o primeiro componente de águas pluviais com $\rho = 0,19$, um valor inferior ao limiar de significância.

Erros, vieses e fatores de confundimento podem ser responsáveis pela dificuldade de se encontrar correlações fortes e estatisticamente válidas. No contexto dos dados do SINISA, em que os indicadores caracterizam uma visão geral de cada município, a sensibilidade dos dados às variações existentes em uma escala intraurbana é baixa. Dentre os municípios estudados, alguns possuem condições de saneamento básico amplamente diversas entre diferentes regiões, havendo bairros com saneamento universalizado e bairros onde esgoto corre a céu aberto (Mendonça, 2022). Essa condição não é exemplificada apenas no contexto estudado, mas é uma limitação essencial a ser considerada em estudos ecológicos de forma geral, sendo denominada “falácia ecológica” (Bonita; Beaglehole; Kjellström, 2010). Ao utilizar valores únicos para cada município, perdem-se os efeitos das dinâmicas existentes em regiões mais suscetíveis à doença, diluindo a capacidade de obtenção de bons resultados em análises de correlação.

Os resultados obtidos neste capítulo reafirmam a grande complexidade existente em estudos sobre relações existentes entre dengue e condições de saneamento básico. Frente à incapacidade dos métodos estatísticos aplicados de encontrar associações claras entre as variáveis analisadas, tornou-se necessária a elaboração de um método mais robusto para a análise da doença. O próximo passo realizado neste trabalho foi a elaboração e calibração

de um modelo mecânico de dengue, que simula o efeito de diversas variáveis no ciclo de vida e disseminação da doença com base em mecanismos ambientais e biológicos.

6.4 MODELO SIR-SI-SI

A etapa final do presente trabalho consistiu no desenvolvimento, calibração e análise de um modelo mecânico de transmissão de dengue. O modelo foi calibrado nos seis municípios do recorte espacial descritos no capítulo 5.2. Neste capítulo, são apresentados os resultados do processo de calibração, o ajuste do modelo aos dados e uma discussão sobre seu desempenho e limitações.

6.4.1 Resultados da calibração do modelo

A otimização foi executada para cada um dos seis municípios, minimizando a soma dos quadrados dos erros entre os valores de casos estimados e observados. Os parâmetros descritos no quadro 5.5 foram estimados alterando-se computacionalmente os seguintes valores: constantes c nas equações quadráticas e de Brière; valor da capacidade de carga K e valor da proporção da população humana sensível à doença α_{Sh} . Os valores encontrados para cada parâmetro, em cada município, estão sintetizados no quadro abaixo, que está dividido em dois cenários: (A) calibração semanal e (B) calibração mensal. Para a mortalidade de mosquitos, a taxa de sobrevivência aquática e a taxa de desenvolvimento, o resultado não é um valor específico, mas um intervalo de valores obtidos ao longo de todo o período. Por isso, esses valores estão expressos no quadro como intervalos.

Quadro 6.4: Resultado da calibração de parâmetros do modelo nos 6 municípios testados, em comparação com intervalos observados na literatura. (A) Valores após calibração semanal. (B) Valores após calibração mensal.

Variável Calibrada	Símbolo	Intervalo na literatura	Valor após calibração semanal (cidade)					
			Campos	Caxias	Macaé	Paraty	Resende	Rio de Janeiro
Proporção da população de hospedeiros suscetíveis	α_{Sh}	0.04 a 1 (2); 0.569 (3); 0.3 a 1 (5);	0.749	0.822	0.974	0.494	0.942	0.859
Capacidade de carga	K	1 a 5 (2); 3 (3); 0,25 a 4 (5)	0.534	0.676	0.601	1.138	0.645	0.578
Mortalidade de mosquitos	μ_v	30^{-1} a $7,7^{-1}$ (1); 0,026 a 0,043 (2); 12^{-1} (4)	0.072 a 0.094	0.084 a 0.138	0.095 a 0.123	0.088 a 0.144	0.096 a 0.206	0.074 a 0.102

Variável Calibrada	Símbolo	Intervalo na literatura	Valor após calibração semanal (cidade)					
			Campos	Caxias	Macaé	Paraty	Resende	Rio de Janeiro
Taxa de sobrevivência aquática	sa	0,0 a 0,76 (5)	1.18 a 2.74	0.049 a 0.338	0.150 a 0.348	0.044 a 0.299	0.000 a 1.805	0.080 a 0.186
Taxa de desenvolvimento	dr	0,169 (1); 0,011 a 0,158 (2); 0,063 a 0,1 (5)	0.013 a 0.052	0.023 a 0.172	0.040 a 0.165	0.004 a 0.030	0.010 a 0.162	0.033 a 0.137

Variável Calibrada	Símbolo	Intervalo na literatura	Valor após calibração mensal (cidade)					
			Campos	Caxias	Macaé	Paraty	Resende	Rio de Janeiro
Proporção da população de hospedeiros suscetíveis	α_{Sh}	0.04 a 1 (2); 0.569 (3); 0.3 a 1 (5);	0.949	0.564	0.905	0.986	0.984	0.502
Capacidade de carga	K	1 a 5 (2); 3 (3); 0,25 a 4 (5)	0.535	1.068	0.550	0.456	0.570	1.040
Mortalidade de mosquitos	μ_v	30^{-1} a $7,7^{-1}$ (1); 0,026 a 0,043 (2); 12^{-1} (4)	0.092 a 0.119	0.082 a 0.135	0.086 a 0.111	0.065 a 0.107	0.088 a 0.190	0.078 a 0.107
Taxa de sobrevivência aquática	sa	0,0 a 0,76 (5)	2.965 a 6.874	0.028 a 0.194	0.270 a 0.626	0.025 a 0.170	0.000 a 1.248	0.092 a 0.214
Taxa de desenvolvimento	dr	0,169 (1); 0,011 a 0,158 (2); 0,063 a 0,1 (5)	0.013 a 0.054	0.019 a 0.148	0.031 a 0.127	0.021 a 0.156	0.009 a 0.132	0.264 a 1.111

Referências nos quadros acima: (1) Aguiar *et al.* (2022); (2) de Vasconcelos *et al.* (2023); (3) Nuraini *et al.* (2021); (4) Chanprasopchai *et al.* (2018); (5) Alves (2021).

Comparando-se os valores calculados com os intervalos levantados na literatura, é possível notar que a maior parte dos parâmetros calibrados se encaixou dentro dos limites esperados. Apesar disso, alguns dos parâmetros levantam debates quanto aos valores finais encontrados. Os valores calculados de capacidade de carga (K) encaixaram-se apenas nos limites descritos por Alves (2021). Os valores encontrados indicam a proporção de mosquitos, em relação à quantidade de humanos, inferior a um, isto é, que há menos mosquitos do que humanos nas cidades analisadas. Este valor não corrobora o levantado pela literatura e aplicado por Nuraini *et al.* (2021) e de Vasconcelos *et al.* (2023), que adotaram valores de até 5 mosquitos por humano.

Em relação à mortalidade de mosquitos (μ_v), em sua maioria, os valores se localizaram dentro do estudado na literatura, com apenas alguns valores ligeiramente

superiores aos limites superiores. Para a taxa de desenvolvimento (dr), todos os valores calculados corroboram com a literatura. Por fim, quanto à taxa de sobrevivência aquática (sa), 4 dos municípios receberam parâmetros dentro dos limites delimitados por Alves (2021). Os municípios de Campos e Resende, por outro lado, receberam valores calibrados amplamente distintos do esperado. Isso pode ter ocorrido devido à utilização de limites amplos demais durante o ato de calibração. Mesmo assim, por este ser um valor pouco explorado na literatura, sendo elaborado por apenas um autor dentre as referências bibliográficas utilizadas, uma justificativa para esta diferença é devida à baixa disponibilidade de análises concretas desta variável em diferentes contextos.

Além das comparações com a literatura, outro aspecto importante observado nos resultados é a diferença entre parâmetros obtidos com a calibração utilizando agregação semanal e mensal. Tomando o Rio de Janeiro como exemplo, a proporção de população suscetível à doença (α_{Sh}) foi estimada em 50,2% da população na calibração mensal e em 86,9% na calibração semanal. Essa diferença significativa sugere a ocorrência de equifinalidade, isto é, que conjuntos distintos de variáveis produzam resultados finais semelhantes

Uma justificativa para isso é devida às diferenças inerentes às escalas temporais adotadas. Uma escala mensal suaviza os picos epidêmicos da doença, que geralmente ocorrem em escalas de dias ou semanas. Nesses casos, o algoritmo, que não tem acesso a informação sobre tais picos, acaba por encontrar valores distintos para reproduzir as tendências observadas na escala temporal de meses. Por outro lado, a agregação semanal tem maior capacidade de captar flutuações mais detalhadas, resultando em parâmetros mais sensíveis às dinâmicas reais do sistema, ao passo que necessitam de maior capacidade computacional.

A partir dos resultados obtidos, afirma-se, de forma geral, que os equacionamentos utilizados para os cálculos dos parâmetros foram eficientes ao resultarem em valores coerentes com a literatura para a maioria dos parâmetros analisados. Os parâmetros que obtiveram valores diferentes dos esperados devem ser reanalisados com mais cuidado, posteriormente. O fenômeno da equifinalidade também deve ser levado em consideração para garantir que os dados levantados sejam coerentes com a realidade. Diante disso, na etapa seguinte, em que são apresentados gráficos comparando os resultados dos modelos em relação aos casos observados, os parâmetros utilizados para a geração de gráficos e demais análises foram obtidos com a calibração semanal.

6.4.2 Ajuste do modelo aos dados observados

Após análise comparativa entre os resultados do ajuste e os intervalos observados na literatura, na sequência é realizada uma comparação entre as curvas estimadas de casos e os casos observados (reais). Conforme explicitado no capítulo 5.6.5, o modelo foi dividido em duas partes: treinamento, com dados de 2014 a 2017, e validação, com dados de 2018 a 2022. Os resultados estão apresentados nas figuras a seguir e nos apêndices do trabalho.

Mesmo com a justificativa elaborada no capítulo anterior para o uso da agregação semanal, também foi de interesse do presente trabalho investigar visualmente as diferenças das duas abordagens de agregação aplicadas no ajuste do modelo. A figura 6.12 apresenta uma comparação dos resultados das calibrações realizadas no município de Caxias.

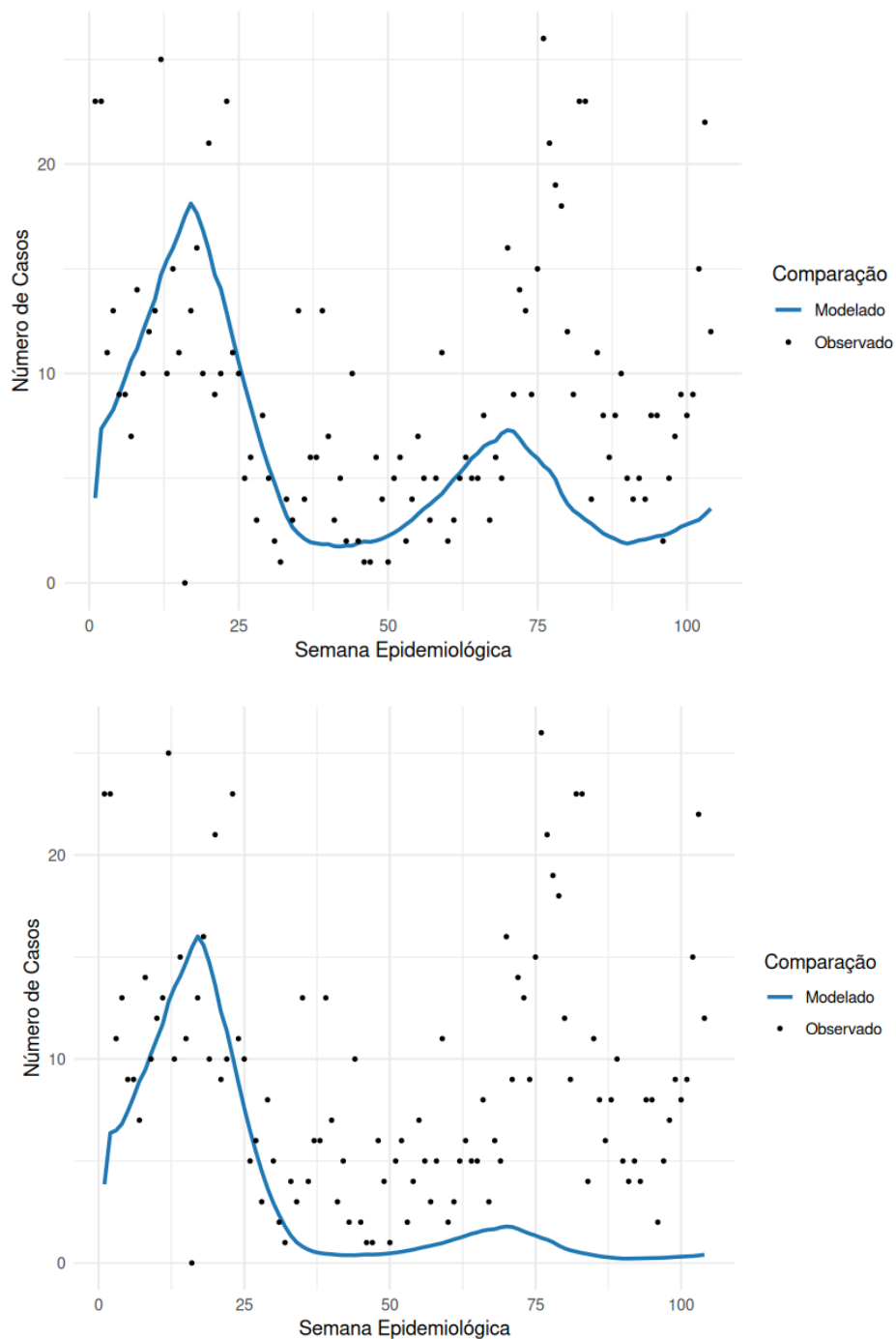


Figura 6.12: Comparação entre calibração do modelo com agregamento semanal (acima) e mensal (abaixo) para o município de Duque de Caxias nas primeiras 100 semanas epidemiológicas.

O caso exposto acima corrobora o elaborado no capítulo anterior quanto à sensibilidade da calibração com dados agrupados semanalmente, em relação ao agrupamento mensal. A calibração semanal (figura 6.12 A) demonstrou maior aderência ao pico observado durante as primeiras semanas epidemiológicas. Ao mesmo tempo, no pico que ocorre no entorno da semana epidemiológica 75, a calibração semanal teve maior aproximação, enquanto a calibração mensal teve uma variação muito pequena.

Outro caso notável que trata da importância dessa comparação foi observado em Paraty, o menor município utilizado no estudo. Diversas semanas epidemiológicas foram registradas com zero casos, dificultando a convergência do algoritmo de calibração. Essa é uma dificuldade encontrada para a calibração em municípios de menor porte, tornando a modelagem mais complexa em cenários como esse. A calibração realizada com agregação mensal não obteve resultados mais contundentes, também não conseguindo modelar o aumento de casos ocorrido durante o surto epidêmico ocorrido a partir da semana epidêmica 50, após semanas sem caso algum (figura 6.13). Isso sugere que a dinâmica da dengue no município pode ser influenciada por outros fatores externos que não foram considerados no modelo SIR-SI-SI, como a importação de casos através da migração de indivíduos contaminados, como debatido no capítulo 6.1 e por Alves (2021).

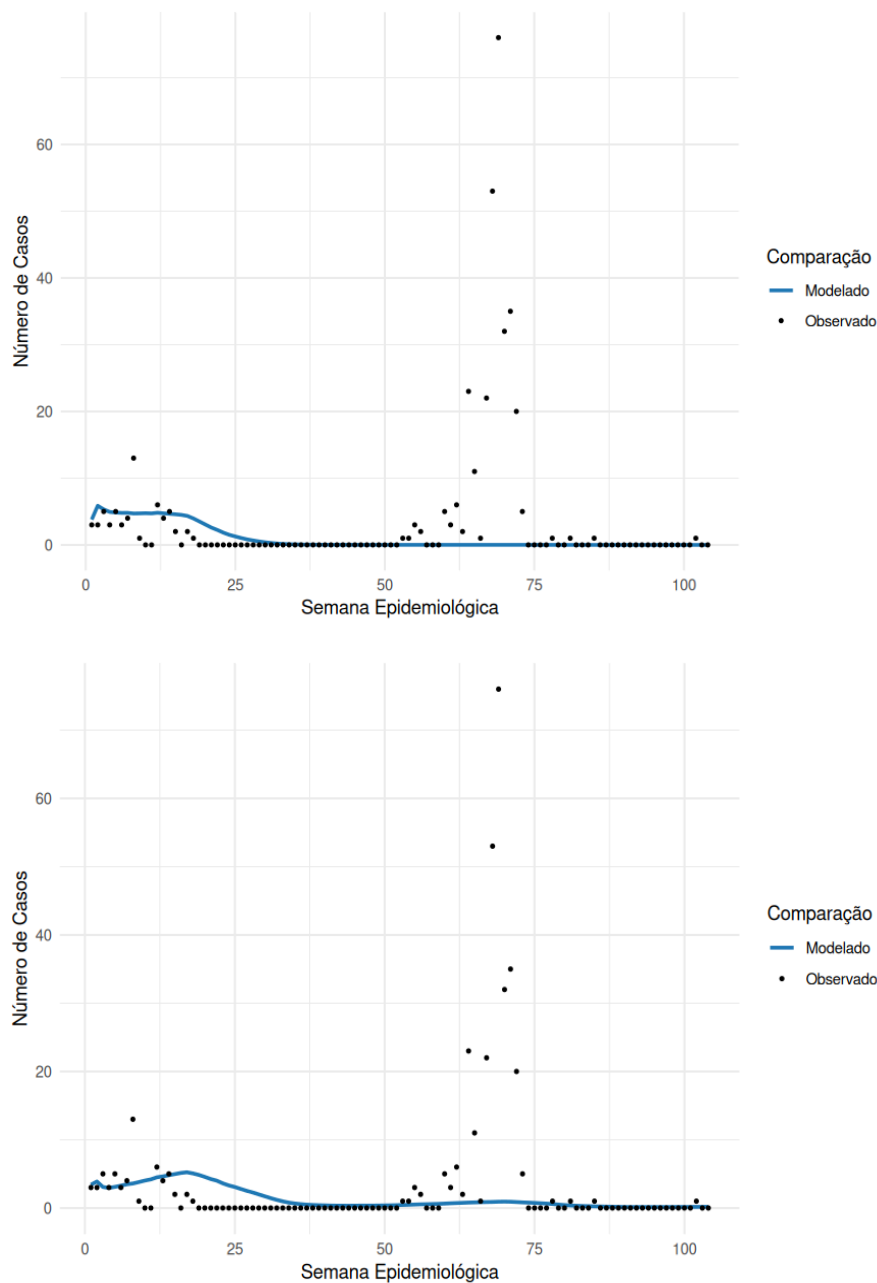


Figura 6.13: Comparação entre calibração do modelo com agregamento semanal (acima) e mensal (abaixo) para o município de Paraty nas primeiras 100 semanas epidemiológicas.

Os surtos abruptos observados em Paraty e em Campos dos Goytacazes exemplificam uma dificuldade existente na modelagem de doenças: a capacidade de captar e reproduzir com fidelidade a magnitude dos picos epidêmicos. Um modelo climático, como o aplicado neste trabalho, possui boa aptidão para captar efeitos da sazonalidade na doença, mas aparentemente não captura a intensidade dos surtos epidêmicos, sugerindo que outros fatores são responsáveis por tais episódios extremos.

A figura 6.14 apresenta resultados da calibração para o município de Macaé em todo o período de treinamento (2014 a 2017). O Brasil passou por grandes epidemias

durante esse período, que estão visíveis na figura através dos casos “Observados”. A magnitude dos picos não foi captada com eficiência pelo modelo, mas a sazonalidade, sim.

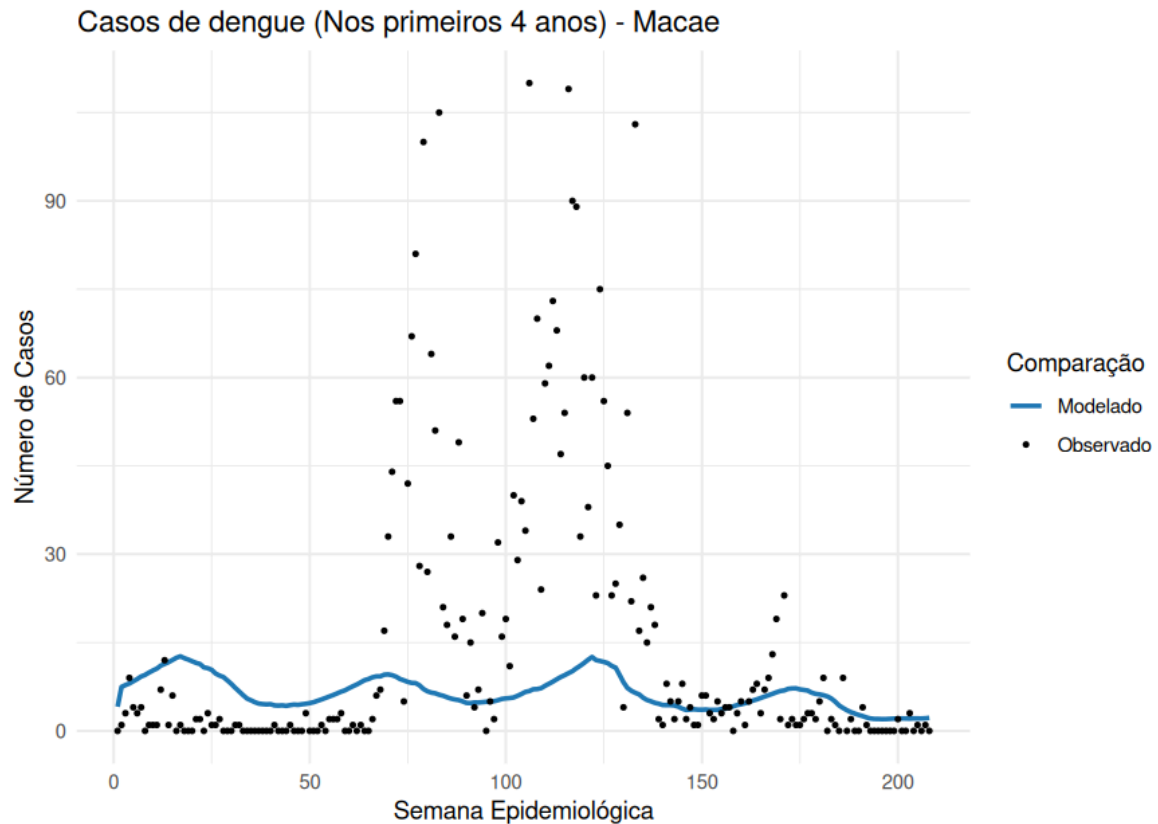


Figura 6.14: Modelo calibrado com agregamento semanal para o município de Macaé. Período de treinamento (2014–2017).

No geral, o modelo demonstrou um bom ajuste durante o período de treinamento (2014-2017). As simulações capturaram a sazonalidade da doença, reproduzindo os picos epidêmicos observados nos meses de verão e os períodos com menos casos durante os meses mais frios. A capacidade de simulação da magnitude dos picos, porém, foi baixa, como esperado, dado que o modelo não incorporou fatores capazes de captar tais efeitos, como introdução de novos sorotipos, efeitos culturais ou comportamentais da população.

Expandindo-se as curvas para todos os 9 anos do período analisado, considerando tanto o período de treinamento quanto o de validação, é possível observar algumas dificuldades do modelo. O exemplo seguinte (figura 6.15) demonstra a capacidade do modelo de captar sazonalidade da doença, estimando aumentos de casos coerentes com os observados, porém em magnitudes diferentes: durante épocas com maiores surtos (em torno das semanas epidemiológicas 60, 110 e 280), o modelo subestimou os casos, enquanto em épocas com baixas de casos, o modelo superestimou a transmissão da doença

(semanas epidemiológicas 180 e 380) estimando mais casos do que os observados. O modelo suavizou os extremos observados na realidade.

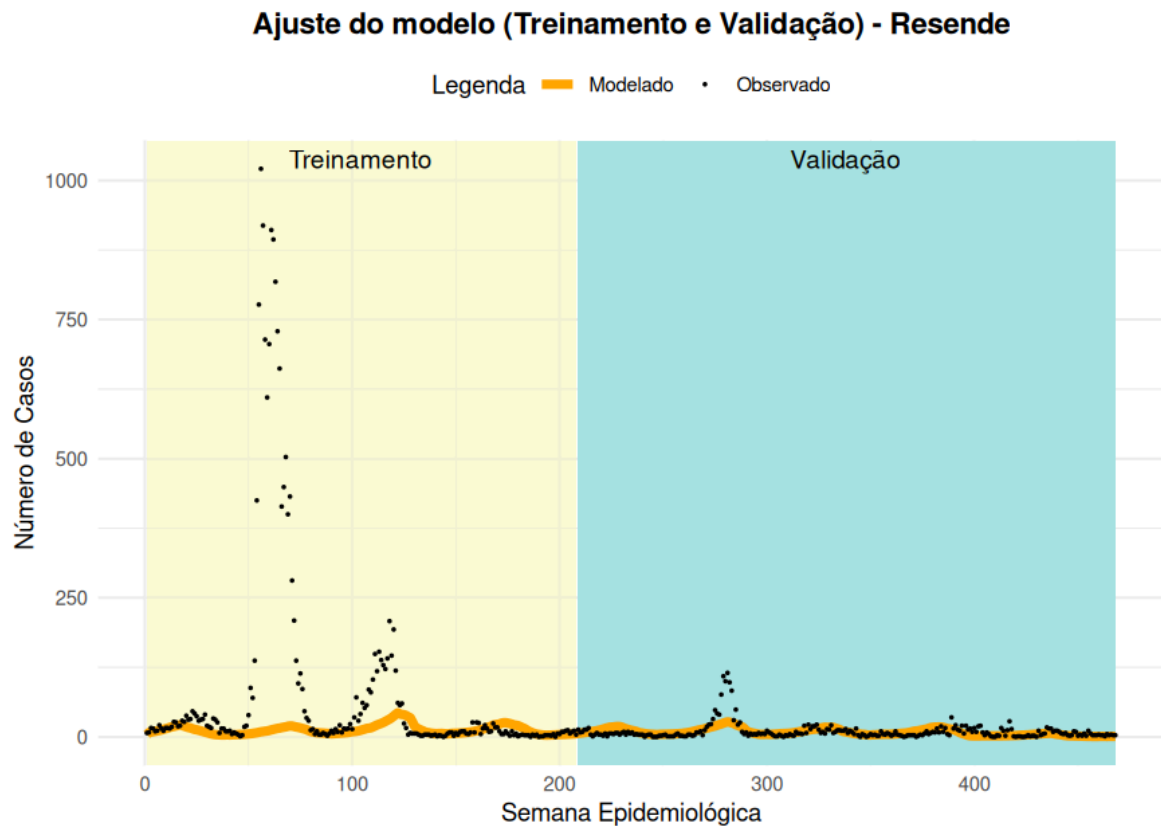


Figura 6.15: Modelo calibrado com agregamento semanal para o município de Resende. Comparação entre período de treinamento (2014–2017) e período de validação (2018–2022).

Em contraste com os resultados obtidos para Resende, que aproximaram-se da realidade mesmo durante o período de validação, o ajuste para Campos dos Goytacazes demonstrou problemas acentuados. A projeção para o futuro aumenta indefinidamente o número de casos durante o período de validação (figura 6.16), gerando um crescimento exacerbado e irreal.

O resultado obtido com a aplicação do modelo em Campos demonstra uma das fragilidades de modelagens mecânicas para a estimativa de casos futuros: a otimização em um período específico pode extrapolar tendências de maneira equivocada. Neste exemplo, é possível que, durante o período de treinamento, o algoritmo de calibração tenha encontrado resultados mais aproximados aos dados reais quando adotou um valor alto para a taxa de transmissão. Ao extrapolar este valor no período de validação, o modelo explodiu. A estrutura matemática adotada no presente modelo não possui gatilhos para reverter comportamentos de crescimento como esse (como por exemplo a mortalidade de humanos e a redução de população suscetível ao longo do tempo, dentre outros), portanto o

comportamento de crescimento se manteve, ocasionando crescimentos irreais como o observado na figura 6.16.

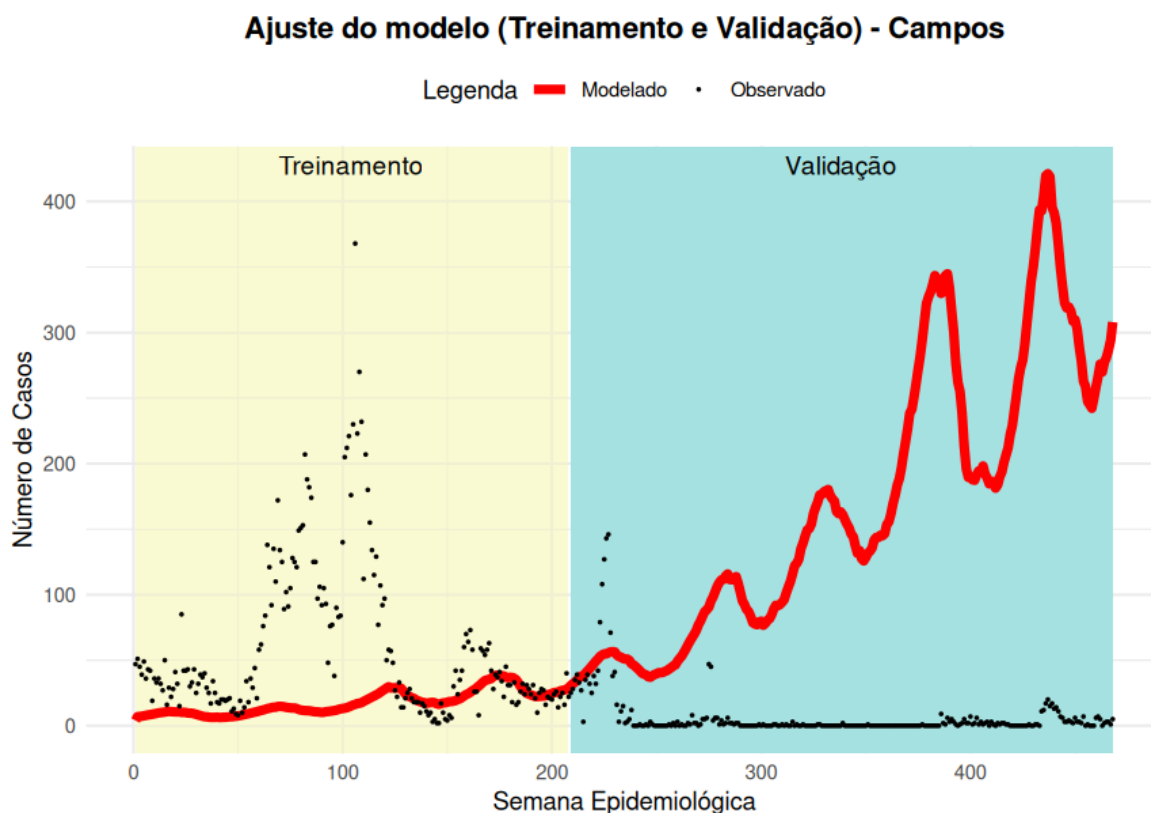


Figura 6.16: Modelo calibrado com agregamento semanal para o município de Campos dos Goytacazes. Comparação entre período de treinamento (2014–2017) e período de validação (2018–2022).

O fato de o modelo encontrar boa aproximação durante o período de treinamento, mas um ajuste menos aproximado durante o período de validação, não é necessariamente uma falha do modelo. Trata-se de uma dificuldade comum em modelos que testam suas capacidades preditivas. A modelagem de doenças infecciosas é complexa e mediada por uma série de determinantes, que têm efeitos variados em diferentes níveis. Fatores não incluídos no modelo, como mudanças no comportamento da população ou a aplicação de medidas de controle, podem ter causado efeitos expressivos na incidência observada, distanciando o modelo da realidade.

Além disso, outros pontos relevantes devem ser considerados quanto aos efeitos das características individuais de cada município nos resultados obtidos. Cidades maiores e com maior PIB, como o Rio de Janeiro, podem ter sistemas mais eficientes de notificação

de casos, o que diminui o ruído dos dados gerais. Por outro lado, cidades menores, como Paraty, podem sofrer com subnotificação.

Outro ponto é quanto à densidade populacional. Uma suposição básica adotada na modelagem de doenças infecciosas é de que há uma mistura homogênea e probabilidade igual de transmissão da doença. Municípios maiores podem ser mais aptos à utilização deste modelo SIR-SI-SI, enquanto os municípios menores podem sofrer mais com a falta de homogeneidade e facilidade de transmissão da doença.

6.4.3 Perspectivas de incorporação de parâmetros de saneamento básico no modelo

Assim como debatido no capítulo 4.2.2, modelos matemáticos podem ser subdivididos a partir das suas finalidades: modelos preditivos buscam prever casos futuros, enquanto modelos de extração de informação buscam inferir efeitos externos na disseminação da doença, com base em dados passados. Nesse sentido, o presente trabalho debruçou-se sobre ambas as perspectivas, tentando compreender o passado, ao passo em que buscou analisar também a sua aplicabilidade para condições futuras.

A escolha de variáveis a serem calibradas está diretamente ligada às informações às quais intenciona-se obter com a aplicação do modelo. As variáveis calibradas neste trabalho foram escolhidas por possuírem relações diretas ou indiretas com os efeitos das condições de saneamento básico no ciclo de vida e disseminação da dengue. Dentre elas, a capacidade de carga (K) e a taxa de sobrevivência aquática (sa_t) se sobressaem como as mais diretamente relacionadas.

A capacidade de carga (K) representa o número máximo de mosquitos que o ambiente consegue sustentar. Este parâmetro está intrinsecamente ligado à disponibilidade de criadouros para os mosquitos. Condições mais precárias de saneamento, como o armazenamento inadequado de água, acúmulo de resíduos sólidos e ineficácia de sistemas de drenagem urbana podem aumentar a disponibilidade de locais para oviposição dos mosquitos *Aedes aegypti*, e, conseqüentemente, resultar em um K mais elevado. Outra condicionante para o crescimento de K é a taxa de urbanização, que pode atuar como fator de confundimento sobre o efeito do saneamento básico na incidência da doença (ver capítulos 6.2 e 6.3.3)

De forma análoga, a taxa de sobrevivência aquática (sa_t) representa a probabilidade das larvas e pupas sobreviverem até a fase adulta. A quantidade e qualidade da água disponível nos criadouros desempenham um papel essencial para a manutenção e sobrevivência das fases iniciais dos *Aedes*. Assim, esta variável também pode possuir

correlações significativas e ser modulada por condições precárias de saneamento, como as descritas no parágrafo anterior, que disponibilizam condições propícias para as fases iniciais do ciclo de vida do mosquito.

Análises quantitativas correlacionando os valores calibrados no presente trabalho com as condições de saneamento devem ser feitas com cuidado. Os valores mais altos de K obtidos com a calibração mensal foram destinados aos municípios do Rio de Janeiro e de Duque de Caxias, o que é coerente, considerando-se suas características de urbanização e saneamento. Por outro lado, o município que recebeu maior valor de K após a calibração semanal foi Paraty, enquanto o Rio de Janeiro e Caxias receberam valores mais baixos. Isso ocorreu devido ao efeito de equifinalidade apresentado anteriormente, uma vez que o modelo chegou a resultados semelhantes mesmo com a adoção de valores distintos para os parâmetros, o que inviabiliza o uso desses dados de forma direta para as análises aqui propostas.

Considerando todos os debates apresentados até aqui, embora os resultados não permitam fazer associações quantitativas diretas entre os indicadores de saneamento e o ciclo da vida do mosquito, eles demonstram a grande complexidade de relações existentes. Assim, o debate realizado sobre K e sa_t demonstra caminhos promissores para investigações futuras que busquem evoluir modelos mecanísticos ao incluir efeitos das condições de saneamento básico na disseminação de arboviroses. Isso abre portas para a utilização de modelos mecanísticos não apenas para inferência, mas também para previsão e avaliação de políticas públicas de saneamento básico e saúde.

A modelagem foi limitada pela sua dependência exclusiva das variáveis climáticas utilizadas. Uma simulação mais eficiente da magnitude dos surtos poderia ser obtida através da incorporação de fatores comportamentais (como a mobilidade urbana e migração de pessoas), socioeconômicos (desigualdade social em escala local), ligados à vigilância sanitária (eficiência de práticas de controle vetorial) ou mesmo de imunização da população ao vírus, em um contexto de início de vacinação.

Nesse contexto, o conjunto de resultados encontrados cumpre o objetivo do trabalho de aprofundar a compreensão sobre relações entre dengue, saneamento básico e perspectivas de aprimoramento de modelagens mecanísticas. No capítulo seguinte, a etapa final do presente trabalho, são apresentadas as conclusões e recomendações para trabalhos futuros.

7. CONSIDERAÇÕES FINAIS E RECOMENDAÇÕES

Este trabalho se propôs a investigar as complexas relações existentes entre condições de saneamento básico, variáveis climáticas e a epidemiologia da dengue no estado do Rio de Janeiro. Além de quantificar as relações, também foram feitas análises qualitativas buscando nutrir um debate sobre vias para incorporação de determinantes de saneamento em modelos mecanísticos da doença.

A metodologia dupla desenvolvida para o reconhecimento e análise de determinantes da dengue, com aplicação de métodos estatísticos e mecanísticos, foi bem-sucedida. O trabalho seguiu uma linha de desenvolvimento lógica, em que padrões da doença foram observados nas análises espaço-temporais, a dimensão de variáveis de saneamento foi reduzida na análise de componentes principais (PCA) e, por fim, um modelo mecanístico de dengue foi desenvolvido.

Dentre os principais resultados obtidos com as análises estatísticas exploratórias, pontuam-se as relações diretas observadas entre a incidência de dengue e condições climáticas nos municípios, com picos da doença acompanhando períodos de maior temperatura e precipitação. Outro resultado relevante foi quanto aos padrões espaciais de epidemias de dengue: foram observados anos em que houve padrões estatísticos de agrupamento das incidências, indicando que epidemias de dengue estavam focadas em regiões específicas do estado. Apesar disso, uma análise com recorte temporal maior demonstrou instabilidade nos agrupamentos, onde municípios assolados pela doença em um período tiveram redução nos casos no período seguinte. Além de fatores de estocasticidade, uma justificativa elaborada para esse fenômeno é a dinâmica reativa de políticas de controle da doença.

Para as análises de correlação, o principal resultado foi a observação de um padrão paradoxal, em que melhores condições de saneamento foram associadas à maior incidência de dengue. Nesse contexto, elaborou-se uma hipótese de que a urbanização funcionou como fator de confusão entre as variáveis.

A aplicação da PCA reduziu a dimensionalidade dos 41 indicadores pré-selecionados de saneamento básico do SINISA em um conjunto menor de componentes sem grandes perdas na capacidade de explicação das variáveis. Com isso, foram reveladas dimensões latentes de cada um dos eixos do SINISA, sendo os três com maior explicação da variância: “Mão de Obra Operacional” (32,3%), “Eficiência da Rede de Água” (30,6%) e “Densidade Urbana e cursos d’água” (20,6%), para os eixos Resíduos Sólidos, Água e Esgoto e Águas Pluviais, respectivamente.

O desenvolvimento do modelo mecanístico SIR-SI-SI foi bem-sucedido e sua calibração em 6 municípios do estado do Rio de Janeiro obteve resultados satisfatórios. O modelo demonstrou ser uma ferramenta eficaz na reprodução da sazonalidade da dengue, alinhando picos simulados com os observados. Em relação aos surtos epidêmicos ocorridos ao longo da série histórica, o modelo consistentemente subestimou a magnitude dos picos, tanto no período de treinamento como na validação, o que é justificado pela própria estratégia metodológica do modelo, composto apenas por variáveis climáticas.

Na comparação entre períodos de treinamento e de validação, o modelo foi capaz de acompanhar os picos de casos com resultados satisfatórios em alguns dos municípios, ao passo em que tendências irrealistas foram observadas em outros municípios, justificadas pela captação de tendências com ajuste excessivamente estrito durante o período de treinamento.

Os debates expostos acima revelam um ponto central do trabalho: o uso de variáveis climáticas em modelos mecanísticos é muito útil para a captação da sazonalidade da dengue, mas é insuficiente na captação dos picos de casos. Para uma captura da mecânica ocorrida em surtos de forma eficiente, outros fatores devem ser adotados. Para capturar de forma mais eficiente a mecânica dos surtos, outros fatores podem ser considerados, como por exemplo: dinâmicas de mobilidade urbana e migração intermunicipal, a heterogeneidade socioeconômica em uma escala local, a eficiência de práticas de controle vetorial e avanços na imunização populacional através da vacinação.

Outra contribuição relevante do modelo não está na sua capacidade de previsão de casos futuros, mas sim na sua capacidade de inferência. A identificação dos parâmetros “Capacidade de Carga” (K) e Taxa de Sobrevivência Aquática (sa_t) pode ser a chave para o debate sobre a incorporação de determinantes de saneamento em modelos mecanísticos da dengue. Dessa forma, o presente trabalho contribui para a proposição de um arcabouço para a análise da doença de forma mais ampla.

Ao considerar variáveis ou dimensões latentes do saneamento básico (como as reveladas pela PCA) em sua estrutura, um modelo mecanístico pode tornar-se não apenas uma ferramenta de inferência com base em condições passadas, mas também simular e avaliar impactos de políticas públicas de saneamento básico na saúde.

Apesar dos resultados alcançados, é fundamental reconhecer as diversas limitações presentes neste trabalho. O uso de dados secundários, tanto do SINISA (com dados faltantes, autodeclarados, que não consideram soluções descentralizadas e possuem grande variabilidade interna), quanto do DATASUS (sujeitos à subnotificação) pode ter efeitos

expressivos nos resultados e conclusões desenvolvidas. Esta é uma limitação inerente a estudos ecológicos de larga escala, que utilizam grandes quantidades de dados secundários em sua composição.

Outra limitação metodológica fundamental está na escala de agregação dos dados utilizados. A análise em nível municipal, embora necessária para compreender as condições da doença em uma escala mais ampla, oculta a heterogeneidade existente na escala intraurbana, tornando a análise sujeita à falácia ecológica.

A estrutura adotada para o modelo mecanístico também foi uma limitação do trabalho. Como debatido ao longo do trabalho, não existe um modelo perfeito, e todos os modelos são simplificações da realidade. O modelo foi elaborado conforme as práticas da literatura, mas fatores significativos, como a circulação de múltiplos sorotipos de dengue e o efeito de outras variáveis socioeconômicas não foram considerados. Além disso, sua construção foi feita com base em outros modelos existentes elaborados para realidades diferentes da delimitada neste estudo.

Com base nos resultados, limitações e percepções autorais ao longo do desenvolvimento do trabalho, algumas sugestões e recomendações para trabalhos futuros foram elaboradas e estão apresentadas a seguir. Uma lista em formato de roteiro, com sugestões mais detalhadas, está presente no apêndice F.

Como recomendação para trabalhos futuros, sugere-se uma análise de sensibilidade focada nos parâmetros que conectam o modelo aos determinantes ambientais. Uma abordagem promissora seria fixar os demais parâmetros biológicos com base nos valores da literatura e testar exclusivamente as variações na taxa de sobrevivência aquática (sa) e na capacidade de carga (K), permitindo isolar e quantificar de forma mais precisa o impacto do clima e do saneamento, respectivamente.

A partir da constatação de que o clima, por si só, não explica a magnitude dos picos epidêmicos, recomenda-se, também, que trabalhos futuros investiguem mais a fundo a incorporação de outras variáveis nos modelos mecanísticos. Fatores como a mobilidade urbana (e importação de casos), indicadores socioeconômicos em escala intraurbana e a eficiência de ações de controle vetorial são candidatos promissores. A incorporação dessas variáveis poderia se dar, inclusive, através de parâmetros já utilizados no modelo, como a taxa de transmissão e a capacidade de carga, transformando-os em funções dinâmicas com o tempo, ao invés de valores fixos.

Uma das principais contribuições da análise espacial deste trabalho foi a revelação de um padrão de agrupamentos de dengue fracos e geograficamente instáveis ao longo do

tempo. Uma hipótese levantada para explicar essa dinâmica é a atuação reativa das políticas de controle vetorial, que concentram esforços nos locais mais afetados por epidemias recentes, deslocando o epicentro da doença para outras áreas nos anos seguintes. Isso abre uma interessante e relevante linha de pesquisa futura: a realização de um estudo que cruze os dados epidemiológicos com dados de investimentos e de ações de vigilância em saúde de cada município. Tal análise poderia verificar se os municípios com picos epidêmicos em um determinado ano de fato recebem maiores recursos e intensificam suas ações no ano seguinte, e se isso se correlaciona com uma subsequente redução de casos e com o deslocamento geográfico da doença, validando quantitativamente a hipótese aqui levantada. Um estudo dessa natureza poderia fornecer subsídios para a avaliação da efetividade de políticas municipais reativas, a longo prazo, em comparação a estratégias integradas de vigilância e controle da doença.

Uma evolução sugerida para o modelo seria quanto à sua capacidade explicativa para predição de casos em cenários futuros. Mesmo que modelos epidemiológicos, em geral, busquem quantificar parâmetros com base em resultados passados, a consolidação de um modelo com certa capacidade preditiva poderia ser útil para estimar efeitos de melhorias em serviços de saneamento básico (como expansão da cobertura de esgoto ou a redução de perdas de água) na diminuição de casos da doença. Com essa abordagem, o modelo poderia se tornar uma ferramenta aplicável para o planejamento e a avaliação de políticas públicas de saúde e saneamento. Além disso, por incorporar variáveis climáticas em suas equações, o modelo mecanístico poderia ser aplicado para simular impactos de cenários distintos decorrentes da emergência climática, como aumentos nas temperaturas médias, ocorrência de ondas de calor e intensificação de chuvas, nas dinâmicas da doença.

Este trabalho debateu e demonstrou as complexas relações existentes entre saneamento, clima e a epidemiologia de dengue no contexto do Rio de Janeiro. Os resultados obtidos demonstram que não há uma solução simplificada para um problema multifatorial que é a disseminação de uma doença infecciosa; são necessárias abordagens multidisciplinares para sua compreensão e controle. Espera-se que a metodologia desenvolvida e os resultados obtidos no presente trabalho possam contribuir para o avanço do conhecimento, apoiando na construção de ferramentas ainda mais eficazes para vigilância e proposição de soluções de saúde coletiva.

8. CONCLUSÕES

Esta pesquisa investigou relações entre saneamento básico e variáveis climáticas com a epidemiologia da dengue em municípios do estado do Rio de Janeiro. Foi aplicada uma abordagem mista — estatística e mecanística — para analisar e modelar tais relações. A partir dos resultados obtidos, conclui-se que:

1. As variáveis climáticas são determinantes fundamentais para explicar o padrão sazonal da dengue, com picos de incidência associados a períodos de maior temperatura e precipitação;
2. Somente a incorporação de variáveis climáticas em modelos mecanísticos não é suficiente para captar a dinâmica epidêmica completa da dengue, não conseguindo reproduzir a magnitude dos picos de casos. Como exemplo de outras variáveis a serem utilizadas, foram levantadas: dinâmicas de mobilidade urbana e migração intermunicipal, entrada de pessoas infectadas no município pelo turismo, aplicação de métodos de controle, imunização populacional, dentre outros;
3. A análise espacial revelou padrões geográficos de agrupamento da incidência da dengue em diferentes municípios, ao longo do tempo, sugerindo que uma análise em escala intermunicipal também pode ser útil na compreensão das dinâmicas da doença;
4. Foi identificado um padrão paradoxal na análise de correlações entre incidência de dengue e indicadores de saneamento básico, onde melhores condições de saneamento foram positivamente associadas à maior incidência. Elaborou-se uma hipótese de que a urbanização atuou como variável confundidora neste contexto;
5. A aplicação da análise de componentes principais (PCA) obteve sucesso em reduzir a dimensionalidade das variáveis de saneamento básico, revelando dimensões latentes do saneamento como “Mão de Obra Operacional”, “Eficiência (índices de perdas) da Rede de Água” e “Densidade Urbana e cursos d’água”.
6. Os parâmetros Capacidade de Carga (K) e Taxa de Sobrevivência Aquática (sa_t) foram identificados como promissores para incorporação de determinantes de saneamento em modelos mecanísticos.

De forma integrada, os resultados confirmam a hipótese inicial de que o saneamento básico atua como modulador dos efeitos do clima na dinâmica da dengue. Embora as variáveis climáticas tenham se mostrado fundamentais para explicar a sazonalidade da doença, evidenciou-se que apenas elas são insuficientes para capturar a

magnitude dos surtos epidêmicos. A identificação de parâmetros-chave, como a Capacidade de Carga (K) e a Taxa de Sobrevivência Aquática (sa), demonstra que dimensões de saneamento — expressas por indicadores de eficiência da rede de água, manejo de resíduos e drenagem urbana — podem ser incorporadas de forma explícita em modelos mecanísticos. Essa constatação abre oportunidades para que gestores públicos utilizem tais modelos como ferramentas de apoio à decisão, permitindo simular cenários de expansão de redes de esgoto, melhoria da coleta de resíduos ou fortalecimento da drenagem pluvial e, assim, antecipar impactos potenciais na redução da incidência da doença. Ao evidenciar essa conexão entre clima, saneamento e risco epidemiológico, este trabalho oferece subsídios concretos para políticas integradas de saúde e saneamento, fortalecendo a vigilância e o planejamento de ações preventivas em escala municipal.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABDULLAH, N. A. M. H. *et al.* The association between dengue case and climate: A systematic review and meta-analysis. **One Health**, [s. l.], v. 15, p. 100452, 2022.
- AGUIAR, M. *et al.* Mathematical models for dengue fever epidemiology: A 10-year systematic review. **Physics of Life Reviews**, [s. l.], v. 40, p. 65–92, 2022.
- ALKHALDY, I. Modelling the association of dengue fever cases with temperature and relative humidity in Jeddah, Saudi Arabia—A generalised linear model with break-point analysis. **Acta Tropica**, [s. l.], v. 168, p. 9–15, 2017.
- ALMEIDA, L. S.; COTA, A. L. S.; RODRIGUES, D. F. Saneamento, Arboviroses e Determinantes Ambientais: impactos na saúde urbana. **Ciência & Saúde Coletiva**, [s. l.], v. 25, n. 10, p. 3857–3868, 2020.
- ALTO, B. W.; JULIANO, S. A. Precipitation and temperature effects on populations of *Aedes albopictus* (Diptera: Culicidae): implications for range expansion. **Journal of Medical Entomology**, [s. l.], v. 38, n. 5, p. 646–656, 2001.
- ALVES, L. D. **Weather-driven mathematical models of dengue transmission dynamics in twelve Brazilian sites**. 2021. Thesis[s. l.], 2021. Disponível em: <https://www.arca.fiocruz.br/handle/icict/52536>. Acesso em: 14 out. 2024.
- ANDRADE, A. C. *et al.* Are Land Use and Cover Changes and Socioeconomic Factors Associated with the Occurrence of Dengue Fever? A Case Study in Minas Gerais State, Brazil. **Resources**, [s. l.], v. 13, n. 3, p. 38, 2024.
- ANDRADE, C. D. R.; LOPES, G. A. H. Brasil República: uma história de surtos, pandemias e epidemias. **Boletim de Conjuntura (BOCA)**, [s. l.], v. 5, n. 14, p. 70–92, 2021.
- ANSELIN, L. Local Indicators of Spatial Association—LISA. **Geographical Analysis**, [s. l.], v. 27, n. 2, p. 93–115, 1995.
- ANSELIN, L.; LI, X.; KOSCHINSKY, J. GeoDa, From the Desktop to an Ecosystem for Exploring Spatial Data. **Geographical Analysis**, [s. l.], v. 54, n. 3, p. 439–466, 2022.
- BADII, M. H. *et al.* Ecología e historia del dengue en las Américas. **Daena: Internacional J Good Consciente**, [s. l.], v. 2, 2007. Disponível em: <https://philpapers.org/rec/BADEEH>. Acesso em: 4 jun. 2025.
- BARCELLOS, C. *et al.* Climate change, thermal anomalies, and the recent progression of dengue in Brazil. **Scientific Reports**, [s. l.], v. 14, n. 1, p. 5948, 2024.
- BHATT, S. *et al.* The global distribution and burden of dengue. **Nature**, [s. l.], v. 496, n. 7446, p. 504–507, 2013.
- BONITA, R.; BEAGLEHOLE, R.; KJELLSTRÖN, T. **Epidemiologia Básica**. 2. ed. Santos: [s. d.], 2010.

BORGES, M. C. P. Desenvolvimento de um índice de vulnerabilidade urbana a alagamentos em municípios brasileiros a partir dos dados do SNIS. [s. l.], 2024. Disponível em: <http://repositorio.unb.br/handle/10482/49301>. Acesso em: 22 jun. 2025.

BORGES, M. C. P. *et al.* The Brazilian National System for Water and Sanitation Data (SNIS): Providing information on a municipal level on water and sanitation services. **Journal of Urban Management**, [s. l.], v. 11, n. 4, p. 530–542, 2022.

BRASIL, M. da S. **NOTA INFORMATIVA Nº 13/2023-CGAR/DEDT/SVSA/MS**. [S. l.]: Secretaria de Vigilância em Saúde e Ambiente, 2023. Disponível em: <https://www.gov.br/saude/pt-br/centrais-de-conteudo/publicacoes/estudos-e-notas-informativas/2023/nota-informativa-no-13-2023-cgarb-dedt-svsa-ms>. Acesso em: 30 nov. 2023.

BRASIL. MINISTÉRIO DA SAÚDE. **Nota técnica número 1/2023**. Brasília, 2023. Disponível em: <https://www.gov.br/saude/pt-br/centrais-de-conteudo/publicacoes/notas-tecnicas/2023/nota-tecnica-no-01-2023-cgarb-deidt-svs-ms/view>.

BRASIL. MINISTÉRIO DA SAÚDE. **Você sabe o que é o Método Wolbachia?**. [S. l.], 2024. Disponível em: <https://www.gov.br/saude/pt-br/assuntos/saude-com-ciencia/noticias/2024/maio/voce-sabe-o-que-e-o-metodo-wolbachia>. Acesso em: 24 out. 2025.

BRASIL. MINISTÉRIO DAS CIDADES. **SNIS - Série Histórica**. [S. l.], 2025. Disponível em: <https://app4.cidades.gov.br/serieHistorica/>. Acesso em: 5 maio 2025.

BRAUER, F.; CASTILLO-CHAVEZ, C.; FENG, Z. **Mathematical Models in Epidemiology**. New York, NY: Springer New York, 2019. (Texts in Applied Mathematics). v. 69 Disponível em: <https://link.springer.com/10.1007/978-1-4939-9828-9>. Acesso em: 21 mar. 2024.

BREIMAN, L. Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author). **Statistical Science**, [s. l.], v. 16, n. 3, p. 199–231, 2001.

BUSS, P. M.; PELLEGRINI FILHO, A. A saúde e seus determinantes sociais. **Physis: revista de saúde coletiva**, [s. l.], v. 17, p. 77–93, 2007.

CAMPOS, G. L. de; PEDROSA, R. de C. de V.; CRISTIANO, C. T. Study of the Temporal Propagation of Arboviruses in the Region of Recife-PE: Analysis of Climatic Influence using the SIR Model and Recurrent Neural Networks. **Ibero-Latin American Congress on Computational Methods in Engineering (CILAMCE)**, [s. l.], 2024. Disponível em: <https://publicacoes.softaliza.com.br/cilamce/article/view/10164>. Acesso em: 7 jul. 2025.

CARLTON, E. J. *et al.* Regional disparities in the burden of disease attributable to unsafe water and poor sanitation in China. **Bulletin of the World Health Organization**, [s. l.], v. 90, n. 8, p. 578–587, 2012.

CHANPRASOPCHAI, P.; TANG, I. M.; PONGSUMPUN, P. SIR Model for Dengue Disease with Effect of Dengue Vaccination. **Computational and Mathematical Methods in Medicine**, [s. l.], v. 2018, p. 9861572, 2018.

COVIGSAL. **Boletim epidemiológico Mercosul**. [S. l.]: [s. d.], 2023. Disponível em: <https://www.gov.br/saude/pt-br/centrais-de-conteudo/publicacoes/boletins/epidemiologicos/especiais/2023/boletim-epidemiologico-mercosul-covigsal-portugues>.

DE SOUSA, S. C. *et al.* Factors associated with the occurrence of dengue epidemics in Brazil: a systematic review. **Revista Panamericana de Salud Pública**, [s. l.], v. 45, p. e84, 2021.

DE VASCONCELOS, A. S. V.; DE LIMA, J. S.; CARDOSO, R. T. N. Multiobjective optimization to assess dengue control costs using a climate-dependent epidemiological model. **Scientific Reports**, [s. l.], v. 13, n. 1, 2023.

DO CARMO, R. F. *et al.* Spatiotemporal dynamics, risk areas and social determinants of dengue in Northeastern Brazil, 2014-2017: an ecological study. **Infectious Diseases of Poverty**, [s. l.], v. 9, n. 1, p. 153, 2020.

DRUMOND, B. *et al.* Dinâmica espaço-temporal da dengue no Distrito Federal, Brasil: ocorrência e permanência de epidemias. **Ciência & Saúde Coletiva**, [s. l.], v. 25, p. 1641–1652, 2020.

FERNANDES, E. A.; SILVA, A. L. da. Saneamento básico e saúde: um estudo para o estado do Rio de Janeiro. **Revista de Desenvolvimento e Políticas Públicas**, [s. l.], v. 4, n. 2, p. 96–109, 2020.

GIMOND, M. **Intro to GIS and spatial analysis**. [S. l.], 2023. Disponível em: <https://mgimond.github.io/Spatial/index.html>.

GUTIÉRREZ, E. M. P.; VALLEJO, D. A. Z.; RIVAS, A. T. B. Metodología de canales endémicos del dengue en Ecuador 2015-2020: Necesidad para planificar y administrar la salud pública. **Bol. malariol. salud ambient**, [s. l.], p. 105–111, 2021.

HELLER, L. **Melhoramento dos serviços de água e saneamento é a resposta ao Zika vírus, diz Léo Heller**. entrevistador: Bruno C. Dias. Rio de Janeiro: ABRASCO, 11 mar. 2016. Disponível em: https://abrasco.org.br/zika_saneamento_leo_heller/. Acesso em: 24 jul. 2024.

IBARRARÁN, M. E. *et al.* Climate Change and Human Health in Mexico: Public Health Trends and Government Strategies. In: AKHTAR, R. (org.). **Climate Change and Human Health Scenarios**. Cham: Springer Nature Switzerland, 2023. (Global Perspectives on Health Geography). p. 399–416. Disponível em: https://link.springer.com/10.1007/978-3-031-38878-1_25. Acesso em: 4 jun. 2025.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Censo Brasileiro de 2022**. [S. l.]: IBGE, 2022.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Censo demográfico 2010. **IBGE: Instituto Brasileiro de Geografia e**, [s. l.], p. 34–35, 2010.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Produto Interno Bruto dos Municípios**. [S. l.]: IBGE, 2021. Disponível em: <https://www.ibge.gov.br/estatisticas/economicas/contas-nacionais/9088-produto-interno-bruto-dos-municipios.html?=&t=downloads>.

INMET. **BDMEP - Dados Históricos**. [S. l.], 2025. Disponível em: <https://portal.inmet.gov.br/servicos/bdmep-dados-hist%C3%B3ricos>. Acesso em: 19 abr. 2025.

INSTITUTO ÁGUA E SANEAMENTO. **Municípios e Saneamento: Fontes e Notas Técnicas**. [S. l.], 2024. Disponível em: <https://www.aguaesaneamento.org.br/municipios-e-saneamento/fontes-notas-tecnicas>. Acesso em: 19 jun. 2025.

JIN, J.; QUINN, B. K.; SHI, P. The Modified Brière Equation and Its Applications. **Plants**, [s. l.], v. 11, n. 13, p. 1769, 2022.

KATZELNICK, L. C. *et al.* Antibody-dependent enhancement of severe dengue disease in humans. **Science**, [s. l.], v. 358, n. 6365, p. 929–932, 2017.

LAPPICY, T.; BELTRÃO, D. A. S. P.; BIGONHA, M. B. Relação entre saneamento básico e a taxa de mortalidade por COVID-19: um estudo de caso do Rio Grande do Sul, Brasil. **Cadernos Técnicos Engenharia Sanitária e Ambiental**, [s. l.], v. 3, n. 2, p. 105–114, 2023.

LOWE, R. *et al.* Combined effects of hydrometeorological hazards and urbanisation on dengue risk in Brazil: a spatiotemporal modelling study. **The Lancet Planetary Health**, [s. l.], v. 5, n. 4, p. e209–e219, 2021.

MACIEL-DE-FREITAS, R.; VALLE, D. Challenges encountered using standard vector control measures for dengue in Boa Vista, Brazil. **Bulletin of the World Health Organization**, [s. l.], v. 92, n. 9, p. 685–689, 2014.

MENDONÇA, T. T. **Ciência, Tecnologia e Sociedade em prol de um saneamento sustentável: um olhar para duas comunidades no Rio de Janeiro**. 2022. 83 f. TCC (graduação) - Universidade Federal de Santa Catarina, [s. l.], 2022. Disponível em: <https://repositorio.ufsc.br/handle/123456789/243393>.

MESSINA, J. P. *et al.* Global spread of dengue virus types: mapping the 70 year history. **Trends in Microbiology**, [s. l.], v. 22, n. 3, p. 138–146, 2014.

MESSINA, J. P. *et al.* The current and future global distribution and population at risk of dengue. **Nature Microbiology**, [s. l.], v. 4, n. 9, p. 1508–1515, 2019.

MORDECAI, E. A. *et al.* Detecting the impact of temperature on transmission of Zika, dengue, and chikungunya using mechanistic models. **PLOS Neglected Tropical Diseases**, [s. l.], v. 11, n. 4, p. e0005568, 2017.

MS - MINISTÉRIO DE SAÚDE. **Informações de Saúde (TABNET) – DATASUS**. [S. l.], [s. d.]. Disponível em: <https://datasus.saude.gov.br/informacoes-de-saude-tabnet/>. Acesso em: 16 jun. 2024.

MULLEN, K. *et al.* DEoptim: An R Package for Global Optimization by Differential Evolution. **Journal of Statistical Software**, [s. l.], v. 40, n. 6, 2011. Disponível em: <http://www.jstatsoft.org/v40/i06/>. Acesso em: 17 jul. 2025.

NOVAES, C.; SILVA PINTO, F.; MARQUES, R. C. *Aedes Aegypti*—Insights on the Impact of Water Services. **GeoHealth**, [s. l.], v. 6, n. 11, p. e2022GH000653, 2022.

NURAINI, N. *et al.* Climate-based dengue model in Semarang, Indonesia: Predictions and descriptive analysis. **Infectious Disease Modelling**, [s. l.], v. 6, p. 598–611, 2021.

OGUNLADE, S. T. *et al.* A Systematic Review of Mathematical Models of Dengue Transmission and Vector Control: 2010–2020. **Viruses**, [s. l.], v. 15, n. 1, p. 254, 2023.

OLIVEIRA, J. B. *et al.* Paradox between adequate sanitation and rainfall in dengue fever cases. **Science of The Total Environment**, [s. l.], v. 860, p. 160491, 2023.

OLIVEIRA, E. A.; LIMA, C. H. R.; CARDOSO, T. Uso do Banco de Dados do Sistema Nacional de Informações sobre Saneamento (SNIS): uma revisão meta analítica. [s. l.], 2023.

ONU - NAÇÕES UNIDAS. **OMS: Para cada dólar investido em água e saneamento, economiza-se 4,3 dólares em saúde global.** [S. l.], 2014. Disponível em: <https://brasil.un.org/pt-br/55290-oms-para-cada-d%C3%B3lar-investido-em-%C3%A1gua-e-saneamento-economiza-se-43-d%C3%B3lares-em-sa%C3%BAde-global>, <https://brasil.un.org/pt-br/55290-oms-para-cada-d%C3%B3lar-investido-em-%C3%A1gua-e-saneamento-economiza-se-43-d%C3%B3lares-em-sa%C3%BAde-global>. Acesso em: 20 jun. 2025.

PINHO, G. S. D. **Análise da influência de fatores climáticos e sanitários na incidência de dengue na Região Metropolitana de Goiânia.** 2023. Monografia de Projeto Final - Departamento de Engenharia Civil e Ambiental, Universidade de Brasília, Brasília, 2023.

PINHO, G. S. de; MENDONÇA, T. T.; LIMA, C. H. R. DENGUE AND SOCIAL VULNERABILITY IN THE FEDERAL DISTRICT: A HISTORICAL STUDY AND THE RELATIONSHIP WITH THE PEAK OF CASES IN 2024. **Boletim de Conjuntura (BOCA)**, [s. l.], v. 19, n. 57, p. 483–506, 2024.

QUEIROZ, J.; SILVA, P. N. Dengue: por que ignora-se o papel do saneamento? *In: OUTRAS PALAVRAS.* 19 jul. 2024. Disponível em: <https://outraspalavras.net/outrasaude/dengue-por-que-ignora-se-o-papel-do-saneamento/>. Acesso em: 24 jul. 2024.

QUEIROZ, J. T. M. de; SILVA, P. N.; HELLER, L. Novos pressupostos para o saneamento no controle de arboviroses no Brasil. **Cadernos de Saúde Pública**, [s. l.], v. 36, p. e00223719, 2020.

REINHOLD, J. M.; LAZZARI, C. R.; LAHONDÈRE, C. Effects of the Environmental Temperature on *Aedes aegypti* and *Aedes albopictus* Mosquitoes: A Review. **Insects**, [s. l.], v. 9, n. 4, p. 158, 2018.

RIBEIRO, M. S. *et al.* Índices larvais de *Aedes aegypti* e incidência de dengue: um estudo ecológico no Estado do Rio de Janeiro, Brasil. **Cadernos de Saúde Pública**, [s. l.], v. 37, p. e00263320, 2021.

SALDANHA, R. D. F.; BASTOS, R. R.; BARCELLOS, C. Microdatasus: pacote para download e pré-processamento de microdados do Departamento de Informática do SUS (DATASUS). **Cadernos de Saúde Pública**, [s. l.], v. 35, n. 9, p. e00032419, 2019.

SALLES, T. S. *et al.* History, epidemiology and diagnostics of dengue in the American and Brazilian contexts: a review. **Parasites & Vectors**, [s. l.], v. 11, n. 1, p. 264, 2018.

SANTOS, C. A. G. *et al.* Correlation of dengue incidence and rainfall occurrence using wavelet transform for João Pessoa city. **Science of The Total Environment**, [s. l.], v. 647, p. 794–805, 2019.

SEBASTIANELLI, A. *et al.* A reproducible ensemble machine learning approach to forecast dengue outbreaks. **Scientific Reports**, [s. l.], v. 14, n. 1, p. 3807, 2024.

SENHORAS, E. M. CORONAVÍRUS E O PAPEL DAS PANDEMIAS NA HISTÓRIA HUMANA. **Boletim de Conjuntura (BOCA)**, [s. l.], v. 1, n. 1, p. 29–32, 2020.

SILVA, W. L. **Caracterização Climatológica e Mudanças Climáticas no Estado do Rio de Janeiro**. 2014. 95 f. Dissertação de Mestrado - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2014. Disponível em: <http://objdig.ufrj.br/10/dissert/821727.pdf>.

SILVA, J. S.; MARIANO, Z. de F.; SCOPEL, I. A DENGUE NO BRASIL E AS POLÍTICAS DE COMBATE AO AEDES AEGYPTI: DA TENTATIVA DE ERRADICAÇÃO ÀS POLÍTICAS DE CONTROLE. **Hygeia - Revista Brasileira de Geografia Médica e da Saúde**, [s. l.], v. 4, n. 6, p. 163–175, 2008.

TAYLOR, L. **Dengue fever: Brazil rushes out vaccine as climate change fuels unprecedented surge**. [S. l.]: British Medical Journal Publishing Group, 2024. Disponível em: <https://www.bmj.com/content/384/bmj.q483.abstract>. Acesso em: 6 jun. 2025.

TEIXEIRA, M. da G.; BARRETO, M. L. Porque Devemos, de Novo, Erradicar o *Aedes Aegypti*. **Ciência & Saúde Coletiva**, [s. l.], v. 1, p. 122–136, 1996.

WHO. **Dengue and severe dengue**. [S. l.], 2023. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue>. Acesso em: 9 out. 2023.

WHO. **Guidelines on sanitation and health**. Geneva: World Health Organization, 2018. Disponível em: <https://apps.who.int/iris/handle/10665/274939>. Acesso em: 26 jul. 2023.

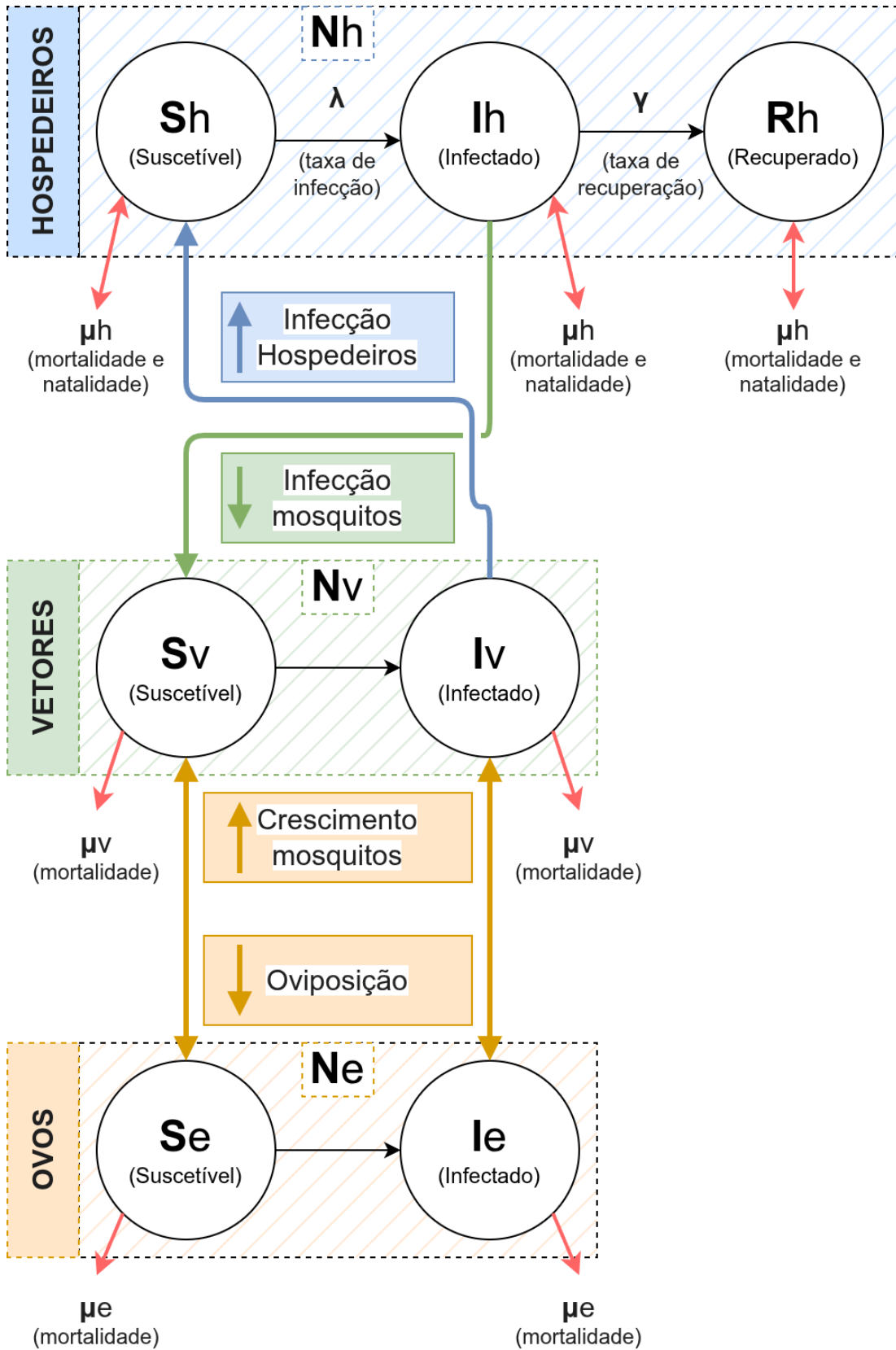
WHO. **Ten threats to global health in 2019**. [S. l.], 2019. Disponível em: <https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019>. Acesso em: 19 nov. 2023.

WONG, J. M. *et al.* Dengue: A Growing Problem With New Interventions. **Pediatrics**, [s. l.], v. 149, n. 6, p. e2021055522, 2022.

ZARA, A. L. de S. A. *et al.* Estratégias de controle do *Aedes aegypti*: uma revisão. **Epidemiologia e Serviços de Saúde**, [s. l.], v. 25, n. 2, p. 391–404, 2016.

APÊNDICES

APÊNDICE A - DIAGRAMA DETALHADO DO MODELO SIR-SI-SI



APÊNDICE B - PARÂMETROS SELECIONADOS PARA ANÁLISE DE COMPONENTES PRINCIPAIS

SELEÇÃO FINAL - RESÍDUOS SÓLIDOS			
Grupo indicador	Indicador	Nome	Unidade
Indicadores sobre despesas e trabalhadores	IN001	Taxa de empregados em relação à população urbana	empreg/1000 hab
	IN002	Despesa média por empregado alocado nos serviços do manejo de rsu	R\$/empreg
	IN006	Despesa per capita com manejo de rsu em relação à população urbana	R\$/hab
Indicadores sobre coleta domiciliar e pública	IN014	Taxa de cobertura do serviço de coleta domiciliar direta (porta-a-porta) da população urbana do município.	%
	IN015	Taxa de cobertura regular do serviço de coleta de rdo em relação à população total do município	%
	IN016	Taxa de cobertura regular do serviço de coleta de rdo em relação à população urbana	%
	IN019	Taxa de empregados (coletadores + motoristas) na coleta (rdo + rpu) em relação à população urbana	empreg/1000 hab
	IN021	Massa coletada (rdo + rpu) per capita em relação à população urbana	Kg/hab/dia
	IN028	Massa de resíduos domiciliares e públicos (rdo+rpu) coletada per capita em relação à população total atendida pelo serviço de coleta	Kg/habitante/dia
Indicadores sobre serviços de varrição, capina e roçada	IN045	Taxa de varredores em relação à população urbana	empreg/1000 hab
	IN051	Taxa de capinadores em relação à população urbana	empreg/1000 hab

SELEÇÃO FINAL - ÁGUA E ESGOTO			
Grupo indicador	Indicador	Nome	Unidade
Indicadores operacionais - água	IN001	Densidade de economias de água por ligação	econ./lig.
	IN009	Índice de hidrometração	percentual
	IN013	Índice de perdas faturamento	percentual
	IN014	Consumo micromedido por economia	m3/mês/econ.
	IN022	Consumo médio per capita de água	l/hab./dia
	IN023	Índice de atendimento urbano de água	percentual
	IN025	Volume de água disponibilizado por economia	m3/mês/econ.
	IN049	Índice de perdas na distribuição	percentual
	IN050	Índice bruto de perdas lineares	m3/dia/Km
	IN051	Índice de perdas por ligação	l/lig./dia
	IN052	Índice de consumo de água	percentual
	IN053	Consumo médio de água por economia	m3/mês/econ.
	IN055	Índice de atendimento total de água	percentual
Indicadores operacionais - esgotos	IN015	Índice de coleta de esgoto	percentual
	IN046	Índice de esgoto tratado referido à água consumida	percentual
Indicadores de qualidade	IN075	Incidência das análises de cloro residual fora do padrão	percentual
	IN076	Incidência das análises de turbidez fora do padrão	percentual
	IN084	Incidência das análises de coliformes totais fora do padrão	percentual

SELEÇÃO FINAL - ÁGUAS PLUVIAIS			
Grupo indicador	Indicador	Nome	Unidade
Dados gerais (Geral)	IN042	Parcela de área urbana em relação à área total	Percentual
	IN043	Densidade Demográfica na Área Urbana	Pessoas por hectares
	IN044	Densidade de Domicílios na Área Urbana	Domicílios por hectares
Dados financeiros (Financeiro)	IN010	Participação da Despesa Total dos Serviços de Drenagem e Manejo das Águas Pluviais Urbanas na Despesa Total do Município	Percentual
	IN048	Despesa per capita com serviços de Drenagem e Manejo das Águas Pluviais Urbanas	Reais por habitante ano
Dados de infraestrutura (Infraestrutura)	IN020	Taxa de Cobertura de Pavimentação e Meio-Fio na Área Urbana do Município	Percentual
	IN021	Taxa de cobertura de vias públicas com redes ou canais pluviais subterrâneos na área urbana	Percentual
	IN026	Parcela de Cursos d'Água Naturais Perenes com Canalização Aberta	Percentual
	IN027	Parcela de Cursos d'Água Naturais Perenes com Canalização Fechada	Percentual
	IN029	Parcela de Cursos d'Água Naturais Perenes com Diques	Percentual
Dados sobre gestão de risco (Gestão de Riscos)	IN040	Parcela de Domicílios em Situação de Risco de Inundação	Percentual
	IN041	Parcela da População Impactada por Eventos Hidrológicos	Percentual

APÊNDICE C - RESULTADOS DAS ANÁLISES DE CORRELAÇÃO ENTRE VARIÁVEIS DE SANEAMENTO BÁSICO E O DESFECHO: “INCIDÊNCIA MÉDIA ANUAL” EM 82 MUNICÍPIOS DO RIO DE JANEIRO.

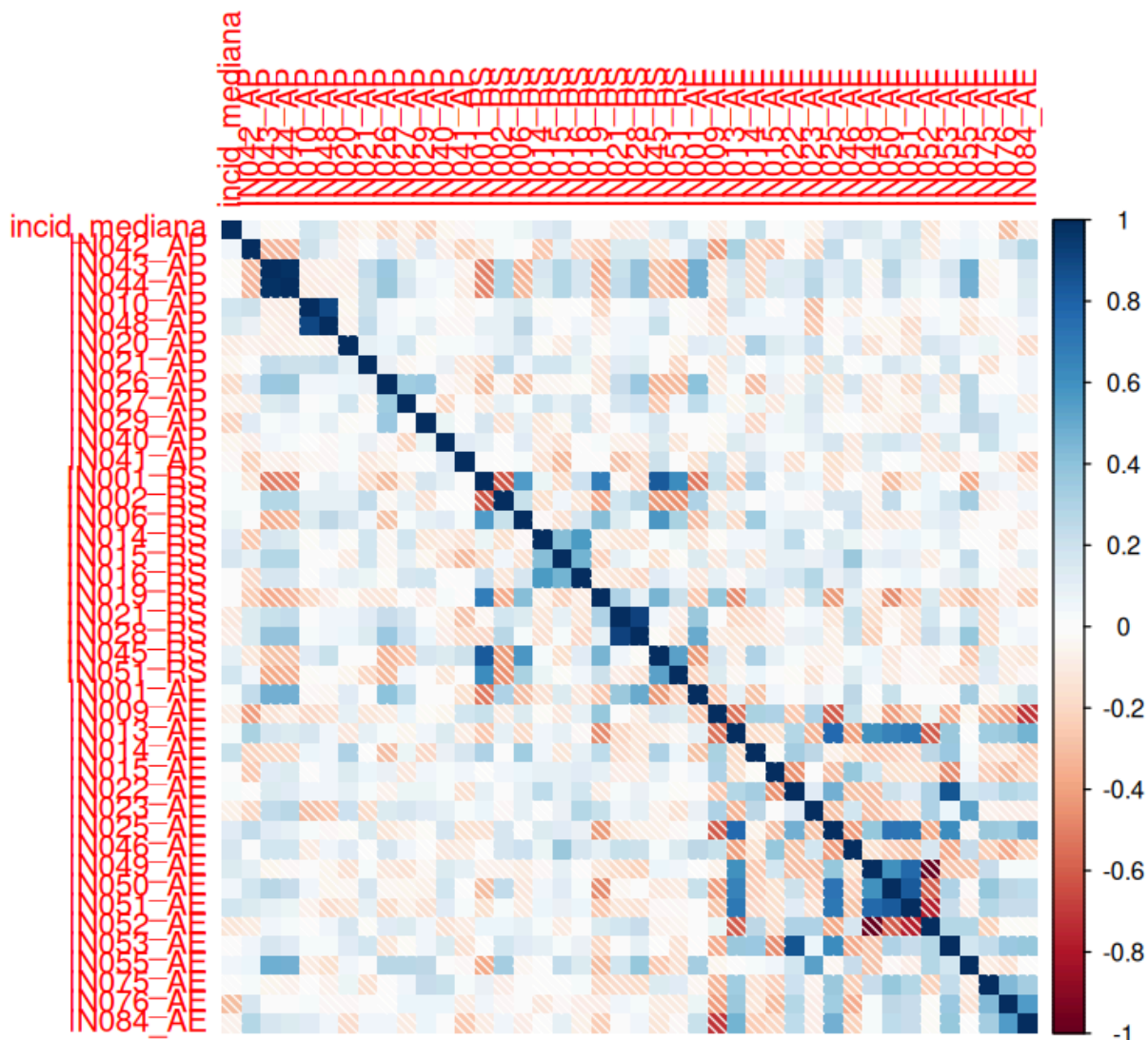


Figura C.1: Matriz de correlações (Pearson) de variáveis do SINISA e a incidência anual mediana de dengue.

Quadro C.1: Valores de correlação entre incidência média mensal e variáveis do SINISA (esquerda) e variáveis do IBGE (direita)

Variável	Correlação de Pearson (r)	Correlação de Spearman (ρ)
incid_mediana	1	1
IN042_AP	-0.0191	0.0573
IN043_AP	-0.1388	-0.0200

Indicador IBGE	Coefficiente de Spearman (ρ)
incid_mediana	1
IBGE_abast_Possui ligação à rede geral e a utiliza como forma principal	0.359
IBGE_abast_Possui ligação à rede	-0.299

Variável	Correlação de Pearson (r)	Correlação de Spearman (ρ)
IN044_AP	-0.1230	-0.0036
IN010_AP	0.0068	0.1747
IN048_AP	-0.0255	0.1342
IN020_AP	0.0170	-0.0927
IN021_AP	-0.1217	-0.0517
IN026_AP	-0.0683	-0.1822
IN027_AP	-0.0430	-0.0782
IN029_AP	-0.1846	-0.2111
IN040_AP	-0.0015	-0.0502
IN041_AP	-0.0896	-0.0319
IN001_RS	-0.0049	0.0396
IN002_RS	-0.0126	0.0390

Indicador IBGE	Coefficiente de Spearman (ρ)
geral, mas utiliza principalmente outra forma	
IBGE_abast_Possui ligação à rede geral, mas utiliza principalmente outra forma - Poço profundo ou artesiano	-0.288
IBGE_abast_Possui ligação à rede geral, mas utiliza principalmente outra forma - Poço raso, freático ou cacimba	-0.053
IBGE_abast_Possui ligação à rede geral, mas utiliza principalmente outra forma - Fonte, nascente ou mina	-0.294
IBGE_abast_Possui ligação à rede geral, mas utiliza principalmente outra forma - Carro-pipa	0.024
IBGE_abast_Possui ligação à rede geral, mas utiliza principalmente outra forma - Água da chuva armazenada	0.084
IBGE_abast_Possui ligação à rede geral, mas utiliza principalmente outra forma - Rios, açudes, córregos, lagos e igarapés	0.01
IBGE_abast_Possui ligação à rede geral, mas utiliza principalmente outra forma - Outra	0.136
IBGE_abast_Não possui ligação com a rede geral	-0.352
IBGE_abast_Não possui ligação com a rede geral - Poço profundo ou artesiano	-0.323
IBGE_abast_Não possui ligação com a rede geral - Poço raso, freático ou cacimba	-0.151
IBGE_abast_Não possui ligação com a rede geral - Fonte, nascente ou mina	-0.261
IBGE_abast_Não possui ligação com a rede geral - Carro-pipa	0.087

Variável	Correlação de Pearson (r)	Correlação de Spearman (ρ)
IN006_RS	0.0710	0.0797
IN014_RS	0.1214	0.1109
IN015_RS	0.1781	0.1625
IN016_RS	0.0992	0.0002
IN019_RS	0.0413	0.0097
IN021_RS	-0.0882	-0.1195
IN028_RS	-0.1067	-0.0808
IN045_RS	-0.0110	0.1022
IN051_RS	0.0892	0.1853
IN001_AE	0.0102	0.0360
IN009_AE	0.1099	-0.1010
IN013_AE	-0.0407	0.0398
IN014_AE	0.4935	0.2161
IN015_AE	-0.0223	0.0155
IN022_AE	0.3339	0.0298
IN023_AE	0.0171	-0.0770
IN025_AE	0.1401	0.1222
IN046_AE	0.0648	0.1700
IN049_AE	-0.0101	0.1360

Indicador IBGE	Coefficiente de Spearman (ρ)
IBGE_abast_Não possui ligação com a rede geral - Água da chuva armazenada	-0.016
IBGE_abast_Não possui ligação com a rede geral - Rios, açudes, córregos, lagos e igarapés	-0.088
IBGE_abast_Não possui ligação com a rede geral - Outra	0.025
IBGE_canal_TOTAL Canalizada até dentro da casa, apartamento ou habitação	0.083
IBGE_canal_TOTAL Canalizada, mas apenas no terreno	-0.065
IBGE_canal_TOTAL Sem água canalizada	-0.086
IBGE_canal_Rede geral de distribuição	0.359
IBGE_canal_Poço profundo ou artesiano	-0.32
IBGE_canal_Poço raso, freático ou cacimba	-0.144
IBGE_canal_Fonte, nascente ou mina	-0.269
IBGE_canal_Carro-pipa	0.061
IBGE_canal_Água da chuva armazenada	0.006
IBGE_canal_Rios, açudes, córregos, lagos e igarapés	-0.096
IBGE_canal_Outra	0.045
IBGE_esg_Coletado	0.087
IBGE_esg_Coletado no domicílio por serviço de limpeza	0.183
IBGE_esg_Depositado em caçamba de serviço de limpeza	-0.19
IBGE_esg_Queimado na propriedade	-0.088
IBGE_esg_Enterrado na propriedade	-0.103

Variável	Correlação de Pearson (r)	Correlação de Spearman (ρ)
IN050_AE	-0.0297	0.0487
IN051_AE	0.1009	0.1726
IN052_AE	0.0101	-0.1360
IN053_AE	0.1520	-0.0184
IN055_AE	0.0161	0.0445
IN075_AE	0.0636	-0.0071
IN076_AE	-0.2250	-0.3082
IN084_AE	-0.0800	-0.0663

Indicador IBGE	Coefficiente de Spearman (ρ)
IBGE_esg_Jogado em terreno baldio, encosta ou área pública	0.022
IBGE_esg_Outro destino	-0.015
IBGE_rs_Rede geral, rede pluvial ou fossa ligada à rede	0.21
IBGE_rs_Rede geral ou pluvial	0.194
IBGE_rs_Fossa séptica ou fossa filtro ligada à rede	-0.103
IBGE_rs_Fossa séptica ou fossa filtro não ligada à rede	-0.103
IBGE_rs_Fossa rudimentar ou buraco	-0.134
IBGE_rs_Vala	-0.113
IBGE_rs_Rio, lago, córrego ou mar	-0.256
IBGE_rs_Outra forma	0.056
IBGE_rs_Não tinham banheiro nem sanitário	0.131

APÊNDICE D - RESULTADOS DETALHADOS DAS ANÁLISES DE COMPONENTES PRINCIPAIS DOS INDICADORES DO SINISA

Quadro D.1: Sumário dos resultados da PCA - Eixo Resíduos Sólidos

Sumário dos resultados da PCA - Eixo Água e Esgoto			
Componente	Desvio Padrão	Proporção da Variância	Proporção Acumulada
PC1	1.884	0.323	0.323
PC2	1.465	0.195	0.518
PC3	1.269	0.146	0.664
PC4	1.047	0.1	0.764
PC5	0.959	0.084	0.848
PC6	0.827	0.062	0.91
PC7	0.661	0.04	0.949
PC8	0.512	0.024	0.973
PC9	0.476	0.021	0.994
PC10	0.209	0.004	0.998
PC11	0.153	0.002	1

Quadro D.2: Componentes principais do eixo Resíduos Sólidos do SINISA

PCA - Eixo Resíduos Sólidos						
Nome indicador	Indicador	PC1	PC2	PC3	PC4	PC5
Taxa de empregados em relação à população urbana	IN001	-0.5180	-0.0350	0.0309	0.0050	-0.0966
Despesa média por empregado alocado nos serviços do manejo de rsu	IN002	0.3040	0.2329	0.1852	-0.5280	-0.2624
Despesa per capita com manejo de rsu em relação à população urbana	IN006	-0.3082	0.1821	0.0750	-0.6105	-0.3036
Taxa de cobertura do serviço de coleta domiciliar direta (porta-a-porta) da população urbana do município.	IN014	-0.0316	-0.1493	0.4161	-0.2109	0.7040
Taxa de cobertura regular do serviço de coleta de rdo em relação à população total do município	IN015	0.1113	-0.1259	0.5768	0.3449	-0.3306
Taxa de cobertura regular do serviço de coleta de rdo em relação à população urbana	IN016	-0.0099	-0.2648	0.5560	-0.1783	0.0446
Taxa de empregados (coletadores + motoristas) na coleta (rdo + rpu) em relação à população urbana	IN019	-0.3721	0.1452	-0.1471	-0.1052	0.3751
Massa coletada (rdo + rpu) per capita em relação à população urbana	IN021	-0.0662	0.6163	0.2353	0.1357	0.1641

PCA - Eixo Resíduos Sólidos						
Massa de resíduos domiciliares e públicos (rdo+rpu) coletada per capita em relação à população total atendida pelo serviço de coleta	IN028	0.0261	0.6286	0.1855	0.2287	0.0209
Taxa de varredores em relação à população urbana	IN045	-0.4661	-0.0680	0.0903	-0.0046	-0.1752
Taxa de capinadores em relação à população urbana	IN051	-0.4123	-0.0499	0.1566	0.2672	-0.1548

Quadro D.3: Sumário dos resultados da PCA - Eixo Água e Esgoto

Sumário dos resultados da PCA - Eixo Água e Esgoto			
Componente	Desvio Padrão	Proporção da Variância	Proporção Acumulada
PC1	2.347	0.306	0.306
PC2	1.966	0.215	0.521
PC3	1.472	0.12	0.641
PC4	1.338	0.099	0.741
PC5	0.98	0.053	0.794
PC6	0.929	0.048	0.842
PC7	0.915	0.047	0.888
PC8	0.74	0.03	0.919
PC9	0.622	0.021	0.94
PC10	0.543	0.016	0.957
PC11	0.444	0.011	0.968
PC12	0.421	0.01	0.978
PC13	0.375	0.008	0.985
PC14	0.349	0.007	0.992
PC15	0.274	0.004	0.996
PC16	0.238	0.003	0.999
PC17	0.102	0.001	1
PC18	0	0	1

Quadro D.4: Componentes principais do eixo Água e Esgoto do SINISA

PCA - Eixo Água e Esgoto						
Nome indicador	Indicador	PC1	PC2	PC3	PC4	PC5
Densidade de economias de água por ligação	IN001	0.0272	0.0419	0.5035	-0.0882	0.1396
Índice de hidrometração	IN009	-0.3264	-0.0709	-0.1867	0.3166	-0.0716
Índice de perdas faturamento	IN013	0.3685	-0.1449	0.0575	0.0359	-0.1465
Consumo micromedido por economia	IN014	-0.0237	0.3482	-0.2572	0.2600	-0.2279

PCA - Eixo Água e Esgoto						
Consumo médio percapita de água	IN022	0.1911	0.3536	-0.1728	0.1565	0.1826
Índice de atendimento urbano de água	IN023	-0.0766	0.2004	0.3520	0.3910	0.0885
Volume de água disponibilizado por economia	IN025	0.3775	0.0793	-0.0579	0.2183	0.0021
Índice de perdas na distribuição	IN049	0.1517	-0.4359	-0.0513	0.1863	-0.1312
Índice bruto de perdas lineares	IN050	0.3119	-0.1080	0.2086	0.1381	0.1100
Índice de perdas por ligação	IN051	0.3386	-0.2003	0.0800	0.2513	0.0260
Índice de consumo de água	IN052	-0.1517	0.4359	0.0513	-0.1863	0.1312
Consumo médio de água por economia	IN053	0.2602	0.3613	-0.0834	0.0986	0.0261
Índice de atendimento total de água	IN055	-0.0187	0.1677	0.5305	0.2646	-0.0102
Índice de coleta de esgoto	IN015	-0.2293	-0.1689	0.1944	-0.0032	-0.3157
Índice de esgoto tratado referido à água consumida	IN046	-0.2576	-0.0930	0.2293	-0.0852	0.2511
Incidência das análises de cloro residual fora do padrão	IN075	0.0475	-0.1891	-0.1708	-0.0265	0.7853
Incidência das análises de turbidez fora do padrão	IN076	0.2467	0.0996	0.0451	-0.4514	-0.0951
Incidência das análises de coliformes totais fora do padrão	IN084	0.2518	0.0789	0.1412	-0.3937	-0.1560

Quadro D.5: Sumário dos resultados da PCA - Eixo Águas Pluviais

Sumário dos resultados da PCA - Eixo Águas Pluviais			
Componente	Desvio Padrão	Proporção da Variância	Proporção Acumulada
PC1	1.572	0.206	0.206
PC2	1.32	0.145	0.351
PC3	1.237	0.128	0.479
PC4	1.086	0.098	0.577
PC5	1.062	0.094	0.671
PC6	1.015	0.086	0.757
PC7	0.908	0.069	0.825
PC8	0.829	0.057	0.883
PC9	0.81	0.055	0.937
PC10	0.696	0.04	0.977
PC11	0.5	0.021	0.998
PC12	0.141	0.002	1

Quadro D.6: Componentes principais do eixo Águas Pluviais do SINISA

PCA - Eixo Águas Pluviais							
Nome indicador	Indicador	PC1	PC2	PC3	PC4	PC5	PC6
Parcela de área urbana em relação à área total	IN042	-0.0515	-0.1410	0.1917	-0.4400	-0.3730	0.3767
Densidade Demográfica na Área Urbana	IN043	-0.5917	-0.0225	0.0952	0.0718	0.0876	-0.1137
Densidade de Domicílios na Área Urbana	IN044	-0.5871	-0.0163	0.0803	0.0843	0.1465	-0.1710
Participação da Despesa Total dos Serviços de Drenagem e Manejo das Águas Pluviais Urbanas na Despesa Total do Município	IN010	0.1377	-0.6387	0.0630	-0.1139	0.0917	-0.1130
Despesa per capita com serviços de Drenagem e Manejo das Águas Pluviais Urbanas	IN048	0.1908	-0.5236	-0.027 1	0.0120	-0.2451	-0.3042
Taxa de Cobertura de Pavimentação e Meio-Fio na Área Urbana do Município	IN020	-0.0863	-0.2972	-0.215 5	0.5251	0.1842	0.2385
Taxa de cobertura de vias públicas com redes ou canais pluviais subterrâneos na área urbana	IN021	-0.0127	-0.2793	0.4170	0.4827	-0.0666	0.1287
Parcela de Cursos d'Água Naturais Perenes com Canalização Aberta	IN026	-0.3395	-0.1598	0.0308	-0.1112	-0.4254	0.2933
Parcela de Cursos d'Água Naturais Perenes com Canalização Fechada	IN027	-0.0759	0.0466	-0.476 4	0.2667	-0.4013	0.3694
Parcela de Cursos d'Água Naturais Perenes com Diques	IN029	-0.3393	-0.2542	-0.3119	-0.3598	0.0690	-0.2006
Parcela de Domicílios em Situação de Risco de Inundação	IN040	0.0058	-0.1482	0.1500	-0.2332	0.5733	0.6088
Parcela da População Impactada por Eventos Hidrológicos	IN041	0.0586	-0.1290	-0.612 4	-0.0708	0.2223	0.0619

APÊNDICE E - RESULTADOS DA ANÁLISE BIBLIOMÉTRICA

Uma análise bibliométrica trata de estudar não apenas os respectivos materiais publicados, mas também as características de publicações ao longo de um período de tempo que tenham relação entre si. Permite a análise de países de origem dos documentos, do número de artigos publicados por determinado autor ou conjunto de autores, palavras-chave que são comumente utilizadas em conjunto, dentre outros fatores. O SCOPUS e a Web of Science, dois dos maiores e mais relevantes bancos de dados de artigos do mundo, podem ser utilizados para a realização de uma análise como essa.

Comparando-se os materiais indexados na base de dados SCOPUS relacionados à dengue, foi observado que há disponíveis um total de 1282 documentos publicados associando palavras-chave relacionadas à "Dengue" e ao "Clima" entre os anos de 2018 e 2023. Por outro lado, estudos que associam a palavras-chave referentes a "Dengue" e "Saneamento Básico" somam um total de 189 documentos. Trata-se de uma diferença de 6,8 vezes menos materiais recentes publicados que estudaram as relações da doença com as condições de saneamento da população em comparação com os que mencionam o clima.

A partir dos dados obtidos com uma pesquisa no SCOPUS com delimitação de data entre 2018 e 2023 e palavras-chave Dengue AND (operador booleano) Sanitation, uma nuvem de palavras foi gerada. As figuras abaixo apresentam as principais palavras-chave encontradas nos documentos, organizadas pelo número de aparições. As palavras que possuem tonalidades mais amareladas são as que aparecem em um maior número de artigos. Nota-se a presença de outras doenças infecciosas e a proximidade entre os termos “dengue” e “saneamento”, mas também a aparição do conceito de mudanças climáticas.

APÊNDICE F - LISTA DE SUGESTÕES (ROTEIRO) PARA PESQUISAS FUTURAS

Sugestões relacionadas às análises espaço-temporais (capítulo 6.1)

- Aplicar as mesmas técnicas de análise espaço-temporais, realizadas com a casuística da dengue, para dados individuais ou agregados de saneamento. Compreender, assim, padrões de distribuição espacial dos dados de saneamento.
- Analisar e compreender efeitos tardios de condições climáticas na incidência da dengue em mais municípios do estado do Rio de Janeiro em busca de padrões. Perguntas relevantes a serem feitas: os lags são sempre iguais? Caso contrário, o que pode alterá-los? Aspectos socioeconômicos? Condições de saneamento?
- Analisar efeitos de ações de controle de dengue em sua incidência, ao longo do tempo. Observar defasagens temporais e duração dos efeitos. Isso pode ajudar a compreender a eficiência das ações e ajudar a propor melhorias.

Sugestões referentes às análises de correlação (capítulo 6.2)

- Analisar correlações utilizando não somente a incidência mediana anual, mas também outras variações dos dados de dengue. Por exemplo: média de incidência; incidência anual ao longo de todo o período; incidência anual máxima durante o período; desvio padrão dos totais anuais.
- Fazer análises de correlação entre dengue e indicadores específicos de urbanização. Caso encontre bons resultados, incorporá-los em um modelo mecanístico.
- Reaplicar as análises de correlação em grupos de municípios estratificados, de forma a controlar o fator de confusão da urbanização. Uma sugestão seria agrupar os municípios por níveis semelhantes de desenvolvimento urbano (utilizando, por exemplo, o componente principal de "Densidade Urbana") e, em seguida, reavaliar as correlações entre os indicadores de saneamento e a dengue dentro de cada estrato. Essa abordagem permitiria isolar com maior precisão o efeito direto do saneamento, separando-o da influência da urbanização.

Sugestões referentes à análise de componentes principais (capítulo 6.3)

- Aplicar uma abordagem *top-down*, como a de Borges (2024), criando categorias prévias para analisar os PCs. Fazer isso tanto para a de cada um dos eixos, como uma análise GERAL que considera todos os dados juntos.

- Utilizar diferentes métodos de agregação de variáveis e analisar suas respectivas eficiências. Escolher e debater o melhor método para compreender se um município tem “boas” condições de saneamento básico.

Sugestões referentes à modelagem mecanística de dengue (capítulo 6.4)

- Investigar mais profundamente a inserção de parâmetros que representam fases iniciais de vida dos mosquitos *Aedes*, em modelos matemáticos. Buscar compreender quantitativamente suas relações com condições de saneamento básico.
- Procurar na bibliografia mais parâmetros aplicados em modelos que têm capacidade de conectar determinantes de saneamento básico na incidência de dengue, como a taxa de sobrevivência aquática (s_a) e a capacidade de carga (K).
- Após obter uma compreensão qualitativa mais profunda entre variáveis de saneamento e dengue, elaborar um modelo que tenha capacidade de predição de casos para ajudar a estimar efeitos de melhorias em condições de saneamento na redução de casos de dengue em longo termo.