



Análise da Variabilidade da Dengue na Cidade do Rio de Janeiro

Analysis of Dengue Variability in the City of Rio de Janeiro

Suellen Araujo Franco dos Santos

Orientador: Hugo Abi Karam, Prof. DSc.

Dissertação de Mestrado

Programa de Pós-graduação em Meteorologia

Instituto de Geociências / Centro de Ciências Matemáticas e da Natureza

Universidade Federal do Rio de Janeiro

PPGM-IGEO/CCMN/UFRJ

Rio de Janeiro, RJ

Abril de 2023

ANÁLISE DA VARIABILIDADE DA DENGUE NA CIDADE DO RIO DE JANEIRO

SUELLEN ARAUJO FRANCO DOS SANTOS

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM METEOROLOGIA DO INSTITUTO DE GEOCIÊNCIAS DO CENTRO DE CIÊNCIAS MATEMÁTICAS E DA NATUREZA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO (PPGM-IGEO-CCMN-UFRJ) COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA À OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS (ÁREA: METEOROLOGIA).



Prof. Dr. Hugo Abi Karam - PPGM/UFRJ - Orientador



Profa. Dra. Claudine Pereira Dereczynski - PPGM/UFRJ



Profa. Dra. Rose Ane Pereira de Freitas - UFPEL

RIO DE JANEIRO, RJ - BRASIL

ABRIL 2023

S586a Araujo Franco dos Santos, Suellen
Análise da variabilidade da dengue na cidade do Rio de Janeiro / Suellen
Araujo Franco dos Santos; orientador: Hugo Abi Karam. -- Rio de Janeiro,
2023.

112 p.

Dissertação (Mestrado - Programa de Pós-graduação em Meteorologia)
-- Instituto de Geociências, Universidade Federal do Rio de Janeiro, 2023.

1. Epidemia de dengue. 2. Cidade do Rio de Janeiro. 3. Análise de
agrupamento. 4. Análise de variância. 5. Transformada de ondeleta.. I.
Karam, Hugo Abi, orient. II. Título.

Dedico a Suellen que habita em mim,
por não ter me deixado desistir,
mesmo quando tudo me levava a querer desistir.

AGRADECIMENTOS

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela bolsa de mestrado recebida e ao Programa de Pós-Graduação em Meteorologia da UFRJ.

RESUMO

Neste trabalho, dados semanais de notificações de agravamento da dengue do Ministério da Saúde brasileiro são utilizados em análises de variância particionadas por bairros, agrupamento k-médias e em modelagem por regressão linear multivariada. Dados das estações meteorológicas mesonet AlertaRio do município do Rio de Janeiro também foram usados na caracterização da distribuição térmica na superfície urbana. Os resultados obtidos mostram que variáveis por bairro padronizadas pela população local melhoram os resultados obtidos como o modelo de regressão linear, para além da confiança estatística de 95%. Quatro grupos homogêneos explicam a variância para um grau de confiança acima de 90%. A incidência de dengue padronizada mostra-se bem distribuída na cidade do Rio de Janeiro, apesar da Ilha de Calor Urbana apresentar estrutura térmica espacial e variação sazonal. Embora a reprodutividade do vetor da Dengue esteja associada a presença de temperaturas tropicais, sua distribuição pelos bairros provavelmente também pode estar associada a outras variáveis, como o transporte aéreo do vetor da dengue promovido por recirculações de mesoescala como brisas.

Palavras chave – Epidemia de dengue; Cidade do Rio de Janeiro; Análise de agrupamento; Análise de variância; Transformada de ondeleta.

ABSTRACT

In this work, weekly data on reports of worsening dengue from the Brazilian Ministry of Health are used in analyzes of variance partitioned by neighborhoods, k-means clustering and in multivariate linear regression modeling. Data from the AlertaRio mesonet meteorological stations in the city of Rio de Janeiro were also used to characterize the thermal distribution on the urban surface. The results show that neighborhood variables standardized by the local population improve the results obtained with the linear regression model, beyond the 95% statistical confidence. Four homogeneous groups explain the variance for a confidence level above 90%. The incidence of standardized dengue is well distributed in the city of Rio de Janeiro, despite the Urban Heat Island presenting spatial thermal structure and seasonal variation. Although the reproducibility of the Dengue vector is associated with the presence of tropical temperatures, its distribution across neighborhoods can probably also be associated with other variables, such as air transport of the Dengue vector promoted by mesoscale recirculations such as breezes.

Keywords – Dengue epidemic; City of Rio de Janeiro; Cluster analysis; Analysis of variance; Wavelet transform.

ÍNDICE DE FIGURAS

Figura 2.1– Série de período igual a 50, amplitude igual a 1 e comprimento igual a 1000 (Fonte: Rösch e Schmidbauer, 2014).....	11
Figura 2.2- Espectro de potência de ondeleta da função periódica com período igual a 50, amplitude igual a 1 e comprimento igual a 1000 (Fonte: Rösch e Schmidbauer, 2014).....	12
Figura 2.3 - Séries temporais das funções x e y de subperíodos igual a 60 e 100 e comprimento igual a 1000 (Fonte: Rösch e Schmidbauer, 2014).....	13
Figura 2.4 - Espectro de ondeleta global relativo a (a) função x e (b) função y (Fonte: Rösch e Schmidbauer, 2014 – adaptado).....	14
Figura 2.5 - Séries temporais das funções x e y de subperíodos 1, 2, 4, 8 e 16 e comprimento igual a 96 (Fonte: Rösch e Schmidbauer, 2014).....	15
Figura 2.6 – Espectro de potência bivariado das funções x e y de subperíodos 1, 2, 4, 8 e 16 e comprimento igual a 96 (Fonte: Rösch e Schmidbauer, 2014 – adaptado).....	16
Figura 2.7 - Espectro de ondeleta global do cruzamento das funções x e y de subperíodos 1, 2, 4, 8 e 16 e comprimento igual a 96 (Fonte: Rösch e Schmidbauer, 2014 – adaptado).....	17
Figura 3.8 - Regiões de rejeição e não-rejeição do teste de hipótese (Fonte: Moraes Júnior, 2015).....	21
Figura 3.9 - Município do Rio de Janeiro e suas zonas geográficas. Fonte: adaptado de IBGE, 2010 e Nossedotti, 2013.....	22
Figura 4.10 – Fluxograma dos dados e métodos utilizados para obtenção dos resultados do estudo. Fonte: Este trabalho.....	30
Figura 4.11 – Agrupamento dos bairros do município do Rio de Janeiro via método <i>k-means</i> a) em função das notificações de dengue e b) em função das notificações de dengue proporcionais à população de cada bairro.....	36
Figura 4.12 - Distribuição espacial dos grupos obtidos na análise de agrupamento via método <i>k-means</i> para a) os totais absolutos de dengue por bairro e para b) os totais proporcionalizados de dengue por bairro no município do Rio de Janeiro.....	38
Figura 4.13 - Agrupamento dos bairros do município do Rio de Janeiro via método hierárquico a) em função das notificações de dengue e b) em função das notificações de dengue proporcionais à população de cada bairro.....	39
Figura 4.14 - Distribuição espacial do teste-F obtido para as RLM completas a) dos totais absolutos de dengue por bairro e b) dos totais proporcionalizados de dengue por bairro no município do Rio de Janeiro.	42
Figura 4.15 - Distribuição espacial da significância estatística a) dos preditores dentro da RLM e b) dos preditores dentro da análise ANOVA dos totais proporcionalizados de dengue por bairro no município do Rio de Janeiro.....	43
Figura 4.16 - Distribuição espacial do teste F obtido para os <i>clusters</i> na análise de agrupamento via método <i>k-means</i> para a) os totais absolutos de dengue por bairro e para b) os totais proporcionalizados de dengue por bairro no município do Rio de Janeiro.....	49
Figura 4.17 - Séries temporais de dados semanais normalizados para o bairro de Irajá – temperatura média (linha preta cheia), temperatura máxima (linha vermelha pontilhada), temperatura mínima (linha verde pontilhada) e casos de dengue (linha azul pontilhada) entre os anos de 2015 e 2019.....	51

Figura 4.18 - (a) Série temporal, (b) espectro de potência de ondeleta contínua e (c) espectro de ondeleta global dos casos de dengue semanais para o bairro de Irajá. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).....	53
Figura 4.19 - (a) Série temporal, (b) espectro de potência de ondeleta contínua e (c) espectro de ondeleta global da temperatura média semanal para o bairro de Irajá. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).....	54
Figura 4.20 - (a) Série temporal, (b) espectro de potência de ondeleta contínua e (c) espectro de ondeleta global da temperatura máxima semanal para o bairro de Irajá. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).....	55
Figura 4.21 - (a) Série temporal, (b) espectro de potência de ondeleta contínua e (c) espectro de ondeleta global da temperatura mínima semanal para o bairro de Irajá. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).....	56
Figura 4.22 - Espectros bivariados de (a) potência de ondeleta contínua e (b) ondeleta global temperatura média semanal com casos de dengue semanais para o bairro de Irajá. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).....	58
Figura 4.23 - Espectros bivariados de (a) potência de ondeleta contínua e (b) ondeleta global temperatura máxima semanal com casos de dengue semanais para o bairro de Irajá. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).....	59
Figura 4.24 - Espectros bivariados de (a) potência de ondeleta contínua e (b) ondeleta global temperatura mínima semanal com casos de dengue semanais para o bairro de Irajá. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).....	60
Figura 4.25 - (a) espectro de potência de ondeleta contínua e (b) espectro de ondeleta global das diferenças semanais de temperatura média para o bairro de Irajá durante o ano de 2016. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).....	62
Figura 4.26 - (a) espectro de potência de ondeleta contínua e (b) espectro de ondeleta global das diferenças semanais de temperatura máxima para o bairro de Irajá durante o ano de 2016. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo	

sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).....	63
Figura 4.27 -(a) espectro de potência de ondeleta contínua e (b) espectro de ondeleta global das diferenças semanais de temperatura mínima para o bairro de Irajá durante o ano de 2016. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).....	64
Figura 4.28 - Espectros bivariados de (a) potência de ondeleta contínua e (b) ondeleta global diferenças semanais de temperatura média com casos de dengue semanais para o bairro de Irajá durante o ano de 2016. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).....	65
Figura 4.29 - Espectros bivariados de (a) potência de ondeleta contínua e (b) ondeleta global diferenças semanais de temperatura máxima com casos de dengue semanais para o bairro de Irajá durante o ano de 2016. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).....	67
Figura 4.30 - Espectros bivariados de (a) potência de ondeleta contínua e (b) ondeleta global diferenças semanais de temperatura mínima com casos de dengue semanais para o bairro de Irajá durante o ano de 2016. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).....	68
Figura 4.31 - (a) espectro de potência de ondeleta contínua e (b) espectro de ondeleta global das diferenças semanais de temperatura média para o bairro de Irajá durante o ano de 2018. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).....	69
Figura 4.32 - (a) espectro de potência de ondeleta contínua e (b) espectro de ondeleta global das diferenças semanais de temperatura máxima para o bairro de Irajá durante o ano de 2018. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).....	70
Figura 4.33 - (a) espectro de potência de ondeleta contínua e (b) espectro de ondeleta global das diferenças semanais de temperatura mínima para o bairro de Irajá durante o ano de 2018. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).....	71
Figura 4.34 - Espectros bivariados de (a) potência de ondeleta contínua e (b) ondeleta global das diferenças semanais de temperatura média com casos de dengue semanais para o bairro de Irajá durante o ano de 2018. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).....	72
Figura 4.35 - Espectros bivariados de (a) potência de ondeleta contínua e (b) ondeleta global das diferenças semanais de temperatura máxima com casos de dengue semanais para o bairro de Irajá durante	

o ano de 2018. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).....74

Figura 4.36 - Espectros bivariados de (a) potência de ondeleta contínua e (b) ondeleta global das diferenças semanais de temperatura mínima com casos de dengue semanais para o bairro de Irajá durante o ano de 2018. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).....75

ÍNDICE DE TABELAS

Tabela 4.1 - Lista dos bairros do município do Rio de Janeiro utilizados na análise de agrupamento e suas respectivas numerações.....	32
Tabela 4.2 - Resultado obtidos para os resíduos associados a RLM com grupos <i>k-means</i> dos totais absolutos de dengue por bairro.....	44
Tabela 4.3 - Medidas estatísticas auxiliares obtidas para a RLM com grupos <i>k-means</i> dos totais absolutos de dengue por bairro.....	44
Tabela 4.4 - Coeficientes lineares e significância estatística na RLM dos grupos <i>k-means</i> dos totais absolutos de dengue por bairro.....	45
Tabela 4.5 - Análise de variância para RLM dos grupos <i>k-means</i> dos totais absolutos de dengue por bairro.	46
Tabela 4.6 - Resultado obtidos para os resíduos associados a RLM com grupos <i>k-means</i> dos totais proporcionalizados de dengue por bairro.....	46
Tabela 4.7 - Medidas estatísticas auxiliares obtidas para a RLM com grupos <i>k-means</i> dos totais proporcionalizados de dengue por bairro.....	47
Tabela 4.8 - Coeficientes lineares e significância estatística na RLM dos grupos <i>k-means</i> dos totais proporcionalizados de dengue por bairro.....	47
Tabela 4.9 - Análise de variância para RLM dos grupos <i>k-means</i> dos totais proporcionalizados de dengue por bairro.....	48

LISTA DE ABREVIATURAS E SIGLAS

ANOVA – Análise de variância

CPTEC - Centro de Previsão de Tempo e Estudos Climáticos

ENOS – El Niño Oscilação Sul

FGV - Fundação Getúlio Vargas

FIOCRUZ - Fundação Oswaldo Cruz

FOE - Função ortogonal empírica

GL - Graus de liberdade

IBGE – Instituto Brasileiro de Geografia e Estatística

IPCC – *Intergovernmental Panel on Climate Change*

IPP – Instituto Pereira Passos

NOAA - *National Oceanic and Atmospheric Administration*

PBMC - Painel Brasileiro de Mudanças Climáticas

RLM – Regressão linear múltipla

SQE - Soma dos quadrados explicados

SQR - Soma dos quadrados dos resíduos

SMS - Secretaria Municipal de Saúde

SOI - *South Oscillation Index*

SUS - Serviço Único de Saúde

TO – Transformada de ondeleta

TOC – Transformada de ondeleta cruzada

WHO – *World Health Organization*

Sumário

1	Introdução.....	1
1.1	Objetivo.....	2
1.2	Estrutura da dissertação.....	3
2.4	Visualização gráfica da transformada de ondeleta.....	11
2.5	Aplicações da TO no estudo da ICU e da dengue.....	16
2.6	Análise de agrupamento e de variância.....	18
3	Materiais e Métodos.....	22
3.1	Área de estudo.....	22
3.2	Dados utilizados na análise.....	23
3.3	Métodos para análise dos dados.....	24
4.	Resultados.....	31
4.1	Análise de agrupamento e de variância.....	31
4.2	Análise de agrupamento e de variância.....	39
4.3	Incidência de dengue e ICU.....	49
4.3.1	Análise ano epidêmico vs. ano não epidêmico.....	60
5	Conclusões.....	77
	Referências.....	79
	Anexo A – Artigo submetido ao ‘Anuário do Instituto de Geociências da UFRJ’ (2023).....	91

Capítulo 1

1 Introdução

Os ambientes urbanos têm como característica marcante a substituição da natureza por novas estruturas (como por exemplo, pavimentação de ruas) e materiais provenientes da criação humana, como o uso de concreto e asfalto, além do crescimento e desenvolvimento populacional, criando assim uma nova caracterização do ambiente local devido às alterações realizadas. Essas alterações, associadas às formas de organização antropocêntrica, devido ao uso indevido de recursos petrolíferos, produção de dejetos e gases poluentes, dentre outros, contribuem a longo prazo na alteração de variáveis como temperatura e precipitação em escala local e global (OLIVEIRA, 2019).

Um efeito resultante dessas alterações e que é de grande importância é a intensificação do aquecimento nos locais urbanos em comparação com seus entornos, caracterizando o fenômeno conhecido como Ilha de Calor Urbana – ICU (OKE et al., 2017). Essa complexidade cria dificuldade no controle, prevenção e manutenção de doenças. Em ambientes urbanos, as arboviroses (doenças causadas por arbovírus e transmitidas por insetos) recebem destaque, sendo essas doenças consideradas ameaças significativas para a saúde pública. Uma das principais arboviroses nos países tropicais é a dengue.

A dengue é uma doença viral transmitida majoritariamente pelo mosquito vetor *Aedes aegypti*, sendo de grande importância para a sociedade, considerada atualmente a mais importante das arboviroses no mundo principalmente nos países de região tropical, onde existem históricos e casos registrados e a existência de condições climáticas, sociais e econômicas que favorecem a proliferação do mosquito (OLIVEIRA, 2019).

As infecções por dengue ocorrem em mais de 100 países na região da Ásia, nas Américas, no Oriente Médio e na África, e os casos de infecção seguem aumentando em todo o mundo, sendo estimadas em torno de 50 milhões de infecções para ocorrer a cada ano (WHO, 2009). A realização de estudos sobre a doença no Brasil ainda se mostra de grande relevância, tendo em

vista o aumento do número de casos e de mortes registradas no país desde os anos 2000 (ARAÚJO et al., 2017).

A Região Sudeste é a de maior incidência de dengue, tendo maior população e recursos para realização de diagnóstico e notificação (BORGES, 2001). Desde a reinvasão do Brasil pelo mosquito vetor, em 1977, o país passou por várias epidemias de dengue (HONÓRIO et al., 2009). A dengue recebeu um caráter epidemiológico quando o Estado do Rio de Janeiro teve uma epidemia no ano de 1986 que se alastrou para outras regiões brasileiras (BRAGA e VALLE, 2007). Dentro do Estado, o maior número de casos são encontrados na Região Metropolitana do Rio de Janeiro (RMRJ), com a maioria na sua capital, Rio de Janeiro, seguido por outras localidades da mesma região - Niterói e Baixada Fluminense, sendo esta última o local onde teve o início da propagação da doença no estado (MIAGOSTOVICH *et al.*, 1993).

O ambiente urbano da cidade do Rio de Janeiro é heterogêneo, com bairros apresentando uma diversidade de microáreas variando entre si no que diz respeito ao uso e cobertura do solo, microclima, saneamento e urbanização, fatores esses que podem influenciar na transmissão de dengue local em cada bairro (XAVIER et al., 2017). A forma de transmissão da dengue pode ser influenciada por diversos fatores de diferentes escalas. Em um âmbito local, podem ser citados como exemplos as estruturas socioeconômica e urbana de um município específico. Barcellos et al. (2005) citam o papel latente de atividades humanas como fluxo populacional para a difusão do mosquito vetor. Desta forma, surge o questionamento: até que ponto a transmissão de dengue na cidade do Rio de Janeiro é influenciada por ocorrência de ICUs ?

Várias ferramentas podem ser utilizadas no estudo da dengue e que podem ajudar a responder o questionamento acima. Uma delas é a transformada de ondeleta. Ela tem como vantagem a análise simultânea de tempo e frequência de séries temporais, sendo aplicada em diversas finalidades, como estudo de sinais geofísicos, turbulência e variabilidade climática.

1.1 Objetivo

O estudo da dengue no município do Rio de Janeiro vem sendo feito ao longo dos últimos anos durante a pesquisa da autora da dissertação. Após verificar a existência de um padrão epidêmico com intervalo de aproximadamente três anos, foi estudada a relação entre os casos de

dengue e o fenômeno de variabilidade climática El Niño Oscilação Sul (ENOS), que foi representado pelo índice *South Oscillation Index* (SOI), visando compreender o nível de influência do ENOS sobre a periodicidade dos casos de dengue na cidade do Rio de Janeiro como um todo (SANTOS et al., 2022). Em seguida, no trabalho de conclusão de curso foi investigada a relação entre dengue e ENOS através de diversos índices disponibilizados pela instituição *National Oceanic and Atmospheric Administration* (NOAA), bem como a previsibilidade da dengue em função dos vários indicadores de fase do ENOS (SANTOS, 2019).

A incidência de dengue no município do Rio de Janeiro pode estar associada ao fenômeno da ICU. Adicionalmente, o estudo sobre a transmissão de dengue na localidade, citado no parágrafo acima, não levou em conta as contribuições individuais de cada bairro do município. Por conta de ambos os fatores citados, este trabalho é uma progressão da pesquisa acadêmica sobre a dengue, tendo como objetivo principal analisar a ICU como um fenômeno local que pode influenciar as ocorrências de dengue no município, tendo a transformada de ondeleta como ferramenta de estudo. Os objetivos secundários são: investigar o papel de variáveis de temperatura nos casos de dengue; verificar as contribuições dos diferentes bairros para o total de dengue no município; aprofundar a contribuição para o estudo da dengue no Rio de Janeiro.

Com isso, espera-se que o presente trabalho possa favorecer futuros desenvolvimentos de pesquisas sobre dengue na cidade do Rio de Janeiro em associação a fenômenos meteorológicos e climáticos que possam exercer influência nos padrões de transmissão da doença, bem como os fatores da distribuição da dengue no município, mostrando-se assim relevante para estudos na área de meteorologia e áreas correlacionadas ao estudo de arboviroses transmitidas pelo *Aedes aegypti*.

1.2 Estrutura da dissertação

O texto está dividido em 6 capítulos. No capítulo 2 é feita uma fundamentação teórica que inclui um apanhado histórico sobre a incidência de dengue, uma abordagem geral sobre o fenômeno Ilha de Calor Urbana (ICU) e estudos dos realizados em diversas localidades sobre a

ICU em si, bem como as possíveis relações entre o fenômeno e registros de casos de dengue. Também é abordado o método de transformada de ondeleta e suas diversas aplicações, com exemplos de análise. No terceiro capítulo são apresentados os materiais e métodos utilizados neste estudo. O capítulo 4 apresenta os resultados obtidos a partir dos métodos citados no capítulo 3 e, finalmente, o capítulo 5 aborda as conclusões obtidas a partir dos resultados.

Capítulo 2

2 Fundamentação teórica

2.1 Ilha de Calor Urbana

A ICU é um fenômeno caracterizado pela diferença de temperatura entre a área urbana, onde ocorre um excesso de produção de calor, e o seu entorno. Howard (1833) foi pioneiro em comprovar que as áreas urbanas apresentam diferença térmica significativa em relação às regiões rurais ao entorno. O termo “ilha de calor” foi estabelecido em 1960 por Gordon Manley (ALCOFORADO, 1993), tendo sua origem no padrão das isotermas localizadas sobre a camada limite urbana, padrão este que se assemelha ao formato de uma ilha (OKE et al., 2017).

Alguns dos principais fatores que tem grande potencial de influenciar a formação de ICUs são: redução do albedo superficial; redução da perda de radiação de onda longa devido ao pequeno fator de visão do céu; uso de materiais com maior capacidade de retenção de calor sensível; baixa umidade da superfície por conta da impermeabilização do solo; efeito estufa associado aos gases poluentes e aerossóis emitidos na atmosfera (OKE, 1982; OKE, 1995; OKE et al., 1991; VOOGT, 2002). Outro aspecto que influencia na ICU, no que diz respeito a sua intensidade, é a ocorrência de condições sinóticas estáveis com ventos de menor intensidade e menor cobertura de nuvens (CARDOSO e AMORIM, 2018).

A ICU é um problema recorrente de diversas localidades no Brasil. Em um estudo simultâneo das metrópoles do país, notou-se que a intensidade da ICU é mais notável durante o dia para as cidades de Manaus, Porto Alegre, Belém e Recife, enquanto que o efeito noturno da ICU é mais presente em Curitiba, Brasília, São Paulo e Rio de Janeiro (MONTEIRO et al., 2021).

Em Manaus, a diferença térmica entre as porções urbana e rural da cidade alcança valores de até 3 °C. Um aspecto ressaltado é a presença de dois picos de intensidade da ICU na cidade, um durante o meio da manhã e outro durante o fim da tarde. Eles são relacionados com o ciclo

diurno em conjunto com a estrutura urbana, que resulta no aquecimento antecipado no início do dia e no atraso no resfriamento em comparação com o ambiente rural (SOUZA e ALVALÁ, 2014).

Até em cidades de pequeno e médio porte, o efeito da ICU pode ser notado. Em Presidente Prudente, Rancharia e Paranavaí, por exemplo, a magnitude média da ICU no verão pode chegar a 3,7 °C, com picos de até 6 °C em condições de tempo calmo (CARDOSO et al., 2017). Ainda mencionando Presidente Prudente, a cidade apresenta maiores contrastes térmicos entre as áreas da cidade com maiores diferenças na cobertura de solo e estrutura local durante as noites de inverno (CARDOSO e AMORIM, 2018).

Paço do Lumiar, um município da Região Metropolitana de São Luís, no estado do Maranhão, sofreu um avanço rápido na urbanização, o que trouxe mudanças nas condições de temperatura devido às novas condições de uso do solo e materiais utilizados nesses ambientes. Isso provocou um aumento nos efeitos da ICU sobre a cidade, diferente dos anos com menor urbanização (SILVA et al., 2018).

O estado do Rio de Janeiro concentra diversas regiões com características relacionadas a ocorrência de ICU. Na RMRJ, por exemplo, as áreas com as temperaturas absolutas mais elevadas foram localizadas sobre áreas urbanas centrais e regiões de subúrbio (LUCENA et al., 2013).

Lucena e Peres (2017), analisando dados de temperatura média em superfície, além de dados de satélite e de modelagem atmosférica, constataram o aumento da variável atmosférica no centro do município do Rio de Janeiro ao longo do século XX, tendo a década de 1990 e a de 2000 como as mais quentes de todo o intervalo de estudo. Adicionalmente, as áreas com maiores temperaturas no município são a Zona Norte e o Centro, que são locais com maior concentração de urbanização.

2.2 ICU e dengue

O aumento da temperatura é uma questão de grande importância, tendo em vista as modificações que esse aumento pode acarretar em diversos aspectos. Há evidências de que maiores temperaturas têm alterado as regiões do planeta suscetíveis a epidemias de dengue, assim como também estão aumentando a capacidade dos mosquitos vetores de se infectarem com o vírus da dengue e transmitirem a doença (IPCC, 2022).

A variável temperatura é um dos fatores que exerce influência no ciclo de vida do mosquito vetor *Aedes aegypti*. Por exemplo, para o mosquito conseguir se manter dentro de um ambiente com água parada a fim de ocorrer o desenvolvimento dos ovos até a fase adulta, o ideal é que a temperatura esteja dentro de um intervalo de 21°C a 29 °C (BESERRA et al., 2006).

Em um levantamento realizado em diversas cidades localizadas no estado da Paraíba, foi detectado que para o período de incubação dos ovos ocorreu uma redução de aproximadamente 27 dias quando os ovos foram mantidos a 25°C e 30°C em relação à incubação realizada a uma temperatura de 15°C (BESERRA et al., 2009). A taxa de eclosão dos ovos, por sua vez, alcança valores acima de 90% para o intervalo de temperaturas de 22°C a 28 °C, tendo um limar de 35°C de temperatura para ocorrer a eclosão em aproximadamente metade dos ovos (FARSINE, 2009).

A ICU tem fatores associados a ocorrência de dengue: urbanização, pouca vegetação e altas temperaturas. Eles fornecem boas condições para a proliferação de vetores *Aedes* (SANTOS et al., 2020). A temperatura influencia na habilidade do mosquito para sobreviver e determina seu desenvolvimento e taxas reprodutivas (JOHANSSON et al., 2009). Quando ocorre o seu aumento, dois importantes processos que acontecem durante o tempo de vida dos mosquitos vetores são afetados: redução do tempo de desenvolvimento das larvas e do tempo que o vírus leva para chegar à glândula salivar do mosquito, que acabam resultando na agilização da transmissão da doença (LIMA-CAMARA, 2016).

Para a transmissão da dengue, zika e chikungunya nas Américas, foi detectado através dos três vírus, modelos e dados de casos registrados, que a transmissão ocorre entre 18 – 34°C, com a transmissão máxima ocorrendo num intervalo entre 26 – 29°C (MORDECAI et al., 2017).

Algo similar foi detectado na cidade de Areia, localizada no estado da Paraíba. Os anos com maior ocorrência de casos foram aqueles com a temperatura dentro do intervalo de 21° a 29 °C, que é considerado o ideal para o desenvolvimento do mosquito (BESSERA et al., 2006). Além disso, foi encontrado outro padrão de comportamento das variáveis: à medida que a temperatura apresentava declínio, o número de casos de dengue também diminuía (BARACHO, 2013).

Foi realizada, para o município do Rio de Janeiro, uma comparação dos primeiros trimestres de anos epidêmicos em relação aos anos não epidêmicos, sendo então observado que a dengue atingiu uma maior quantidade de habitantes quando as temperaturas estavam mais elevadas, principalmente quando a temperatura mínima esteve acima de 22 °C (CÂMARA et al., 2009). Além disso, um aumento da temperatura mínima foi associado a um aumento significativo do número de casos de dengue no mês posterior (GOMES et al., 2012). Ao se analisar a distribuição espacial de temperatura média da superfície para a mesma localidade, foi encontrada uma correlação significativa entre a variável meteorológica e a distribuição de larvas do mosquito *Aedes aegypti* (LEMOS et al., 2021).

Em São Paulo, foi evidenciado na capital que a incidência da doença se mostrou maior nas áreas com maior temperatura e menor cobertura vegetal. Além disso, foi constatado que a fase larval do mosquito vetor se desenvolve melhor dentro do intervalo de 28 – 32°C de temperatura (ARAUJO et al., 2015). Na cidade de Taubaté, a correlação se mostrou alta entre a temperatura da superfície terrestre e as áreas do município com maior quantidade de casos registrados de dengue (ALEXANDRINO, 2017).

2.3 Transformada de ondeleta

A transformada de ondeleta (TO) se mostrou ao longo dos anos uma ferramenta de grande utilidade no estudo de transmissão de doenças epidêmicas como a dengue por permitir a decomposição de sinais não estacionários, tendo entre suas vantagens as variações em frequência e a detecção de estruturas temporal e/ou espacialmente localizadas (DOMINGUES et al., 2003).

A nomenclatura ondeleta (em inglês, *wavelet*) se refere ao conjunto de pequenas ondas formadas por dilatação e translação de uma função $\psi(t)$, quadraticamente integrável no intervalo dos números reais, ou seja, deve possuir energia finita (REBOITA, 2004). Esta função é denominada ondeleta mãe, e as funções geradas através dela são as ondeletas filhas (WENG E LAU, 1994).

A transformada de ondeleta (TO) permite analisar periodicidade de eventos em diferentes escalas de variabilidade temporal e não necessita de série estacionária (SANTOS et al., 2013). Sendo assim, pode ser utilizada para analisar séries temporais com potência não estacionária em diferentes frequências (DAUBECHIES, 1990). Para ser admissível como ondeleta, a função deve ter média zero e localizada em ambos espaços de tempo e frequência (FARGE, 1992). Uma função que satisfaz esta condição é a ondeleta de Morlet, que consiste em uma onda exponencial complexa (onda seno) multiplicada por um envelope gaussiano:

$$\psi_0(t) = \pi^{-1/4} e^{i\omega t} e^{-1/2t^2} \quad (2.1)$$

onde t é um parâmetro de tempo adimensional e ω é a frequência adimensional ou parâmetro da ondeleta de Morlet e deve ser escolhido de forma a garantir que a condição de admissibilidade seja mantida. Aqui o valor utilizado é $\omega=6$, a fim de satisfazer esta propriedade (TORRENCE E COMPO, 1998). A função se mostra adequada para a análise de sinais geofísicos e séries temporais por sua quantidade grande de oscilações.

A TO contínua é obtida pela convolução entre a série temporal da variável de estudo com a função ondeleta $\psi(t)$ dilatada e transladada:

$$W(\tau, s) = \sum_t X_t \frac{1}{\sqrt{s}} \psi^* \left[\frac{t - \tau}{s} \right] \quad (2.2)$$

Onde X_t representa a função da série temporal, $\frac{1}{\sqrt{s}}$ é o fator normalizador de energia das ondeletas, (*) representa o conjugado complexo, τ é o parâmetro de translação e s é o parâmetro de dilatação (escala). A convolução deve ser feita, para cada escala, de acordo com o número total de pontos da série temporal.

O espectro de potência da ondeleta, representa a densidade de energia da ondeleta em tempo e frequência e é obtido pelo quadrado do módulo (ou amplitude) da transformada, já

corrigida com o parâmetro de dilatação ($1/s$) para evitar resultados subestimados para altas frequências (LIU et al., 2007):

$$P(\tau, s) = \frac{1}{s} |W(\tau, s)|^2 \quad (2.3)$$

A transformada de ondeleta cruzada (TOC), do inglês *crosswavelet transform*, é uma ferramenta poderosa para testar conexões propostas entre diferentes variáveis/séries temporais. De acordo com a versão retificada de Veleda et al. (2012), a TOC para duas séries temporais distintas x e y é dada por:

$$W_{x,y}(\tau, s) = \frac{1}{s} W_x(\tau, s) W_y(\tau, s) \quad (2.4)$$

onde W_x e W_y são as TOs das variáveis x e y , respectivamente, que são obtidas pela equação 2.2, e $\frac{1}{s}$ é o parâmetro de dilatação utilizado para correção. A TOC é considerada um análogo da covariância, indicando a interdependência numérica entre duas variáveis.

A potência da ondeleta cruzada é obtida pelo módulo da equação 2.4, sendo uma avaliação da potência das duas séries nos domínios do tempo e frequência:

$$P_{x,y}(\tau, s) = |W_{x,y}(\tau, s)| \quad (2.5)$$

A diferença de fase de uma série x em relação a uma outra série y é dada pelo argumento da TOC, que é o ângulo entre o eixo positivo da parte real da função complexa e o seu módulo (amplitude):

$$Ang_{x,y}(\tau, s) = Arg(W_{x,y}(\tau, s)) \quad (2.6)$$

Esta fórmula é equivalente à diferença de fases individuais convertida para um ângulo dentro do intervalo $[-\pi, \pi]$. Valores absolutos menores que $\pi/2$ caracterizam séries em fase, sendo o sinal um indicativo da série x liderando na região positiva e y na negativa. As séries se encontram defasadas para valores absolutos maiores que $\pi/2$, podendo estar em antifase.

O espectro de ondeleta global é uma forma de representar graficamente o comportamento da variância na ondeleta, exibindo os períodos com maior energia dentro da série, isto é, a potência média ao longo do tempo.

2.4 Visualização gráfica da transformada de ondeleta

Tendo como base a apostila guiada de Rösch e Schmidbauer (2014) para o pacote WaveletComp, são apresentados nesta seção exemplos de séries periódicas e suas saídas gráficas dos espectros de potência uni e bivariado, além dos espectros globais.

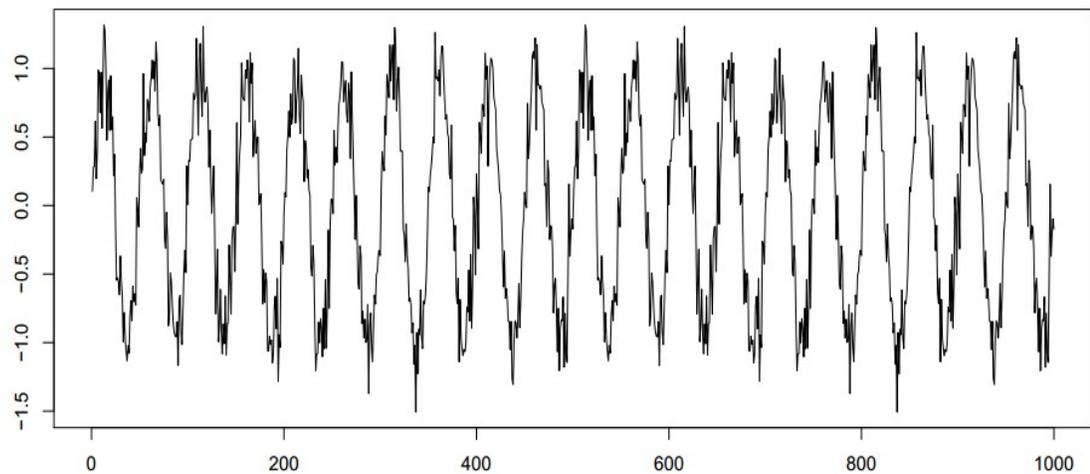


Figura 2.1– Série de período igual a 50, amplitude igual a 1 e comprimento igual a 1000 (Fonte: Rösch e Schmidbauer, 2014).

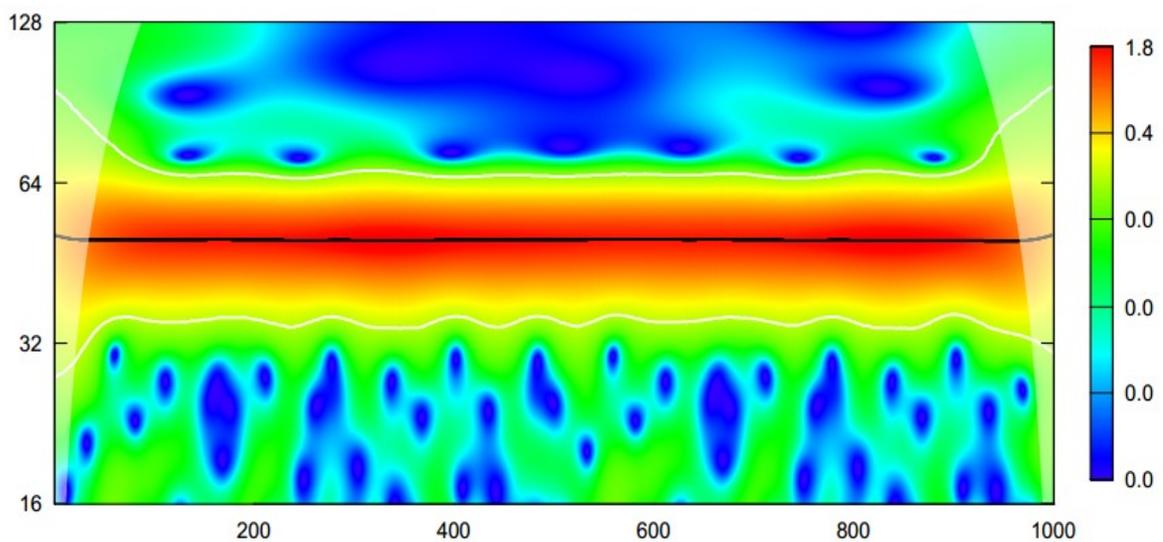


Figura 2.2- Espectro de potência de ondeleta da função periódica com período igual a 50, amplitude igual a 1 e comprimento igual a 1000 (Fonte: Rösch e Schmidbauer, 2014).

i) Série periódica simples com espectro de potência de ondeleta univariada

O primeiro exemplo é uma função periódica com valores estabelecidos de período igual a 50, amplitude igual a 1 e comprimento igual a 1000. A função é representada na figura 2.1. Por conta da função ter um valor constante de período atribuído, isto se traduz no seu espectro de potência (Figura 2.2). O valor máximo de potência, caracterizado pela cor vermelha de acordo com a escala posicionada à direita, se localiza em uma única região do gráfico de forma horizontal ao longo de toda a série, abrangendo a reta correspondente ao período igual a 50, de acordo com a escala vertical.

É importante ressaltar que há um cone de influência representado pelo sombreado branco nas bordas, indicando que as transformadas que estão inseridas nele não são influenciadas por efeitos de borda e possuem significância estatística. Adicionalmente, tem-se contornos em linhas brancas sólidas que ressaltam onde há maior significância dentro o que se encontra no cone, com valor padrão igual a 0,1.

ii) Duas séries periódicas com espectros de ondeleta global

O segundo exemplo retrata duas séries distintas x e y . Ambas possuem o comprimento igual a 1000 e subséries de períodos iguais a 60 e 100, porém distribuídos de formas diferentes (Figura 2.3). A série x possui o período maior (100) somente na sua primeira parte, com a menor periodicidade (60) na segunda parte. Em contrapartida, a série y apresenta as duas subséries de forma distribuída por todo seu comprimento.

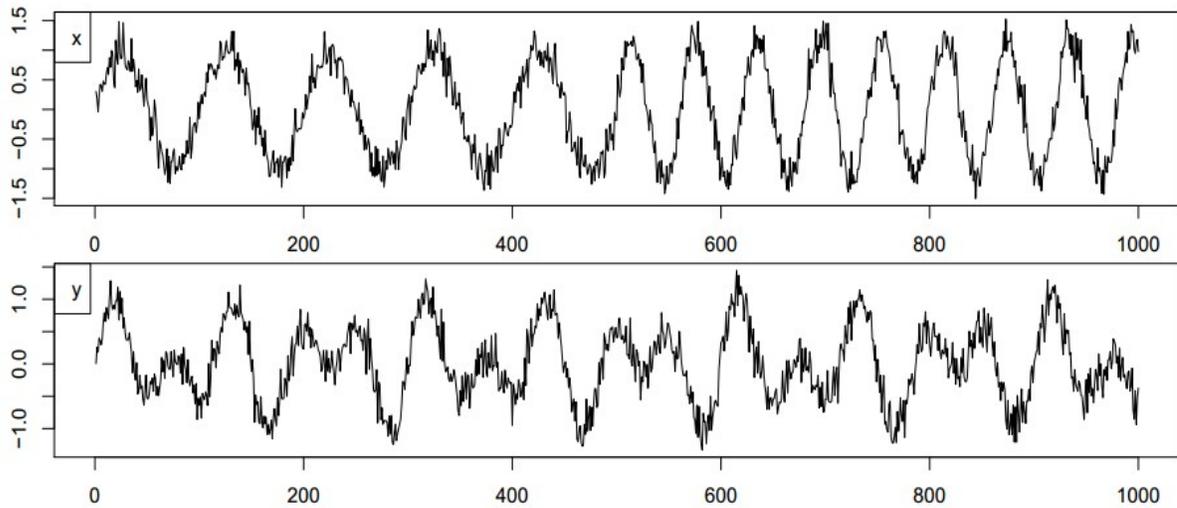


Figura 2.3 - Séries temporais das funções x e y de subperíodos igual a 60 e 100 e comprimento igual a 1000 (Fonte: Rösch e Schmidbauer, 2014).

O espectro de ondeleta global representa a média da potência da série ao longo do tempo (Figura 2.4). O seu eixo horizontal representa a potência média e o eixo vertical representa o período. Existem pontos azuis e vermelhos associados a níveis de significância estatística de 0.01 e 0.05 (99 e 95%), respectivamente, indicando onde estão as seções mais potentes da série ao todo. O que se pode notar é que as duas funções possuem a potência média mais intensa associada aos subperíodos de 60 e 100, apesar das diferenças de distribuição em cada série.

iii) Duas séries periódicas distintas com espectro de potência bivariada e espectro global

Este terceiro exemplo diz respeito a um outro par de funções x e y. Ambas representam observações horárias de um intervalo de 96 dias e possuem períodos 1, 2, 4, 8 e 16 iguais, sendo o período igual a 4 de maior amplitude para ambas as funções. O que difere as séries é a aplicação de maior amplitude para o período igual a 16 somente para a série y de forma adicional (Figura 2.5).

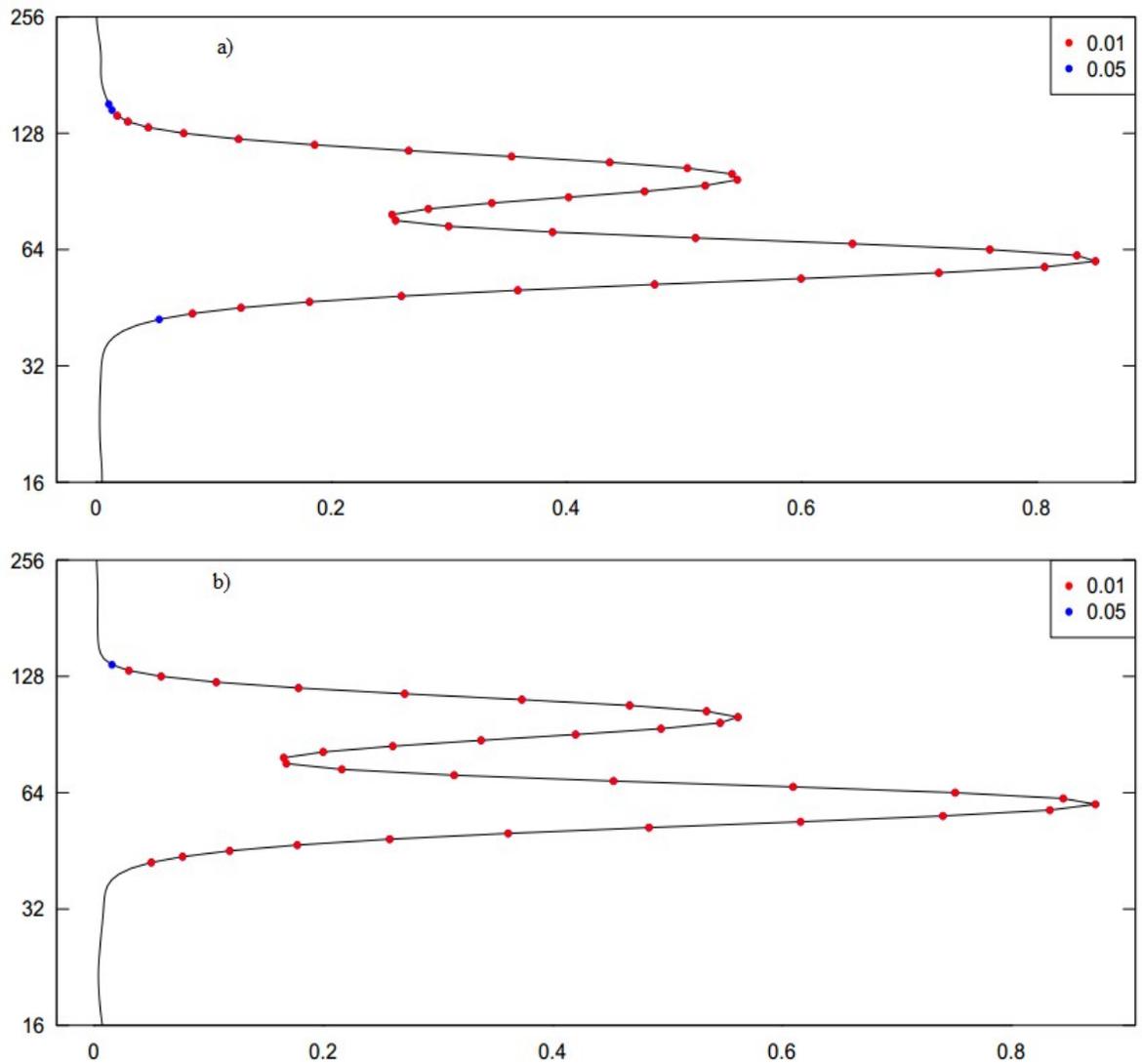


Figura 2.4 - Espectro de ondeleta global relativo a (a) função x e (b) função y (Fonte: Rösch e Schmidbauer, 2014 – adaptado).

O espectro de potência bivariado mostra o nível de covariância entre as séries que se quer avaliar (Figura 2.6). Aqui são utilizados os mesmos critérios e valores de significância estatística do primeiro exemplo desta subseção. É possível observar dentro do cone de influência que os períodos fixados (1, 2, 4, 8 e 16) para as duas funções são os que possuem maiores valores no

espectro. Os que mais se destacam são os períodos iguais a 4 e 16; o primeiro se repete mais intenso nas duas funções, apresentando então maior potência cruzada, enquanto que o segundo valor é bem mais intenso somente para a função y, o que torna evidente a ligeira diferença no sinal de potência.

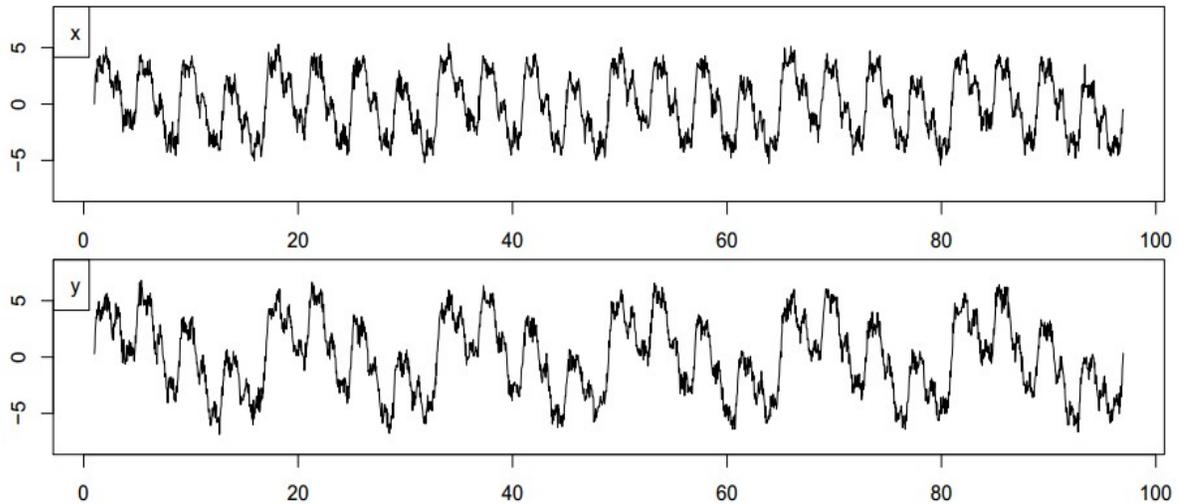


Figura 2.5 - Séries temporais das funções x e y de subperíodos 1, 2, 4, 8 e 16 e comprimento igual a 96 (Fonte: Rösch e Schmidbauer, 2014).

Esta análise bivariada traz, além do espectro de potência de ondeleta, a análise das possíveis diferenças de fase, representadas pelos vetores pretos. É notável que neste caso, eles estão posicionados exatamente para a direita em todas as faixas dos períodos fixos. De acordo com a abordagem apresentada na seção 2.4 sobre este item, isto significa que as funções estão em fase, ou seja, sem atrasos ou adiantos entre si.

O espectro de ondeleta global do cruzamento das funções x e y (Figura 2.7) reforça o significado da saída de potência anterior, mostrando o retorno para todos os períodos fixos, além de mostrar os períodos 4 e 16 como os mais potentes, respectivamente, dentro da significância estatística de 0.1 e 0.05 (90 e 95%).

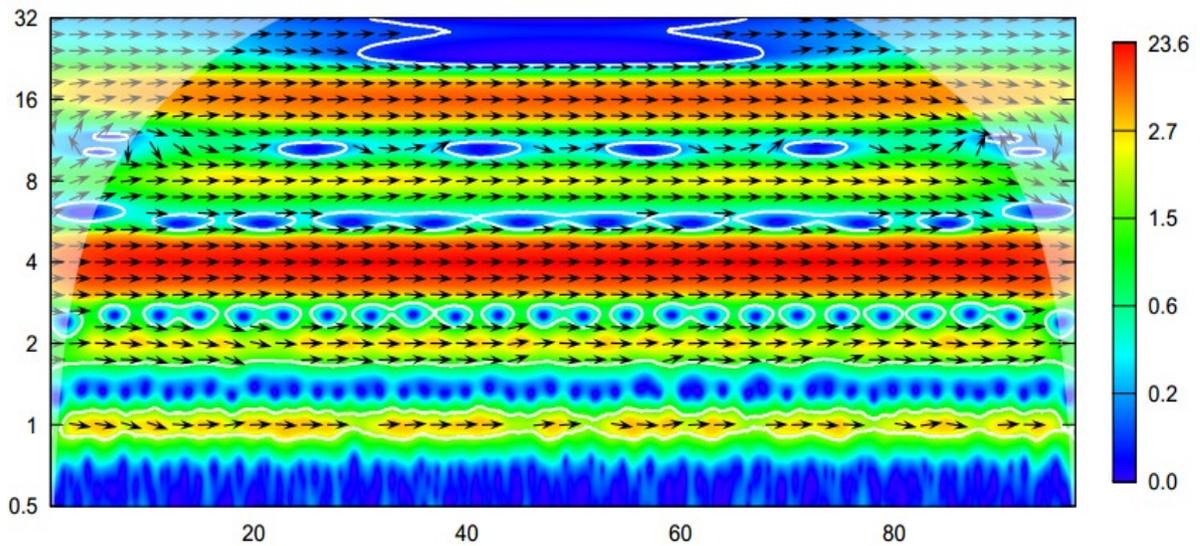


Figura 2.6 – Espectro de potência bivariado das funções x e y de subperíodos 1, 2, 4, 8 e 16 e comprimento igual a 96 (Fonte: Rösch e Schmidbauer, 2014 – adaptado).

2.5 Aplicações da TO no estudo da ICU e da dengue

TO é uma ferramenta multifuncional e que pode ser aplicada em diversos estudos. Os casos presentes na literatura apresentados a seguir exemplificam o seu uso na investigação da periodicidade da dengue e da ICU em diversas localidades e escalas temporais.

Os padrões epidêmicos de dengue foram analisados também simultaneamente em diversas regiões do Sri Lanka, a partir de notificações semanais de dengue entre 2009 e 2014. A dinâmica da dengue mostrou padrões periódicos de 2-8 semanas, 26-32 semanas e 52 semanas. O primeiro padrão mostrou uma característica intermitente, enquanto que os outros apresentaram uma significância inconsistente apesar do espectro de potência indicar altos valores (TALAGALA E LOKUPITIYA, 2015).

Enduri e Jolad (2017) analisaram os surtos de dengue pós-moção na cidade de Ahmedabad, no oeste da Índia, de 2005 a 2012. Foi constatado que há uma defasagem de 20 semanas entre o pico de temperatura média e o pico das incidências do dengue a partir da correlação cruzada da dengue com a variável meteorológica.

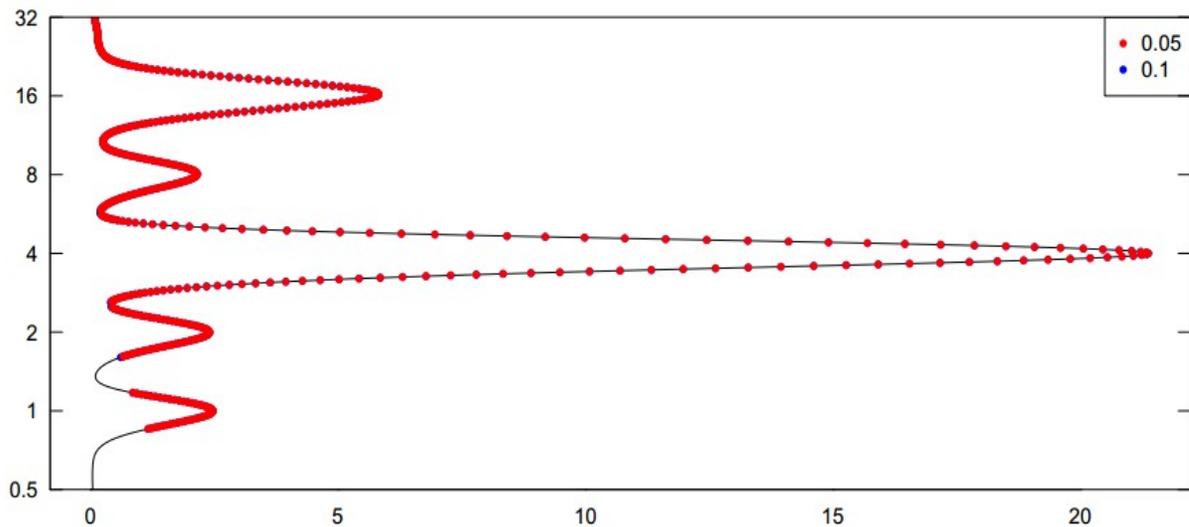


Figura 2.7 - Espectro de ondeleta global do cruzamento das funções x e y de subperíodos 1, 2, 4, 8 e 16 e comprimento igual a 96 (Fonte: Rösch e Schmidbauer, 2014 – adaptado).

Johansson et al. (2009), por exemplo, estudaram a transmissão da dengue em Porto Rico, México e Tailândia utilizando a TO e concluíram que nas áreas tropicais e subtropicais ela tipicamente segue um padrão sazonal pontuado por intervalos de alguns anos por uma grande epidemia. Estas áreas geralmente possuem características climáticas de temperaturas altas e grandes volumes de precipitação, sendo assim ambientes propícios para o crescimento da doença.

A ferramenta TO também se mostra útil nas investigações relacionadas à ICU. Ao se analisar uma série temporal longa de temperatura do ar média mensal na República Tcheca, foram encontrados padrões cíclicos de maior energia nos períodos de 8 a 14 anos, tendo uma tendência de aumento da variável no século XX, sendo mais pronunciado a partir da segunda metade do século (PIŠOFT et al., 2004).

Um país que utiliza a TO em diversas pesquisas de evolução da ICU é a China. Em Xangai, por exemplo, foi notado que a ICU se mostrou bem caracterizada nas escalas semidiurna e diurna, tendo maior intensidade do fenômeno durante o inverno (SUN et al., 2007).

A cidade de Xian, outra localização urbana de destaque, também teve a variabilidade da ICU avaliada. A partir de dados horários de observação de temperatura, foi constatado que ao

comparar as estações inverno e verão, o período de 24h predomina em ambas as estações. Entretanto, dentro da escala de alguns dias, existem diferenças nos períodos principais de ocorrência da ICU para cada estação. Outro ponto a se destacar é a mudança de intensidade da ICU, sendo mais forte nos dias de semana e mais fraca nos finais de semana (HU et al., 2010).

Em Pequim, a ICU apresentou períodos dominantes de 10 e 24 anos para a série de temperatura média observada entre 1960 e 2006. O período de 10 anos se mostra mais caracterizado na primeira metade da série temporal e só volta a mostrar maior potência no final, enquanto que o período de 24 anos é bem demarcado ao longo de toda a série (ZHANG et al., 2010).

Na cidade de Guangzhou, a ICU apresentou padrões diários, semanais e bimestrais como principais, sendo justificados por diversos fatores como ações antropológicas, trocas de calor na superfície terrestre e atuação de sistemas meteorológicos (JIANG et al., 2007).

No Brasil, Vilani e Sanches (2011) fizeram uma comparação da variável temperatura do ar entre o centro urbano de Cuiabá e a área rural do entorno e tiveram como resultado um pico energético na escala de 24 horas para ambas localizações. Entretanto, a energia média encontrada no espectro global foi ligeiramente maior para a localização urbana, sendo associada aos valores de temperatura mais elevados para a região urbana em comparação com a rural.

2.6 Análise de agrupamento e de variância

A análise de agrupamento, do termo em inglês *cluster analysis*, é um método estatístico que tem como finalidade dividir os dados de acordo com os níveis de semelhança e diferença entre as n observações presentes dentro da série. Por exemplo, ela pode ser aplicada para entender os padrões de velocidade do vento na região Nordeste (SANTANA e DA SILVA, 2019), investigar similaridades nos padrões de precipitação para o Rio Grande do Norte em associação com padrões de temperatura da superfície do mar (AMORIM et al., 2020), analisar ocorrências de secas em diversas bacias hidrográficas (SILVA E COSTA, 2020; MELO e SOUSA, 2021), verificar a homogeneidade na distribuição de pontos de queimadas e incêndios (SILVA et al., 2021), entre outros.

De acordo com Wilks (2019), os pontos que compõem os grupos devem possuir distâncias pequenas entre si em comparação com as distâncias entre cada grupo. O método de distanciamento mais utilizado é o euclidiano, que representa a menor distância entre dois objetos em um plano multidimensional (MACHADO, 2011 apud BRAZ et al., 2020).

A análise de agrupamento pode ser realizada através de diversos métodos de separação, que podem ser de origem hierárquica ou não. Dois métodos que são utilizados na literatura são o k-média (do inglês *k-means*) e o hierárquico.

O método *k-means* tem o perfil de agrupamento iterativo, no qual os objetos são realocados para o *cluster* que possui a média mais próxima do agrupamento. Além disso, o método *k-means* é considerado um algoritmo de agrupamento não supervisionado, por gerar agrupamentos a partir de números de classes pré-determinadas (BRAZ et al., 2020). O algoritmo depende de um número *k* de grupos que é escolhido pelo usuário antes da implementação do método. Uma das formas de se realizar essa escolha do número de grupo é através da soma dos quadrados intragrupos. O total de grupos *k-means* a ser utilizado corresponde ao total de grupos em que a diferença entre a soma dos quadrados da distância entre o objeto e o *cluster* central seja a menor possível.

Após a escolha no número *k* de grupos, é feita a seguinte sequência (WILKS, 2019):

- cálculo dos centróides/médias vetoriais de cada grupo;
- cálculo da distância entre o ponto atual e o centróide;
- se o ponto atual já pertencer ao grupo cuja média é mais próxima, é feito um novo cálculo de distância para o próximo ponto; se o ponto atual não pertencer a esse grupo, ele é realocado para o grupo com a média mais próxima.

Essa sequência é realizada para todos os pontos até que eles sejam realocados em um grupo que tenha a média mais próxima a cada um deles, atingindo o patamar de não haver mais nenhuma realocação dos pontos para outro grupo.

O método hierárquico, por sua vez, não permite a realocação dos objetos de análise uma vez que eles são posicionados dentro de um grupo. Como o próprio nome diz, nesse método há uma hierarquia pré-determinada de conjuntos de grupos, cada um dos quais é formado através da junção de um par da coleção de grupos previamente definidos (WILKS, 2019).

A análise hierárquica pode ser realizada de duas formas: aglomerativa e divisiva. No formato aglomerativo, cada ponto inicia como seu próprio *cluster* e novos pares semelhantes de *clusters* são adicionados à medida que se eleva a hierarquia. Já no formato divisivo o contrário acontece, ou seja, todos os dados iniciam em um grupo único e passam a ser divididos em grupos diferentes. O formato divisivo praticamente não é usado em agrupamentos hierárquicos por ser altamente complexo, dificultando sua aplicabilidade em bases de dados mais extensas (METZ, 2006).

No método hierárquico, é utilizado o conceito de distância entre grupos como método de aglomeração de dados. Existem diversas formas de se calcular essa distância. As principais formas de calcular a distância entre grupos citadas por Wilks (2019) são: completa/máxima (considera o maior valor calculado dentre as dissimilaridades existentes entre dois grupos como critério de ligação entre os grupos), mínima (considera o menor valor calculado dentre as dissimilaridades existentes entre dois grupos como critério de ligação entre os grupos), média (considera a média de todas as dissimilaridades existentes entre dois grupos como critério de ligação entre os grupos) e centróide (considera a dissimilaridade existente entre dois centróides como critério de ligação entre os grupos).

A análise de variância, conhecida na literatura pela sigla ANOVA, é um método estatístico utilizado para testar as diferenças entre três ou mais médias populacionais, baseado na análise da variância entre as amostras das populações.

Para ser realizada a ANOVA, devem ser levadas em conta as seguintes condições: as amostras devem ser independentes entre si, a distribuição dos dados deve se assemelhar a uma distribuição normal e as populações devem ter valores de variância próximos entre si (VIEIRA, 2006 apud MORAES JÚNIOR, 2015).

Essa metodologia é usada na literatura com diversas finalidades, como por exemplo: testar distribuições de probabilidade em séries pluviométricas, estudar a possível relação entre a temperatura da superfície do mar e ocorrências de linhas de instabilidade no Oceano Atlântico, verificar a possível tendência de aumento da temperatura média no oeste do Pará (PEREIRA et al., 2014; SILVA et al., 2017; DOS SANTOS GOMES et al., 2015).

O método ANOVA baseia-se em duas hipóteses. Na primeira hipótese, H_0 , assume-se que todas as médias das populações ($\mu_1, \mu_2, \mu_3, \dots, \mu_k$) são aproximadamente iguais, indicando que não há diferenciação entre os grupos (MORAES JÚNIOR, 2015). A segunda hipótese, H_1 , chamada de hipótese alternativa, assume que há diferença entre as médias das populações ou em pelo menos uma parte delas. Para se testar qual das duas hipóteses é válida para os dados que se quer estudar é realizado o teste F, que leva em conta um nível α de significância estatística pré-estabelecido (Figura 3.8).

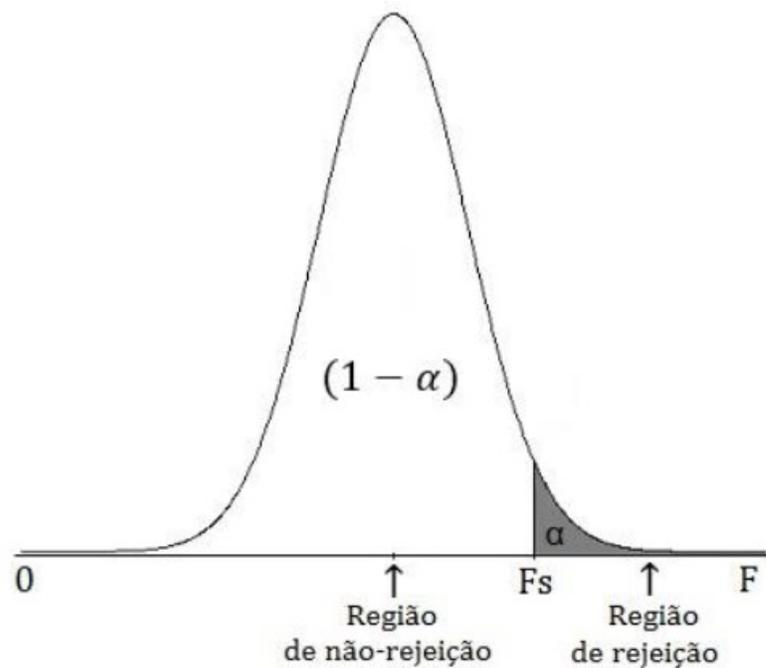


Figura 3.8 - Regiões de rejeição e não-rejeição do teste de hipótese (Fonte: Moraes Júnior, 2015).

Quando o teste F resulta em um valor acima do limiar de rejeição de H_0 , com um valor-p de significância estatística menor que α , a hipótese nula é rejeitada. As definições do teste F e do valor-p são explicitadas na seção 3.3 deste texto.

Capítulo 3

3 Materiais e Métodos

3.1 Área de estudo

A cidade do Rio de Janeiro, capital do estado do Rio de Janeiro, está situada a 22°54'23"S e 43°10'21"W. Ao norte faz limite com vários municípios da Baixada Fluminense, ao sul é banhada pelo Oceano Atlântico, a leste pela Baía de Guanabara e a oeste pela Baía de Sepetiba. Localizada na região Sudeste do Brasil, com uma área de aproximadamente 1.197 km², a cidade do Rio de Janeiro possui o total populacional estimado em 6.747.815 habitantes (IBGE, 2020). O município é dividido em quatro zonas geográficas principais: norte, sul, oeste e centro (Figura 3.9).

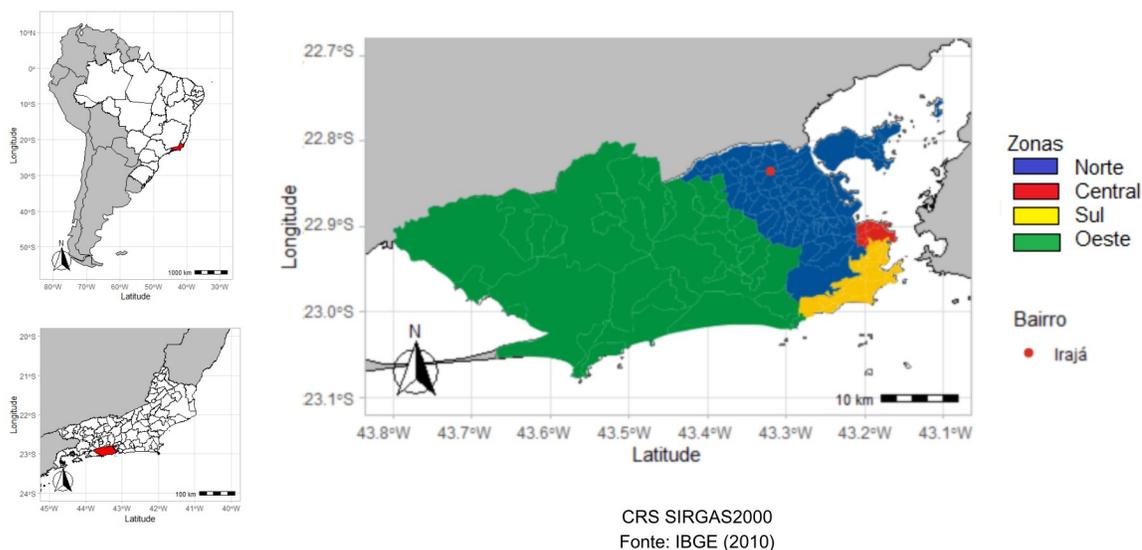


Figura 3.9 - Município do Rio de Janeiro e suas zonas geográficas. Fonte: adaptado de IBGE, 2010 e Nossedotti, 2013.

O município do Rio de Janeiro é o segundo maior do Brasil em total populacional, sendo caracterizado por forte segregação social e espacial, além da diversidade do uso e ocupação do

solo e das paisagens. Um aspecto importante na localidade é a presença de remanescentes de Mata Atlântica, a influência da costa atlântica e seu relevo acidentado (SANTOS et al., 2020).

O clima do município é tropical (quente e úmido), com temperatura variando entre 20 °C e 27 °C. Os meses mais quentes são os compreendidos entre novembro e abril e os mais frios, entre maio e outubro. Durante o verão, a temperatura pode chegar a 40 °C, e nas noites de inverno pode atingir 15 °C (CÂMARA et al., 2009).

No ano de 2020, o bairro de Irajá, denotado na Figura 3.9 com um ponto vermelho, recebeu destaque por ter sido considerado o mais quente do município (ALTINO, 2021). Além disso, o bairro veio tendo uma crescente na representatividade dos valores de temperatura mais altos dentro do município ao longo da última década, principalmente a partir de 2018, segundo os boletins de temperaturas máximas mensais emitidos pelo Sistema Alerta Rio da Prefeitura do Rio de Janeiro (disponíveis em <http://alertario.rio.rj.gov.br/temp-maxs-mins-mensais/>). A população do bairro de Irajá em 2010 foi contabilizada no total de 96382 habitantes (IBGE, 2010). O bairro fica localizado na Zona Norte, área ocupada principalmente por áreas residenciais e comerciais; por outro lado, a cobertura arbórea é praticamente inexistente (IPP, 2018). Além disso, ele se localiza dentro de uma das áreas do município que satisfazem os critérios para classificação da ICU (LUCENA e PERES, 2017). Por conta desses fatores, o bairro se demonstrou uma localização com grande potencial para a análise da ICU no município do Rio de Janeiro.

3.2 Dados utilizados na análise

Os dados de notificação de infecção de dengue serão obtidos online através da Secretaria Municipal de Saúde (SMS). Eles são de domínio público, divulgados a partir de tabelas de dados para cada ano, apresentando os dados totais do município, além dos dados separados por bairro (<http://www.rio.rj.gov.br/web/sms/exibeconteudo?id=2815389>). Os dados a serem utilizados abrangem o período de 2000 a 2019, possuindo intervalo semanal entre os anos de 2000 a 2010 e

de 2015 a 2019. Entre os anos de 2011 e 2014 as notificações de dengue são disponibilizadas com totais mensais.

Para a análise de variância da transmissão de dengue dentro do município, foram utilizados os dados obtidos pela SMS no período entre 2000 e 2019, sendo feitas as somas dos totais semanais dos anos disponibilizados com intervalo semanal para que todos os anos tivessem intervalo mensal. Já para a análise de ICU, foram utilizados os dados de dengue entre 2015 e 2019, com intervalo semanal. Todas as notificações de dengue foram organizadas em planilhas do Microsoft Excel, descartando-se os casos classificados como “Ignorados”. Durante a organização dos dados, foram excluídos três bairros por ausência de registros de dengue: Gericinó, Vasco da Gama e Parque Colúmbia.

Os dados de temperatura do ar serão provenientes das estações meteorológicas do Sistema AlertaRio, disponíveis em

<<http://www.sistema-alerta-rio.com.br/dados-meteorologicos/download/dados-meteorologicos>>.

Os dados meteorológicos abrangem o período de 2015 a 2019, possuindo intervalo de 15 em 15 minutos. A partir desses dados, foram obtidos os valores de temperatura média semanal e temperatura mínima semanal, calculados no Microsoft Excel.

3.3 Métodos para análise dos dados

Um aspecto expressivo abordado dentro do estudo da ICU é a variação de respostas da transmissão de dengue para diferentes variáveis térmicas. Câmara et al. (2009) detectaram uma relação significativa entre dengue e temperatura mínima, como já citado na fundamentação teórica. Isso traz uma motivação para pesquisar se existe essa influência dentro da localização que será classificada como referência para ICU. No trabalho presente, é proposta a avaliação das temperaturas média e máxima em conjunto com a temperatura mínima para verificar possíveis diferenças de influência sobre a dengue.

Além da comparação das variáveis térmicas semanais, propõe-se nesse trabalho o uso da diferença entre a temperatura entre as semanas, tanto para a temperatura média quanto temperaturas mínima e máxima semanais, a fim de se tentar observar como o comportamento das

temperaturas entre diferentes semanas pode afetar a transmissão de dengue nessa escala temporal. A relação é explicada na equação abaixo:

$$Dif = T_i - T_{i-1} \quad (3.1)$$

Onde T_i é a temperatura da semana atual e T_{i-1} é a temperatura da semana anterior.

Adicionalmente, Vilani e Sanches (2011) fizeram uso da metodologia de TO no estudo da ICU para Cuiabá – MT, obtendo resultados relevantes para a localidade e mostrando a utilidade da ferramenta TO em estudos de investigação de possíveis padrões de variáveis térmicas.

Ditas as observações acima, a relação entre dengue e ICU será avaliada através da verificação do nível de influência das variáveis temperatura média, temperatura máxima e temperatura mínima na transmissão de dengue dentro da localização de referência para os extremos térmicos no município do Rio de Janeiro, com o auxílio do método de TO para análise de frequência.

Além da análise de TO uni e bivariada, também é feita uma análise estatística a partir da construção de equações de regressão linear múltipla. A primeira equação é construída tendo a dengue como variável dependente e a incidência por bairros como variáveis explicativas, enquanto que a segunda equação tem como variáveis independentes as saídas dos grupos gerados através da análise de agrupamento citada acima.

Sequencialmente, é feita a análise da estatística descritiva e da variância através de diversas medidas: resíduos, erro padrão residual, erro padrão, valor t, valor p, $\Pr(>|t|)$, R^2 , teste F, soma dos quadrados, valor F, $\Pr(>F)$, RSS (*Residual Sum of Squares*). A variância é avaliada com o objetivo de averiguar quais bairros possuem uma relação de maior significância estatística com a dengue, para cada forma de RLM abordada. A seguir, são listadas as medidas estatísticas utilizadas para a análise dos resultados.

- Resíduos: são a diferença entre a previsão do modelo e o resultado real da variável dependente. Valores pequenos são considerados melhores para o resultado do modelo. A sua saída na sumarização do modelo RLM apresenta os valores dos quantis e mediana, representando como estes resíduos estão distribuídos.

- Erro padrão residual: é um dos resultados obtidos na estatística descritiva aplicada ao modelo de regressão linear, que consiste na divisão entre a soma dos quadrados dos resíduos e um fator $N - (K+1)$, onde N é o total de variáveis e K é o número total de variáveis subtraído de 1, devido à constante de interceptação. Representa o desvio padrão dos resíduos. Quanto menor seu valor, melhor o modelo de regressão linear. Ele é utilizado na obtenção dos resultados da coluna “Erro padrão” na seção 4.3. Sua fórmula é representada na equação 3.1:

$$EPR = \frac{\sqrt{\sum (y - \hat{y})^2}}{(N - (K + 1))} \quad (3.2)$$

onde y é o valor previsto pelo modelo e \hat{y} é o valor observado.

- Erro padrão: é feito para cada índice a partir da razão entre o erro padrão residual e a raiz quadrada da soma dos quadrados dos resíduos (SQR) de cada um deles. Indica o quão bem foram estimados os coeficientes de regressão. Utilizado para calcular o valor t de cada variável da regressão. A equação 3.2 mostra como esta medida é obtida.

$$EP = EPR / \sqrt{SQR} \quad (3.3)$$

- Valor t: cada coeficiente estimado é dividido pelo seu respectivo erro padrão. Com este valor, é possível verificar se os coeficientes possuem valores próximos de zero. Quanto maior em módulo, mais distante de zero está o coeficiente.

- Valor p: conhecido como probabilidade de significância, é definido como a probabilidade de se obter uma estatística de teste igual ou mais extrema que aquela observada em uma amostra, sob a hipótese nula. O seu limiar é fixado em 0,05. Quando o valor obtido pelo modelo for menor que este limiar, indica significância estatística. Amostras maiores geram menores valores p pequenos, enquanto que amostras menores geram valores p grandes. Seu valor final na sumarização da RLM é utilizado em conjunto com o F teste para a verificação do nível de significância das variáveis preditoras do modelo RLM.

- $\Pr(>|t|)$: indica o nível de significância do valor t para cada variável. Em outras palavras, é o valor p para cada valor t obtido por coeficiente. O módulo de t é utilizado pois o valor t pode ser negativo. O seu limiar é fixado em 0,05. Sendo $\Pr(>|t|)$ menor que este limiar, maior a significância da variável para o modelo.

- R^2 : conhecida como coeficiente de determinação, esta medida mostra o quanto a variância é explicada pelo modelo ou, em outras palavras, o quão bem o modelo se ajusta aos dados observados. Varia entre 0 e 1. Quanto mais próximo de 1, melhor o desempenho do modelo com relação aos dados. Sua fórmula é dada pela equação 3.3:

$$R^2 = 1 - \frac{\sum (y - \hat{y})^2}{\sum (y - \bar{y})^2} \quad (3.3)$$

onde y é o valor previsto pelo modelo, \hat{y} é o valor observado e \bar{y} é a média das observações.

- Teste F: teste de análise global para auxiliar na análise do modelo. O teste F avalia os diversos coeficientes de forma simultânea, comparando o modelo de RLM com um modelo baseado somente na constante de intercepção, ou seja, sem variáveis preditoras. Este modelo é a hipótese nula H_0 do teste. Com isto, tenta estimar se a inserção dessas variáveis é significativa para o modelo. Caso a hipótese nula H_0 seja verdadeira, elas não fazem diferença em conjunto no resultado do modelo RLM. Caso contrário, o modelo exibe melhor ajuste com a adição de variáveis preditoras. Neste teste, o R^2 do modelo obtido anteriormente é utilizado (equação 3.4).

$$F = \frac{R^2 / (K - 1)}{(1 - R^2) / (N - K)} \quad (3.4)$$

onde N representa o número total de observações e K , o número de variáveis do modelo+1 para o interceptador.

O número de graus de liberdade (GL) para o numerador é representado pelo fator $k-1$, enquanto que para o denominador ele é representado pelo fator $N-K$. Estes valores de GL são associados ao próprio resultado do teste F, sendo úteis para a análise em conjunto com o valor p do modelo, para rejeitar ou não a hipótese nula H_0 . Caso o valor p para os GL do teste F seja muito pequeno em comparação ao limiar de 0,05, H_0 é rejeitada, demonstrando que diferentes variáveis influenciam na qualidade do modelo de regressão.

- Soma dos quadrados explicados (SQE): é a soma do quadrado da diferença entre o valor previsto pelo modelo e a média das observações (equação 3.5). Quanto maior seu valor, melhor a performance do modelo. Esta medida estatística faz parte da saída de resultados presentes na tabela de variância (tabela ANOVA).

$$SQE = \sum (\hat{y} - \bar{y})^2 \quad (3.5)$$

onde \hat{y} é o valor previsto e \bar{y} é a média dos valores observados.

- Valor F: definido pela razão entre a variação das médias amostrais e a variação dentro do próprio grupo da variável em questão. Sendo o valor F grande, a hipótese nula H_0 de que as médias sejam iguais é descartada, ou seja, as variáveis preditoras conseguem descrever a variável dependente. Esta medida é utilizada na análise de variância do modelo linear (tabela ANOVA).

- $\Pr(>F)$: indica o nível de significância do valor F para cada variável (ou seja, é o valor p para cada valor F na análise de variância). O seu limiar é fixado em 0,05. Sendo $\Pr(>|t|)$ menor que este limiar, maior a sua significância para o modelo.

- Soma dos quadrados dos resíduos (SQR): A soma dos quadrados dos resíduos é considerada uma medida da discrepância entre o modelo e os dados observados (equação 3.6). Este valor é calculado para cada um dos índices e é utilizado na obtenção do erro padrão dos coeficientes da RLM completa. Quanto menor o seu valor, melhor o ajuste do modelo com relação ao dados.

$$SQR = \sum (y - \hat{y})^2 \quad (3.6)$$

Em suma, os resultados serão exibidos na próxima seção com o auxílio das saídas gráficas listadas a seguir:

a) Espectro de potência de ondeleta - É apresentado tanto na versão uni quanto na bivariada. Para ambas, o eixo vertical mostra os períodos em semanas, o horizontal mostra a contagem de tempo, também em semanas. Uma curva de secção de um cone limita a região com significância estatística da análise de ondeletas, indicando que os resultados plotados na área externa (sombreado branco) estão suscetíveis a efeitos de borda e não possuem significância estatística. Há também linhas brancas para períodos específicos (inserida a posteriori) para indicar os períodos em que a potência da ondeleta apresenta máximo. A barra de cores representa o nível de potência do sinal; cores mais quentes indicam maiores valores de potência, cores mais frias indicam menores valores de potência. A versão bivariada apresenta a diferença de fase entre as séries no caso de ondeletas cruzadas, representada por vetores (quanto mais apontado para a direita indica variáveis em fase e quanto mais para a esquerda, anti-fase). Também é realizada a

análise por ondeleta global, que possui o mesmo tipo de representação gráfica indicado acima. O eixo vertical mostra os períodos em semanas. O eixo horizontal mostra o tempo, ordenado pelo número de semanas. A função de distribuição das potências das ondeletas tem valores apresentados em escala de cores. As transformadas descritas acima foram obtidas a partir do pacote WaveletComp do software R versão 4.2.2, como descrita por Rösch e Schmidbauer (2014).

b) Análise de agrupamento (“*cluster analysis*”) – Uma análise de agrupamentos (*cluster analysis*) foi implementada para os dados de notificações de agravamento por dengue do SUS no seu absoluto, bem como em proporção com os totais populacionais dos bairros da cidade do Rio de Janeiro. O número de grupos considerados representativos foi determinado através da soma dos quadrados intra *clusters* para o método *k-means*, sendo também esse número aplicado para o corte do dendrograma no método hierárquico, mantendo-se a mesma quantidade de grupos em ambos os métodos para fins de comparação. Neste trabalho foram utilizados scripts de linguagem R versão 4.2.2 para análise de agrupamentos, customizados para a análise em curso.

c) Análise da variância particionada (“*variance by parts*”) – Essa metodologia ANOVA por partes foi utilizada para caracterizar a contribuição à explicação da variância global das ocorrências de dengue por bairro da cidade do Rio de Janeiro, bem como a explicação da variância da dengue através de cada grupo gerado na saída da análise de agrupamento *k-means*. Para ambos os casos, também foram utilizados scripts R versão 4.2.2.

d) Espacialização de dados – Os resultados das análises de variância e de agrupamento foram organizados em mapas do município do Rio de Janeiro, com os respectivos valores das diferentes variáveis estatísticas para cada bairro. Para isso, foi utilizado o pacote “geobr” dentro dos scripts de linguagem R versão 4.2.2.

A figura 4.10 sumariza os passos realizados para a obtenção dos dados e as metodologias utilizadas neste presente trabalho.

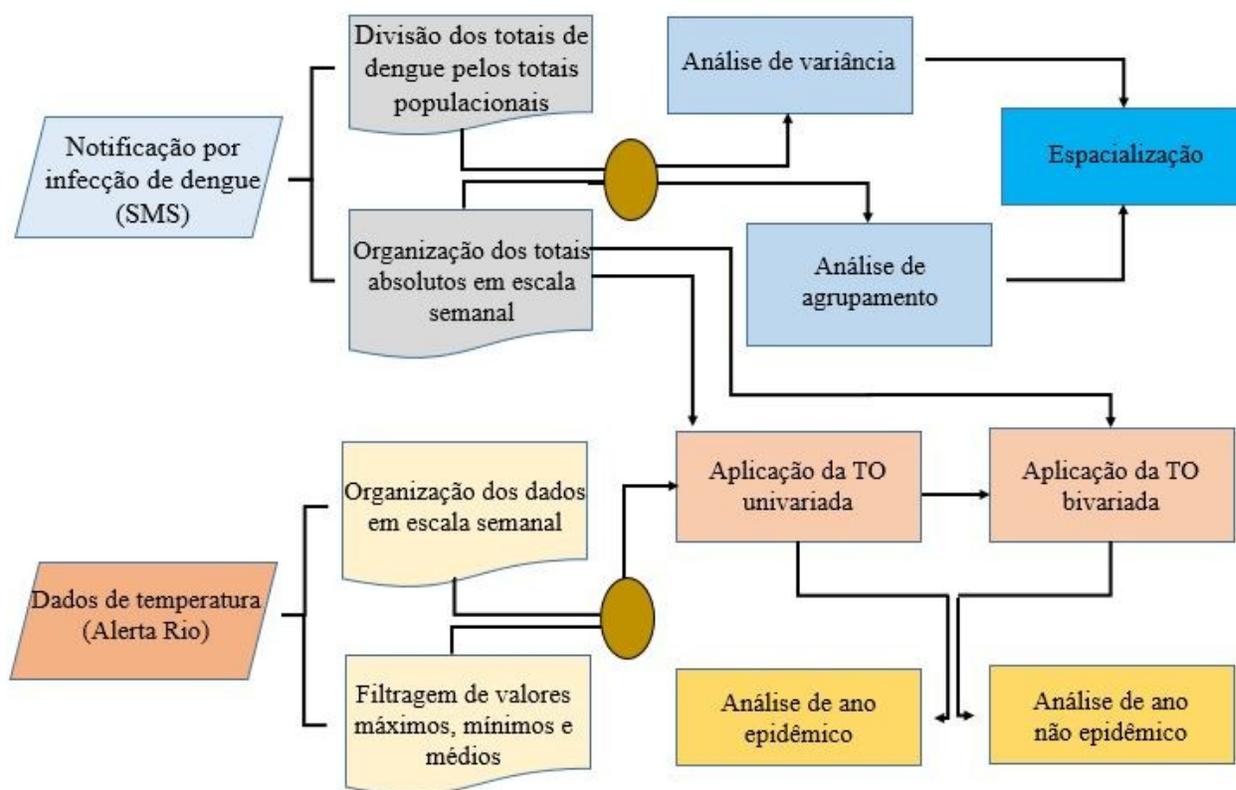


Figura 4.10 – Fluxograma dos dados e métodos utilizados para obtenção dos resultados do estudo. Fonte: Este trabalho.

Capítulo 4

4. Resultados

Para a geração dos resultados das TOs e das análises de variância e de agrupamento, tanto a dengue quanto as temperaturas foram normalizadas a fim de se ter todas as variáveis com magnitudes similares para possibilitar as comparações entre elas.

4.1 Análise de agrupamento e de variância

A partir da soma dos quadrados intragrupos, nota-se que os bairros puderam ser divididos em quatro grupos principais. Isso ocorre tanto para o agrupamento *k-means* feito a partir do total absoluto de cada bairro (Figura 4.11 a), quanto para o agrupamento feito a partir das notificações proporcionalizadas a partir do total populacional de cada bairro (Figura 4.11 b). Os bairros são representados por uma numeração dentro da ordem das tabelas de incidências de dengue disponibilizadas pela SMS (Tabela 4.1).

Na figura 4.11 a, temos em destaque o grupo 3, representado pela cor amarela. Ele possui somente quatro bairros, que possuem os maiores números absolutos de ocorrências de dengue. Os bairros são Bangu (número 145), Campo Grande (número 148), Santa Cruz (número 156) e Realengo (número 143). Estes bairros ficam localizados na Zona Oeste da cidade, possuindo alguns dos maiores totais populacionais do município.

O grupo 4, representado pela cor roxo, possui quantidades significativas de casos de dengue, entretanto ligeiramente menores em comparação com o grupo 3. Ele é formado pelos bairros Senador Camará (número 146), Guaratiba (número 152), Paciência (número 155), Inhoaíba (número 150), Santíssimo (número 147), Pedra de Guaratiba (número 154), Sepetiba (número 157) e Complexo do Alemão (número 66). Todos os bairros também se localizam na Zona Oeste da cidade, tendo proximidade com os bairros do grupo 3. A exceção fica por conta do bairro Complexo do Alemão, que é um conjunto de comunidades localizado na Zona Norte da cidade.

Tabela 4.1 - Lista dos bairros do município do Rio de Janeiro utilizados na análise de agrupamento e suas respectivas numerações.

Nome	Número	Nome	Número	Nome	Número
Saúde	1	Pitangueiras	54	Oswaldo Cruz	107
Gamboa	2	Praia da Bandeira	55	Bento Ribeiro	108
Santo Cristo	3	Cocotá	56	Marechal Hermes	109
Caju	4	Bancários	57	Guadalupe	110
Centro	5	Freguesia (Ilha)	58	Anchieta	111
Catumbi	6	Jardim Guanabara	59	Parque Anchieta	112
Rio Comprido	7	Jardim Carioca	60	Ricardo de Albuquerque	113
Cidade Nova	8	Tauá	61	Coelho Neto	114
Estácio	9	Moneró	62	Acari	115
São Cristóvão	10	Portuguesa	63	Barros Filho	116
Mangueira	11	Galeão	64	Costa Barros	117
Benfica	12	Cidade Universitária	65	Pavuna	118
Paquetá	13	Complexo do Alemão	66	Parque Colúmbia	119
Santa Teresa	14	Maré	67	Jacarepaguá	120
Flamengo	15	Higienópolis	68	Anil	121
Glória	16	Maria da Graça	69	Gardênia Azul	122
Laranjeiras	17	Del Castilho	70	Curicica	123
Catete	18	Inhaúma	71	Freguesia (Jacarepaguá)	124
Cosme Velho	19	Engenho da Rainha	72	Pechincha	125
Botafogo	20	Tomás Coelho	73	Taquara	126
Humaitá	21	São Francisco Xavier	74	Tanque	127

Urca	22	Rocha	75	Praça Seca	128
Leme	23	Riachuelo	76	Vila Valqueire	129
Copacabana	24	Sampaio	77	Cidade de Deus	130
Ipanema	25	Engenho Novo	78	Joá	131
Leblon	26	Jacaré	79	Itanhangá	132
Lagoa	27	Lins de Vasconcelos	80	Barra da Tijuca	133
Jardim Botânico	28	Méier	81	Camorim	134
Gávea	29	Todos os Santos	82	Vargem Pequena	135
Vidigal	30	Cachambi	83	Vargem Grande	136
São Conrado	31	Engenho de Dentro	84	Recreio dos Bandeirantes	137
Rocinha	32	Água Santa	85	Grumari	138
Praça da Bandeira	33	Encantado	86	Deodoro	139
Tijuca	34	Piedade	87	Vila Militar	140
Alto da Boa Vista	35	Abolição	88	Campo dos Afonsos	141
Maracanã	36	Pilares	89	Jardim Sulacap	142
Vila Isabel	37	Jacarezinho	90	Magalhães Bastos	143
Andaraí	38	Vila Kosmos	91	Realengo	144
Grajaú	39	Vicente de Carvalho	92	Padre Miguel	145
Manguinhos	40	Vila da Penha	93	Bangu	146
Bonsucesso	41	Vista Alegre	94	Senador Camará	147
Ramos	42	Irajá	95	Santíssimo	148
Olaria	43	Colégio	96	Campo Grande	149
Penha	44	Campinho	97	Senador Vasconcelos	150
Penha Circular	45	Quintino Bocaiúva	98	Inhoaíba	151

Brás de Pina	46	Cavalcanti	99	Cosmos	152
Cordovil	47	Engenheiro Leal	100	Guaratiba	153
Parada de Lucas	48	Cascadura	101	Barra de Guaratiba	154
Vigário Geral	49	Madureira	102	Pedra de Guaratiba	155
Jardim América	50	Vaz Lobo	103	Paciência	156
Ribeira	51	Turiaçu	104	Santa Cruz	157
Zumbi	52	Rocha Miranda	105	Sepetiba	158
Cacuaia	53	Honório Gurgel	106		

Por sua vez, o grupo 2 (cor bordô) é formado pelos bairros Taquara (número 125), Tijuca (número 34), Copacabana (número 24), Rocinha (número 32), Jacarepaguá (número 119), Barra da Tijuca (número 132), Bonsucesso (número 41), Curicica (número 122), Irajá (número 95), Padre Miguel (número 144), Pavuna (número 118), Vila Isabel (número 37), Humaitá (número 20), Penha (número 44), Ramos (número 42), Méier (número 81), Anil (número 120), Guadalupe (número 110), Anchieta (número 111), Jardim Sulacap (número 141), Cidade de Deus (número 129) e Praça Seca (número 127). Similarmente ao grupo 4, o grupo 2 apresenta boa parte dos bairros com totais significativos de casos de dengue, principalmente os bairros localizados na porção mais à esquerda da pepita do grupo. Um dos bairros em destaque nesse grupo é o bairro de Irajá, que apresenta uma variância significativa nos números de casos absolutos de dengue durante as epidemias.

No grupo 2, alguns bairros são da Zona Oeste (Taquara, Jacarepaguá, Barra da Tijuca, Curicica, Anil, Padre Miguel, Jardim Sulacap, Cidade de Deus e Praça Seca), outros pertencem à Zona Sul (Copacabana, Rocinha, Humaitá) e uma porção se localiza na Zona Norte (Tijuca, Bonsucesso, Irajá, Pavuna, Vila Isabel, Penha, Ramos, Méier, Guadalupe, Anchieta). Com isso, é possível perceber que ao se ter uma diminuição gradual dos totais absolutos de dengue, o agrupamento deixa de ser voltado para a Zona Oeste e começa a incluir bairros mais afastados geograficamente e com outras configurações territoriais.

O último grupo a ser comentado (grupo 1 – azul), apresenta os bairros Cosmos (número 151), Senador Vasconcelos (número 149), Jacarezinho (número 90), Colégio (número 96), Rocha Miranda (número 105), Engenheiro Leal (número 100), Brás de Pina (número 46), Jardim América (número 50), dentre os demais bairros da Zona Norte e Zona Sul da cidade, sendo eles a maioria neste grupo. Também fazem parte alguns bairros menores da Zona Oeste como Joá (número 130) e Grumari (número 137). O grupo 1 possui os menores totais absolutos de dengue dentro do município do Rio de Janeiro.

Por outro lado, ao se analisar a figura 4.11 b, é visível a mudança de configuração de agrupamento *k-means* ao se utilizar os totais de dengue proporcionalizados pela população de cada bairro. O grupo 2 (cor bordô) é composto pelos bairros Campo dos Afonsos (número 140), Zumbi (número 52), Barra de Guaratiba (número 153), Camorim (número 133) e Pedra de Guaratiba (número 154). O grupo 4 (cor roxo) é composto pelos bairros Caju (número 4), Anil (número 120), Curicica (número 122), Bonsucesso (número 41) e Saúde (número 1). O grupo 3, por sua vez, é composto somente por dois bairros: Joá (número 130) e Grumari (número 137).

Os três grupos são os que apresentam os bairros com populações pequenas e as maiores proporcionalidades de dengue por população. O maior destaque fica por conta do grupo 3, que apresenta dois dos menores bairros em relação ao total populacional, porém com as maiores proporções de infecção por dengue. No bairro do Joá, por exemplo, que possui o total de 818 habitantes de acordo com o Censo 2010, houve um total de 448 infectados durante um episódio epidêmico no município do Rio de Janeiro.

Por último, o grupo 1 (azul) apresenta todos os outros bairros do município, que possuem menores proporcionalidades de infecção de dengue por total populacional. Esse grupo inclui os bairros que possuem os maiores totais absolutos na figura 4.11 a (Campo Grande, Bangu, Santa Cruz e Realengo). Apesar desses bairros terem esses máximos de dengue, seus totais populacionais também são elevados, o que leva a proporção a ser menor em relação aos bairros dos grupos 2, 3 e 4, principalmente em relação ao grupo 3.

A Figura 4.12 mostra a distribuição espacial dos quatro grupos obtidos na análise de agrupamento *k-means* para cada bairro do município do Rio de Janeiro tanto para os totais absolutos quanto para os totais proporcionalizados de dengue por bairro, a fim de facilitar a

associação do número de cada grupo com a localização geográfica de cada bairro listado. Os bairros em branco (Gericinó, Parque Colúmbia e Vasco da Gama) são os bairros que não foram incluídos por terem os totais de dengue zerados em todos os anos.

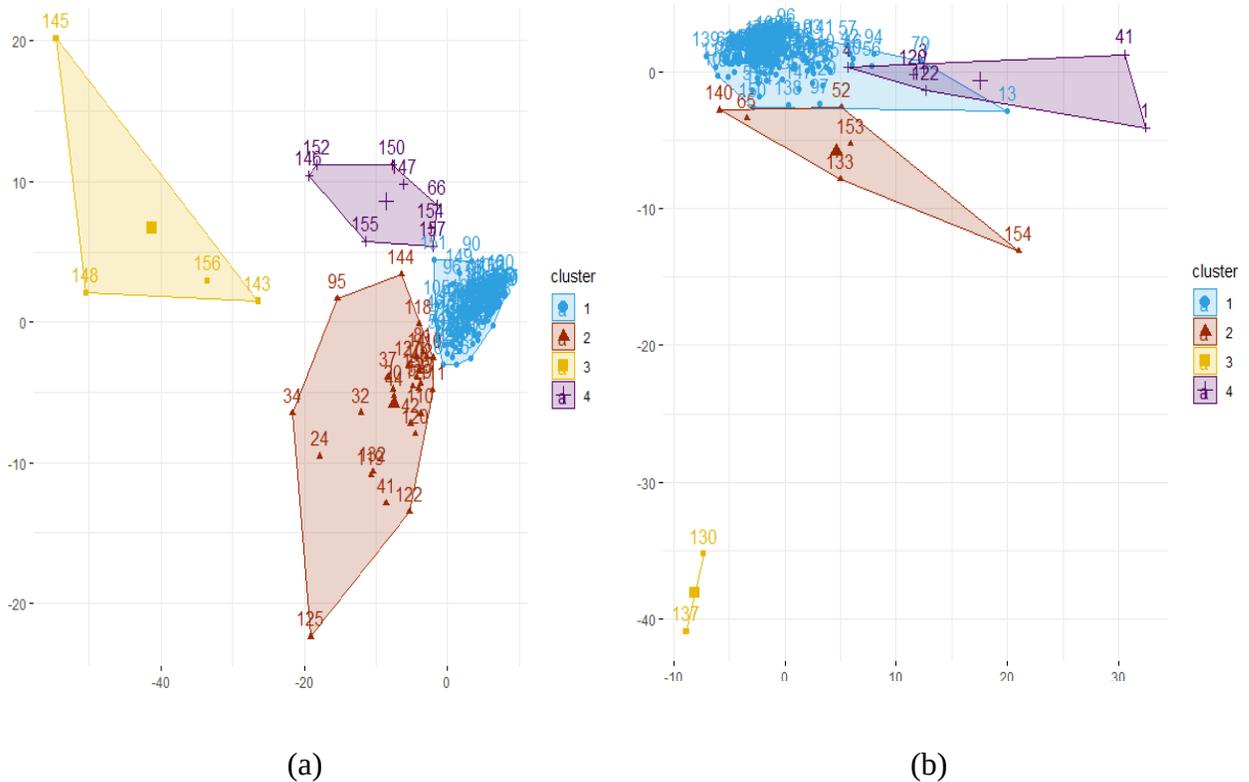


Figura 4.11 – Agrupamento dos bairros do município do Rio de Janeiro via método *k-means* a) em função das notificações de dengue e b) em função das notificações de dengue proporcionais à população de cada bairro.

A partir da Figura 4.12a fica mais fácil notar que o grupo 3, que apresenta os maiores totais de casos de dengue no município, se concentra na Zona Oeste da cidade. O grupo 4, que apresenta totais de dengue menores em comparação aos bairros do grupo 3, porém ainda significativos, também se localiza na Zona Oeste. Também compondo o grupo 4, tem-se o

Complexo do Alemão, que fica na Zona Norte. Os grupos 1 e 2 se distribuem tanto nas Zonas Oeste, como na Zona Norte e Zona Sul. Entretanto, o grupo 2 apresenta bairros com totais de dengue significativos, como por exemplo o bairro de Irajá. As zonas Oeste e Norte são as zonas com os maiores totais populacionais da cidade, tendo em ambas a existência de problemas como falta de saneamento básico e restrição nos serviços de saúde (OLIVEIRA, 2019).

Na zona Oeste há uma exceção para os bairros da Barra da Tijuca e Recreio dos Bandeirantes, por serem bairros com uma população que possui maior poder aquisitivo em comparação aos moradores dos outros bairros que compõem a Zona Oeste. A Zona Sul também apresenta, de forma geral, um perfil de bairros constituídos por moradores de classes sociais mais altas. Esses bairros mais favorecidos da cidade, mesmo que apresentando totais significativos de dengue, tendem a sofrer menos com os impactos da doença pela existência de melhor infraestrutura e saneamento, que dificulta a formação de espaços onde o mosquito vetor possa ter seu ciclo de vida efetivado, isto é, locais onde possa haver o acúmulo de água e o mosquito possa realizar a deposição dos seus ovos (CONSOLI e LOURENÇO-DE-OLIVEIRA, 1994).

Por outro lado, ao analisar a distribuição da dengue proporcionalizada (Figura 4.12 b), os bairros que antes estavam incluídos nos grupos com os maiores totais absolutos de dengue agora passam a estar incluídos no grupo referente aos menores valores proporcionais, tendo então praticamente o município do Rio de Janeiro inteiro proporções menores entre as ocorrências de dengue e a quantidade de habitantes por bairro. Os outros grupos são distribuídos por alguns bairros da Zona Oeste e Centro, que não possuem um grande número de habitantes, como por exemplo Joá, Grumari e Saúde. Nesses grupos também há a presença de outros bairros da Zona Oeste e da Zona Norte com totais populacionais um pouco maiores, como Curicica, Barra de Guaratiba, Pedra de Guaratiba e Bonsucesso.

Ao se usar a abordagem de agrupamento hierárquica (Figura 4.13), é perceptível a diferença de resultados. Enquanto que no método anterior foi possível realizar algum tipo de separação significativa entre grupos, na técnica hierárquica a separação entre grupos é praticamente nula. Na Figura 4.13a, todos os bairros acabam sendo agrupados praticamente em um único *cluster*, exceto pelos bairros Bangu (número 145), Campo Grande (número 148), Santa Cruz (número 156) e Inhoaíba (número 150). Os três primeiros bairros compõem o grupo 3 na

abordagem *k-means*, enquanto que o bairro 150 faz parte do grupo 4 na mesma abordagem. Entretanto, no método hierárquico os quatro bairros se configuram em um grupo de somente dois bairros, além dos outros dois bairros sendo tratados como dois grupos individuais, com o bairro Bangu (número 145) ficando isolado. Este bairro, inclusive, teve o maior valor absoluto de infectados por dengue em uma das epidemias.

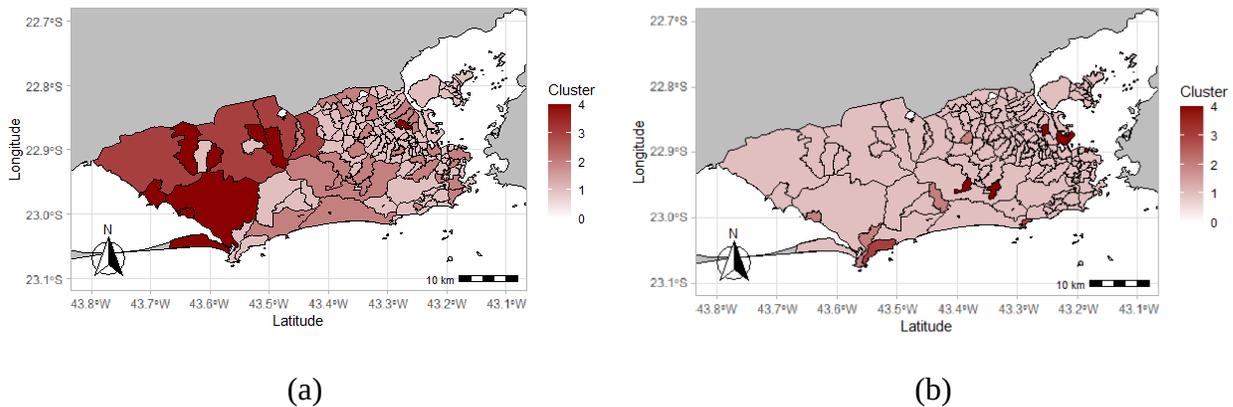
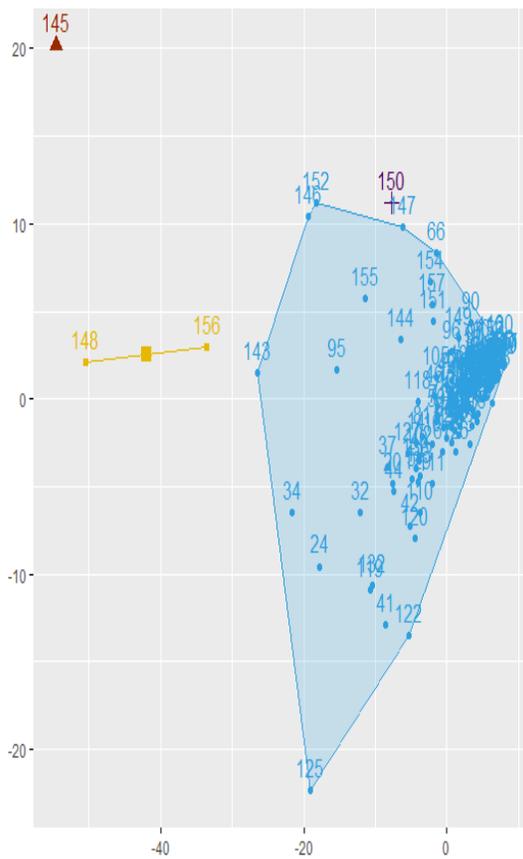
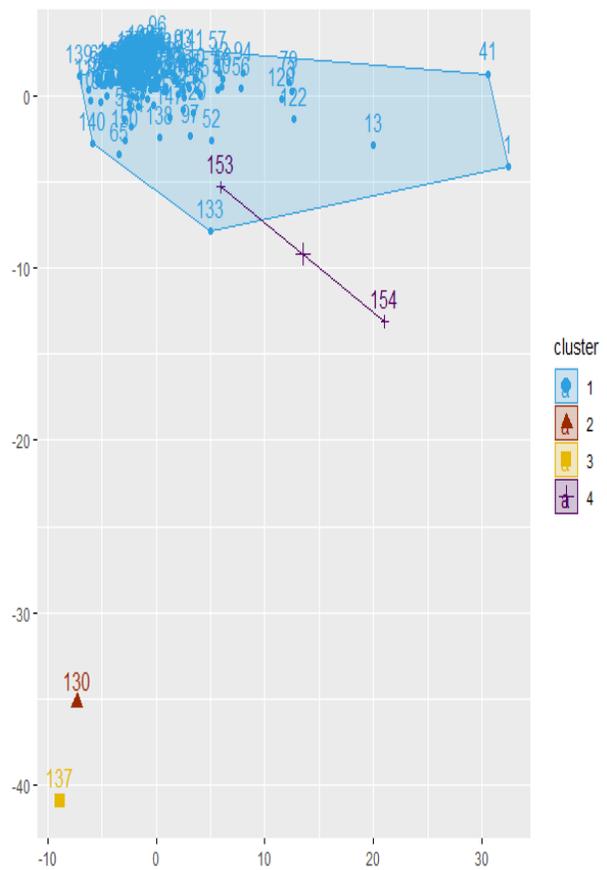


Figura 4.12 - Distribuição espacial dos grupos obtidos na análise de agrupamento via método *k-means* para a) os totais absolutos de dengue por bairro e para b) os totais proporcionalizados de dengue por bairro no município do Rio de Janeiro.

Até no agrupamento hierárquico para a dengue proporcionalizada (Figura 4.13 b) um resultado relativamente similar é percebido no que se diz respeito à falta de separação efetiva entre os bairros. Também há a configuração de um grupo concentrando quase todos os bairros (grupo 1 – azul) e a presença de outros três grupos, onde dois deles são constituídos somente por um bairro. O grupo 2 contém somente pelo bairro Joá (número 130), enquanto que o grupo 3 é composto somente pelo bairro Grumari (número 137). Isso indica que a separação entre os bairros através do método hierárquico, utilizando o mesmo número de *clusters* do método *k-means*, não trouxe um resultado com grupos mais definidos. Isso pode ser um indicativo da necessidade de se realizar testes com um número maior de grupos para o agrupamento hierárquico e verificar se há a melhora na separação dos bairros.



(a)



(b)

Figura 4.13 - Agrupamento dos bairros do município do Rio de Janeiro via método hierárquico a) em função das notificações de dengue e b) em função das notificações de dengue proporcionais à população de cada bairro

4.2 Análise de agrupamento e de variância

Após a análise de agrupamento, foi gerada a análise de variância em relação aos casos de dengue no município do Rio de Janeiro utilizando primeiramente todos os bairros do município e, em seguida, as saídas das médias normalizadas de cada grupo do método *k-means*. Esse procedimento foi realizado tanto para os totais absolutos de dengue por bairro quanto para os

totais de dengue proporcionalizados pela população total de cada bairro. Com isso, foram geradas quatro equações de regressão linear múltipla no total. Duas utilizaram todos os bairros como as variáveis explicativas da variável dependente “dengue”, e as outras duas utilizaram os grupos *k-means* como as variáveis explicativas da variável dependente “dengue”.

Os resultados comentados a seguir se referem às saídas de RLM e análise ANOVA com todos os bairros, tanto na versão de totais absolutos quanto na versão de totais proporcionalizados. Ao se gerar a RLM e a análise de variância associada aos totais absolutos de todos os bairros em relação ao total de dengue no município, todos os bairros preditores são considerados estatisticamente relevantes, apresentando significância estatística dentro da RLM acima de 99%, sendo enfatizado pelos altos valores F de cada bairro (Figura 4.14 a). Esses valores F indicam que estatisticamente os preditores inclusos na regressão se mostram relevantes.

Os valores dos resíduos para o modelo de RLM absoluto possuem valores pequenos e menores que zero, abrangendo um intervalo de $-5,438 \cdot 10^{-16}$ a $3,424 \cdot 10^{-16}$, dando um outro indício inicial de que o modelo pode ter um ajuste satisfatório devido à dimensão dos resíduos, pois isso mostra que as diferenças são mínimas entre o previsto e observado.

O erro padrão residual da RLM completa é igual a $1,56 \cdot 10^{-16}$, sendo esse valor usado para calcular os valores-t na próxima etapa de análise da regressão. Em seguida, tem-se o coeficiente de determinação da RLM, o R^2 , que apresenta o valor exatamente igual a 1. Em complemento ao R^2 , há também o teste F descrito no Capítulo 3.

Para a RLM completa dos totais absolutos de dengue por bairro, o teste F de resultado $6,256 \cdot 10^{31}$ se aplica aos GL 157 e 82 do modelo. A sua relevância é testada pelo valor-p associado ($< 2,2 \cdot 10^{-16}$). Por conta do valor-p estar bem abaixo do limiar de 0,05 usado no seu critério de avaliação, pode-se dizer que de acordo com os valores obtidos na análise ANOVA há significância estatística no teste e, conseqüentemente, a hipótese nula H_0 é rejeitada e considera-se que a adição de variáveis ao modelo RLM gera melhorias na sua resposta final. Todavia, deve ser observado que em nenhum momento da análise aplicada sobre os totais absolutos de dengue por bairro houve a possível rejeição de alguma variável ou um menor valor de significância estatística associado a alguma variável.

Em suma, na análise ANOVA aplicada para os totais absolutos de dengue por bairro, tem-se que, a partir dos critérios utilizados, todos os bairros apresentam relevância estatística com um modelo bem ajustado.

Entretanto, ao se aplicar a RLM sobre os totais proporcionalizados pelo número de habitantes de cada bairro, o resultado se mostra diferente. Os resíduos são ligeiramente maiores em comparação com os resíduos obtidos para a RLM completa dos totais absolutos, apresentando um intervalo de $-5.471 \cdot 10^{-7}$ a $7.263 \cdot 10^{-7}$. O coeficiente de determinação da RLM, o R^2 , também apresenta o valor igual a 1, indicando uma situação similar a que ocorreu com o caso dos totais absolutos de dengue.

No que se diz respeito ao teste F, o valor obtido para a variável *F-statistics* é igual a $1,841 \cdot 10^7$, e que se aplica aos GL 157 e 82 do modelo. O valor-p associado também apresenta valores menores que $2,2 \cdot 10^{-16}$, estando bem abaixo do limiar de 0,05 usado no critério de avaliação. Isso indica que há significância estatística no teste e, conseqüentemente, a hipótese nula H_0 é rejeitada e considera-se que a adição de variáveis ao modelo RLM gera melhorias na sua resposta final. Os valores F para cada bairro apresentam valores elevados na RLM proporcionalizada (Figura 4.14 b), mesmo que menores em comparação aos valores obtidos na RLM absoluta.

Entretanto, em relação aos preditores da RLM proporcionalizada, alguns bairros não apresentam mais uma relevância estatística elevada. Os bairros Cidade Nova, Cosme Velho, Ribeira, Praia da Bandeira, Cocotá, Del Castilho, Riachuelo, Água Santa, Cavalcanti, Grumari e Vila Militar passam a ter significância de 90% ou menos. Esses são alguns dos bairros que estão no grupo de menor variância dentro das análises de agrupamento apresentadas na seção 4.1, tanto para a análise com os totais absolutos quanto para a análise com totais proporcionalizados – salvo a exceção para Grumari, que teve destaque na análise proporcionalizada. Adicionalmente, os bairros Leme, Gávea, Zumbi, São Francisco Xavier, Rocha, Pilares, Parque Anchieta, Campo dos Afonsos e Magalhães Bastos tem a significância entre 95 e 99%. A distribuição espacial dessas significâncias estatísticas podem ser visualizadas na Figura 4.15a.

Na análise ANOVA da RLM proporcionalizada, os únicos bairros que apresentam menor significância estatística são Itanhangá (95 – 99%) e São Francisco Xavier (menor que 90%). Os

valores F da análise ANOVA associados a esses bairros são os menores em comparação aos demais bairros, indicando uma menor relação entre os totais de dengue proporcionalizados e o total de dengue no município do Rio de Janeiro. É possível observar na Figura 4.15b a coloração mais clara em ambos os bairros em comparação com o restante da cidade, indicando a menor significância estatística. O bairro São Francisco Xavier foi o único que teve uma significância moderada na RLM ao mesmo tempo que também apresentou menor significância na análise ANOVA, sendo então numa visão geral o bairro com menor influência nos totais de dengue proporcionalizados na cidade do Rio de Janeiro.

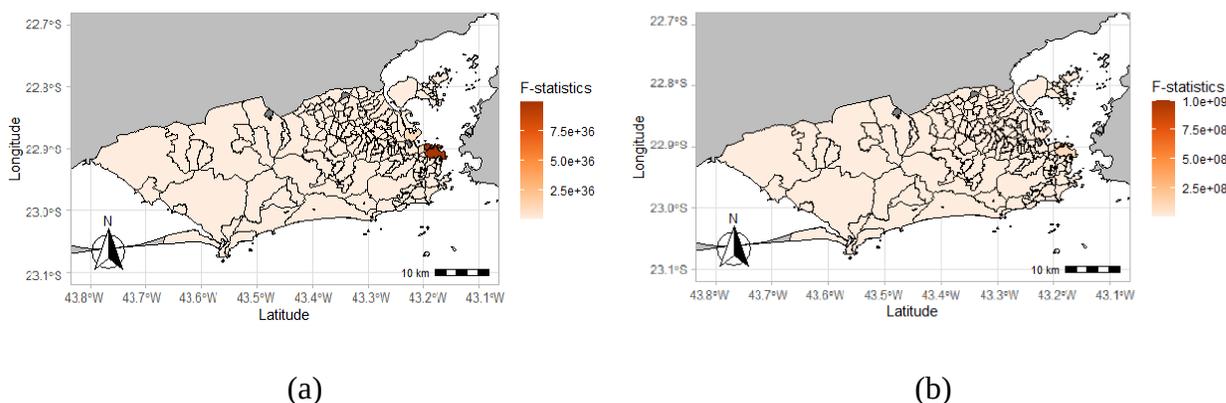


Figura 4.14 - Distribuição espacial do teste-F obtido para as RLM completas a) dos totais absolutos de dengue por bairro e b) dos totais proporcionalizados de dengue por bairro no município do Rio de Janeiro.

É importante frisar que essa análise exibida na Figura 4.15, referente às significâncias dos preditores da RLM e análise ANOVA dos totais proporcionalizados de dengue por bairro, não foi realizada para os totais absolutos de dengue por bairro pois em ambos os casos todos os bairros houve significância estatística acima de 99,9%.

Os resultados a seguir se referem às respostas da RLM e da análise ANOVA aplicadas às saídas das médias geradas dentro dos *clusters* do agrupamento *k-means*, tanto para os totais absolutos por bairro quanto para os totais proporcionalizados por bairro. O método *k-means* foi selecionado para a aplicação da análise ANOVA por ter tido uma melhor resposta, em

comparação com o método hierárquico, na separação por grupos dentro do número de *clusters* escolhidos a partir do método da variância intragrupos.

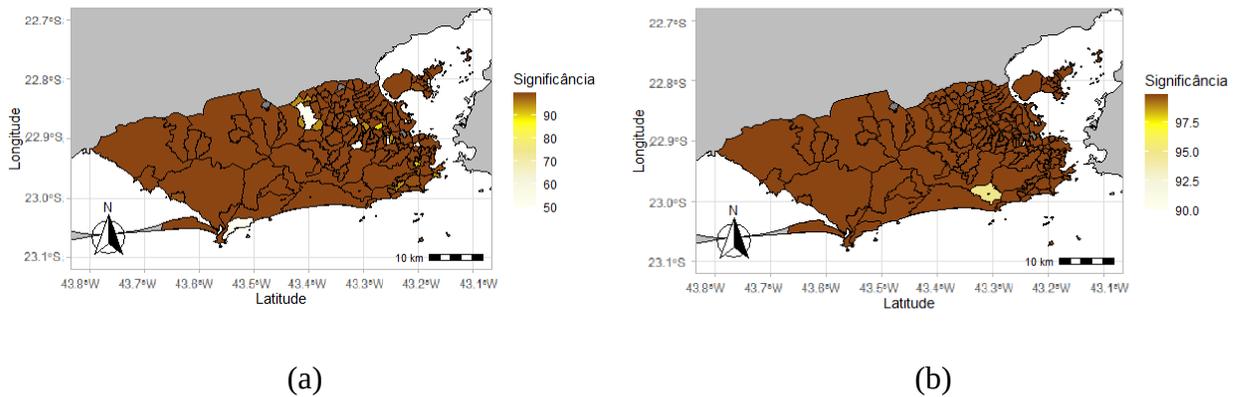


Figura 4.15 - Distribuição espacial da significância estatística a) dos preditores dentro da RLM e b) dos preditores dentro da análise ANOVA dos totais proporcionalizados de dengue por bairro no município do Rio de Janeiro.

Os resíduos da RLM dos grupos *k-means* dos totais absolutos de dengue por bairros apresentados na Tabela 4.2 apresentam um intervalo entre -0,7631 e 5,5427, com mediana próxima a zero (-0,2378). Resíduos com esse tipo de padrão podem indicar um bom ajuste no modelo em comparação aos dados observados.

Com as medidas auxiliares exibidas na Tabela 4.3, é possível observar que o coeficiente de determinação da RLM apresenta um valor baixo de 0,0956, indicando uma relação entre valores previstos e observados de 9,56%. Complementando a análise do desempenho da RLM com o auxílio do teste F, tem-se que o valor *F-statistics* é de 6,21 nos graus de liberdade 4 e 235 e o valor-p associado ao teste F igual a $9,119 \cdot 10^{-5}$. Esse valor está abaixo do limiar de 0,05 associado ao teste p, indicando que mesmo com um coeficiente de determinação apresentando um valor pequeno, a hipótese H_0 é rejeitada e a adição de variáveis ao modelo RLM pode melhorar a sua resposta final.

Tabela 4.2 - Resultado obtidos para os resíduos associados a RLM com grupos *k-means* dos totais absolutos de dengue por bairro.

Resíduo	Valor obtido
Mínimo	-0.7631
Quartil 1	-0.4581
Mediana	-0.2378
Quartil 3	0.0769
Máximo	5.5427

Tabela 4.3 - Medidas estatísticas auxiliares obtidas para a RLM com grupos *k-means* dos totais absolutos de dengue por bairro.

Medida estatística	Valor obtido
Erro padrão residual	0,9591
R ²	0,0956
Teste F	6,21 nos GL 4 e 235
Valor p	9,119*10 ⁻⁵

Dentre os coeficientes lineares associados a cada grupo *k-means* (Tabela 4.4) dentro da regressão, os que apresentam maiores valores são os associados aos grupos 1 e 2. Entretanto, todos os grupos têm significância estatística entre 0 e 90%, não tendo nenhum grupo específico em destaque na RLM. Os valores-t estão próximos de zero em módulo, além dos valores-p associados estarem acima do limiar de 0,05, indicando baixa significância estatística dos grupos dentro da RLM.

Por outro lado, a análise ANOVA para o modelo de grupos *k-means* dos totais absolutos de dengue por bairro (Tabela 4.5) indica que os grupos com maior significância estatística são o 1 e o 2, tendo o grupo 1 uma significância acima de 99%.

Tabela 4.4 - Coeficientes lineares e significância estatística na RLM dos grupos *k-means* dos totais absolutos de dengue por bairro.

Grupos <i>k-means</i>	Coeficiente linear	Erro padrão	Valor-t	Pr (> t)	Intervalo de significância (%)
1	$6,443 \cdot 10^5$	$2,368 \cdot 10^6$	0,272	0,786	0 - 90
2	$1,542 \cdot 10^5$	$5,668 \cdot 10^5$	0,272	0,786	0 - 90
3	$2,203 \cdot 10^4$	$8,097 \cdot 10^4$	0,272	0,786	0 - 90
4	$4,406 \cdot 10^4$	$1,619 \cdot 10^5$	0,272	0,786	0 - 90

O grupo 1 é aquele que engloba a maior parcela de bairros da cidade do Rio de Janeiro dentro da análise de agrupamento via *k-means*, sendo esses bairros os que possuem os menores totais absolutos de casos de dengue. Esses dois grupos também apresentam os maiores valores de SQE, sendo que maiores valores de SQE indicam melhores performances das variáveis dentro de um modelo. Os grupos destacados também tem valores F maiores que os valores F dos grupos 3 e 4. Em tese, quanto maior for um valor F, maior a chance de a variação de médias dos índices ser significativa e, conseqüentemente, a relevância que o próprio índice tem para explicar a variável dependente do modelo. A Figura 4.16a mostra a espacialização dos testes F de cada cluster dos totais absolutos de dengue.

No caso das saídas dos grupos *k-means* para os totais proporcionalizados por bairro, os resultados se mostram diferentes. Por exemplo, os resíduos obtidos nesse caso (Tabela 4.6) se apresentam bem menores ao serem comparados com os resíduos da RLM dos grupos *k-means* dos totais absolutos de dengue por bairro, tendo um intervalo da ordem de 10^{-4} .

Em relação às medidas complementares para avaliar o desempenho da RLM exibidas na Tabela 4.7, tem-se que o erro padrão residual do modelo igual a $7,766 \cdot 10^{-5}$. Quanto menor o valor do erro padrão residual, menor o desvio dos resíduos e, portanto, melhor o desempenho da RLM.

Tabela 4.5 - Análise de variância para RLM dos grupos *k-means* dos totais absolutos de dengue por bairro.

Grupos <i>k-means</i>	SQE	Valor F	Pr(>F)	Intervalo de significância (%)
1	19,192	20,8654	$7,956 \cdot 10^{-6}$	99,9 - 100
2	3,580	3,892	0,04969	95 - 99
3	0,008	0,0091	0,92394	0 - 90
4	0,068	0,0740	0,78582	0 - 90

Tabela 4.6 - Resultado obtidos para os resíduos associados a RLM com grupos *k-means* dos totais proporcionalizados de dengue por bairro.

Resíduo	Valor obtido
Mínimo	$-4,482 \cdot 10^{-4}$
Quartil 1	$-2,210 \cdot 10^{-6}$
Mediana	$8 \cdot 10^{-6}$
Quartil 3	$1,275 \cdot 10^{-5}$
Máximo	$3,541 \cdot 10^{-4}$

Adicionalmente, o coeficiente de determinação R^2 apresenta o valor de 0,9957, representando uma relação estatística de 99,57% entre valores previstos e observados. Esses já são os primeiros indícios de que a RLM dos grupos *k-means* proporcionalizados entrega melhores resultados em comparação a RLM dos grupos *k-means* referentes aos totais absolutos. Complementando a análise do modelo, o teste F apresenta o valor de $1,346 \cdot 10^4$ nos graus de liberdade 4 e 235, com o valor-p associado menor que $2,2 \cdot 10^{-16}$. Esse valor-p é bem menor que o limiar de 0,05, indicando que os grupos *k-means* se mostram relevantes dentro da construção da RLM, contribuindo para melhores resultados da regressão.

Tabela 4.7 - Medidas estatísticas auxiliares obtidas para a RLM com grupos *k-means* dos totais proporcionalizados de dengue por bairro.

Medida estatística	Valor obtido
Erro padrão residual	$7,766 \cdot 10^{-5}$
R^2	0,9957
Teste F	$1,346 \cdot 10^4$ nos GL 4 e 235
Valor p	$< 2,2 \cdot 10^{-16}$

Na análise da RLM dos *k-means* proporcionalizados, onde os grupos são os preditores da regressão, os grupos apresentam relevância estatística, diferente do que ocorreu com a RLM dos grupos *k-means* dos totais absolutos. Através da Tabela 4.8, é possível ver que todos os grupos apresentam significância de pelo menos 99%. O grupo 2, que possui a menor significância em comparação com os restantes, é o que contém os bairros Campo dos Afonsos, Zumbi, Barra de Guaratiba, Camorim e Pedra de Guaratiba na análise de agrupamento (Figura 4.11 b). O erro padrão dos grupos é pequeno, sendo o maior deles na escala de 10^{-2} , indicando que os coeficientes foram bem calculados. Além disso, os valores-t se distanciam de zero em módulo, principalmente para o grupo 1. Em conjunto, os valores Pr ($>|t|$) associados estão abaixo do limiar de 0,05, o que reforça a validade dos valores dos coeficientes obtidos na regressão.

Tabela 4.8 - Coeficientes lineares e significância estatística na RLM dos grupos *k-means* dos totais proporcionalizados de dengue por bairro.

Grupos <i>k-means</i>	Coeficiente linear	Erro padrão	Valor-t	Pr ($> t $)	Intervalo de significância (%)
1	$9,278 \cdot 10^{-1}$	$1,248 \cdot 10^{-2}$	74,325	$< 2 \cdot 10^{-16}$	99,9 - 100
2	$1,324 \cdot 10^{-2}$	$3,991 \cdot 10^{-3}$	3,317	0,00105	99 – 99,9
3	$-2,211 \cdot 10^{-3}$	$4,584 \cdot 10^{-4}$	-4,822	$2,55 \cdot 10^{-6}$	99,9 - 100
4	$9,376 \cdot 10^{-3}$	$1.816 \cdot 10^{-3}$	5,164	$5,16 \cdot 10^{-7}$	99,9 - 100

A análise ANOVA da RLM dos grupos *k-means* proporcionalizados (Tabela 4.9) também indica que os grupos apresentam significância estatística, todos com valores acima de 99%. Além da alta significância das variáveis, os valores F associados apresentam valores distantes de zero. Quanto maior for um valor F dentro da análise ANOVA, maior a chance de a variação de médias dos grupos ser significativa e, conseqüentemente, a relevância que o próprio grupo tem para explicar a variável dependente. Os $Pr(>F)$ associados aos valores F estão bem abaixo do limiar de 0,05, reforçando a ideia de que os grupos apresentam uma relação relevante com os totais de dengue proporcionalizados no município do Rio de Janeiro. A Figura 4.16b mostra a espacialização dos testes F de cada *cluster* dos totais proporcionalizados de dengue.

Ao se realizar uma comparação final, é possível notar que no caso das RLMs completas (com a dengue de todos os bairros como preditores da dengue no município do Rio de Janeiro), há uma resposta mais criteriosa para a seleção dos bairros mais relevantes ao se aplicar a proporção de totais de dengue por habitantes de cada bairro. Algo similar acontece com os grupos da análise de agrupamento via método *k-means*, tendo uma melhor resposta estatística no resultado da RLM e da análise ANOVA ao se utilizar os totais de dengue proporcionais à população de cada bairro.

Tabela 4.9 - Análise de variância para RLM dos grupos *k-means* dos totais proporcionalizados de dengue por bairro.

Grupos k-means	SQE	Valor F	Pr(>F)	Intervalo de significância (%)
1	0,00032393	69,113	$< 2.2*10^{-16}$	99,9 - 100
2	0,00000035	58,602	$4,984*10^{-13}$	99,9 - 100
3	0,00000016	27,312	$3,819*10^{-7}$	99,9 - 100
4	0,00000016	26,662	$5,163*10^{-7}$	99,9 - 100

Na Figura 4.16, é possível observar que em ambos os casos de distribuição espacial dos valores de teste F são compatíveis com os resultados obtidos para as análises de agrupamento abordadas na seção anterior (Figura 4.12).

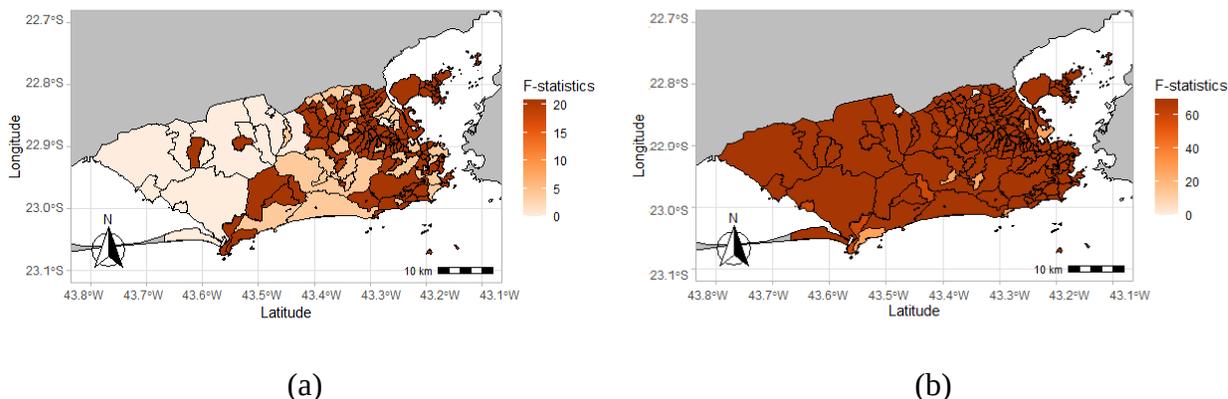


Figura 4.16 - Distribuição espacial do teste F obtido para os *clusters* na análise de agrupamento via método *k-means* para a) os totais absolutos de dengue por bairro e para b) os totais proporcionalizados de dengue por bairro no município do Rio de Janeiro.

4.3 Incidência de dengue e ICU

Nesta seção é feita a abordagem da análise do comportamento da dengue em Irajá, escolhido como bairro representativo do fenômeno ICU, conforme detalhado na seção 3.1, para tentar observar as possíveis relações entre a resposta térmica do bairro e os casos de dengue registrados nele.

Para auxiliar na análise da ICU, foram extraídas as temperaturas média, máxima e mínima semanais para o bairro de Irajá. A Figura 4.17 exibe as séries temporais dessas variáveis em conjunto com os totais de dengue durante o período de 2015 a 2019, com todas as variáveis normalizadas para fins de comparação, deixando todas dentro da mesma escala. Durante esse intervalo, o bairro de Irajá começou a ganhar destaque nos recordes de calor registrados, como

mencionado na seção 3.1. É possível notar duas epidemias ocorridas em 2015-2016 e 2019, sendo a de 2019 um pouco mais sutil na série normalizada.

No início do ano de 2015, tem-se uma ligeira elevação das temperaturas média e mínima logo antes do aumento súbito das notificações de dengue para Irajá, que atingem seu pico enquanto as temperaturas diminuem em direção ao meio do ano. Logo após a diminuição de casos em 2015, ocorre um aumento da temperatura máxima próximo do final do ano, seguido da diminuição das temperaturas e um novo aumento da temperatura média e da temperatura mínima logo no início de 2016. Na sequência, observa-se um novo máximo de dengue. Com o decorrer do ano de 2016, todas as temperaturas diminuem, bem como os totais de dengue.

Ao longo de 2017 e 2018, ocorrem alguns casos de dengue mas não o suficiente para ocorrer um surto epidêmico, além de acontecer durante o ano de 2018 uma diminuição da amplitude dos máximos das temperaturas em comparação aos outros anos, bem como uma diminuição ligeiramente mais pronunciada da temperatura máxima em meados de 2018 em comparação com o mesmo período dos outros anos.

Por fim, na virada de 2018 para 2019 ocorre uma nova elevação considerável das temperaturas médias e mínima. Algumas semanas depois, a temperatura máxima se eleva e começa a diminuir mais uma vez junto com as temperaturas média e mínima. Algumas semanas depois desses picos, ocorre um máximo de dengue de menores proporções em comparação com os de 2015 e 2016, acompanhado de uma ligeira elevação das temperaturas.

Nos anos de 2015, 2016 e 2019 ocorreram epidemias de dengue no município do Rio de Janeiro, que tiveram os totais de 17451 casos, 23946 casos e 14343 casos, respectivamente. Essas epidemias possuem menor magnitude em comparação a outras epidemias que ocorreram na localidade. Por exemplo, em 2002 foram registrados oficialmente 139400 casos; já em 2008 teve-se o total de 120440 registros; no ano de 2012 houveram 50184 casos de dengue no município todo.

Após a ocorrência de duas epidemias de grandes dimensões, foram publicadas diversas ações para o combate à dengue no Rio de Janeiro, que podem ter influenciado na redução gradual dos totais de dengue durante as epidemias que ocorreram a partir de 2012 no município. Por exemplo, o documento “Diretrizes Nacionais para a Prevenção e Controle de Epidemias de

Dengue” foi divulgado pelo Ministério da Saúde (BRASIL, 2009), com ações de vigilância e combate à dengue no Brasil a serem aplicadas nos estados e municípios.

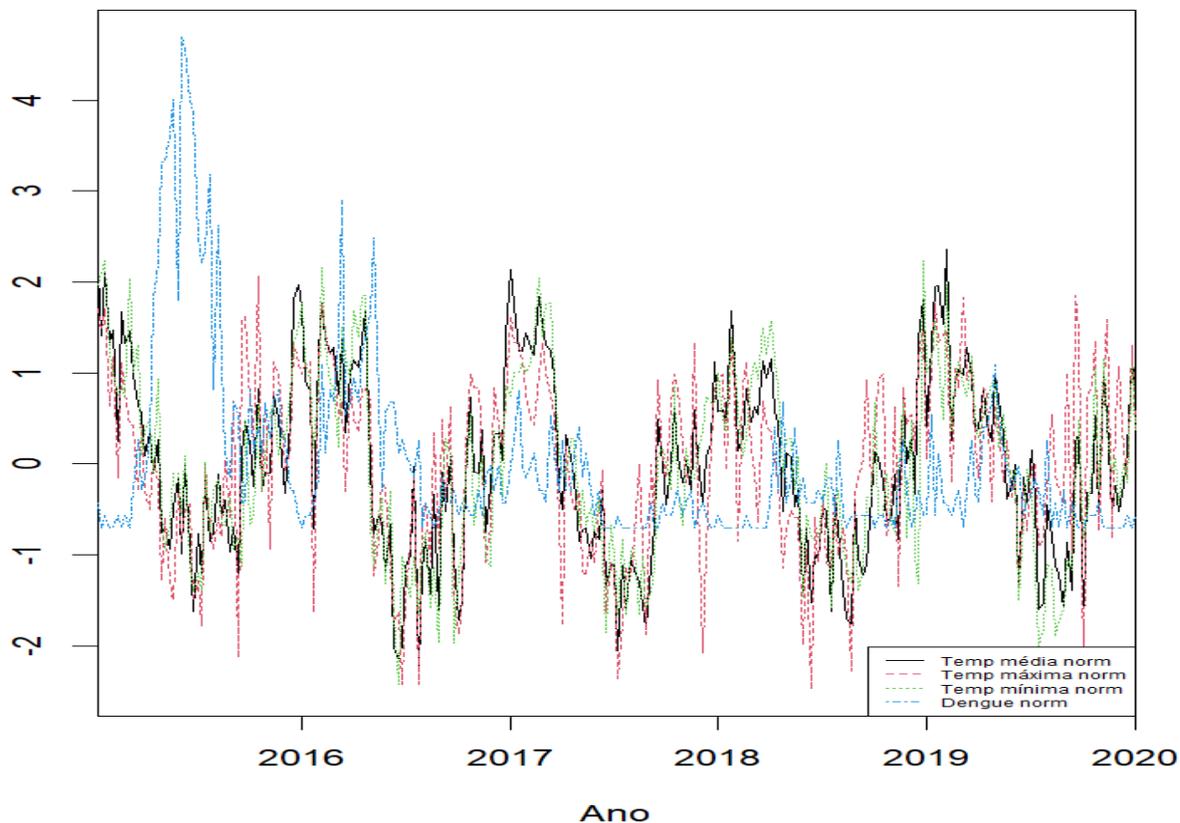


Figura 4.17 - Séries temporais de dados semanais normalizados para o bairro de Irajá – temperatura média (linha preta cheia), temperatura máxima (linha vermelha pontilhada), temperatura mínima (linha verde pontilhada) e casos de dengue (linha azul pontilhada) entre os anos de 2015 e 2019.

Ações locais também foram tomadas e implementadas no Rio de Janeiro, como por exemplo a publicação do Decreto Municipal nº 34377/2011, de 31 de agosto de 2011, que instituiu estado de alerta contra a dengue e dispôs sobre a prevenção e controle da doença na cidade do Rio de Janeiro (RIO DE JANEIRO (RJ), 2011). Além disso, também foi realizado o lançamento de um sistema de monitoramento de ocorrências de dengue no município numa ação

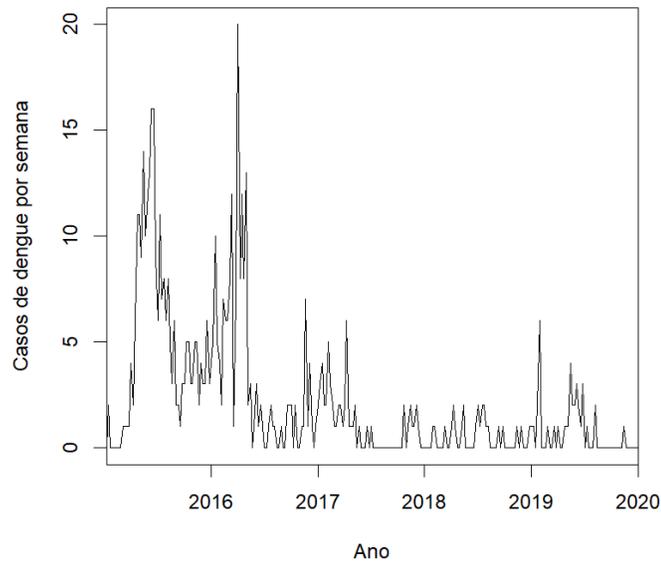
conjunta da Fundação Oswaldo Cruz – FIOCRUZ com a Fundação Getúlio Vargas – FGV (FIOCRUZ, 2015).

Trazendo à luz brevemente a questão da variabilidade climática, as epidemias de 2015 e 2016 ocorreram durante episódios de El Niño (SANTOS et al., 2022), bem como a epidemia de 2019, de acordo com os índices publicados pela NOAA, disponível em <<https://www.esrl.noaa.gov/psd/enso/dashboard.html>>. Durante esses anos foram registradas ocorrências de altas temperaturas, contando inclusive com quebra de recordes para o ano de 2015 durante a primavera para o município como um todo (PBMC, 2015) e altas temperaturas durante o verão, tendo os maiores valores de temperatura máxima durante o final do outono e durante o inverno no bairro de Irajá, de acordo com o boletim de temperaturas máximas e mínimas mensais do Sistema Alerta Rio para o ano de 2019.

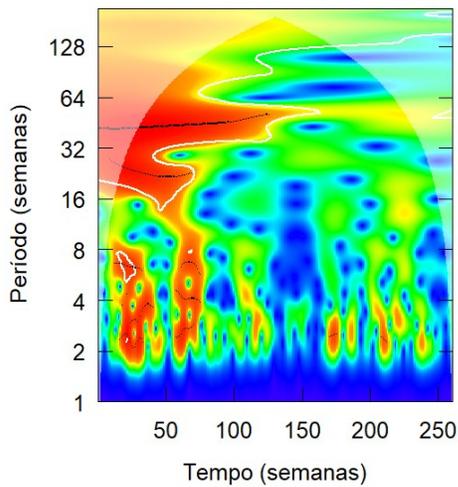
O espectro de potência de ondeleta (Figura 4.18 b) enfatiza a ocorrência das epidemias no bairro de Irajá. É possível observar sinais mais fortes no período anual (aproximadamente 52 semanas), indicando a ocorrência de dengue em todos os anos, sendo mais evidente em 2015 e 2016 devido aos maiores totais registrados nesses anos. Esse sinal anual se propaga para períodos menores. Entre 16 e 32 semanas, há um sinal com uma linha de tendência localizada em aproximadamente 26 semanas (período semestral), indicando o distanciamento entre os picos epidêmicos.

Entre 2 e 8 semanas, são vistos sinais mais fortes que coincidem com os máximos de epidemia durante os anos de 2015 e 2016, além de um sinal um pouco mais sutil durante o ano de 2019. A potência média do sinal da dengue (Figura 4.18 c) reforça o período aproximado de 52 semanas como o principal ao longo de toda a série, tendo o período aproximado de 26 semanas como o destaque secundário. Há também o destaque da potência média para períodos maiores, porém eles estão fora do cone de influência do espectro de potência.

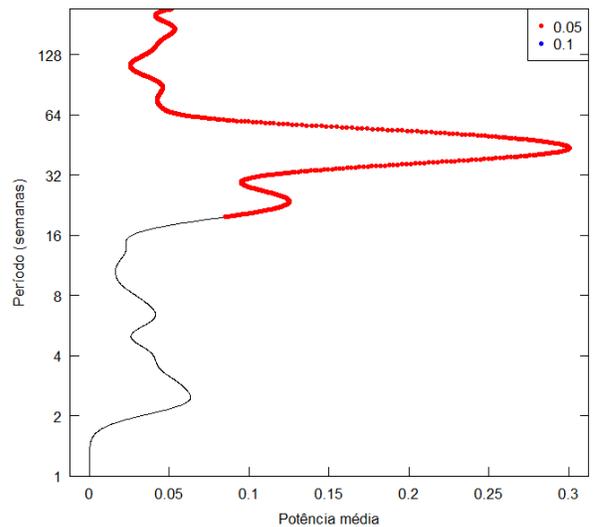
As temperaturas média, máxima e mínima semanais (Figuras 4.19, 4.20 e 4.21) apresentam sinais predominantemente anuais (em torno de 52 semanas) nos seus espectros de potência contínuo e na potência global, demarcando os ciclos anuais característicos de aumento e declínio de temperatura.



(a)

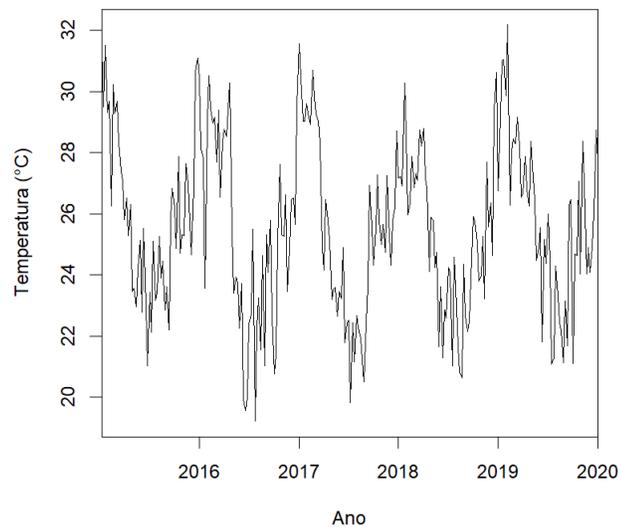


(b)

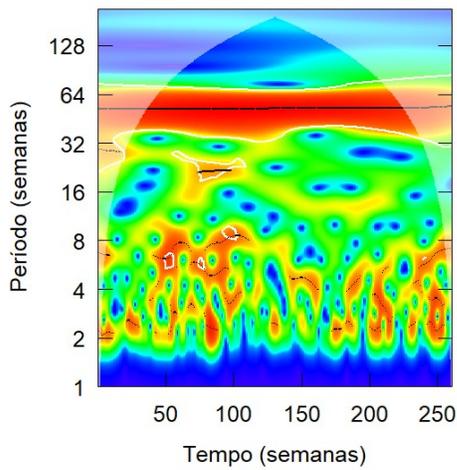


(c)

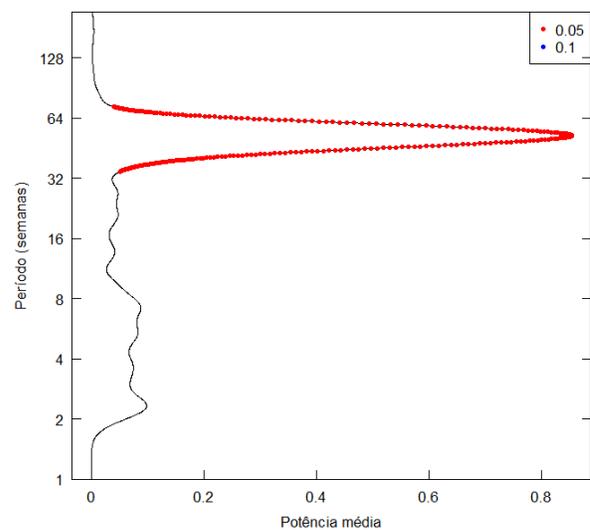
Figura 4.18 - (a) Série temporal, (b) espectro de potência de ondeleta contínua e (c) espectro de ondeleta global dos casos de dengue semanais para o bairro de Irajá. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).



(a)

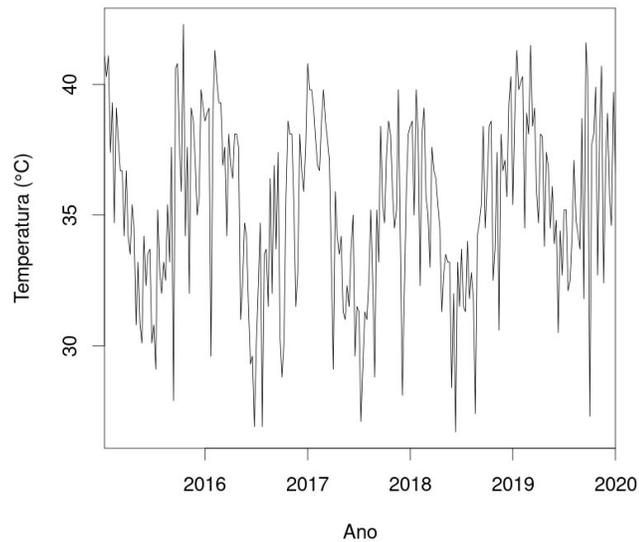


(b)

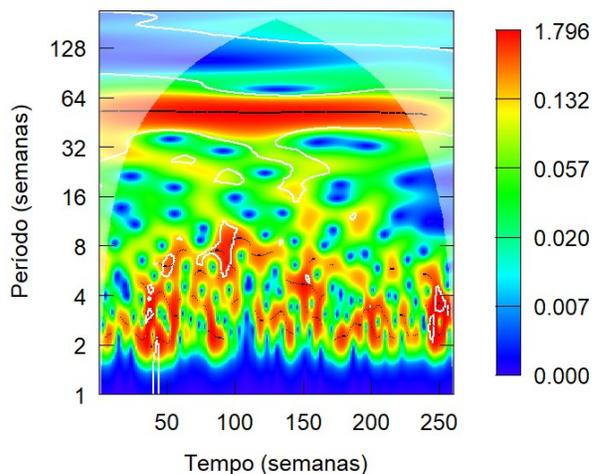


(c)

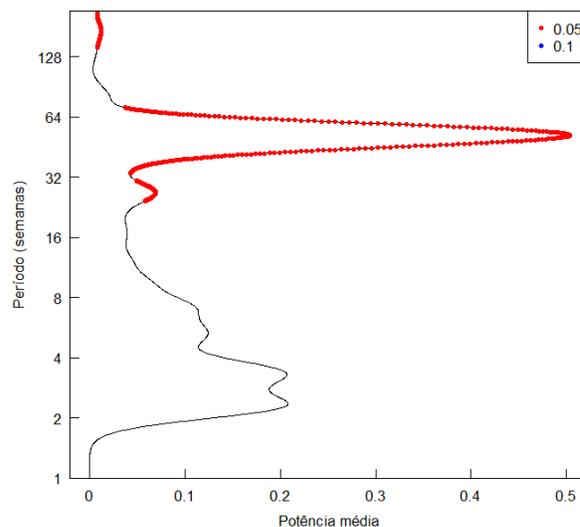
Figura 4.19 - (a) Série temporal, (b) espectro de potência de ondeleta contínua e (c) espectro de ondeleta global da temperatura média semanal para o bairro de Irajá. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).



(a)

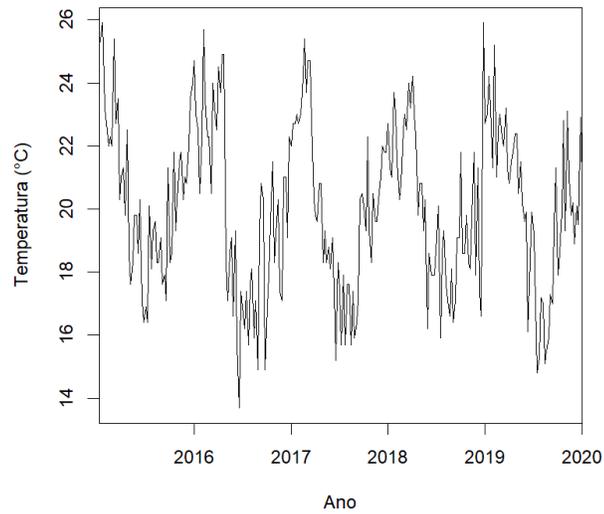


(b)

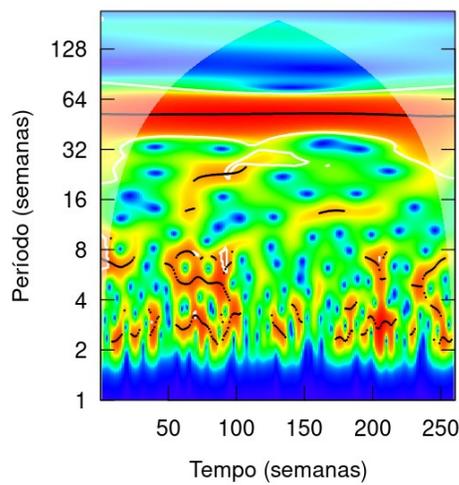


(c)

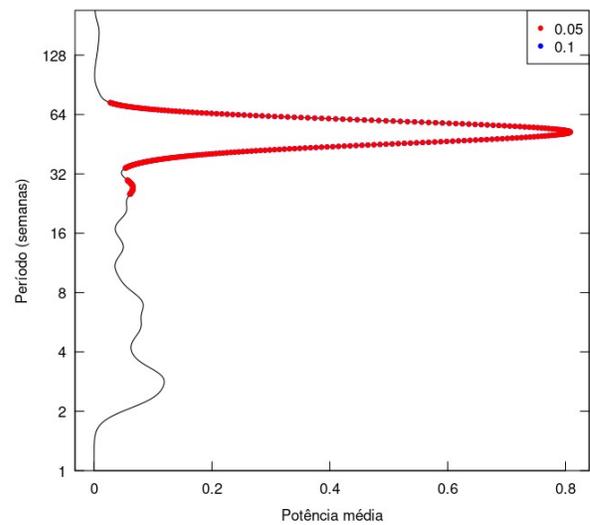
Figura 4.20 - (a) Série temporal, (b) espectro de potência de ondeleta contínua e (c) espectro de ondeleta global da temperatura máxima semanal para o bairro de Irajá. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).



(a)



(b)



(c)

Figura 4.21 - (a) Série temporal, (b) espectro de potência de ondeleta contínua e (c) espectro de ondeleta global da temperatura mínima semanal para o bairro de Irajá. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).

Nota-se também há a presença de sinais de potência mais pronunciados no intervalo entre os períodos de 2 a 8 semanas em todas as temperatura. Entretanto, no espectro de potência contínuo da temperatura máxima semanal (Figura 4.20 b) os sinais de potência são mais evidentes nos períodos menores comparados com os espectros de potência das temperaturas média e temperatura mínima semanais, sendo esses máximos de potência da temperatura máxima demarcados pela linha branca de 90% de significância estatística. Isso é reforçado pelo espectro global (Figura 4.20 c), onde a temperatura máxima apresenta uma potência média maior em comparação às outras temperaturas apresentadas no trabalho.

O espectro de potência bivariada apresenta dois aspectos importantes: a potência do sinal do cruzamento entre as duas variáveis, representada pela barra de cores e a diferença de fase, representada por vetores.

Assim como os espectros univariados das temperaturas, os espectros bivariados de potências entre cada temperatura e as ocorrências de dengue apresentam resultados similares de forma geral (Figuras 4.22 a, 4.23 a e 4.24 a).

Os cruzamentos mais significativos ocorrem durante o período anual (aproximadamente 52 semanas) e semestral (aproximadamente 26 semanas). Analisando a posição dos vetores no período anual, é possível ver que eles estão em sua maioria apontando próximo a $\pi/2$ (para cima), indicando que as temperaturas e a dengue estão em fases próximas entre si, com as temperaturas (variáveis x na aplicação da TOC) estando um pouco mais à frente da série das ocorrências de dengue. O sinal predominante dos espectros de ondeleta globais (Figuras 4.22 b, 4.23 b e 4.24 b) é o do período anual, condizendo com os espectros contínuos.

Ocorrem também diferenças pontuais no espectro de potência contínuo de cada temperatura. Por exemplo, no intervalo entre os períodos de 16 e 32 semanas para a temperatura média semanal (Figura 4.22 a), ao se observar o gráfico de baixo para cima, ocorre um giro dos vetores da posição próxima a $-\pi/2$ passando pelo zero trigonométrico e chegando próximo a $\pi/2$. Isso indica um início com fases próximas entre as variáveis, com a dengue liderando. Em seguida, tem-se o zero trigonométrico, indicando que as variáveis estão em fase, finalizando com um perfil similar ao que ocorre no período anual.

Já para a temperatura máxima (Figura 4.23 a), no mesmo intervalo de período entre 16 e 32 semanas, o padrão dos vetores é predominantemente na direção de $\pi/2$, indicando fases próximas de temperatura e dengue, com a temperatura liderando. Também aparece, nesse mesmo espectro de potência contínuo, um sinal mais moderado por volta da faixa de 30 semanas que também apresenta vetores próximos de $\pi/2$. Adicionalmente, há um pequeno sinal vetorial no início da série temporal (em 2015), dentro dos períodos de 2 e 4 semanas, onde ocorre uma das epidemias de dengue. Em ambos os casos, os vetores estão posicionados entre $\pi/2$ e $-\pi$, indicando uma defasagem entre a temperatura máxima e a dengue nestes períodos, com a dengue liderando.

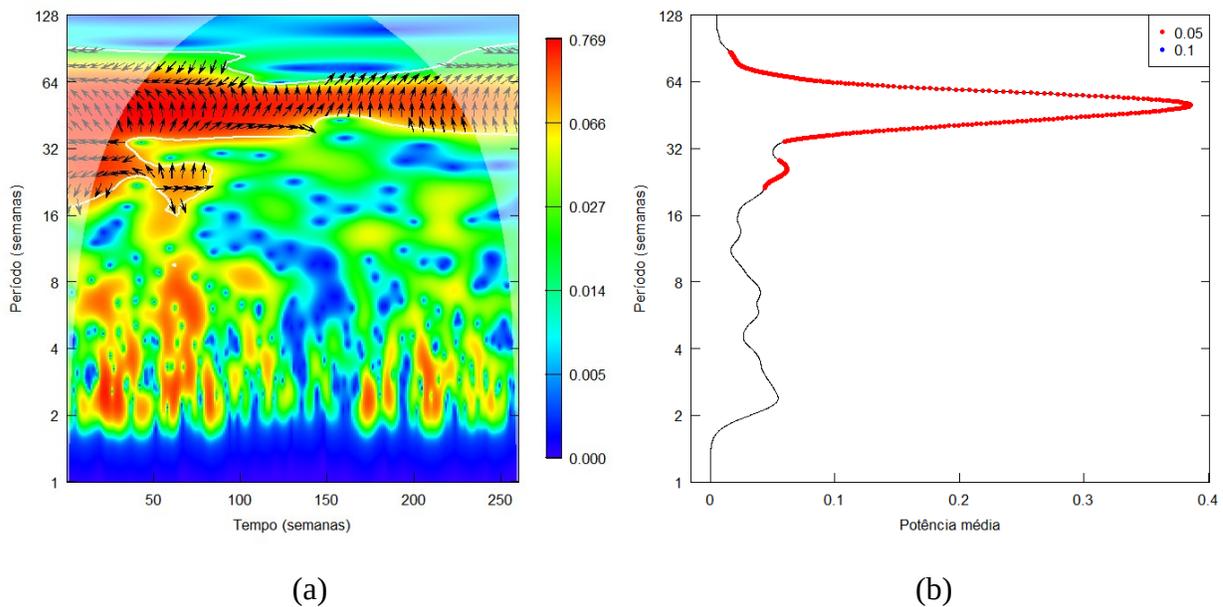


Figura 4.22 - Espectros bivariados de (a) potência de ondeleta contínua e (b) ondeleta global temperatura média semanal com casos de dengue semanais para o bairro de Irajá. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).

Em relação ao espectro da temperatura mínima com a dengue (Figura 4.24 a), o padrão dos vetores é relativamente similar ao que é encontrado no espectro bivariado da temperatura média, exceto pela ausência de vetores posicionados próximos de $\pi/2$. Um sinal moderado aparece no final da série temporal (em 2019), com o vetores posicionados próximos de $-\pi/2$. Isso indica uma defasagem entre a temperatura mínima e a dengue, com a dengue liderando em detrimento da temperatura mínima.

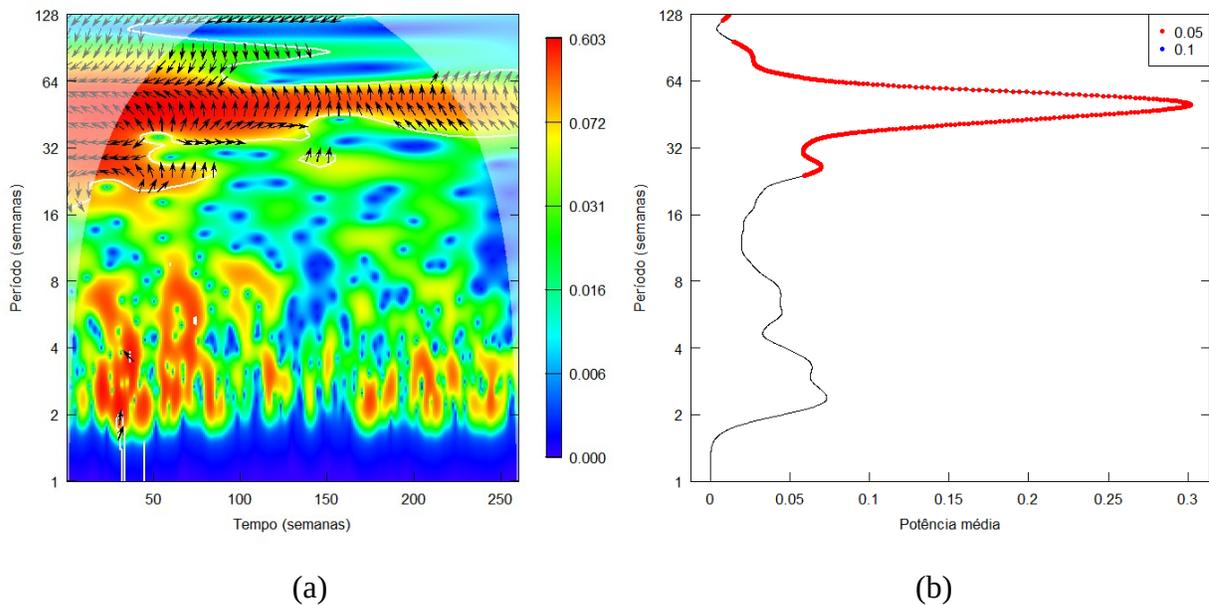


Figura 4.23 - Espectros bivariados de (a) potência de ondeleta contínua e (b) ondeleta global temperatura máxima semanal com casos de dengue semanais para o bairro de Irajá. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).

4.3.1 Análise ano epidêmico vs. ano não epidêmico

Nesta subseção, em vez de utilizar as séries de temperaturas semanais em si, são usadas as diferenças calculadas através da equação 3.1. Tendo em vista que os resultados da subseção anterior trouxeram destaque para os períodos semestral e anual, são então utilizadas nessa subseção as diferenças semanais das variáveis térmicas com ênfase em um ano considerado epidêmico e um ano considerado não epidêmico na cidade do Rio de Janeiro. Isso é feito com o objetivo de observar se as tendências semanais de alguma variável térmica apresentam alguma relação significativa com os diferentes comportamentos da dengue no bairro de Irajá, representativo da ICU neste trabalho, dentro de alguma periodicidade menor.

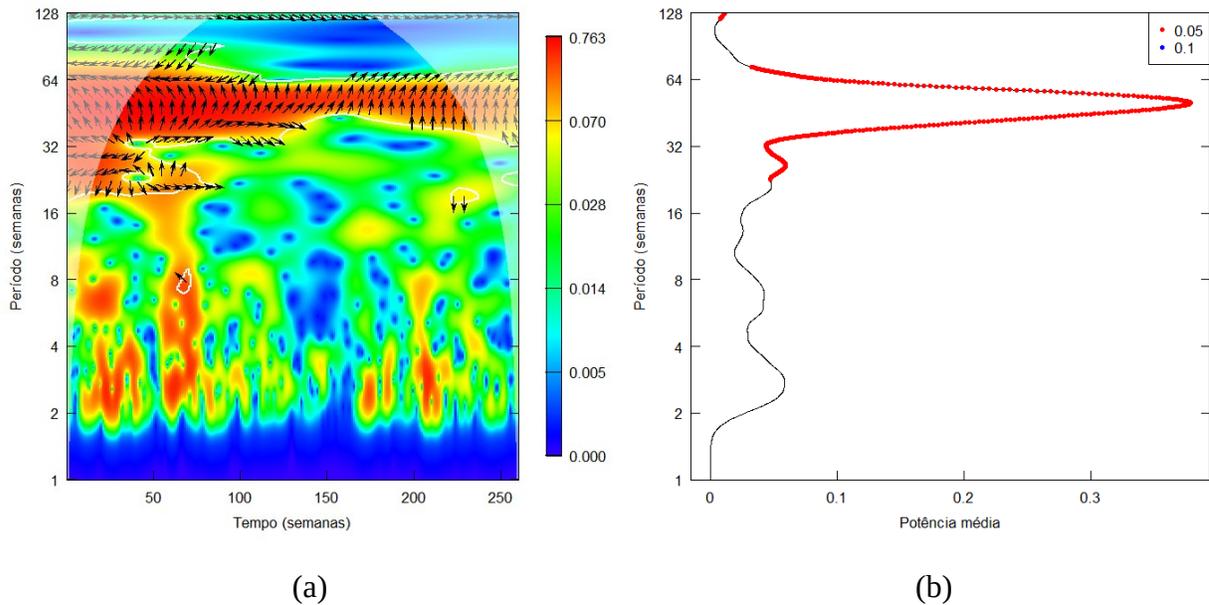


Figura 4.24 - Espectros bivariados de (a) potência de ondeleta contínua e (b) ondeleta global temperatura mínima semanal com casos de dengue semanais para o bairro de Irajá. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).

Para o panorama epidêmico foi considerado o ano de 2016, enquanto que para o panorama não epidêmico foi considerado o ano de 2018. A fim de observar o comportamento das variáveis térmicas antes dos anos selecionados para análise de um ano epidêmico e um ano não epidêmico de dengue, a equação 3.1 foi aplicada sobre os seis meses anteriores em cada ano. Em outras palavras, para o ano epidêmico foram analisadas as diferenças térmicas semanais desde o segundo semestre de 2015 até o final de 2016, enquanto que para o ano não epidêmico foram analisadas as diferenças térmicas semanais desde o segundo semestre de 2017 até o final de 2018.

Os resultados a seguir são referentes ao ano de 2016, representativo de uma das epidemias ocorridas no município do Rio de Janeiro. As diferenças de temperatura média (Figura 4.25) apresentam sinais significativos na escala de duas semanas, tendo potências mais pronunciadas entre o final de 2015 e início de 2016 e em meados de 2016, com um sinal menos intenso entre esses máximos. Esses máximos marcam as maiores diferenças semanais de temperatura média majoritariamente positivas.

Também é notável a conexão desses máximos de duas semanas com alguns picos de potência no período de aproximadamente seis semanas também no final de 2015 e em meados de 2016, indicando um intervalo de aumento de potência em seis semanas com um pico interno de duas semanas.

Para a temperatura máxima, o espectro contínuo (Figura 4.26 a) mostra os sinais de potência mais intensos em direção ao fim de 2015 e no início de 2016, entre os períodos de duas a quatro semanas, onde ocorre seu aumento pré-epidemia. Os outros máximos ocorrem em meados de 2016 no período de aproximadamente duas semanas e no fim de 2016 no período de seis a oito semanas. Durante esses máximos, ocorrem grandes diferenças térmicas entre algumas semanas, com aumentos de mais de 3°C na temperatura máxima em alguns casos. O espectro global (Figura 4.26 b) mostra os períodos de duas, quatro e oito semanas como os de maior potência média no espectro.

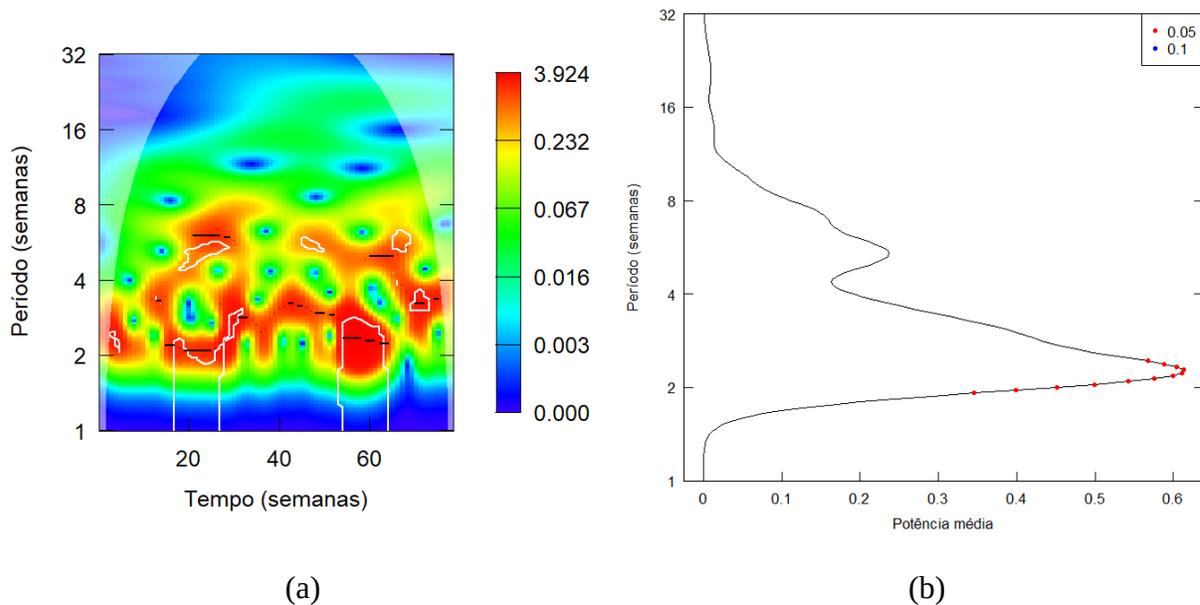


Figura 4.25 -(a) espectro de potência de ondeleta contínua e (b) espectro de ondeleta global das diferenças semanais de temperatura média para o bairro de Irajá durante o ano de 2016. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).

Em relação às diferenças de temperatura mínima (Figura 4.27), há também a predominância do sinal mais intenso na faixa de aproximadamente três semanas e de um máximo secundário em torno de seis semanas, tanto para o espectro de potência contínua quanto para o espectro global. Entretanto, para a temperatura mínima, os sinais mais intensos e significativos estatisticamente no espectro contínuo no período de duas semanas são encontrados logo no início e um pouco antes do final de 2015, durante o primeiro semestre de 2016 e no final de 2016. Os sinais presentes em 2015 e no final de 2016 caracterizam os episódios de aumento de temperatura mínima anteriormente citados na seção 4.3, enquanto que o sinal que ocorre no primeiro semestre de 2016 marca uma diminuição considerável da temperatura mínima em direção ao meio do ano.

O espectro bivariado das diferenças semanais de temperatura média e ocorrências de dengue (Figura 4.28 a) mostra uma covariância com significância estatística logo no início da escala temporal, ou seja, em meados de 2015. Os vetores se mostram presentes entre os períodos de uma a três semanas, com um posicionamento entre $\pi/2$ e π , indicando que as diferenças semanais de temperatura média e a dengue se encontram defasados entre si com a dengue liderando. Há também um pequeno núcleo de vetores próximo ao final de 2015 que apontam na direção entre o zero trigonométrico e $\pi/2$, mais próximo de zero, indicando que as variáveis estão em fases próximas, com as diferenças semanais de temperatura média liderando. Isto indica que antes da epidemia de dengue ocorrida em 2016, a temperatura média que vinha em elevação no final de 2015 esteve à frente do aumento das ocorrências de dengue.

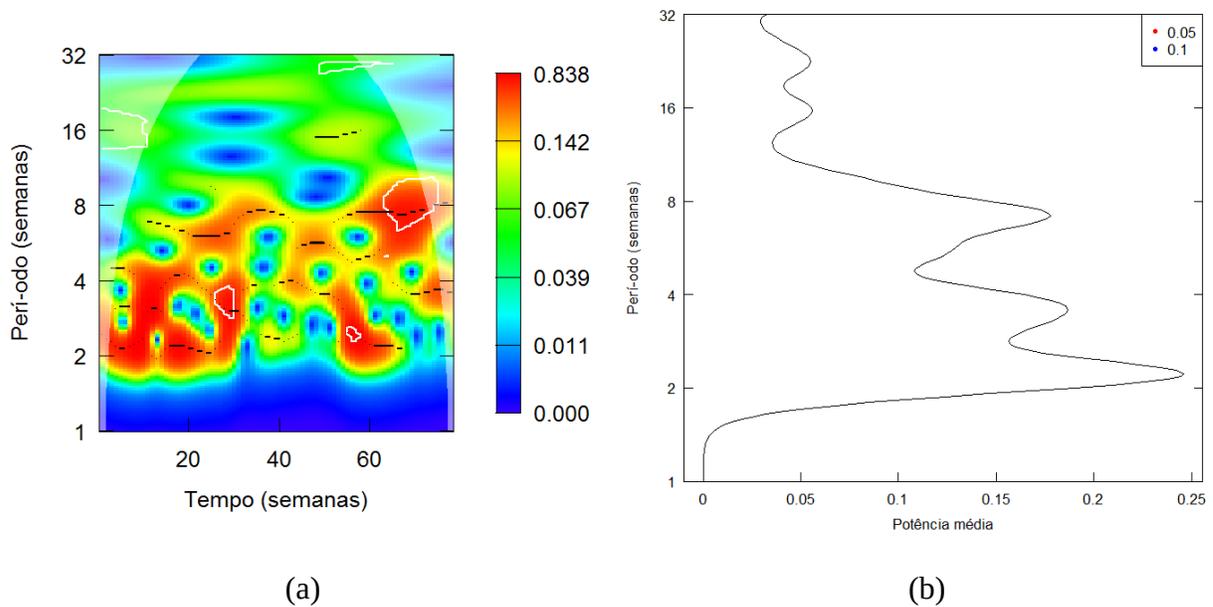


Figura 4.26 - (a) espectro de potência de ondeleta contínua e (b) espectro de ondeleta global das diferenças semanais de temperatura máxima para o bairro de Irajá durante o ano de 2016. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).

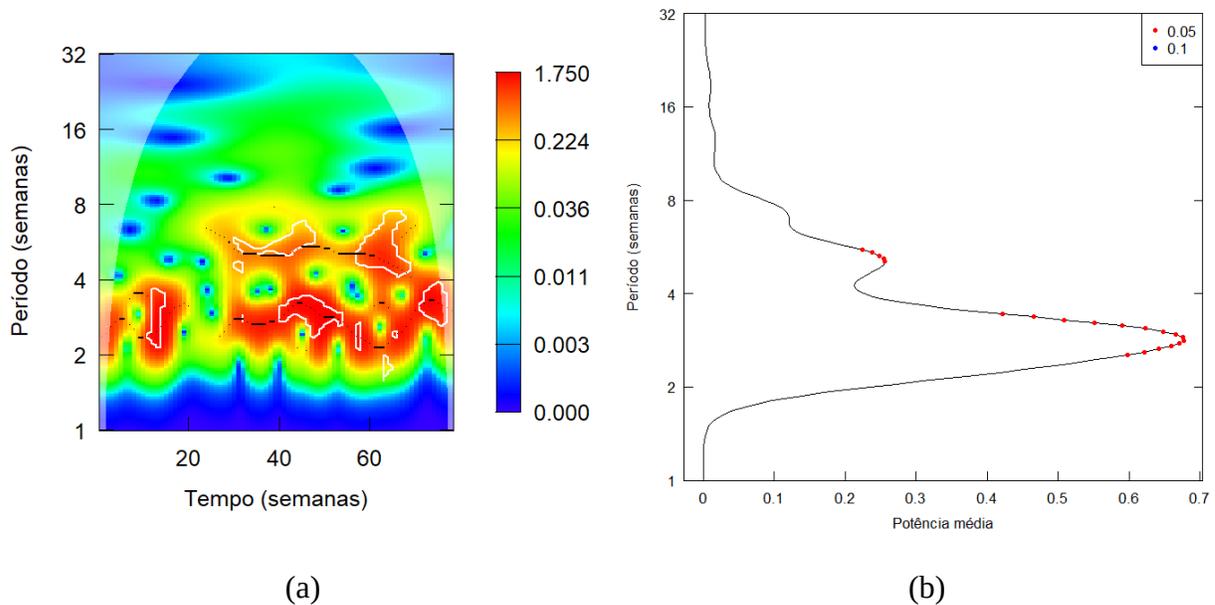


Figura 4.27 -(a) espectro de potência de ondeleta contínua e (b) espectro de ondeleta global das diferenças semanais de temperatura mínima para o bairro de Irajá durante o ano de 2016. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).

No espectro global (Figura 4.28 b), apesar de não haver a presença dos marcadores de significância de 95%, os períodos predominantes são os de três semanas e de seis semanas.

No espectro bivariado da temperatura máxima com a dengue (Figura 4.29 a), os vetores marcadores de fase apresentam significância estatística durante o segundo semestre de 2015 em três regiões diferentes, com dois núcleos no período de duas semanas e um terceiro núcleo em torno de quatro semanas.

No início do segundo semestre de 2015 há um grupo de vetores posicionados entre $\pi/2$ e π , indicando que nesse período as variáveis estão fora de fase, com a dengue liderando. Logo após há um outro grupo de vetores próximo ao fim de 2015, com posicionamento entre 0 e $\pi/2$,

indicando que a temperatura máxima e a dengue estão com fases próximas tendo a temperatura máxima liderando. Por último, tem-se um núcleo no período de aproximadamente quatro semanas entre os dois núcleos anteriores e que está conectado com esses núcleos através do sinal da potência. Os vetores presentes apresentam também um posicionamento entre $\pi/2$ e π , estando as variáveis também fora de fase e com a dengue liderando.

Os principais períodos no espectro de potência global (Figura 4.29b) são os de duas e quatro semanas, tendo em vista que esses são os períodos onde mais tem a presença de máximos de potência no espectro contínuo.

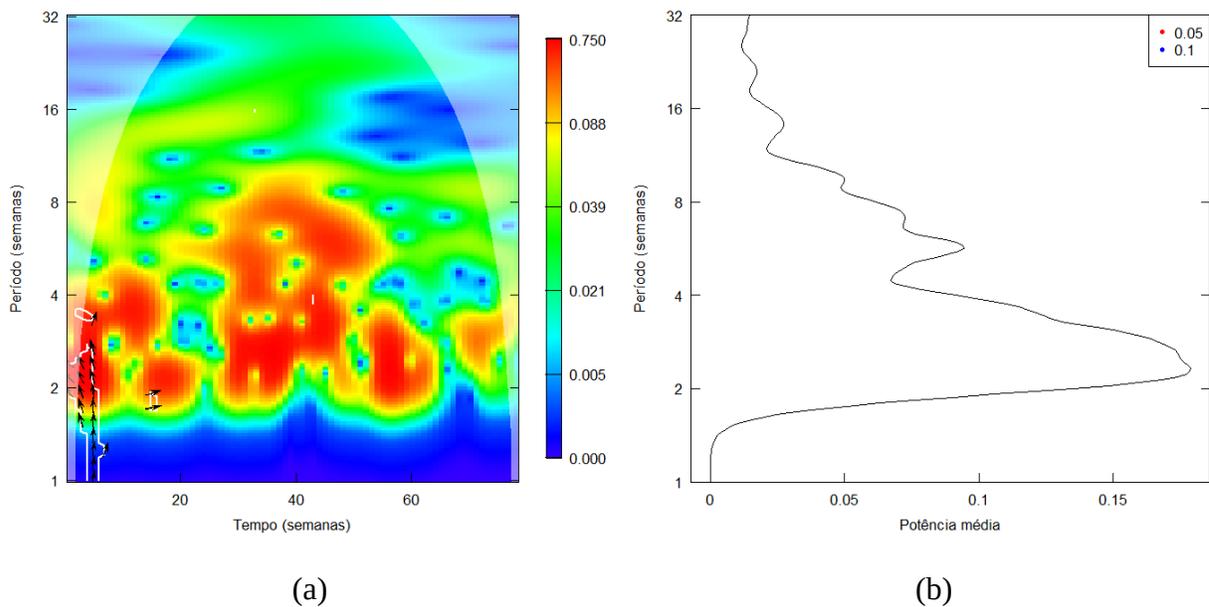


Figura 4.28 - Espectros bivariados de (a) potência de ondeleta contínua e (b) ondeleta global diferenças semanais de temperatura média com casos de dengue semanais para o bairro de Irajá durante o ano de 2016. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).

A TOC das diferenças semanais de temperatura mínima com a dengue para Irajá (Figura 4.30 a) apresenta vetores com significância estatística no período de aproximadamente três semanas em duas áreas diferentes: uma no início de 2016 e outra no segundo trimestre de 2016, onde ocorrem a maior parte dos registros de dengue (principalmente no início do ano), bem como as maiores diferenças semanais de temperatura mínima. Em ambos os casos os vetores apontam entre $-\pi/2$ e zero, indicando fases próximas entre as variáveis, com a dengue liderando.

Há também um sinal menor de covariância próximo ao final de 2015, com um vetor apontando entre $\pi/2$ e π , indicando que nesse momento as variáveis se encontram defasadas, também com a dengue liderando.

Um terceiro núcleo de vetores também se pronuncia no período de dezesseis semanas, entre o fim de 2015 e o começo de 2016, onde o sinal da covariância possui um perfil mais moderado e que se conecta com o sinal mais forte dos períodos menores que aparecem em 2016. Nesse caso, os vetores apontam em direção ao zero trigonométrico, indicando que as variáveis estão em fase entre si no período de dezesseis semanas após estarem com uma pequena diferença de fase no período de três semanas.

Agora serão comentados os resultados referentes ao ano de 2018, escolhido como referência para o ano não epidêmico no bairro de Irajá. Assim como na análise do ano epidêmico, aqui as variáveis térmicas e a dengue também foram analisadas a partir do semestre anterior ao ano escolhido como referência, ou seja, desde o segundo semestre de 2017 até o final de 2018.

O espectro de potência da diferença semanal da temperatura média (Figura 4.31 a) mostra sinais mais intensos no período de duas semanas logo no início da série temporal - meados de 2017, próximo ao fim de 2017, em meados de 2018 e durante o fim de 2018. Durante os três primeiros picos citados ocorrem aumentos significativos de temperatura média, porém o último aumento ocorrido no fim de 2018 precedeu a epidemia que ocorreu no início de 2019. No período de quatro a cinco semanas, também há a ocorrência de alguns máximos de potência contínua que se conectam com os máximos de duas semanas que ocorrem próximo ao fim de 2017, meados de 2018 e no fim de 2018. No espectro global (Figura 4.31 b), os períodos de maior potência são os de duas semanas e quatro semanas.

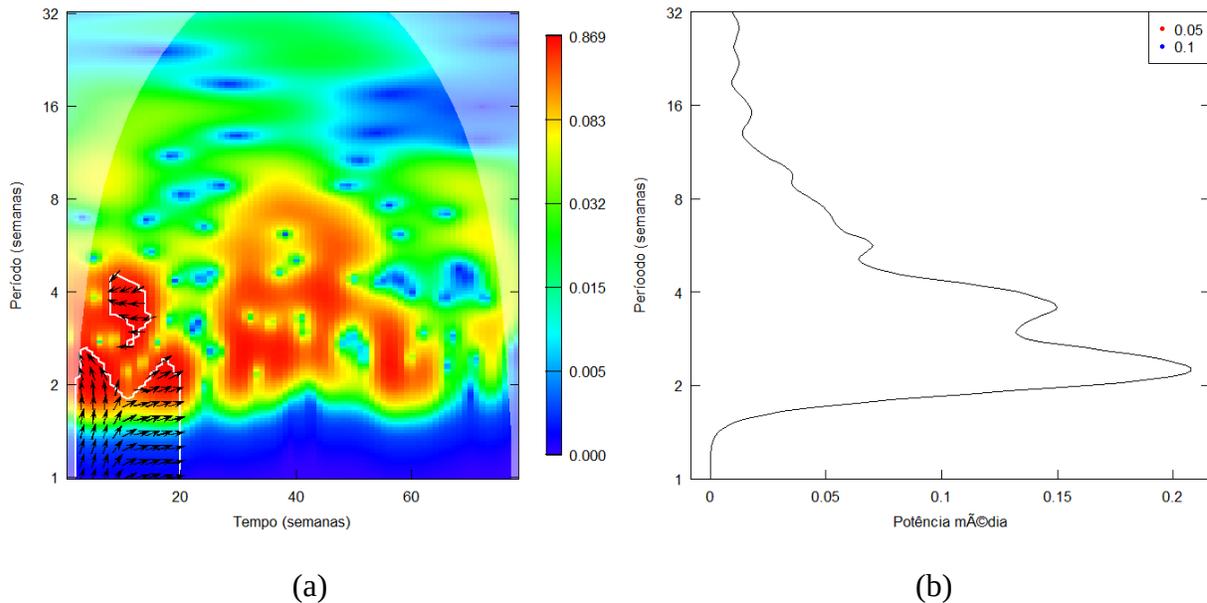


Figura 4.29 - Espectros bivariados de (a) potncia de ondeleta contnua e (b) ondeleta global diferenas semanais de temperatura mxima com casos de dengue semanais para o bairro de Iraj durante o ano de 2016. As linhas cheias brancas no espectro contnuo representam o nvel de confiana estatstica de 0,1 (90%) e as regies delimitadas pelo sombreado branco indicam o cone de influncia. Os pontos vermelhos no espectro global representam o nvel de significncia estatstica de 0,05 (95%).

As diferenas semanais de temperatura mxima, por sua vez, apresentam um padro diferente no espectro contnuo (Figura 4.32 a). Alm dos sinais que aparecem no terceiro trimestre e final de 2017, bem como os do incio, meados e final de 2018 no perodo aproximado de trs semanas, que so relativamente similares aos sinais principais das diferenas de temperatura mdia. Entretanto, esses mximos no apresentam a demarcao estatstica de 90%, exceto pelo mximo presente no final de 2018, que precede a epidemia que ocorre em 2019. Tambm h a presena de sinais em torno de seis semanas durante todo o segundo semestre de 2017, com marcao de significncia prximo ao fim de 2017, onde ocorrem declnio de temperatura com diferenas superiores a 5 °C.

Adicionalmente, nesse espectro contínuo de potência, há dois sinais de potência moderados em períodos maiores: um no segundo semestre de 2017 no período de aproximadamente vinte semanas e outro sinal no segundo semestre de 2018 no período de aproximadamente doze semanas, demarcando a presença de ciclos maiores de temperatura máxima. Durante ambos os sinais ocorre o aumento da temperatura máxima, sendo o de 2018 o mais pronunciado dentre os dois, no qual ocorre um novo aumento de temperatura máxima que precede a epidemia de 2019.

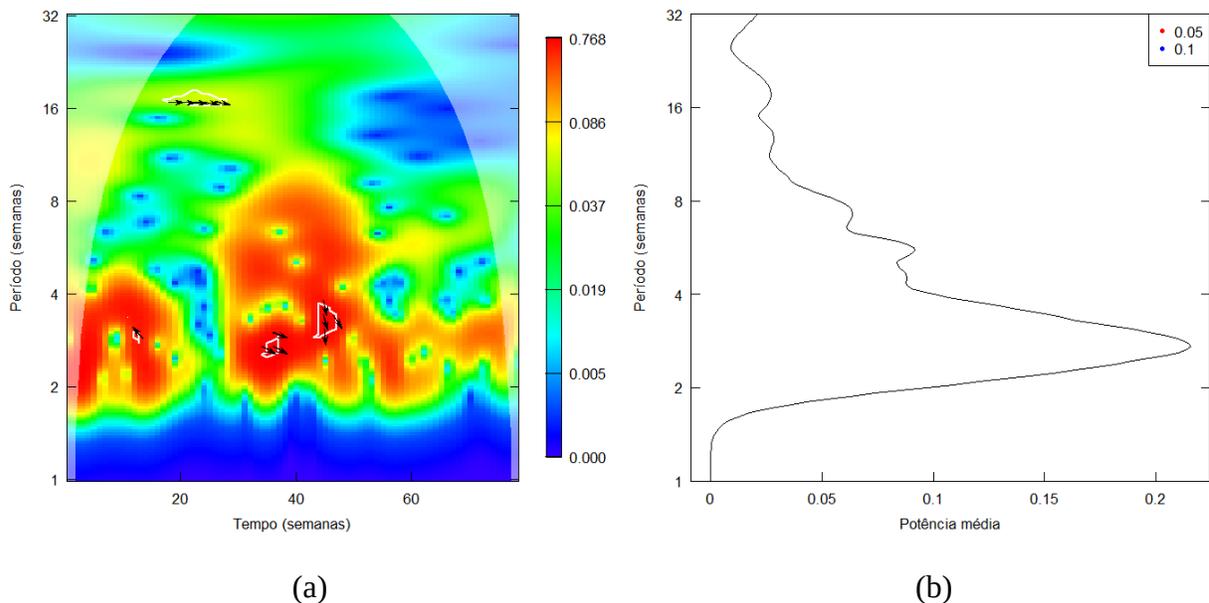


Figura 4.30 - Espectros bivariados de (a) potência de ondeleta contínua e (b) ondeleta global diferenças semanais de temperatura mínima com casos de dengue semanais para o bairro de Irajá durante o ano de 2016. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).

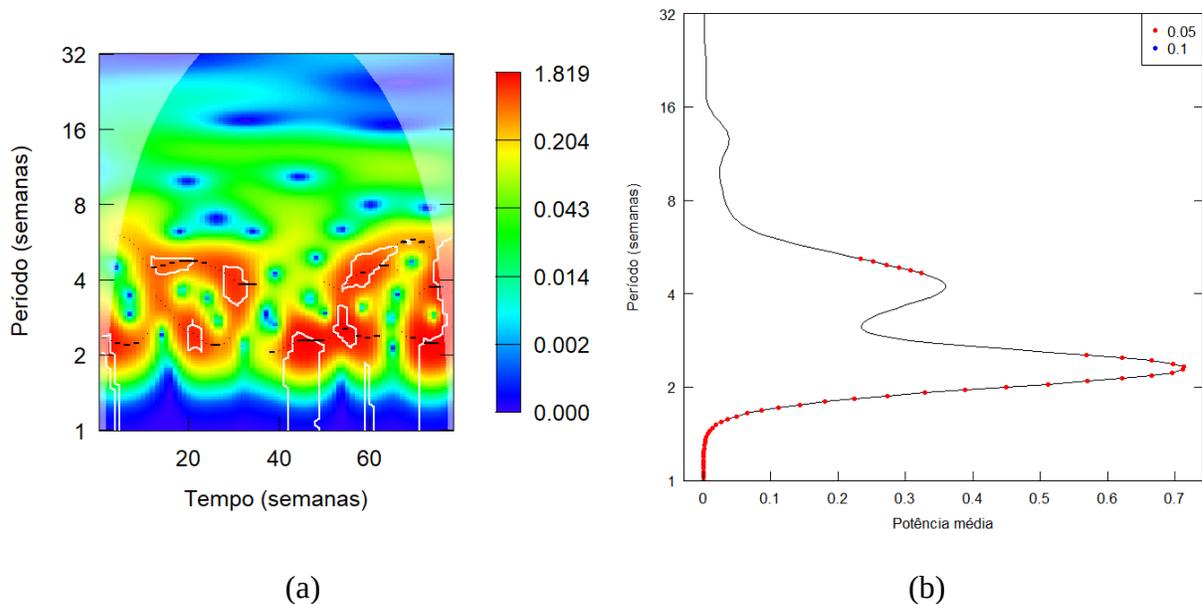


Figura 4.31 - (a) espectro de potência de ondeleta contínua e (b) espectro de ondeleta global das diferenças semanais de temperatura média para o bairro de Irajá durante o ano de 2018. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).

Os períodos principais no espectro global (Figura 4.32 b) são os de três e seis semanas, além dos secundários em doze e vinte semanas, sendo este último demarcado pela significância estatística de 95%.

No espectro contínuo de temperatura mínima para o ano não epidêmico (Figura 4.33 a), é possível também observar os máximos de potência durante o período de aproximadamente três semanas nos meados de 2017, meados de 2018 e final de 2018, marcando as maiores diferenças semanais de temperatura mínima.

O máximo localizado ao final de 2018 apresenta a linha de significância de 90%, além de ser o mais proeminente dentre todos os máximos de potência, sendo conectado a um máximo que

ocorre em torno de oito semanas e que se propaga até em torno de duas semanas. O período com maior potência, num panorama geral, é o de três semanas, sendo enfatizado pelo espectro de potência global (Figura 4.33 b).

Um ponto a se observar nos espectros de potência contínua das variáveis térmicas durante o ano não epidêmico é o sinal mais evidente que ocorre no final de todas as séries, principalmente no caso das diferenças semanais de temperatura mínima. Esse sinal é explicado pelas diferenças majoritariamente positivas entre as semanas do final de 2018, indicando um aumento das temperaturas nesse intervalo de tempo. Logo após esse período de elevação térmica, ocorre uma nova epidemia no início de 2019, que teve um total de 14343 infectados dentro do município do Rio de Janeiro.

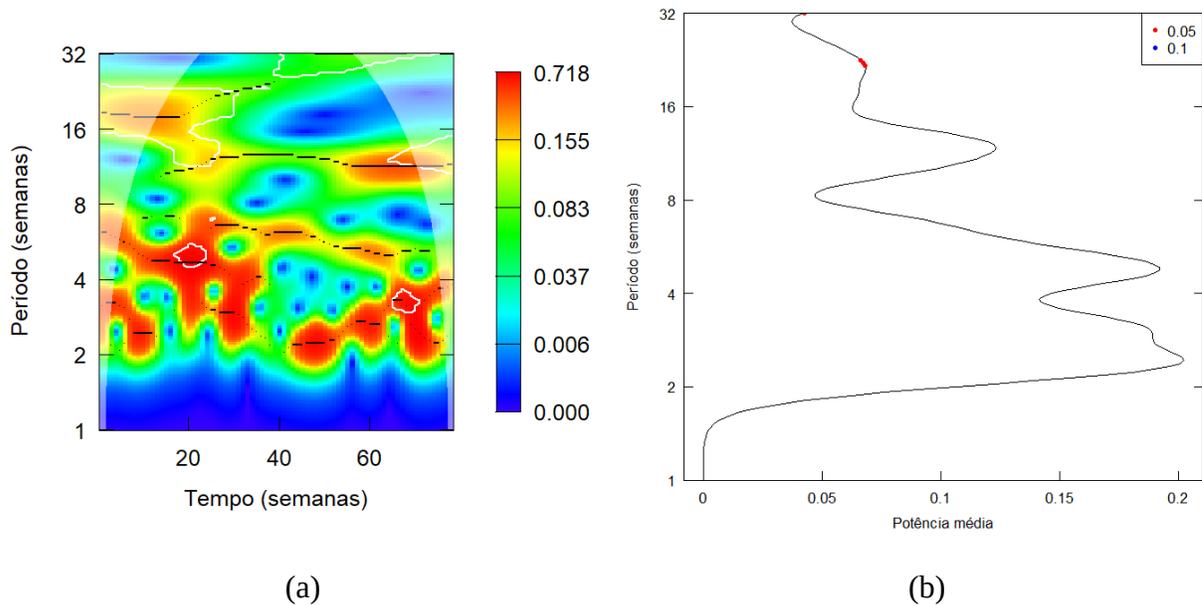


Figura 4.32 - (a) espectro de potência de ondeleta contínua e (b) espectro de ondeleta global das diferenças semanais de temperatura máxima para o bairro de Irajá durante o ano de 2018. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).

O espectro bivariado das diferenças semanais de temperatura média com os casos semanais de dengue (Figura 4.34 a) apresenta núcleos de cruzamento com significância estatística durante o primeiro trimestre e meados de 2018 no período aproximado de duas a três semanas. No primeiro núcleo, os vetores estão posicionados entre π e $-\pi/2$, indicando que no primeiro trimestre de 2018 a temperatura média e a dengue estão fora de fase e com a temperatura média liderando. Já para o segundo grupo de vetores, eles estão posicionados entre $\pi/2$ e π , indicando que durante o meio de 2018 a temperatura média e a dengue estão também fora de fase, com a dengue liderando.

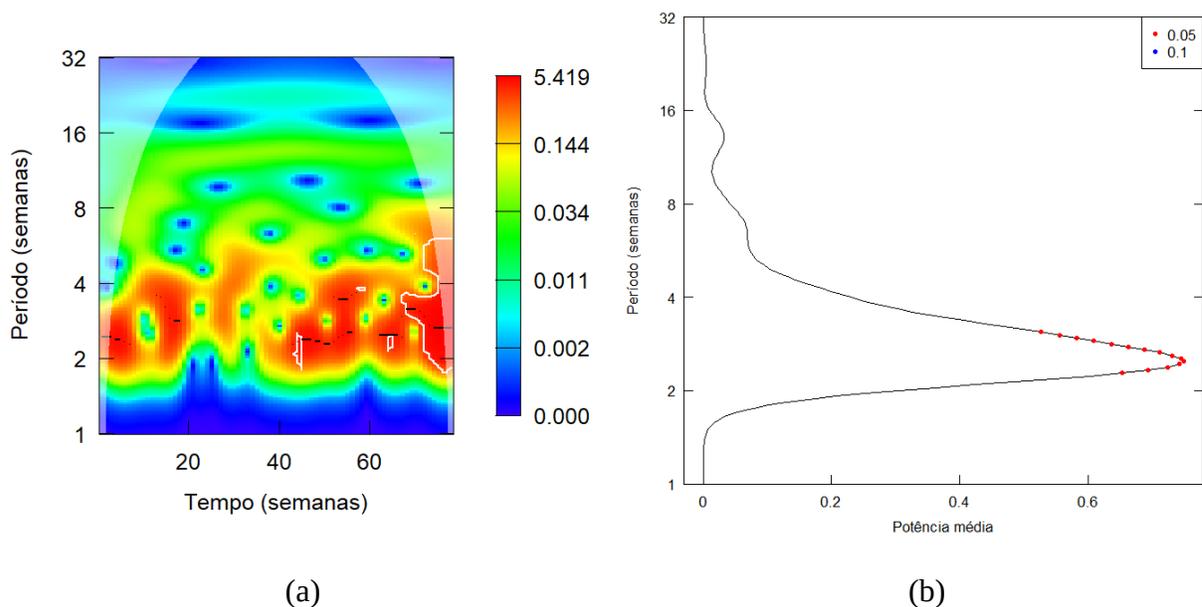


Figura 4.33 - (a) espectro de potência de ondeleta contínua e (b) espectro de ondeleta global das diferenças semanais de temperatura mínima para o bairro de Irajá durante o ano de 2018. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).

No período de aproximadamente quatro semanas há outros dois núcleos, um também nos meados de 2018 e outro no fim de 2018. Para o primeiro caso, os vetores estão posicionados entre

0 e $\pi/2$, indicando que as variáveis estão com fases próximas durante o meio de 2018, tendo a temperatura média liderando. Já para o final de 2018 os vetores estão posicionados entre $-\pi/2$ e 0, mostrando que as variáveis também se encontram em fase, porém com a dengue liderando.

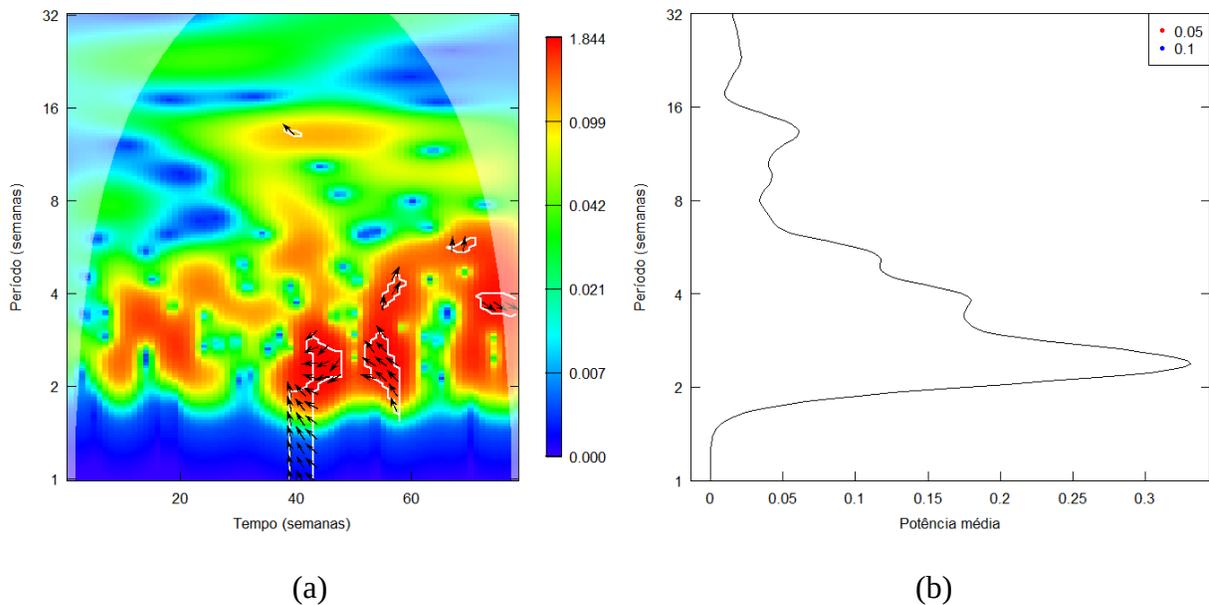


Figura 4.34 - Espectros bivariados de (a) potência de ondeleta contínua e (b) ondeleta global das diferenças semanais de temperatura média com casos de dengue semanais para o bairro de Irajá durante o ano de 2018. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).

Para a temperatura máxima, os núcleos de vetores que aparecem no espectro bivariado contínuo (Figura 4.35 a) também estão localizados no primeiro trimestre, meados e final de 2018, no período aproximado de duas a três semanas, sendo diferenciados pela posição dos vetores. No núcleo do primeiro trimestre os vetores iniciam entre $\pi/2$ e π , ficando próximos de π até ficarem entre π e $-\pi/2$ com o decorrer da série. Isso quer dizer que as variáveis, neste intervalo de tempo,

iniciam fora de fase com a liderança da dengue, para em seguida ter a liderança da diferença da temperatura máxima. Os núcleos de meados de 2018 também apresentam o padrão entre π e $-\pi/2$, indicando que as variáveis continuam fora de fase e com a liderança das diferenças de temperatura máxima. Por último, o núcleo do final de 2018 apresenta vetores posicionados entre $-\pi/2$ e 0, demonstrando que as variáveis estão em fase nesse trecho da série temporal, com a liderança da dengue.

Há também dois núcleos próximos a vinte e seis semanas (período semestral) que estão parcialmente dentro do cone de influência, localizados no fim de 2017 e em meados de 2018. O primeiro apresenta vetores posicionados em π , indicando que as variáveis estão fora de fase, com a liderança da dengue. Já o segundo núcleo apresenta vetores entre π e $-\pi/2$, indicando que as variáveis continuam fora de fase, porém com a temperatura máxima liderando. Por fim, o espectro de potência global (Figura 4.35 b) denota o período de aproximadamente três semanas como o de maior potência média, com um máximo secundário de aproximadamente seis semanas.

Ao se avaliar o espectro cruzado das diferenças de temperatura mínima com a dengue (Figura 4.36 a), as áreas com significância para os vetores também se encontram no primeiro trimestre de 2018, meados de 2018 e fim de 2018, no período de aproximadamente três semanas. O posicionamento dos vetores no primeiro semestre de 2018 aqui se assemelha com o caso da temperatura média, estando então a temperatura mínima e a dengue também fora de fase e com a temperatura mínima liderando.

Nos meados de 2018 também ocorre algo parecido em relação à posição dos vetores em comparação com a temperatura média, onde os vetores se posicionam entre 0 e $\pi/2$ no período de quatro semanas, virando em direção ao posicionamento entre $\pi/2$ e π durante o período de três semanas. Isso indica que durante o meio de 2018 as variáveis temperatura mínima e dengue estão com fases próximas, tendo a temperatura mínima liderando. Adicionalmente, em direção às periodicidades menores, as variáveis apresentam comportamento fora de fase, com a dengue liderando.

A maior diferença, mais uma vez, é encontrada no último núcleo de vetores posicionado no final da série temporal. Além da maior intensidade no espectro de potência, o posicionamento

dos vetores apresenta um padrão entre $-\pi/2$ e 0 no período de quatro semanas, indo em direção ao padrão entre $\pi/2$ e π no período de três semanas. Isso significa que as variáveis iniciam em fases próximas com a liderança da dengue numa periodicidade maior, mudando para uma situação onde as variáveis ficam fora de fase, também com a dengue liderando, durante um período menor.

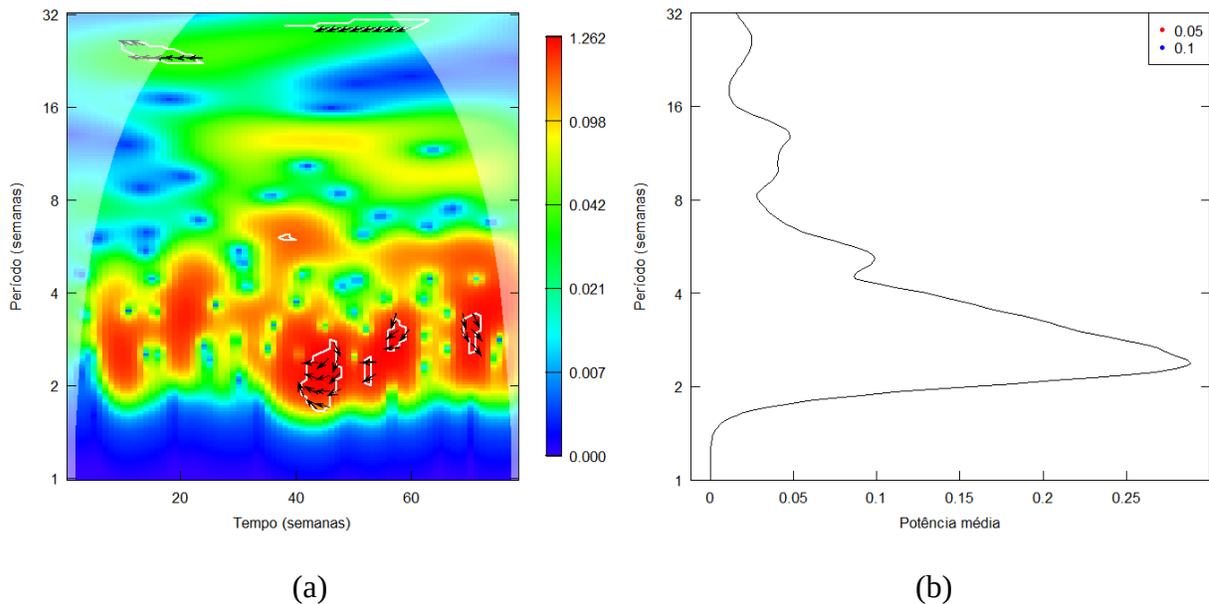


Figura 4.35 - Espectros bivariados de (a) potência de ondeleta contínua e (b) ondeleta global das diferenças semanais de temperatura máxima com casos de dengue semanais para o bairro de Irajá durante o ano de 2018. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).

O período de maior potência média no espectro bivariado é o de aproximadamente três semanas, como demonstrado na Figura 4.36b. Também há a presença de um pico secundário de aproximadamente quatorze semanas, porém ele não apresenta nenhum núcleo sobreposto representativo de sinal cruzado entre as variáveis no espectro contínuo bivariado.

Grandes cidades como o Rio de Janeiro, que apresentam totais populacionais de grandes dimensões, e que apresentam características climáticas como altas temperaturas e grande aporte de umidade, tem as arboviroses como uma problemática de alta revelância para os órgãos públicos de diversas esferas, levando em conta além dos fatores climáticos as questões que envolvem a vulnerabilidade socioambiental, como falta de infraestrutura e ausência de saneamento básico, que também contribuem para a manutenção do ciclo de vida do mosquito vetor de transmissão da dengue (OLIVEIRA, 2019).

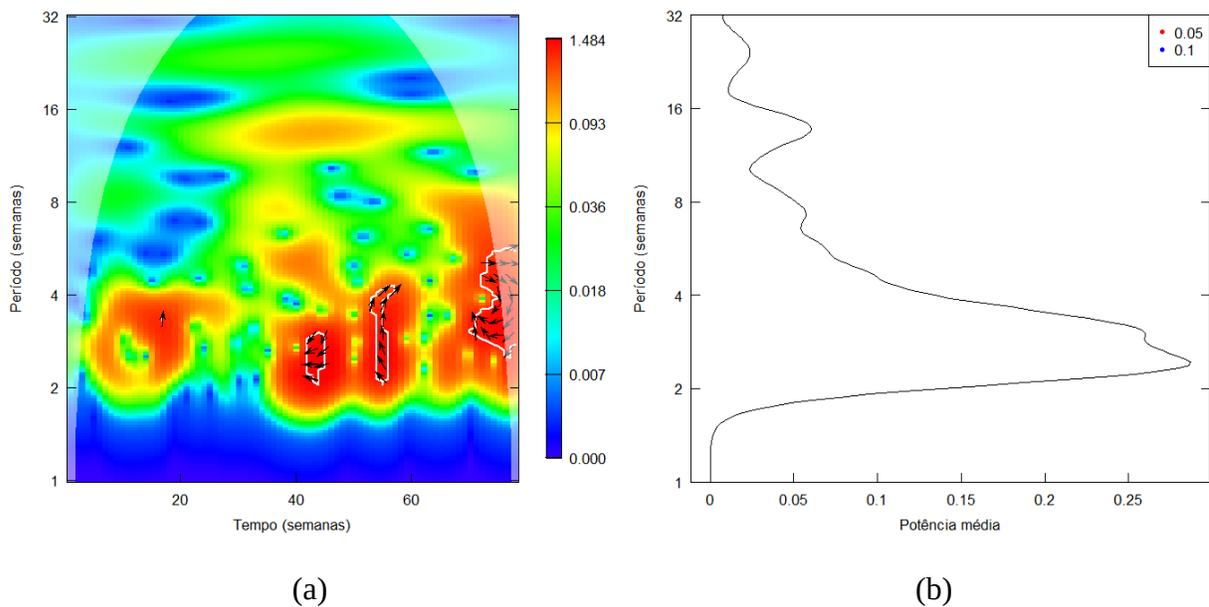


Figura 4.36 - Espectros bivariados de (a) potência de ondeleta contínua e (b) ondeleta global das diferenças semanais de temperatura mínima com casos de dengue semanais para o bairro de Irajá durante o ano de 2018. As linhas cheias brancas no espectro contínuo representam o nível de confiança estatística de 0,1 (90%) e as regiões delimitadas pelo sombreado branco indicam o cone de influência. Os pontos vermelhos no espectro global representam o nível de significância estatística de 0,05 (95%).

Ao se verificar todas as análises trazidas nesse capítulo, percebe-se que o bairro de Irajá, escolhido como representativo da ICU atendendo os critérios citados por Lucena e Peres (2017), possui uma representatividade considerável no total de casos de dengue para o município do Rio de Janeiro ao se analisar os valores absolutos. Entretanto, ao se levar em conta o número de habitantes, o bairro passa a ser menos representativo em relação à dengue no município.

Adicionalmente, ao se analisar as ocorrências epidêmicas registradas durante o período em que o bairro passou a ser considerado o mais quente do município, tendo recordes de temperatura registrados a partir da segunda metade da década passada, tem-se que essas epidemias ocorridas após o aumento das temperaturas possuem menor quantidade de registros em comparação às epidemias anteriores que aconteceram no município.

O bairro de Irajá contendo os máximos de temperatura durante os anos avaliados, não apresentou os maiores totais de casos de dengue. Isso traz a constatação de que a parte térmica da ICU por si só exerce influência na transmissão de dengue até um certo ponto, dando espaço para novas investigações no bairro levando em conta outros fatores que influenciam na transmissão da dengue, como por exemplo a precipitação e a influência de ações governamentais no combate e prevenção à doença.

Capítulo 5

5 Conclusões

A dengue é uma doença viral de grande destaque nas localidades tropicais, sendo bem expressiva na cidade do Rio de Janeiro, que é uma das mais importantes metrópoles do Brasil. Existem diversos estudos na literatura que enfatizam a relação entre a dengue e seus fatores influenciadores, como, por exemplo, variáveis meteorológicas medidas no local de pesquisa ou até fenômenos de variabilidade em diversas escalas temporais.

Nos últimos anos vem sendo feitos diversos estudos sobre a dengue, incluindo efeitos do risco de epidemia na saúde humana, que podem estar crescendo por conta de fatores como aquecimento global, ilhas de calor e episódios de ondas de calor, bem como a forma de organização da população.

Neste estudo, é investigada a possível relação entre a ocorrência de dengue e as variáveis térmicas no bairro de Irajá, localizado na cidade do Rio de Janeiro, que vem sendo considerado um dos bairros mais quentes do município nos últimos anos e que atende os critérios de caracterização de ilha de calor. Em conjunto, também é investigada a variância da dengue no município do Rio de Janeiro em função dos totais de dengue por bairro.

Durante as epidemias investigadas para o bairro de Irajá é verificado através das diferenças térmicas semanais que antes dos surtos de dengue houve um aumento de, pelo menos, uma das variáveis térmicas avaliadas no intervalo de uma a duas semanas. Antes das epidemias de 2016 e 2019 as ocorrências de dengue e as variáveis térmicas locais se encontram em fase, com o diferencial das temperaturas média e máxima estarem liderando em relação à dengue no fim de 2015, enquanto que para o fim de 2018 a dengue liderou em relação às temperaturas.

Em relação às contribuições por bairro para o total de dengue no município, tem-se através da análise de agrupamento que os bairros com maiores valores absolutos se localizam majoritariamente em bairros na Zona Oeste e na Zona Norte, que também apresentam grandes

totais populacionais, inclusive o bairro de Irajá, que possui um total de quase 100000 habitantes. Ao aplicar na análise a proporção entre os totais de dengue e os totais populacionais de cada bairro, os bairros com menor quantidade de habitantes acabam ganhando maior destaque com relação aos menos populosos.

Foram geradas equações de regressão múltipla para verificar a relação estatística entre a dengue no município do Rio de Janeiro e a dengue nos bairros e, a partir das equações de regressão foram realizadas análises de variância para todos os bairros e para os grupos gerados na análise de agrupamento do método *k-means*. Em ambos os casos, os resultados com maior caráter discriminatório entre as contribuições dos bairros e com melhor resposta estatística dos grupos foram relacionados à dengue proporcionalizada, reforçando a necessidade de não se avaliar somente os totais absolutos.

Este trabalho traz mais uma contribuição no estudo da dengue, dentre os diversos estudos epidemiológicos realizados, na tentativa de auxiliar ainda mais na compreensão da doença. O estudo também abre a possibilidade de trazer outros aspectos a serem abordados e que interferem na transmissão da doença e conseqüentemente no surgimento de epidemias na cidade do Rio de Janeiro. Podem ser utilizadas outras variáveis meteorológicas locais (precipitação, por exemplo) ou variáveis sociais relacionadas a vulnerabilidade e estruturas locais que auxiliam no ciclo de vida do vetor transmissor da doença.

Referências

- ALCOFORADO, M. J. *O clima da região de Lisboa: contrastes e ritmos térmicos*. Centro de estudos geográficos Universidade de Lisboa, 1993.
- ALEXANDRINO, F. I. G. *Ilhas de calor e casos de dengue na área urbana de Taubaté-SP*. Dissertação (Mestrado em Engenharia Mecânica), Faculdade de Engenharia do Campus de Guaratinguetá, Universidade Estadual Paulista. 2017.
- ALTINO, L. *Entenda como Irajá se tornou o bairro mais quente do Rio*. Extra, Rio de Janeiro, 29 de jan. De 2021. Disponível em: <<https://extra.globo.com/noticias/rio/entenda-como-iraja-se-tornou-bairro-mais-quente-do-rio-24859930.html>>. Acesso em: 20 de abr. De 2021.
- AMORIM, A. C. B., LUCIO, P. S., SILVA, C. M. S., RODRIGUES, D. T., BRITO NETO, F. A. D. *Regionalização e análise da tendência da precipitação do Rio Grande do Norte associados a padrões de TSM*. Revista Brasileira de Meteorologia, v. 35, p. 269-280, 2020.
- ARAÚJO, R. V., ALBERTINI, M. R., COSTA-DA-SILVA, A. L., SUESDEK, L., FRANCESCHI, N. C. S., BASTOS, N. M., KATZ G., CARDOSO V. A., CASTRO B. C., CAPURRO M. L., ALLEGRO, V. L. A. C. *São Paulo urban heat islands have a higher incidence of dengue than other urban areas*, The Brazilian Journal of Infectious Diseases, São Paulo, v. 19, p.146-155, 2015.
- ARAÚJO, V. E. M.; BEZERRA, J. M. T.; AMÂNCIO, F. F.; PASSOS, V. M. D. A.; CARNEIRO, M. *Aumento da carga de dengue no Brasil e unidades federadas, 2000 e 2015: análise do Global Burden of Disease Study 2015*. Revista Brasileira de Epidemiologia, v. 20, p. 205-216, 2017.

- BARACHO, R. C. M. *Influência de variáveis meteorológicas sobre a incidência dos casos de dengue no município de Areia-PB. Trabalho de conclusão de curso, Universidade Federal da Paraíba. 2013.*
- BAJSANSKI, I.V., Milošević, D.D. and Savić, S.M. Evaluation and improvement of outdoor thermal comfort in urban areas on extreme temperature days: Applications of automatic algorithms. *Building and Environment*, v. 94, p. 632-643, 2015.
- BARCELLOS, C. D. C., PUSTAI, A. K., WEBER, M. A., BRITO, M. R. V. *Identificação de locais com potencial de transmissão de dengue em Porto Alegre através de técnicas de geoprocessamento. 2005.*
- BESERRA, E. B. JUNIOR, F. P. C; SANTOS, T.S.; FERNANDES, C. R. M. *Biologia e exigências térmicas de Aedes aegypti (L.) (Díptera: culicidae) provenientes de quatro regiões bioclimáticas da Paraíba. Revista Neotropical Entomology. V. 35, n.6, p. 853-860. 2006.*
- BESERRA, E. B., FERNANDES, C. R., SILVA, S. A. D. O., SILVA, L. A. D., SANTOS, J. W. D. *Efeitos da temperatura no ciclo de vida, exigências térmicas e estimativas do número de gerações anuais de Aedes aegypti (Diptera, Culicidae). Iheringia. Série Zoologia, v. 99, p. 142-148, 2009.*
- BLENKINSOP, S., E. Lewis, S. C. Chan and H. J. Fowler. *Quality-control of an hourly rainfall dataset and climatology of extremes for the UK. International Journal of Climatology. 37: 722–740, 2017.*
- BORGES, S. M. A. A. *Importância epidemiológica do Aedes albopictus nas Américas. Tese de Doutorado. Universidade de São Paulo – USP, 2001.*
- BRAGA, I. A.; VALLE, D. *Aedes aegypti: histórico do controle no Brasil. Epidemiologia e Serviços de Saúde, v. 16, n. 2, p. 113-118, 2007.*

- BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Diretrizes Nacionais para a Prevenção e Controle de Epidemias de Dengue. Brasília, 2009.
- BRAZ, A. M., OLIVEIRA, I. J. D., CAVALCANTI, L. C. D. S., ALMEIDA, A. C. D., CHÁVEZ, E. S. Análise de agrupamento (cluster) para tipologia de paisagens. Mercator (Fortaleza), v. 19, 2020.
- CÂMARA, F. P., GOMES, A. F., SANTOS, G. T. D., CÂMARA, D. C. P. *Clima e epidemias de dengue no Estado do Rio de Janeiro. Revista da Sociedade Brasileira de Medicina Tropical*, v. 42, n. 2, p. 137-140, 2009.
- CARDOSO, R. D. S., DORIGON, L. P., TEIXEIRA, D. C. F., AMORIM, M. C. D. C. T. *Assessment of urban heat islands in small-and mid-sized cities in Brazil. Climate*, v. 5, n. 1, p. 14, 2017.
- CARDOSO, R. D. S.; AMORIM, M. C. D. C. T. *Urban heat island analysis using the 'local climate zone'scheme in Presidente Prudente, Brazil. 2018.*
- CONSOLI, A. G. B. R.; LOURENÇO-DE-OLIVEIRA, R. Principais mosquitos de importância sanitária no Brasil. Rio de Janeiro: Fiocruz, 1994.
- DOS SANTOS GOMES, A. C., DA SILVA COSTA, M., COUTINHO, M. D. L., DO VALE, R. S., DOS SANTOS, M. S., DA SILVA, J. T., FITZJARRALD, D. R. ANÁLISE ESTATÍSTICA DAS TENDÊNCIAS DE ELEVAÇÃO NAS SÉRIES DE TEMPERATURA MÉDIA MÁXIMA NA MAZÔNIA CENTRAL: ESTUDO DE CASO PARA A REGIÃO DO OESTE DO PARÁ.(STATISTICAL ANALYSIS OF THE TRENDS OF ELEVATION IN MAXIMUM AVERAGE TEMPERATURE IN CENTRAL AMAZONIA: CASE STUDY FOR REGION IN WEST OF PARÁ STATE). *Revista Brasileira de Climatologia*, v. 17, 2015.

DOMINGUES, M. O., MENDES JR, O., DA COSTA, A. M. *Algumas aplicações wavelet na análise de sinais atmosféricos*. 2003. *Série Arquimedes, Volume 2, Anais do DINCON 2003. 2o Congresso Temático de Aplicações de Dinâmica e Controle da Sociedade Brasileira de Matemática Aplicada e Computacional (SBMAC)*. Disponível em <<http://www.lac.inpe.br/~margarete/Dincon2003Wavelet.pdf>>. Acesso em Mai. 2017.

FARNESI, L. C. et al. *Desenvolvimento embrionário de Aedes aegypti (Diptera: Culicidae): influência de diferentes temperaturas constantes*, 2009.

FUNDAÇÃO OSWALDO CRUZ – FIOCRUZ. *FGV e Fiocruz lançam sistema de alerta para monitorar dengue no Rio de Janeiro*. 09 de fev. De 2015. Disponível em <<https://www.canalsaude.fiocruz.br/noticias/noticiaAberta/fgv-e-fiocruz-lancam-sistema-de-alerta-para-monitorar-dengue-no-rio-de-janeiro-2015-02-09>>. Acesso em 10 de jan. de 2023.

GOMES, A. F.; NOBRE, A. A.; CRUZ, O. G. *Temporal analysis of the relationship between dengue and meteorological variables in the city of Rio de Janeiro, Brazil, 2001-2009*. *Cadernos de Saúde Pública*, v. 28, n. 11, p. 2189-2197, 2012.

HONÓRIO, N.A.; CODEÇO, C.T.; ALVES, F.C.; MAGALHÃES, M.A.F.M.; LOURENÇO-DE-OLIVEIRA, R. *Temporal Distribution of Aedes aegypti in Different Districts of Rio De Janeiro, Brazil, Measured by Two Types of Traps*. *J. Med. Entom.*, v. 46, p. 1001–1014, 2009.

HOWARD, L. *The Climate of London, Vols. I - III*, London. 1833.

HU, L.; MONAGHAN, A. J. *Investigation of Urban Air Temperature and Humidity Patterns during Extreme Heat Conditions Using Satellite-Derived Data*. *JAM-AMS*. DOI: 10.1175/JAMC-D-15-0051.1, 2015.

IGUCHI, J. A.; SEPOSO, X. T.; HONDA, Y. *Meteorological factors affecting dengue incidence in Davao, Philippines*. *BMC public health*, v. 18, n. 1, p. 629, 2018.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. *Censo Demográfico 2010*. Disponível em https://biblioteca.ibge.gov.br/visualizacao/periodicos/93/cd_2010_caracteristicas_populacao_domicilios.pdf. Acesso em 02 ago. de 2021.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. *Estimativas da população residente com data de referência 1o de julho de 2020*. Disponível em <https://www.ibge.gov.br/estatisticas/sociais/populacao/9103-estimativas-de-populacao.html?=&t=resultados>. Acesso em 02 ago. de 2021.

INSTITUTO PEREIRA PASSOS – IPP. *Atlas Escolar da Cidade do Rio de Janeiro – Edição 2018*. Disponível em: https://portalgeo.rio.rj.gov.br/RioAtlas/PDF/ATLAS_ESCOLAR_2018_COMPLETO.pdf. Acesso em 02 ago. de 2021.

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE – IPCC. *Sixth Assessment Report: Impacts, Adaptation and Vulnerability*. 2022.

JIANG, X. D., XIA, B. C., GUO, L., Li, N. *Characteristics of multi-scale temporal-spatial distribution of urban heat island in Guangzhou*. *Ying yong sheng tai xue bao= The journal of applied ecology*, v. 18, n. 1, p. 133-139, 2007.

JOHANSSON, M. A.; CUMMINGS, D. AT; GLASS, G. E. *Multiyear climate variability and dengue—El Nino southern oscillation, weather, and dengue incidence in Puerto Rico, Mexico, and Thailand: a longitudinal data analysis*. *PLoS medicine*, v. 6, n. 11, p. e1000168, 2009.

LIMA-CAMARA, T. N. *Emerging arboviruses and public health challenges in Brazil*. *Revista de saúde publica*, v. 50, p. 36, 2016.

LIU Y., LIANG X.S., WEISBERG R.H. Rectification of the Bias in the Wavelet Power Spectrum. *Journal of Atmospheric and Oceanic Technology*, v. 24, p. 2093–2102, 2007.

LEMOS, L. D. O.; JÚNIOR, A. C. O.; MENDONÇA, F. D. A. Urban climate maps as a public health tool for urban planning: The case of dengue fever in Rio De Janeiro/Brazil. *Urban Climate*, v. 35, p. 100749, 2021.

LOUGHNAN, M., Nicholls, N. and Tapper, N.J., 2012. Mapping heat health risks in urban areas. *International Journal of Population Research*, 2012.

LUCENA, A. J., ROTUNNO FILHO, O. C., DE ALMEIDA FRANÇA, J. R., DE FARIA PERES, L., XAVIER, L. N. R. Urban climate and clues of heat island events in the metropolitan area of Rio de Janeiro. *Theoretical and applied climatology*, v. 111, n. 3, p. 497-511, 2013.

LUCENA, A. J., PERES, L. F. Métodos em clima urbano aplicados à cidade do Rio de Janeiro (Brasil) e sua região metropolitana. *The Overarching Issues of the European Space: Society, Economy and Heritage in a Scenario. Porto: FLUP*, pp. 312-326. 2017.

MELO, V. S.; SOUSA, F. A. S. Análise de frequência de secas utilizando técnica de agrupamento e distribuições de probabilidades. *Engenharia Sanitaria e Ambiental*, v. 26, p. 53-60, 2021.

METZ, J. Interpretação de clusters gerados por algoritmos de clustering hierárquico. Dissertação (Mestrado em Ciências de Computação e Matemática Computacional) – Instituto de Ciências de Computação e Matemática Computacional, Universidade de São Paulo, 2006.

MIAGOSTOVICH, M. P., NOGUEIRA, R. M., CAVALCANTI, S., MARZOCHI, K. B., & SCHATZMAYR, H. G. Dengue epidemic in the state of Rio de Janeiro, Brazil: virological and epidemiological aspects. *Revista do Instituto de Medicina Tropical de São Paulo*, 35(2), 149-154. 1993.

MISHRA, V., Ganguly, A.R., Nijssen, B. and Lettenmaier, D.P., 2015. Changes in observed climate extremes in global urban areas. *Environmental Research Letters*, 10(2), p.024005.

MONTEIRO, F. F., GONÇALVES, W. A., ANDRADE, L. D. M. B., VILLAVICENCIO, L. M. M., DOS SANTOS SILVA, C. M. Assessment of Urban Heat Islands in Brazil based on MODIS remote sensing data. *Urban Climate*, v. 35, p. 100726, 2021.

MORAIS JÚNIOR, R. R. Estudo do desempenho da soldagem por eletrodo revestido em diferentes posições por meio de análise de variância. Trabalho de Conclusão de Curso, Universidade Federal do Pampa. 2015.

MORDECAI, E. A., COHEN, J. M., EVANS, M. V., GUDAPATI, P., JOHNSON, L. R., LIPPI, C. A., MIAZGOWICZ, K., MURDOCK, C. C., ROHR, J. R., RYAN, S. J., SAVAGE, V., SHOCKET, M. S., IBARRA, A. S., THOMAS, MATTHEW. B., WEIKEL, D. P. Detecting the impact of temperature on transmission of Zika, dengue, and chikungunya using mechanistic models. *PLoS neglected tropical diseases*, v. 11, n. 4, p. e0005568, 2017.

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION – NOAA. El Niño Index Dashboard. Disponível em <<https://www.esrl.noaa.gov/psd/enso/dashboard.html>>. Acesso em 15 Dez. 2022.

NOSSEDOTTI, C. C. BY-SA 3.0. Zoneamento da Cidade do Rio de Janeiro. 2013. Disponível em <https://commons.wikimedia.org/wiki/File:Zoneamento_da_Cidade_do_Rio_de_Janeiro.svg>Nossedotti, <<https://creativecommons.org/licenses/by-sa/3.0/>>CC BY-SA 3.0, via Wikimedia Commons. Acesso em 24 Jan. 2023.

OKE, T. R. The energetic basis of the urban heat island. *Quarterly Journal of the Royal Meteorological Society*, v. 108, n. 455, p. 1-24, 1982.

OKE, T. R. *The heat island of the urban boundary layer: characteristics, causes and effects. Wind climate in cities. Springer, Dordrecht, 1995. p. 81-107.*

OKE, T. R., JOHNSON, G. T., STEYN, D. G., WATSON, I. D. *Simulation of surface urban heat islands under 'ideal' conditions at night part 2: Diagnosis of causation. Boundary-layer meteorology, v. 56, n. 4, p. 339-358, 1991.*

OKE, T. R., MILLS, G., CHRISTEN, A., VOOGT, J. A. *Urban climates. Cambridge University Press, 2017.*

OLIVEIRA, R. F. O papel do clima na ocorrência da Chikungunya, Dengue e Zika entre 2007 e 2017 nos bairros de Copacabana, Guaratiba, Jardim Botânico e Santa Cruz da cidade do Rio de Janeiro. Dissertação (Mestrado em Geografia) – Programa de Pós-graduação em Geografia, Faculdade de Formação de Professores, Universidade do Estado do Rio de Janeiro, p. 165. 2019.

PAINEL BRASILEIRO DE MUDANÇAS CLIMÁTICAS – PBMC. Rio registra terceira maior temperatura em 100 anos, diz Climatempo. Disponível em <<http://pbmc.coppe.ufrj.br/index.php/en/news/494-rio-registra-3-maior-temperatura-em-100-anos-diz-climatempo-quase-43c>>. Acesso em 03 Nov. 2022.

PEREIRA, M. Al. F.; KOBAYAMA, M.; CASTRO, N. M. R. *Análise de variâncias pluviométricas na bacia hidrográfica do rio Taboão-RS. Revista Brasileira de Meteorologia, v. 29, p. 409-419, 2014.*

PIŠOFT, P.; KALVOVÁ, J.; BRÁZDIL, R. *Cycles and trends in the Czech temperature series using wavelet transforms. International Journal of Climatology: A Journal of the Royal Meteorological Society, v. 24, n. 13, p. 1661-1670, 2004.*

REBOITA, M. S. *Elementos da Variabilidade Climática no Extremo Sul do Brasil, no Período de 1990 a 2001. Dissertação, Mestrado. Curso de Pós-Graduação em Engenharia Oceânica da Fundação Universidade Federal do Rio Grande, 2004.*

RIO DE JANEIRO. Decreto Municipal nº 34377/2011, de 31 de agosto de 2011. Institui estado de alerta contra a dengue e dispõe a prevenção e o controle da transmissão e a atenção primária à saúde nos casos de dengue na Cidade do Rio de Janeiro e dá outras providências. Disponível em <http://www.rio.rj.gov.br/dlstatic/10112/5125745/4132389/DecretoMunicipaln34377.pdf>. Acesso em 24 Jan. 2023.

RÖSCH, A.; SCHMIDTBAUER, H. *WaveletComp: Computational Wavelet Analysis. R package version 1.0. Disponível em* http://www.hs-stat.com/projects/WaveletComp/WaveletComp_guided_tour.pdf. 2014. Acesso em 12 Jul. de 2020.

SANTANA, L. V. R.; DA SILVA, A. S. A. *Análise de agrupamento da velocidade do vento no Nordeste do Brasil. Sigmae, v. 8, n. 2, p. 481-489, 2019.*

SANTOS, J. P. C., HONÓRIO, N. A., BARCELLOS, C., NOBRE, A. *A perspective on inhabited urban space: land use and occupation, heat islands, and precarious urbanization as determinants of territorial receptivity to dengue in the city of Rio de Janeiro. International Journal of Environmental Research and Public Health, v. 17, n. 18, p. 6537, 2020.*

SANTOS, S. A. F. *Variabilidade e previsibilidade interanual da dengue no município do Rio de Janeiro. 2019. Monografia (Bacharelado em Meteorologia) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2019.*

- SANTOS, S. A. F., KARAM, H. A., JOSÉ, A., ROJAS, J. L. F., SUAZO, J. M. A., VÁSQUEZ, I. L., PEÑA, C. A. S. Dengue Climate Variability in Rio de Janeiro City with Cross-Wavelet Transform. *Journal of Environmental Protection*, v. 13, n. 3, p. 261-276, 2022.
- SCHATZ, J. and Kucharik, C.J., 2015. Urban climate effects on extreme temperatures in Madison, Wisconsin, USA. *Environmental Research Letters*, 10(9), p.094024.
- SECRETARIA MUNICIPAL DE SAÚDE - SMS. Dengue – dados epidemiológicos. Disponível em <<http://www.rio.rj.gov.br/web/sms/exibeconteudo?id=2815389>>. Acesso em 10 jul. De 2021.
- SILVA, M. T., ALCÂNTARA, C. R., SOUZA, E. P. D., OLINDA, R. A. D., GONÇALVES, W. A. Influência da temperatura da superfície do mar na ocorrência de linhas de instabilidade na costa Norte e Nordeste do Brasil. *Revista Brasileira de Meteorologia*, v. 32, p. 260-268, 2017.
- SILVA, A. F. A.; COSTA, V. A. F. Caracterização espacial de secas meteorológicas e hidrológicas na porção mineira da bacia do São Francisco. *III Simpósio da Bacia Hidrográfica do Rio São Francisco. Belo Horizonte, Minas Gerais, 2020.*
- Silva, E. M. D. S., Correia, W. L. F., Oliveira, J. F. D., Barros, H. G., Costa, M. D. S., Gois, G. D., Falcão, N. A. D. M. Espaço-temporalidade dos focos de calor na região metropolitana de Maceió. *Revista Brasileira de Meteorologia*, v. 35, p. 1029-1043, 2021.
- SILVA, J. S., DA SILVA, R. M., SANTOS, C. A. G. Spatiotemporal impact of land use/land cover changes on urban heat islands: A case study of Paço do Lumiar, Brazil. *Building and Environment*, v. 136, p. 279-292, 2018.
- SISTEMA ALERTA RIO DA PREFEITURA DO RIO DE JANEIRO. Temperaturas maxs/mins mensais. Disponível em <<http://alertario.rio.rj.gov.br/temp-maxs-mins-mensais/>>. Acesso em 12 dez. 2022.

- SOUZA, D. O. D.; ALVALÁ, R C D. S. *Observational evidence of the urban heat island of Manaus City, Brazil. Meteorological Applications*, v. 21, n. 2, p. 186-193, 2014.
- SUN, J., SHU, J., YUE, Q., LU, X. Q. *Multi-scale Characteristics of the Urban Heat Island Intensity in Shanghai. Journal-East China Normal University Natural Science*, v. 2, n. 132, p. 36, 2007.
- TALAGALA, T.; LOKUPITIYA, R. *Wavelet analysis of dengue transmission pattern in Sri Lanka. International Journal of Mosquito Research*, v.2, p.13-18, 2015.
- TORRENCE, C.; COMPO, G. P. *A practical guide to Wavelet analysis. Bull. Amer.Meteor. Soc.*, v.79, p.61-78, 1998.
- VELEDA, D.; MONTANGNE, R.; ARAUJO, M. *Cross-wavelet bias corrected by normalizing scales. J. Atmos. Oceanic Technol.*, v. 29, p. 1401–1408, 2012. doi:10.1175/JTECH-D-11-00140.1.
- VILANI, M. T.; SANCHES, L. *Transformada wavelet aplicada à variável temperatura do ar para área urbana e rural. In: XVII Congresso Brasileiro de Agrometeorologia. 2011.*
- VOOGT, J.A. *Urban Heat Island. Munn, T., Ed., Encyclopedia of Global Environmental Change, Wiley, Chichester*, v. 3, p. 660-666, 2002.
- XAVIER, D. R., MAGALHÃES, M. D. A. F. M., GRACIE, R., REIS, I. C. D., MATOS, V. P. D., BARCELLOS, C. *Difusão espaço-tempo do dengue no Município do Rio de Janeiro, Brasil, no período de 2000-2013. Cadernos de Saúde Pública*, v. 33, 2017.
- XIAO, J. *et al. Weather variables and the El Niño Southern Oscillation may drive the epidemics of dengue in Guangdong Province, China. Science of The Total Environment*, v. 624, p. 926-934, 2018.

- WANG, C., Zhang, Z., Zhou, M., Wang, P., Yin, P., Ye, W. and Zhang, L., *Different response of human mortality to extreme temperatures (MoET) between rural and urban areas: a multi-scale study across China. Health & place, 50, pp.119-129. 2018.*
- WENG, H., LAU, K. M. *Wavelets, period doubling, and time-frequency localization with application to organization of convection over the tropical western Pacific. Journal of the Atmospheric Sciences, v. 51, n. 17, p. 2523-2541, 1994.*
- WILKS, D. *Statistical Methods in the Atmospheric Sciences, 4th Edition, Elsevier, 840 p., 2019.*
- WORLD HEALTH ORGANIZATION - WHO. *Dengue: Guidelines for Diagnosis, Treatment, Prevention and Control. 2009. pp. 1-160. Disponível em http://whqlibdoc.who.int/publications/2009/9789241547871_eng.pdf. Acesso em 02/2018.*
- ZHANG J., ZHANG X., WANG X., ZHANG J., LIU Y., XIAO Y. *Análise multi-escala da temperatura e efeito da ilha de calor na área de Pequim. Geografia de Áreas Áridas , n. 1, p. 51-58, 2010.*

Anexo A – Artigo submetido ao ‘Anuário do Instituto de Geociências da UFRJ’ (2023)

Author’s version

Partitioned Variance of Dengue Disease Incidence in the City of Rio de Janeiro
Variância Particionada da Incidência da Dengue no Município do Rio de Janeiro

Suellen Araujo Franco dos Santos *et al.*

Abstract

In this work, weekly data on reports of worsening dengue from the Brazilian Ministry of Health are used in analyzes of variance partitioned by neighborhoods, k-means clustering and in multivariate linear regression modeling. Data from the AlertaRio mesonet meteorological stations in the city of Rio de Janeiro were also used to characterize the thermal distribution on the urban surface. The results show that neighborhood variables standardized by the local population improve the results obtained with the linear regression model, beyond the 95% statistical confidence. Four homogeneous groups explain the variance for a confidence level above 90%. The incidence of standardized dengue is well distributed in the city of Rio de Janeiro, despite the Urban Heat Island presenting spatial thermal structure and seasonal variation. Although the reproducibility of the dengue vector is associated with the presence of tropical temperatures, its distribution across neighborhoods can probably also be associated with other variables, such as air transport of the Dengue vector promoted by mesoscale recirculations such as breezes.

Key words: Dengue distribution in tropical metropolis; ANOVA of Dengue; Urban Heat Island

Resumo

Neste trabalho, dados semanais de notificações de agravamento da dengue do Ministério da Saúde brasileiro são utilizados em análises de variância particionadas por bairros, agrupamento k-médias e em modelagem por regressão linear multivariada. Dados das estações meteorológicas mesonet AlertaRio do município do Rio de Janeiro também foram usados na caracterização da

distribuição térmica na superfície urbana. Os resultados obtidos mostram que variáveis por bairro padronizadas pela população local melhoram os resultados obtidos como o modelo de regressão linear, para além da confiança estatística de 95%. Quatro grupos homogêneos explicam a variância para um grau de confiança acima de 90%. A incidência de dengue padronizada mostra-se bem distribuída na cidade do Rio de Janeiro, apesar da Ilha de Calor Urbana apresentar estrutura térmica espacial e variação sazonal. Embora a reprodutividade do vetor da dengue esteja associada a presença de temperaturas tropicais, sua distribuição pelos bairros provavelmente também pode estar associada a outras variáveis, como o transporte aéreo do vetor da dengue promovido por recirculações de mesoescala como brisas.

Palavras-chave: *Distribuição da dengue em metrópole tropical; ANOVA de Dengue; ilha de calor urbana*

1 Introduction

Dengue received an epidemiological character in the State of Rio de Janeiro (RJ) in 1986, having spread to other Brazilian regions (Braga & Valle, 2007). In RJ, the largest number of cases is found in the Metropolitan Region of Rio de Janeiro (RMRJ) (Miagostovich et al. 1993). The urban environment of the city of Rio de Janeiro is heterogeneous, with neighborhoods presenting a diversity of micro areas that vary among themselves in terms of topographic relief, land use and cover, urban microclimates, UHI, sanitation and urbanization, vulnerability, which influence the incidence and development of dengue in each neighborhood (Xavier et al., 2017).

There are different approaches to the problem, such as data analysis (statistics) and modeling (epidemiological). Among the most promising tools of relatively recent development is the wavelet transform, which is an advance over the Fourier transform for signal analysis. It has the advantage of allowing simultaneous analysis of power and frequency variations (or characteristic periods) over time. In general, the observational approach is based on wavelet transform analysis and is preceded by modeling based on linear regression models, assuming that the distribution errors are normal or even generalized, taking synchronous or lagged data. Although WLT-based modeling results in per-scale diagnoses and prognoses and still measures

coherence and phase with respect to predictor variables, using LRM still has the advantage of simplicity, i.e., it generates average results without bias.

There is evidence that higher temperatures expected in the future may increase the regions susceptible to dengue epidemics, in addition to increasing the ability of vector mosquitoes to become infected with the dengue virus and transmit the disease (IPCC, 2022). Variable temperature is one of the factors that influence the life cycle of the mosquito vector *Aedes aegypti*. For the mosquito to be able to remain in an environment with still water for the eggs to develop until the adult stage, the ideal is that the temperature is in the range of 21°C to 29°C (Beserra et al. 2006). In a survey carried out in several cities located in the state of Paraíba, Brazil, it was detected that for the incubation period of the eggs there was a reduction of approximately 27 days when the eggs were kept at 25°C and 30°C in relation to the incubation performed at a temperature of 15°C (Beserra et al., 2009). The hatching rate of the eggs, in turn, reaches values above 90% for the temperature range from 22°C to 28°C, with a limit temperature of 35°C for hatching to occur in approximately half of the eggs (Farsine , 2009).

The urban heat island (UIC) (Oke et al. 2017) has been associated a priori with the occurrence of dengue in tropical cities, as it offers more favorable conditions for the proliferation of *Aedes* vectors. Temperature influences the mosquito's ability to survive and determines its development and reproductive rates (Johansson et al. 2009). This increase is associated with two relevant processes that occur during the life of the vectors, which are: the reduction of larval development time and the reduction of the incubation period for the virus to be found in the salivary gland of the mosquito (Lima-Câmara, 2016).

For dengue, zika, and chikungunya transmission in the Americas, all three viruses, models, and recorded case data have detected transmission to occur between 18 and 34°C, with peak transmission occurring between 26 and 29°C (Mordecai et al. 2017). Something similar was detected in the city of Areia, located in the state of Paraíba, Brazil. The years with the highest occurrence of cases were those with temperatures between 21° and 29°C, considered ideal for mosquito development (Bessera et al. 2006). In addition, another pattern of behavior of the variables was found: as the temperature decreased, the number of dengue cases also decreased (Baracho, 2013).

For the municipality of Rio de Janeiro, a comparison was made between the first quarters of epidemic and non-epidemic years, so it was observed that dengue reached a greater number of inhabitants when temperatures were higher, especially when the minimum temperature was above of 22°C (Câmara et al. 2009). Furthermore, an increase in minimum temperature was associated with a significant increase in the number of dengue cases in the following month (Gomes et al. 2012). When analyzing the spatial distribution of the average surface temperature for the same location, a significant correlation was found between the meteorological variable and the distribution of *Aedes aegypti* mosquito larvae (Lemos et al., 2021).

In São Paulo, it was highlighted in the capital that the incidence of the disease was higher in areas with higher temperature and lower vegetation cover. Furthermore, it has been found that the larval stage of the mosquito vector develops best in the temperature range of 28-32°C (Araujo et al. 2015). In the city of Taubaté, Brazil, there was a strong correlation between temperature and the areas of the municipality with the highest number of registered cases of dengue (Alezandrino, 2017).

In addition to the effects of temperature distribution on the *Aedes aegypti* life cycle, socioeconomic and urbanization factors also influence dengue infection rates. For example, Knoblauch et al. (2023) show the important role of the spatial distribution of dengue vector breeding sites (for example, semi-open water tanks) in explaining differences in infection rates between neighborhoods in the city of Rio de Janeiro.

The objective of this work is to analyze the spatial distribution of dengue worsening notifications in the neighborhoods of the city of Rio de Janeiro using a partitioned ANOVA, using the thermal distribution of the UHI as a discussion reference.

2 Materials and Methods

Cluster analysis, from the English term cluster analysis, is a statistical method that aims to divide the data according to the levels of similarity and difference between the n observations present in the series. For example, it can be applied to understand wind speed patterns in the Northeast region (Santana & Da Silva, 2019), investigate similarities in rainfall patterns for Rio

Grande do Norte in association with sea surface temperature patterns (Amorim et al. 2020), analyze occurrences of droughts in several river basins (Silva & Costa, 2020; Melo & Sousa, 2021), verify the homogeneity in the distribution of outbreaks of fires and fires (Silva et al, 2021), among others. According to Wilks (2019), the points that make up the groups must have small distances between them in relation to the distances between each group. The most widely used distance method is the Euclidean, which represents the shortest distance between two objects in a multidimensional plane (Braz et al. 2020). Cluster analysis can be performed using different separation methods, which may or may not have a hierarchical origin. Two methods that are used in the literature are k-means and hierarchical.

The k-means method has the iterative clustering profile, in which objects are reallocated to the cluster that has the closest clustering mean. Furthermore, the k-means method is considered an unsupervised clustering algorithm, as it generates clusters from predetermined class numbers (Braz et al. 2020). The algorithm depends on a number k of groups that is chosen by the user before implementing the method. One of the ways to carry out this choice of group number is through the intragroup sum of squares. The total number of k-means groups to be used corresponds to the total number of groups in which the difference between the sum of squares of the distance between the object and the central cluster is the smallest possible. After choosing the number k of groups, the following sequence is performed (Wilks, 2019): calculation of the centroids/averages of the vectors of each group; calculation of the distance between the current point and the centroid; if the current point already belongs to the group whose mean is closest, a new distance calculation is made for the next point; if the current point does not belong to this group, it is reallocated to the group with the closest mean. This sequence is performed for all points until they are reallocated in a group that has the closest average of each one of them, reaching the limit of no longer reallocating points to another group.

The hierarchical method, in turn, does not allow the relocation of analysis objects once positioned within a group. As the name implies, in this method there is a predetermined hierarchy of sets of groups, each of which is formed by merging a pair of the collection of previously defined groups (Wilks, 2019).

Hierarchical analysis can be performed in two ways: agglomerative and divisive. In the agglomerative format, each point starts out as its own cluster, and new, similar pairs of clusters are added as you move up the hierarchy. In the divisive format, the opposite happens, that is, all data starts in a single group and is divided into different groups. The divisive format is practically not used in hierarchical groupings because it is highly complex, making its applicability in more extensive databases difficult (Metz, 2006).

In the hierarchical method, the concept of distance between groups is used as a data agglomeration method. There are several ways to calculate this distance. The main ways to calculate the distance between groups cited by Wilks (2019) are: complete/maximum (considers the highest value calculated between the dissimilarities existing between two groups as a criterion for linking the groups), minimum (considers the smallest value calculated between the existing dissimilarities between two groups as a linking criterion between the groups), mean (considers the average of all existing dissimilarities between two groups as a linking criterion between the groups) and centroid (considers the existing dissimilarity between two centroids as a linking criterion between groups).

Analysis of variance (ANOVA) is a statistical method used to test differences between three or more population means, based on analysis of variance between population samples. To carry out the ANOVA, the following conditions must be considered: the samples must be mutually independent, the data distribution must resemble a normal distribution and the populations must have variance values close to each other (Vieira, 2006 apud Moraes Júnior, 2015). This methodology is used in the literature for several purposes, such as: testing probability distributions in rainfall series, studying the possible relationship between sea surface temperature and occurrences of squall lines in the Atlantic Ocean, verifying the possible tendency of temperature increase average in western Pará (Pereira et al., 2014; Silva et al., 2017; Dos Santos Gomes et al., 2015).

The ANOVA method is based on two assumptions. In the first hypothesis, H_0 , it is assumed that all population means ($\mu_1, \mu_2, \mu_3, \dots, \mu_k$) are approximately equal, indicating that there is no differentiation between groups (Moraes Júnior, 2015). The second hypothesis, H_1 , called the alternative hypothesis, assumes that there is a difference between the means of the populations or

at least in part of them. To test which of the two hypotheses is valid for the data to be studied, the F test is performed, which takes into account a pre-established α level of statistical significance. If the F-test results in a value above the rejection limit of H_0 , with a p-value of statistical significance less than α , the null hypothesis is rejected.

2.1 Study Area

The city of Rio de Janeiro is located around 22°54'23"S and 43°10'21"W. To the north it borders several municipalities in the Baixada Fluminense where the Serra do Mar escarpments are located, to the south it is bathed by the Atlantic Ocean, to the east by Guanabara Bay and to the west by Sepetiba Bay. Located in the southeastern region of Brazil, with an environmental area of 1,197 km², with a total population of 6,747,815 (IBGE, 2020).

The municipality is divided into four main geographic zones: north, south, west and centre. Being the second largest population center in Brazil, it is characterized by strong socio-spatial segregation, in addition to the diversity of land use and occupation and different landscapes.

In the absence of synoptic scale disturbances, land-sea breezes, as well as topographical effects (plain-mountain breezes and wind convergences) and urban breezes define the airflow at the diurnal scale (ie, at the mesoscale).

An important aspect for the local climate is the presence of (secondary) Atlantic forest, the influence of the Atlantic coast and its relief with rocky massifs dividing the conurbation (Santos et al. 2020). The tropical climate (hot and humid in summer), with temperatures between 20°C and 27°C. The rainiest period extends from November to April. The maximum temperature can reach 40°C in extreme events (Câmara et al. 2009). It is a transition zone between the tropical climate in the north and the subtropical climate in the south.

In 2020, the neighborhood of Irajá in the City of Rio de Janeiro stood out for being considered the hottest in the city (Altino, 2021). The population of Irajá in 2010 was accounted for in a total of 96,382 inhabitants (IBGE, 2020). The neighborhood is located in the North Zone, an area occupied mainly by residential and commercial areas, with virtually no trees (IPP, 2018).

In addition, it is located in an area that meets the UHI classification criteria, as defined by Lucena & Peres (2017), being considered a hotspot of the UHI.

2.2 Dengue worsening data

Dengue infection notification data will be obtained online through the Municipal Health Department (SMS). They are in the public domain, released from data tables for each year, presenting total data for the municipality, in addition to data separated by neighborhood (<http://www.rio.rj.gov.br/web/sms/exibeconteudo?id=2815389>). The data to be used cover the period from 2000 to 2019, with a weekly interval between the years 2000 to 2010 and from 2015 to 2019. Between the years 2011 to 2014, dengue notifications are available with monthly totals. For the variance analysis of dengue transmission within the municipality, the data obtained by SMS in the period between 2000 and 2019 were used, with the sum of the weekly totals of the years with weekly interval being available so that all years had a monthly interval.

2.3 Statistics and Metrics

This work presents a statistical analysis based on Multiple Linear Regression. Two regressions were performed, the first considered the incidence of dengue in the municipality as a whole as a dependent variable and the incidence per neighborhood as explanatory variables. The second regression considered the mean of neighborhoods using the k-means cluster analysis result as explanatory variables. Only four groups were enough to explain the total variance. Descriptive statistics, linear regressions and partitioned analysis of variance (ANOVA) considered the test result on the null hypothesis (H_0). The following statistics were used in the F test of the null hypothesis: p-value, F-value and R_2 (Wilks, 2019):

- **p-value:** significance defined as the probability of obtaining a test statistic (F-value) with a greater p-value, thus rejecting the null hypothesis (H_0), which considers equal variances and different means in the subsamples, in the case of partitioned ANOVA. Its limit is set here at 0.05. When the p-value obtained by the model is below this limit, it indicates

statistical significance. Larger (smaller) samples produce smaller (larger) p-values. In the RLM, it is used in conjunction with the F-test to check the significance level of the predictor variables.

- **F-value:** defined by the ratio of the variation of the sample means to the within-group variation of the independent variable. The F-value is usually found with ANOVA tables. If it is large, the null hypothesis H_0 can be rejected.
- **R²:** the coefficient of determination indicates to what extent the observed variance is explained by the model variance.

3 Discussion of Results

All variables were scaled (i.e., subtracting the mean and dividing by the standard deviation), prior to statistical analysis and obtaining the results presented below.

Cluster analysis was implemented both for data on absolute values of notifications of worsening by neighborhoods, as well as for the same values weighted by population. The number of groups for the cluster analysis was determined from the analysis of the explanatory variance curve as a function of the number of groups. From this analysis, we selected only 4 groups, which together explained more than 60% of the total variance.

Partitioned analysis of variance (ANOVA) is a type of piecewise ANOVA generally applied to assess the explanatory contribution of subgrouped data (i.e., for all neighborhoods one by one or for the clusters).

Scripts in R were customized for cluster analysis and mapping (with geobr library).

3.1 Clustering and Partitioned ANOVA

Clustering analysis k-means resulted in four groupes (Figure 1). The 3rd group in Figure 1a (indicated in yellow) shows the highest absolute numbers of dengue cases, including the neighborhoods: Bangu (number 145), Campo Grande (148), Santa Cruz (156) and Realengo

(143). These neighborhoods are located in the West Zone of the city, with some of the highest population totals in the municipality.

The 4th group (in purple) has significant numbers of dengue cases, but less than the 3rd group. It is formed by the neighborhoods: Senador Camará (146), Guaratiba (152), Paciência (155), Inhoaíba (150), Santíssimo (147), Pedra de Guaratiba (154), Sepetiba (157) and Complexo do Alemão (66). All these neighborhoods are also located in the West Zone of the city, close to the neighborhoods of the 3rd group. The exception is Complexo do Alemão, made up of favelas located in the North Zone.

The 2nd group (in burgundy) is formed by Taquara (number 125), Tijuca (34), Copacabana (24), Rocinha (32), Jacarepaguá (119), Barra da Tijuca (132), Bonsucesso (number 41), Curicica (122), Irajá (95), Padre Miguel (144), Pavuna (118), Vila Isabel (37), Humaitá (20), Penha (44), Ramos (42), Méier (81), Anil (120), Guadalupe (110), Anchieta (111), Jardim Sulacap (141), Cidade de Deus (129) and Praça Seca (127). Like the 4th group, the 2nd has most neighborhoods with significant totals of dengue cases, especially the neighborhoods located in the leftmost portion of the group's prism. One of the neighborhoods highlighted in this group is the neighborhood of Irajá, which presents a significant variation in the absolute numbers of dengue cases during epidemics.

In the 2nd group, some neighborhoods are in the West Zone (Taquara, Jacarepaguá, Barra da Tijuca, Curicica, Anil, Padre Miguel, Jardim Sulacap, Cidade de Deus and Praça Seca), others belong to the South Zone (Copacabana, Rocinha, Humaitá) and a part in the North Zone (Tijuca, Bonsucesso, Irajá, Pavuna, Vila Isabel, Penha, Ramos, Méier, Guadalupe, Anchieta). With this, it is possible to perceive that when a gradual decrease is observed, the grouping ceases to be centered in the West Zone and starts to include geographically more distant neighborhoods and with other territorial configurations.

The last group obtained to be commented on (Figure 1a) is the 1st group (in blue), composed of the neighborhoods of: Cosmos (number 151), Senador Vasconcelos (149), Jacarezinho (90), Colégio (number 96), Rocha Miranda (105), Engenheiro Leal (100), Brás de Pina (46), Jardim América (50), and others in the North and South Zones of the city, being the later majority in this group. There are also some smaller neighborhoods in the West Zone, such as

Joá (130) and Grumari (137). Group 1 has the lowest absolute totals of dengue in the city of Rio de Janeiro.

On the other hand, when analyzing Figure 1b, based on the analysis of proportional numbers, there is a change in the configuration of the k-means groups. The 2nd Group (in burgundy) is formed by the neighborhoods Campo dos Afonsos (number 140), Zumbi (52), Barra de Guaratiba (153), Camorim (133) and Pedra de Guaratiba (154). The 4th group (in purple) is formed by the neighborhoods Caju (number 4), Anil (120), Curicica (122), Bonsucesso (41) and Saúde (1). Differently, the 3rd group comprises only two neighborhoods: Joá (number 130) and Grumari (137).

Three groups are those with neighborhoods with small population and the highest proportions of dengue per population. The biggest highlight is the 3rd group, which has two of the smallest neighborhoods in relation to the total population, but with the highest proportions of dengue infection. In the neighborhood of Joá, for example, which has a total of 818 inhabitants according to the 2010 Census, a total of 448 people were infected (54.8%) during an epidemic episode in the city of Rio de Janeiro.

Finally, the 1st group (blue) presents all the other neighborhoods in the municipality, which have lower proportions of dengue infection in the total population. This group includes the neighborhoods that show the highest absolute totals in Figure 1a (Campo Grande, Bangu, Santa Cruz and Realengo). Although these neighborhoods have these local maximums of dengue, their population totals are also high, which leads to a smaller proportion in relation to the neighborhoods of the other groups (2nd, 3rd and 4th), but particularly in relation to the 3rd group.

Figure 2 shows the spatial distribution of the four groups obtained with the k-means cluster analysis for each neighborhood in the city of Rio de Janeiro, both for absolute totals (Figure 2a) and for proportionalized totals of dengue per neighborhood (Figure 2b). The blank neighborhoods (Gericinó, Parque Colúmbia and Vasco da Gama) are the neighborhoods that were not included because the total number of notifications of dengue worsening was zero in all years, probably due to underreporting.

From Figure 2a, it is easier to see that the 3rd group, which has the highest total number of dengue cases in the municipality, is concentrated in the West Zone of the city. The 4th group,

which has lower dengue totals when compared to the neighborhoods of the 3rd group, although still significant, is also located in the West Zone. Also composing the 4th group, the Complexo do Alemão neighborhood is located in the North Zone. The 1st and 2nd groups have neighborhoods distributed across the West, North and South Zones. However, the 2nd group has neighborhoods with significant totals of dengue, such as the neighborhood of Irajá.

The West and North zones are the zones with the highest population totals in the city, both with problems such as lack of basic sanitation and restriction in health services (Oliveira, 2019). But one cannot establish a direct relationship between these factors and the high rates of dengue infection a priori, because as the cluster analysis for proportional values indicates, the incidence seems to be generalized throughout the city.

In the West Zone, there is an exception for the neighborhoods of Barra da Tijuca and Recreio dos Bandeirantes, neighborhoods with a population that has greater purchasing power compared to residents of the other neighborhoods that make up the West Zone. The South Zone also presents, in general, a profile of neighborhoods formed by residents of higher social classes. These more favored neighborhoods in the city, even if they have significant amounts of dengue, tend to suffer less from the impacts of the disease due to the existence of better infrastructure and sanitation. These factors hinder the occurrence of spaces where the vector mosquito can use as breeding grounds in its complete life cycle, that is, places where water can accumulate and the mosquito can lay its eggs (Consoli & Lourenço-de-Oliveira, 1994).

Where the population is more concentrated, there is a greater number of potable water tanks, which may have sealing problems, thus serving as breeding grounds for *Aedes aegypti*. Another issue is the removal of urban waste and its disposal, which can be more or less effective depending on the neighborhood.

On the other hand, when analyzing the proportional distribution of dengue (Figure 2), the neighborhoods that were in the groups with the highest absolute totals of dengue (Figure 2a) became the groups with the lowest proportional values (Figure 2b), balancing the proportions across the city. Other groups have particular characteristics such as some neighborhoods in the West and Center Zone, but do not have a large number of inhabitants, such as Joá, Grumari and Saúde and other neighborhoods in the West and North Zones, with slightly higher population

totals, such as Curicica, Barra de Guaratiba , Pedra de Guaratiba and Bonsucesso. Generally, a smaller population indicates a decrease in statistical reliability, in this case, it cannot be stated a priori that the neighborhoods have a different endemic characteristic.

After grouping, the ANOVA was generated in relation to dengue cases in the city of Rio de Janeiro, first with data from all neighborhoods and then using only the normalized means of each group. This was done for both the absolute dengue totals by neighborhood and the dengue totals proportional to population by neighborhood. Thus, four multiple linear regressions were used, two using the values of each neighborhood and the others using k-means groupings as predictor variables.

The statistical confidence of the Linear Regression Models (LRM) was evaluated resulting in the distribution of the F-value by neighborhood (figure 3). The F-value distribution is shown both for the absolute predictors (Figure 3a) and for the rescaled predictors as a function of the population of each neighborhood (Figure 3b). In both cases, the highest F-values of the distribution indicate a high degree of statistical reliability (i.e., above 99%).

The values of the LRM residuals show very small absolute values between -5.438×10^{-16} and 3.424×10^{-16} , giving another indication that the model has a satisfactory fit. The residual standard error of the full RLM (using all neighborhoods) equals 1.56×10^{-16} , which is used in calculating the t-values. In association, R_2 is practically equal to unity. As the *p-value* is much lower than the required limit of 0.05, it can be said that there is statistical significance. Consequently, the null hypothesis H_0 is rejected, i.e., all predictors (once weighted by the population) contribute to the explanation of the total variance. Another way of saying this is that neighborhood population differences can explain most of the differences in numbers of notifications for dengue aggravation.

Therefore, data from the most populous neighborhoods are statistically relevant and contribute to the modeled results. However, in relation to the proportional predictors of the RLM, some neighborhoods do not show statistical significance greater than 90%, probably due to underreporting (e.g., Cidade Nova, Cosme Velho, Ribeira, Praia da Bandeira, Cocotá, Del Castilho, Riachuelo, Água Santa, Cavalcanti, Grumari and Vila Militar) (Figure 3b). Therefore,

caution should be exercised when generalizing the results presented here. Other local factors besides population density may be important and should be investigated.

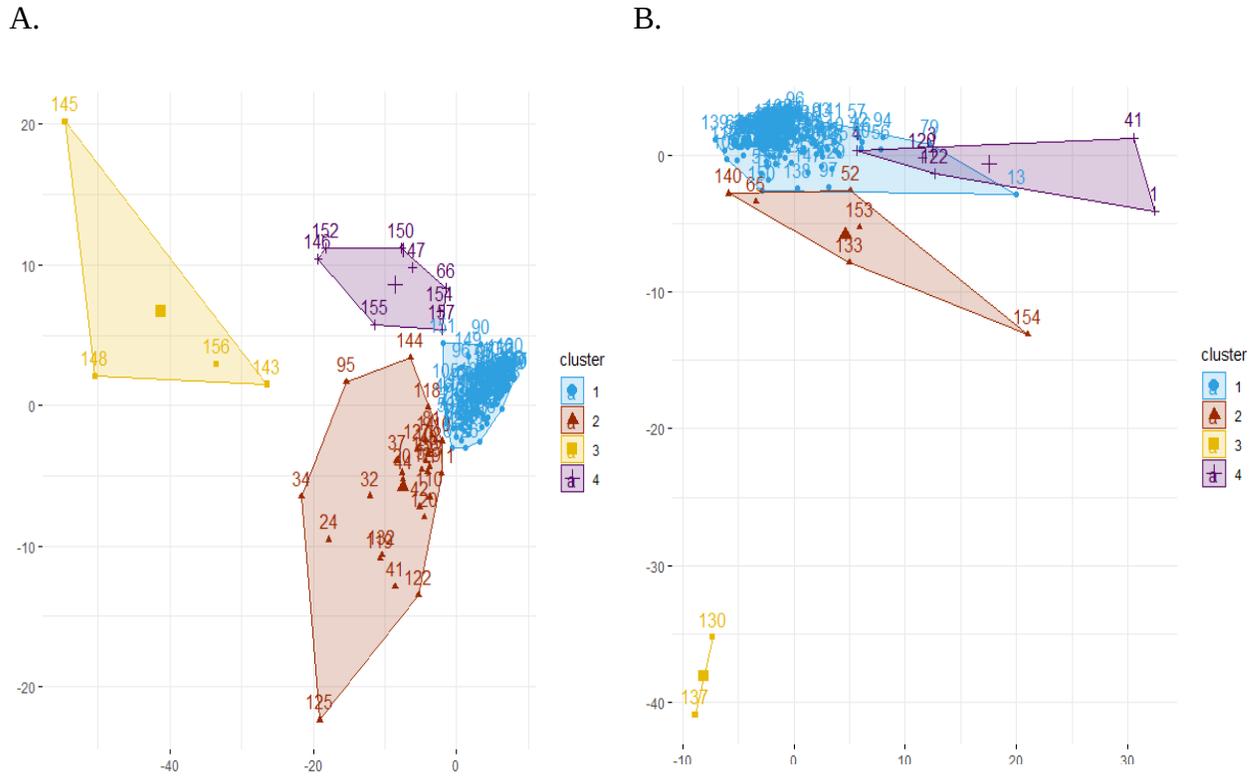


Figure 1 Prisms of clusters in Rio de Janeiro city using k-means method. The numbers indicate different neighborhoods: A. clusters using absolute numbers of dengue worsening notifications; B. Clusters using scaled data by neighborhood population.

3.2 Joint incidence of dengue and ICU

This subsection addresses the analysis of the behavior of dengue in the Irajá neighborhood, chosen as representative of a hot spot in Rio de Janeiro's UHI, as a way of evaluating the relationships between the thermal response and the cases of dengue fever registered. To assist in the analysis of the ICU, the average, maximum and minimum weekly

temperatures of the Irajá neighborhood were extracted for time series of these variables together with the dengue totals in the period from 2015 to 2019.

All thermal variables were normalized for comparison purposes. Between 2015 and 2019, the neighborhood of Irajá began to gain prominence in the high temperature records from the AlertaRio weather stations. It was noticed two dengue epidemics in 2015-2016 and 2019, being the last almost subtle in the normalized series.

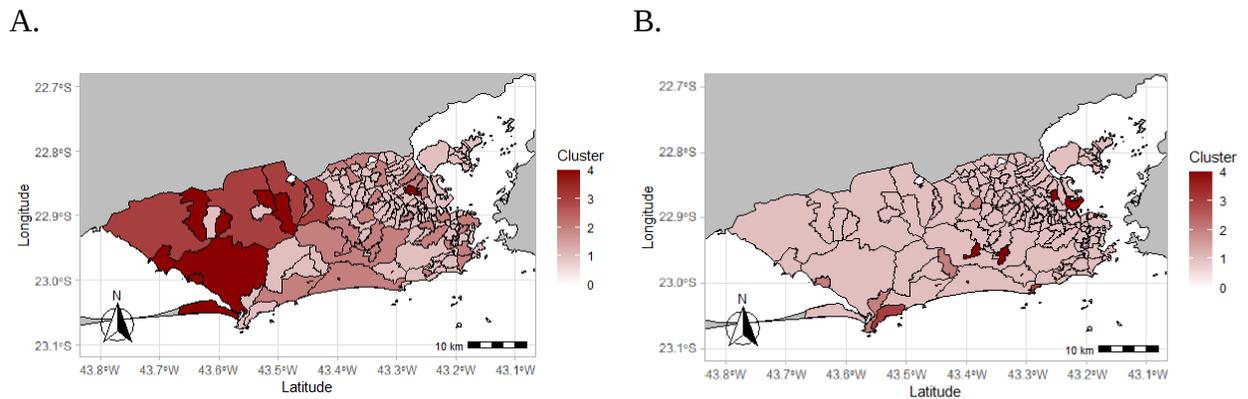


Figure 2 Spatial distribution of neighborhoods through statistical clusters obtained by k-means method: A. distribution to the absolute totals of dengue by neighborhood; B. Distribution to the scaled totals by neighborhood population.

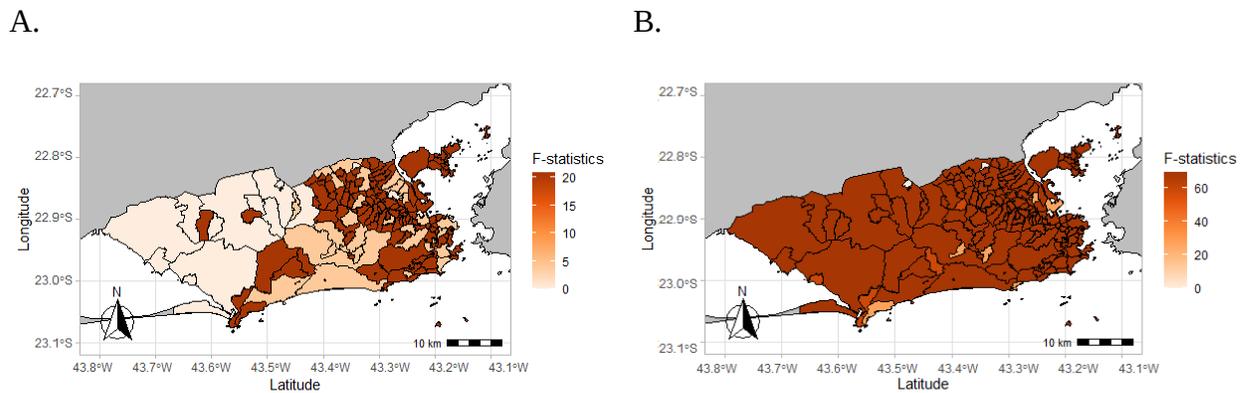


Figure 3 Spatial distribution of the F test obtained for the clusters in the cluster analysis using the k-means method: A. the absolute totals of dengue by neighborhoods; B. the proportionalized totals of dengue by population in each neighborhood, in city of Rio de Janeiro.

In early 2015, there is a slight increase in average and minimum temperatures just before the sudden increase in dengue notifications for Irajá, which peak as temperatures decrease towards mid-year. Shortly after the decrease in cases in 2015, there is an increase in the maximum temperature at the end of the year, followed by a drop in temperatures and a new increase in the average temperature and minimum temperature right at the beginning of 2016. In the sequence, there is a new maximum of dengue worsening notifications. As 2016 progresses, all temperatures (maximum, mean and minimum) decrease, as well as the dengue totals.

Throughout 2017 and 2018, there were some cases of dengue, but not enough for an epidemic outbreak to occur, in addition to the fact that during 2018 there was a decrease in the amplitude of maximum temperatures in relation to other years, as well as a slightly larger decrease in maximum temperature in mid-2018 compared to the same period in other years.

At the turn of 2018 to 2019, there is a new increase in average and minimum temperatures. A few weeks later, the maximum temperature rises and starts to decrease once again along with the average and minimum temperatures. A few weeks after these peaks, there is a dengue maximum of smaller proportions in comparison to 2015 and 2016, accompanied by a slight increase in temperatures.

On the other hand, the average weekly temperatures observed in the Irajá neighborhood (corresponding to a hotspot in the Urban Heat Island of Rio de Janeiro) were compared with reports of worsening dengue. Increase in daily maximum (minimum) temperature seem to precede the increase in reports of worsening by dengue in a week, during epidemics (endemics).

In 2015, 2016 and 2019, dengue epidemics occurred in the city of Rio de Janeiro, which totaled 17,451 cases, 23,946 cases and 14,343 cases, respectively. These epidemics have a smaller magnitude compared to other epidemics that occurred in the locality (i.e., 2002 and 2008). For example, in 2002, a number of 139,400 cases were officially registered; in 2008 there were a total of 120,440 registrations; in 2012, there were 50,184 cases of dengue worsening were recorded throughout the municipality.

After the occurrence of two large epidemics, several actions to combat dengue were published in Rio de Janeiro, which may have influenced the gradual reduction of dengue totals

during the epidemics that occurred from 2012 onwards in the municipality. For example, the document “National Guidelines for Prevention and Control of Dengue Epidemics” was released by the Ministry of Health (Brasil, 2009), with surveillance and dengue control actions in Brazil to be applied in states and municipalities. Additionally, the El Niño-La Niña Southern Oscillation (ENSO) was out of phase in relation of Dengue epidemics period of 3 to 4 years in the city of Rio de Janeiro (Santos et al. 2022).

Local government actions were carried out and implemented in Rio de Janeiro with base in the Municipal Decree nº 34.377/2011 of 31-Aug-2011, which defined the state of alert against dengue and stated need actions for the prevention and control of the disease (Rio de Janeiro-RJ, 2011). In addition, a system for monitoring dengue occurrences in the municipality was launched in a joint effort of Oswaldo Cruz Foundation and Getúlio Vargas Foundation (FioCruz 2015).

Briefly bringing up the issue of climate variability, the 2015 and 2016 epidemics occurred during El Niño episodes (Santos et al. 2022), as well as the 2019 epidemic, according to ENSO indexes published by NOAA (<https://www.esrl.noaa.gov/psd/enso/dashboard.html>). Occurrences of high temperatures were recorded, including breaking the records of the year 2015 during the spring in the Metropolitan Area of Rio de Janeiro (PBMC 2015) and high temperatures during the summer, with the highest values of maximum during the end of autumn and winter in the neighborhood of Irajá (AlertaRio, 2019).

4. Conclusions

Regarding the contributions of the neighborhoods to the total number of dengue cases in the municipality, the cluster analysis shows that the neighborhoods with the highest absolute values are located, for the most part, in neighborhoods in the West and North Zones, which also have large population totals. After an initial linear regression by neighborhood, ANOVA were performed for all neighborhoods. In fact, neighborhoods could be divided in four explanatory groups of total variance, each at a confidence level above 90%. The proportional incidence of dengue seem well distributed among the population of the neighborhoods of Rio de Janeiro

(characterizing endemic conditions), despite the thermal structure of the urban heat island being heterogeneous in space and also showing a daily and seasonal cycle.

For the future development of the work, it should be noted that the UHI is associated with the Urban Boundary Layer, which appears as a response to the surface changes that urbanization causes in the energy and water balance. The flow of breezes (sea-land, urban and plain-mountain) in the city of Rio de Janeiro can be associated with advection of the dengue vector, leading to similar incidence rates throughout the city. This supports the hypothesis that hot spots in the UHI may be associated with greater vector reproductive capacity, while breeze recirculation may increase vector dispersal.

References

- Amorim, A. C. B., Lucio, P. S., Silva, C. M. S., Rodrigues, D. T. & Brito Neto, F. A. D. 2020, 'Regionalização e análise da tendência da precipitação do Rio Grande do Norte associados a padrões de TSM', *Revista Brasileira de Meteorologia*, vol. 35, pp. 269–280.
- Alexandrino, F. I. G., 2017, 'Ilhas de calor e casos de dengue na área urbana de Taubaté-SP', Master Thesis, Universidade Estadual Paulista.
- Altino, L. 2021, *Entenda como Irajá se tornou o bairro mais quente do Rio*, Extra, Rio de Janeiro, 29 Jan 2021, viewed 20 Apr. 2021, <<https://extra.globo.com/noticias/rio/entenda-como-iraja-se-tornou-bairro-mais-quente-do-rio-24859930.html>>.
- Araujo, R.V., Albertini, M.R., Costa-da-Silva, A.L., Suesdek, L., Franceschi, N.C.S., Bastos, N.M., Katz, G., Cardoso, V.A., Castro, B.C., Capurro, M.L. & Allegro, V.L.A.C. 2015, 'São Paulo urban heat islands have a higher incidence of dengue than other urban areas', *The Brazilian Journal of Infectious Diseases*, São Paulo, vol. 19, pp. 146–155.
- Baracho, R.C.M. 2013, 'Influência de variáveis meteorológicas sobre a incidência dos casos de dengue no município de Areia-PB', TCC, Universidade Federal da Paraíba.
- Beserra, E.B., Junior, F.P.C, Santos, T.S. & Fernandes, C.R.M. 2006, 'Biologia e exigências térmicas de *Aedes aegypti* (L.) (Díptera: culicidae) provenientes de quatro regiões bioclimáticas da Paraíba', *Revista Neotropical Entomology*, vol. 35, no. 6, pp. 853–860.

- Beserra, E.B., Fernandes, C.R., Silva, S.A.D.O., Silva, L.A.D & Santos, J.W.D. 2009, 'Efeitos da temperatura no ciclo de vida, exigências térmicas e estimativas do número de gerações anuais de *Aedes aegypti* (Diptera, Culicidae)', *Iheringia, Série Zoologia*, vol. 99, pp. 142–148.
- Braga, I.A. & Valle, D. 2007, '*Aedes aegypti*: histórico do controle no Brasil', *Epidemiologia e Serviços de Saúde*, vol. 16, no. 2, pp. 113–118.
- Brasil 2009, *Diretrizes Nacionais para a Prevenção e Controle de Epidemias de Dengue* Ministério da Saúde, Secretaria de Vigilância em Saúde, Brasília.
- Braz, A.M., Oliveira, I.J.D., Cavalcanti, L.C.D.S., Almeida, A.C.D. & Chávez, E.S. 2020, 'Análise de agrupamento (cluster) para tipologia de paisagens', *Mercator (Fortaleza)*, vol. 19.
- Câmara, F.P., Gomes, A.F., Santos, G.T.D. & Câmara, D.C.P. 2009, 'Clima e epidemias de dengue no Estado do Rio de Janeiro', *Revista da Sociedade Brasileira de Medicina Tropical*, vol. 42, no. 2, pp. 137–140.
- Cardoso, R.D.S., Dorigon, L. P., Teixeira, D.C.F. & Amorim, M.C.D.C.T. 2017, 'Assessment of urban heat islands in small-and mid-sized cities in Brazil', *Climate*, vol. 5, p. 14.
- Cardoso, R.D.S. & Amorim, M.C.D.C.T. 2018, 'Urban heat island analysis using the 'local climate zone scheme in Presidente Prudente, Brazil', *Investigaciones Geográficas*, vol. 69, no 1-6, pp. 107–118.
- Consoli, A.G.B.R. & Lourenço-de-Oliveira, R. 1994, *Principais mosquitos de importância sanitária no Brasil*. Rio de Janeiro: Fiocruz.
- Dos Santos Gomes, A.C., Da Silva Costa, M., Coutinho, M.D.L., Do Vale, R. S., Dos Santos, M.S., Da Silva, J.T. & Fitzjarrald, D.R. 2015, 'Análise estatística das tendências de elevação nas séries de temperatura média máxima na Amazônia Central: Estudo de caso para a região do Oeste do Pará', *Revista Brasileira de Climatologia*, vol. 17., no. 7-12, pp. 82–97.
- Farnesi, L.C., Martins, A.J., Valle, D. and Rezende, G.L., 2009, 'Embryonic development of *Aedes aegypti* (Diptera: Culicidae): influence of different constant temperatures', *Memórias do Instituto Oswaldo Cruz*, vol. 104, pp. 124-126.
- Fiocruz 2015, *FGV e Fiocruz lançam sistema de alerta para monitorar dengue no Rio de Janeiro*, Fundação Oswaldo Cruz, viewed 10 Jan 2023, <<https://www.canalsaude.fiocruz.br/noticias>>

- Gomes, A.F., Nobre, A.A. & Cruz, O.G. 2012, 'Temporal analysis of the relationship between dengue and meteorological variables in the city of Rio de Janeiro, Brazil, 2001-2009', *Cadernos de Saúde Pública*, vol. 28, no. 11, p. 2189–2197.
- IBGE, 2010. Instituto Brasileiro de Geografia e Estatística, Censo demográfico 2010, viewed 02 Aug 2021, <https://biblioteca.ibge.gov.br/visualizacao/periodicos/93cd_2010_caracteristicas_/populacao_domicilios.pdf>.
- Knoblauch, S., Li, H., Lautenbach, S., Elshiaty, Y., Rocha, A.A.D.A., Resch, B., Arifi, D., Jänisch, T., Morales, I. and Zipf, A. 2023, 'Semi-supervised water tank detection to support vector control of emerging infectious diseases transmitted by *Aedes Aegypti*', *International Journal of Applied Earth Observation and Geoinformation*, vol. 119, p. 103304.
- Lemos, L.D.O., Oscar Júnior, A.C. & Assis Mendonça, F.D. 2021, 'Urban climate maps as a public health tool for urban planning: The case of dengue fever in Rio De Janeiro/Brazil', *Urban Climate*, vol. 35:100749, doi: 10.1016/j.uclim.2020.100749.
- Lima-Câmara, T.N. 2016, 'Emerging arboviruses and public health challenges in Brazil', *Revista de Saúde Pública*, vol. 50, no. 36, <https://doi.org/10.1590/S1518-8787.2016050006791>.
- Lucena, A.J., Rotunno Filho, O.C., França, J.R.A., Peres, L.F. & Xavier, L.N.R. 2013, 'Urban climate and clues of heat island events in the Metropolitan area of Rio de Janeiro', *Theor. and Appl. Climat.*, vol. 111, no. 3, pp. 497–511.
- Lucena, A.J. & Peres, L.F. 2017, 'Métodos em clima urbano aplicados à cidade do Rio de Janeiro (Brasil) e sua região metropolitana', *The Overarching Issues of the European Space: Society, Economy and Heritage in a Scenario. Porto: FLUP*, pp. 312–326.
- Melo, V.S. & Sousa, F.A.S. 2021, 'Análise de frequência de secas utilizando técnica de agrupamento e distribuições de probabilidades', *Engenharia Sanitaria e Ambiental*, vol. 26, p. 53-60.
- Metz, J. 2006, 'Interpretação de clusters gerados por algoritmos de clustering hierárquico', Doctoral dissertation', Universidade de São Paulo.
- Miagostovich, M.P., Nogueira, R.M., Cavalcanti, S., Marzochi, K.B. & Schatzmayr, H.G. 1993, 'Dengue epidemic in the state of Rio de Janeiro, Brazil: virological and epidemiological aspects' *Revista do Instituto de Medicina Tropical de São Paulo*, vol. 35, no. 2, pp. 149–154.

- Monteiro, F.F., Gonçalves, W.A., Andrade, L.D.M.B., Villavicencio, L.M.M. & Dos Santos, C.M. 2021, 'Assessment of Urban Heat Islands in Brazil based on MODIS remote sensing data', *Urban Climate*, vol. 35, p.100726.
- Morais Júnior, R.R. 2015, 'Estudo do desempenho da soldagem por eletrodo revestido em diferentes posições por meio de análise de variância', Trabalho de Conclusão de Curso, Universidade Federal do Pampa.
- Mordecai, E.A., Cohen, J.M., Evans, M.V., Gudapati, P., Johnson, L.R., Lippi, C.A., Miazgowicz, K., Murdock, C.C., Rohr, J.R., Ryan, S.J. & Savage, V. 2017, 'Detecting the impact of temperature on transmission of Zika, dengue, and chikungunya using mechanistic models', *PLoS neglected tropical diseases*, vol. 11, no. 4, p.e0005568.
- Oke, T. R., Mills, G., Christen, A. & Voogt, J. A. 2017, *Urban climates*. Cambridge Univ. Press.
- Oliveira, R. F. 2019, 'O papel do clima na ocorrência da Chikungunya, Dengue e Zika entre 2007 e 2017 nos bairros de Copacabana, Guaratiba, Jardim Botânico e Santa Cruz da cidade do Rio de Janeiro'. Masters Thesis, Universidade do Estado do Rio de Janeiro (UERJ).
- PBMC 2022, *Rio registra terceira maior temperatura em 100 anos, diz Climatempo*. Painel Brasileiro de Mudanças Climáticas. Viewed 10 Apr 2023, <<http://pbmc.coppe.ufrj.br/index.php/en/news/>>.
- Pereira, M.A.F, Kobiyama, M. & Castro, N.M.R. 2014, 'Análise de variâncias pluviométricas na bacia hidrográfica do rio Taboão-RS', *Revista Brasileira de Meteorologia*, vol. 29, pp. 409–419.
- Rio de Janeiro 2011, Decreto Municipal nº 34377/2011, de 31 de agosto de 2011. Institui estado de alerta contra a dengue e dispõe a prevenção e o controle da transmissão e a atenção primária à saúde nos casos de dengue na Cidade do Rio de Janeiro e dá outras providências.
- Santana, L.V.R. & Da Silva, A.S.A. 2019, 'Análise de agrupamento da velocidade do vento no Nordeste do Brasil', *Sigmae*, vol. 8, no. 2, pp. 481–489.
- Santos, J.P.C., Honório, N.A., Barcellos, C. & Nobre, A. 2020, 'A perspective on inhabited urban space: land use and occupation, heat islands, and precarious urbanization as determinants of territorial receptivity to dengue in the city of Rio de Janeiro', *International Journal of Environmental Research and Public Health*, vol. 17, no. 18, pp. 6537.

- Santos, S.A.F., Karam, H.A., José, A., Rojas, J.L.F., Suazo, J.M.A., Vásquez, I.L. & Peña, C.A.S. 2022, 'Dengue Climate Variability in Rio de Janeiro City with Cross-Wavelet Transform', *Journal of Environmental Protection*, vol. 13, no. 3, pp. 261–276.
- Silva, M.T., Alcântara, C.R., Souza, E.P.D., Olinda, R.A.D. & Gonçalves, W.A. 2017, 'Influência da temperatura da superfície do mar na ocorrência de linhas de instabilidade na costa Norte e Nordeste do Brasil', *Revista Brasileira de Meteorologia*, vol. 32, pp. 260–268.
- Silva, A.F.A. & Costa, V.A.F. 2020, 'Caracterização espacial de secas meteorológicas e hidrológicas na porção mineira da bacia do São Francisco', *III Simpósio da Bacia Hidrográfica do Rio São Francisco*, Belo Horizonte-MG, Brasil.
- Silva, E.M.D.S., Correia, W.L.F., Oliveira, J.F.D., Barros, H.G., Costa, M.D.S., Gois, G.D. & Falcão, N.A.D.M. 2021, 'Espaço-temporalidade dos focos de calor na região metropolitana de Maceió', *Revista Brasileira de Meteorologia*, vol. 35, pp. 1029–1043.
- Xavier, D.R., Magalhães, M.D.A.F.M., Gracie, R., Reis, I.C.D., Matos, V.P.D. & Barcellos, C. 2017, 'Difusão espaço-tempo do dengue no Município do Rio de Janeiro, Brasil, no período de 2000-2013', *Cadernos de Saúde Pública*, vol. 33, no. 2 (:e00186615), pp. 1–13.
- Wilks, D. 2019, *Statistical Methods in the Atmospheric Sciences*, 4th Edition, Elsevier, 840 p.
- WMO 2009, *Dengue: Guidelines for Diagnosis, Treatment, Prevention and Control*, World Health Organization, pp. 1–160.